ISSN 1726-5479

SENSORS & TRANSDUCERS

vol. 259
5/22

Sensors: Technologies and Applications

International Frequency Sensor Association Publishing

IFSA

Sensors & Transducers

International Official Open Access Journal of the International Frequency Sensor Association (IFSA) Devoted to Research and Development of Sensors and Transducers

Volume 259, Issue 5, October 2022

IFSA Publishing: Barcelona • Toronto

Published by International Frequency Sensor Association (IFSA) Publishing. Printed in the USA.

Sensors & Transducers

Volume 259, Issue 5
October 2022

www.sensorsportal.com

e-ISSN 1726-5479
ISSN 2306-8515

Sensors & Transducers journal is an open access, peer review international journal published monthly by International Frequency Sensor Association (IFSA).
Available in both: print and electronic (printable pdf) formats.

Sensors & Transducers Journal

Contents

Volume 259
Issue 5
October 2022

www.sensorsportal.com

ISSN 2306-8515
e-ISSN 1726-5479

Research Articles

Authors are encouraged to submit article in MS Word (doc) and Acrobat (pdf) formats
by e-mail: editor@sensorsportal.com. Please visit journal's webpage with preparation instructions:
http://www.sensorsportal.com/HTML/DIGEST/Submition.htm

International Frequency Sensor Association (IFSA).

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 1-11

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Capacitance: The Transducer that Converts Resistors into Electromagnetic Sensors

José-Ignacio IZPURA
Dept. of Aerospace Systems, Air Transport & Airports, Universidad Politécnica de Madrid (UPM)
ETS Ingeniería Aeronáutica y del Espacio (ETSIAE),
Plaza Cardenal Cisneros 3, 28040, Madrid, Spain
* E-mail: joseignacio.izpura@upm.es

Received: 5 September 2022 /Accepted: 10 October 2022 /Published: 31 October 2022

Abstract: It is widely known since 1928 that each resistor at temperature T shows a voltage noise of spectral density 4kTR V^2/Hz, where k is the Boltzmann constant. However, the way in which this voltage arises from the exchanges of energy of this device with its thermal bath remains still unclear. Leaving aside strange features ascribed to its resistance R in order to justify how this random voltage is generated, we have considered for this purpose the electrical capacitance C of each resistor of volume V_Q sandwiched between its two equipotential terminals (plates). Using a new, Fluctuation-based noise model that excels the Dissipation-based model in use today, we have shown that Johnson noise is the measurable effect caused by truly impulsive charge noise that takes place in this reactive element with a mean square value of 4kT/R Coulombs^2/second. Hence, capacitance is the transducer that converts thermal fluctuations of electric field within V_Q into this random voltage called Johnson noise. The C^2/s units that emerge from this advanced noise model for the familiar value of the Nyquist density 4kT/R A^2/Hz and its partition into 2kT/R ultra wideband and 2kT/R band-limited noises, are two new results of this advanced model.

Keywords: Charge noise and fluctuations, Conduction current and dissipation, Displacement currents and shot noise, Noise out of thermal equilibrium, Two-terminal devices and complex immitances.

1. Introduction

Even though the voltage noise between terminals of any resistor (i. e. its Johnson noise) and its thermal origin are well-known since the pioneering works of J. B. Johnson [1] and H. Nyquist [2], the noise model developed from their results that is currently in use should be improved. We refer to the dissipation-based model (DBM) where capacitors and inductors do not generate noise. Concerning resistors this model has led to some myths like "Ohmic" resistances taken as noisy whereas dynamical ones are taken as noiseless without any proof. This strange situation in regard to the uniqueness of the magnitude resistance is well summarized by the voltage noise given by the circuit simulator PSPICE for a Schottky diode in thermal equilibrium (TE) at room temperature for example.

People wishing measurements to compare with the simulations could consider that the measurement of the voltage noise of a broad area Schottky diode (e. g. a power diode to have a low dynamical resistance r_d in the MΩ range for the diode in TE) is at the reach of today's instrumentation. This rather low $r_d \approx 1$ MΩ for the diode in TE at room T facilitates the measurement of its voltage noise by a typical Low Noise Amplifier (LNA) of much higher input resistance. If $R_{in} \geq 20$ MΩ the noise contribution of R_{in} becomes much lower than that of this diode under open circuit conditions. This is

so because the noise density S_{Imeas}=(4kT/R_{in}+ 4kT/r_d) A^2/Hz at the input of the LNA is very close to that of the diode without R_{in} (S_{Idiode}=4kT/r_d see below).

For R_{in}=20 MΩ and in the low frequency region, the flat density of voltage noise we would measure is: S_{Vmeas}=4kT×(R_{in}×r_d)/(R_{in}+r_d) V^2/Hz that is 5 % lower than the S_{Vdiode} we are looking for. This departure that comes from the power divider created by R_{in} and r_d connected in parallel ($R_{in}||r_d$), is unavoidable but easy to consider. More important is the fact that the noise power measured would have one part coming from the resistor of R_{in} ohms (device1) and twenty parts coming from the diode (device2) given the null correlation of these two noises that are generated in devices whose volumes are in different regions of space. Hence, the voltage noise we would measure is that of the diode "slightly contaminated by noise of R_{in}".

Concerning the voltage noise given by PSPICE for these two devices, we should simulate the voltage noise between terminals of our diode shunted by R_{in} and shunted also by a dc current source I_{dc}. This source driving the diode allows the presence of the fluctuating voltage that we want to simulate between terminals. Using a low bias current like $I_{dc}=10^{-25}$ A is a way to benefit from the above feature at the cost of adding $S_{Ishot}=2qI_{dc}\approx3\times10^{-44}$ A^2/Hz to S_{Idiode}=4kT/r_d= $1{,}7\times10^{-26}$ A^2/Hz, thus an irrelevant effect. Note that q is the electronic charge. Strictly speaking this $I_{dc}\neq0$ puts the diode "slightly out of TE", but the voltage offset it produces between its terminals that is: $\Delta V_{dc}\approx\pm(I_{dc}\times r_d)$ volts, where the sign depends on the forward/reverse sense of I_{dc}, becomes irrelevant. Since this offset only is: $\Delta V_{dc}=\pm10^{-19}$ volts, we can assume that, for practical purposes, we are simulating the voltage noise of these two devices "both in TE" at room T.

Running PSPICE we could check all the above, but we would see that the density of voltage noise S_{Vsimul} it would give would not be S_{Vmeas}, but twenty-one times lower: S_{Vmeas}=4kT×($R_{in}||r_d$)/21. This low value given by PSPICE appears because PSPICE takes R_{in} as "Ohmic" and ascribes a density 4kT/R_{in} A^2/Hz to it, but concerning r_d (that PSPICE calculates perfectly, see below) it is taken as noiseless. Due to this, the noise that PSPICE uses is: S_{Isimul}=(4kT/R_{in}+0) =S_{Imeas}/21 and this explains why it is unable to give the voltage noise of the diode in TE. The right value of the r_d that PSPICE obtains is observed in the cut-off frequency of the low-pass filtering action that it shows perfectly for the small noise S_{Vsimul}=S_{Vmeas}/21 V^2/Hz that it gives. This filtering is due to the junction capacitance C_j of the diode that shunts the resistance ($R_{in}||r_d$).

Despite this filtering action suggesting that r_d is "actual resistance" (i. e. *the ratio between sinusoidal voltage and current amplitudes found in phase in the circuit at each frequency*) PSPICE considers r_d as if it was a "different" noiseless resistance. Due to this myth that flaws PSPICE, the noise it gives is not enough "to keep thermal equipartition in C_j", which is a notion of TE that should prevail (see below) over myths like "different resistances depending on the technology of the resistors where they come from". The magnitude "resistance", defined uniquely by the first sentence in italics of this paragraph, does not give room for two types of resistances. Hence, the striking "notch in TE" that S_{Vsimul} shows around $\Delta V_{dc}\approx0$ when the sign and the magnitude of I_{dc} is varied to sweep between ±100 mV the dc voltage of the diode, is the "signature" of the aforementioned myth flawing the DBM undergoing this widely used circuit simulator.

For readers wishing to check PSPICE concerning voltage noise in diodes we shall give some numerical data after saying that PSPICE works splendidly for reverse voltages below four thermal voltages V_T. This means that diode voltages are: $\Delta V_{dc}=V\leq$-100 mV at room T ($V_T\approx$25.9 mV at T=300 K). And moreover: as the reverse bias of the diode is higher, the accuracy increases. Though accordingly to the DBM, capacitors do not give noise, we shall consider C_j=100 pF as the depletion capacitance of this diode in TE at room T and whose I-V characteristic will be this one:

$$I = I_{sat} \times \left[\exp\left(\frac{V}{mV_T}\right) - 1\right], \qquad (1)$$

Taking a realistic ideality factor m=1.05 to show what its effect is and $I_{sat}=2.7\times10^{-8}$ A at room T to have $r_d\approx1$ MΩ in TE, the dynamical resistance would be:

$$r_d(V) = \left[\frac{dI}{dV}\right]^{-1} = \left[\frac{I_{sat}}{mV_T}\exp\left(\frac{V}{mV_T}\right)\right]^{-1}, \qquad (2)$$

Eq. (1) with these values of m and I_{sat} for the diode plus the capacitance C_j=100 pF in parallel, would form the model we need for simulations. Since the voltage of the diode in TE is V=0, we obtain:

$$r_d(0) = r_d = \frac{mV_T}{I_{sat}} = \frac{1.05\times0.0259}{2.7\times10^{-8}} = 1M\Omega, \qquad (3)$$

In this way we could simulate the voltage noise of our diode to compare it with measured values (e.g. in a commercial power Schottky diode to get a low r_d) or with values that we can justify theoretically from well-known basis. In any case, this disappointing "state of affairs" concerning the futility of the DBM for noise in junction diodes led us to develop a new noise model from a different basis. We mean a fluctuation-based model (FBM) that we developed to overcome the limitations of the DBM (that strictly speaking is valid only in TE) and to remove misconceptions and myths on this subject that still remain in regard to the voltage noise that we measure in two-terminal devices (2 TDs) like resistors, capacitors or L-C resonators.

Thus, the notion of a 2 TD is essential and in regard to measurements, the magnitude we can measure is a random voltage of spectral density in V^2/Hz, thus not random current that is never measured directly but inferred from a voltage caused by such current. The unawareness of this fact shows that voltage noise is not well-known today, despite the Fluctuation-Dissipation (F-D) framework that H. B. Callen and T. A. Welton

proposed from a quantum physics approach for noisy processes in 1951 [3].

From a circuit point of view, the F-D framework means that conductance or resistance for dissipation of electrical energy, as well as reactance or susceptance to hold this energy that fluctuates, are both required to model a noisy device. When this need of a complex immitance to study electrical noise is accepted, the kT/C noise of capacitors and the Johnson noise of resistors become the same type of voltage noise that can be studied by a common fluctuation-based model (FBM) for voltage noise. It is worth noting that kT/C V^2 is the integral in frequency *f* of the Johnson noise of the resistor that any capacitor becomes as $f \rightarrow 0$ [4].

Resistors and capacitors are 2TDs that differ from inductors, whose two terminals vanish when they are shorted to hold magnetic energy liable to fluctuate within them. By contrast, the energy that fluctuates and is dissipated in a 2TD is electrical energy *U(t)* defined by its voltage *v(t)* at each instant of time and its capacitance C between terminals. Using *Q(t)* for the time-varying charge in one of the plates of C, this fluctuating energy is:

$$U(t) = \frac{1}{2} C v^2(t) = \frac{Q^2(t)}{2C}, \qquad (4)$$

Hence, Johnson noise *v(t)* reflects the energy *U(t)* that exists at any instant of time in the capacitance $C=(C_{str}+C_{mat})$ between terminals of a resistor. This is why C includes any stray value due to wiring (C_{str}) and the inherent term C_{mat} due to the non-null dielectric permittivity of its material between terminals [5]. The FBM for electrical noise that we refer to can be found in [6] and to show its starting point we shall consider the possible shot noise of our Schottky diode in TE.

2. Using Shot Noise to Remove a Myth

Eq. (1) states that our Schottky diode in TE is a 2TD where two currents, each of mean value I_{sat}, run with opposed senses. Hence, they cancel one another on average, but not at each instant of time because they come from different regions of the device. This means that they are uncorrelated and their associated effects like their shot noises will add in power. Let us do this sum manually because to calculate shot noise in a 2TD PSPICE uses its net current, thus being useless for our diode whose net current in TE is null: $I_{dc}=(I_{sat}-I_{sat})=0$.

Then, the total shot noise density of our diode in TE will be: $S_{Ishot}=2qI_{sat}+2qI_{sat}=4qI_{sat}$ A^2/Hz. Since the magnitude we measure is not a current noise density like this but its effect as voltage between terminals, we must obtain the spectral density S_{Vdiode} of voltage noise (in V^2/Hz) due to S_{Ishot}. Multiplying S_{Ishot} by $(r_d)^2$ we obtain this spectral density of voltage noise:

$$S_{Vdiode} = 4qI_{sat} \times (r_d)^2 = m4kTr_d \qquad (5)$$

Since m=1.05 we have: $S_{Vdiode} \approx 4kTr_d$ V^2/Hz, that is the noise density that PSPICE is unable to give due to the myth about electrical noise giving rise to two types of resistances. Hence, removing from PSPICE (i. e. from the DBM) this harmful myth, our knowledge about electrical noise will improve. Added to the above, we observe that working with shot noise linked with the displacement currents of single electrons in a capacitive device can be a good way to obtain cogent results in regard to electrical noise.

Considering that shot noise comes from electrons that are displaced between the terminals of a 2TD, we envisaged years ago that displacement currents (like those I_{sat} due to electrons crossing the depleted region of our Schottky diode) could give rise to the voltage noise of 2TDs like resistors and capacitors. It is worth noting that our Schottky diode in TE is both a resistor and a capacitor. From the cut-off frequency of its r_d-C_j parallel circuit: $f_c=1/(2\pi r_d C_j) \approx 1.6$ kHz, we can say that for $f<160$ Hz, our diode is a quite good resistor of 1 MΩ and for $f>16$ kHz it is a quite good capacitor of 100 pF.

The above paragraph shows the way we envisaged the FBM published in [6] that we had used previously to explain electronic noises out of TE like the flicker noise of vacuum devices [4] and the $1/f$ resistance noise of solid-state devices [5]. Concerning the density $S_{Vdiode}=4kTr_d$ V^2/Hz we have given as the true voltage noise of the diode, we will say that this value is its flat amplitude below f_c. This is so because the truly flat density of shot noise $S_{Ishot}=4qI_{sat}$ A^2/Hz is low-pass filtered by the r_d-C_j parallel circuit of the diode to give this Lorentzian spectrum of voltage noise:

$$S_{Vdiode}(f) = \frac{4kTr_d}{1 + \left(\frac{f}{f_c}\right)^2} \qquad (6)$$

Since Eq. (6) means that "we are keeping thermal equipartition in C_j" because its integral from $f \rightarrow 0$ to $f \rightarrow \infty$ gives the mean square voltage noise kT/C_j V^2 [6], we have said confidently that the S_{Vsimul} of PSPICE for our Schottky diode in TE was untrue.

The reason why PSPICE works splendidly when our diode has reverse voltages below V = -100 mV is because in this case, the net current of the diode tends to be: $I_{dc} \approx -I_{sat}$. From Eq. (1), the condition $I_{dc}=-0{,}99I_{sat}$ is found for V=-125 mV. This means that PSPICE will take $0.99I_{sat}$ as the net current it uses to obtain the shot noise density of the diode as $S_{Ishot}=2q \times 0{,}99I_{sat}$ A^2/Hz. Since $I_{dc}=-0.99\ I_{sat}=I_{sat} \times (0.01-1)$, the two displacement currents giving shot noise in this case are: the reverse saturation current -Isat and a fraction of the forward current I_{sat} that has been weakened down to $0{,}01I_{sat}$ by the reverse voltage (V=-125 mV) of the diode biased by $I_{dc}=-0{,}99I_{sat}$. The relative error for the amount of shot noise is only: (1.01-0.99)/1.01≈2 %.

More striking, however, is the exponential change of the resistance of the diode $r_d(V)$ with the voltage V set by its bias current I_{dc}. We mean Eq. (2) that works perfectly in PSPICE showing that the $r_d(V=0) \approx 1$ MΩ

of the diode in TE, now is: r_d(V=-125 mV)=4.6 MΩ. Since C_j=100 pF does not vary, the cut off frequency of the diode with I_{dc}=-0.99I_{sat} (thus with V=-125 mV) is 4.6 times lower f_c≈350 Hz. All in all, PSPICE would be showing us that the mean square voltage noise of our diode out of TE is half its value in TE. This is so because this noise comes from a Lorentzian spectrum with 4.6 lower bandwidth and only 2.3 times higher amplitude than these two values of the Lorentzian noise spectrum of the diode in TE (V=0 V).

Previous paragraph not only shows the basis of the noise "tunability" we used in [5] to show the synthesis of the 1/*f* spectrum of resistance noise in solid-state devices, but also the suitability of shot noise to handle the voltage noise of 2TDs that are out of TE. Besides, it reinforces the idea of shot noise as a good starting point for a new noise like our FBM. The key role of the capacitance C_j that we have shown in this Section and 4kT/R C^2/s of mean square charge noise given in the Abstract, should pave the way for the new noise model that we discuss in the next Section.

3. A Fluctuation-Dissipation Noise Model

Fig. 1-a sketches a resistor made from a rod of conductive material with two tinned iron caps where its two access wires are soldered. Fig. 1-b is a first order circuit between its two terminals. For the notions to come, stray inductance for its wires is not needed. Taking the impedance *Z(jω)* to study the noise voltage of this 2TD leads to the circuit of Fig. 1-c, where two circuit elements give rise to two orthogonal voltages from the single current they share. The sum of these voltages done by their series connection gives *v(t)*, the voltage of Johnson noise at each frequency *f*=*ω*/(2π), but the annoying frequency-dependent resistance R(*ω*) and capacitance C*(*ω*) of this circuit, are not easy to link with physical phenomena. Hence, a clever choice for its noise *v(t)* is the admittance *Y(jω)* =1/*Z(jω)*.

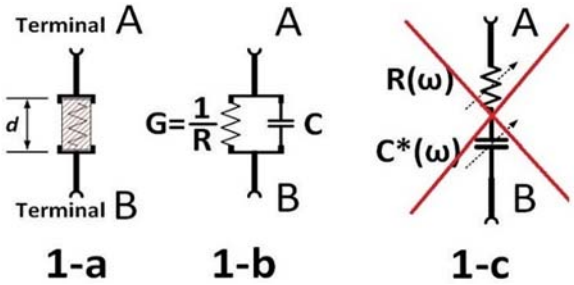


Fig. 1. a) Longitudinal cross section of a cylindrical resistor; b) A simple circuit to study its voltage noise; c) The series circuit of its complex impedance *Z(jω)*.

Admittance is a parallel notion where two mutually orthogonal currents add to give rise to a voltage shared by its conductance *G* and its susceptance *B*=*ωC*. The shared voltage *v(t)* is the Johnson noise and these two currents (the displacement and conduction ones linked to *C* and *G*) are mutually orthogonal due to the time derivative *∂v(t)/∂t* of currents in *C*. From these ideas, we proposed that random passages of single electrons between terminals of a resistor, would give rise to its Johnson noise [6]. The random series of tiny voltage steps of amplitude Δ*V*=±*q*/*C* caused by such passages mean fluctuations of energy accumulated in C that are subsequently dissipated by R as *v(t)* (the voltage in C) decays with time-constant τ_{relax}=R×C.

This Fluctuation-Dissipation dynamics for voltage noise agrees with the F-D framework of [3]. The full agreement found if each electron passed with a null transit time (τ_t=0), led us to look for a cogent proof about such instantaneous translocations. This proof is the key result of [7] that has completed our FBM of [6]. To prove that this null τ_t is possible due to the capacitance *C* of any 2TD, we will study in the next Section how we measure electrical current. This result converts our advanced FBM into a powerful tool to study electrical and electronic noises.

4. Displacement Currents and Shot Noise

The direct measurement of small voltages is a well-known task in noise studies. Those people believing that small currents are measured directly and not from voltages they produce, could find useful this Section with regard to the design of *i-v* converters to measure small currents in noise experiments [8]. The reason of this usefulness is that the main task of this Section is to review the measurement of low-level currents in 2TDs like resistors and capacitors.

"Thermal agitation of electric charge" entitling [2] suggests what causes Johnson noise. Hence, finding how electric charge agitates in the volume V_Q of a 2 TD will likely show us how Johnson noise is generated in resistors. Since electrical measurements are the tool to monitor this agitation, we will consider them briefly. Whereas a voltmeter gives directly a voltage value, the value of any current is always inferred from a voltage one. Using a sensing resistor, a Hall probe or a current clamp (transformer) current is converted into voltage that we measure to infer the value of the current that has caused this voltage. Hence, electrical current is known after measuring the effect it produces: a voltage between two terminals. This leads to consider the way we measure very low currents (e. g. femtoAmps).

Let us think of a current involving a few electrons per second on average. "Average" comes from the idea of current as a flow of discrete, corpuscular electrons. Focusing in our task, let us consider again the resistor made from a rod of conductive material of dielectric relaxation time τ=ε/σ, cladded by its two metal caps at a distance *d* (terminals) that is sketched in Fig. 1-a. To better follow the notions to come, let us use Fig. 2 for a resistor made from a rectangular rod of material instead of the cylindrical one of Fig. 1-a.

Used to charge-matter interactions that take place by discrete packets of charge of *q*=-1,6×10^{-19} C each that we call electrons, let us show how electrons that

"jump" between terminals of the 2 TD of Fig. 2 can agree with measurements. By the verb "to jump" we mean that if an electron travels from one terminal to the other, it will travel as a whole because to the best of our knowledge, pieces of electron are not found in low energy experiments like noise measurements. This leads to consider the possible behavior of the quantum of charge in a 2 TD and the fluctuations of energy that it can produce in it. Bluntly speaking, let us use quantum electronics for our task.

Each time we say "the current *i(t)*", we are actually considering two simultaneous currents at each instant of time. One enters the 2TD at position *x* in space by terminal A, while the other leaves the 2TD at position (*x*+*d*) by terminal B. Simultaneous measurements at two points separated by a distance *d* in space requires using Especial Relativity (just basic skills). Consider two observers at plates A and B of Fig. 2, whose clocks were synchronized by an electromagnetic (EM) signal going from the observer at point *x* (plate A) to the observer at point (*x*+*d*) (plate B). Needless to say, an EM signal going from observer in plate B to that in plate A would synchronize equally well their clocks following Einstein's lections.

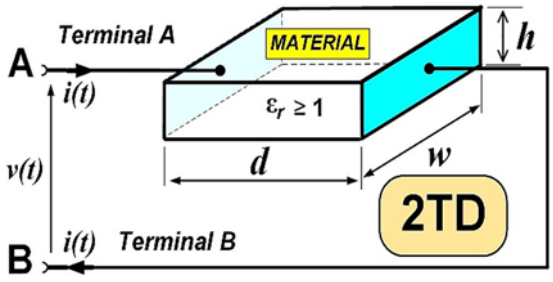


Fig. 2. Physical structure of a two-terminal device of volume $V_Q=w\times h\times d$ called resistor or capacitor, depending on frequency. Note its instantaneous "voltage" between its terminals and its two equal currents at each instant of time.

Although the notion of a corpuscle-like electron that leaves one plate at instant t_0 and arrives in the other at instant $(t_0+\tau_t)$, after a transit time $\tau_t>0$ seems possible, we will discard it for two reasons. The first one is because it makes impossible a cogent measurement of *i(t)*. With this corpuscular approach our two observers would never find the two simultaneous currents *i(t)* of equal magnitude at positions *x* and *x*+*d* that we need. If a current *i(t)* starts to bring some charge $+\partial Q$ in plate A, another current of equal magnitude *i(t)* also will start to extract charge $+\partial Q$ out of plate B.

Corpuscular electrons passing with transit time $\tau_t\neq 0$ would puzzle our observers. If the observer at plate A said: "a quantum of charge is entering the 2TD at $t=t_0$", the observer at plate B should say: "Really? Nothing occurs here at instant $t=t_0$." Then, if an electron passes between terminals, its electrical transit time will be null. This is better understood by taking Fig. 2 as a current sensing resistor of resistance R_s whose voltage *v(t)* will allow us to infer its current by Ohm's Law $i(t)=v(t)/R_s$. The voltmeter giving the difference of electrical potential in plates A and B at each instant of time that is: *v(t)*, is what synchronizes as described our two imaginary observers. The right operation of myriads of ammeters that work in this way suggests that if the observer at plate A said: "a quantum of charge is entering this 2TD at $t=t_0$", the observer at plate B should have reasons to say: "a quantum of charge is exiting this 2TD at $t=t_0$".

The above situation is possible if an electron arrival in plate A created a simultaneous disturbance in plate B of Fig. 2 to warn its observer and vice versa. For an observer within V_Q and next to plate A, an electron arriving there through terminal A would create an electric field ΔE pointing inwards plate A that would reach plate B as an electric field pointing outwards plate B. The inward pointing field in plate A would tell observer A that an electron has just arrived there whereas this same field pointing outwards in plate B, would tell observer B that an electron has just left the 2TD through terminal B. Since this field that is born in a plate (no matter if inwards or outwards) reaches the other plate at the speed of the EM wave in V_Q, the arrival of an electron in one plate and its departure from the other become two simultaneous events. The quantum of charge translocated in this way is what equates the two currents of the 2TD.

Reviewing the measurement of weak currents in a 2TD we find that its capacitance could help electrons to jump suddenly between its terminals. To show the easiness of an electron to pass between terminals of resistors as a displacement current of null dwell-time ($\tau_t=0$), we will say that the amount of energy needed for this passage is truly small ($\Delta U<<kT$). Making $Q=q$ in Eq. (4) we get $\Delta U=1.6\times10^{-7}$ eV for $C_{stray}=0.5$ pF. For kT≈26 meV at room T we obtain $\Delta U=0.000006kT$, or less if we add the unavoidable $C_{mat}>0$ [5] to C_{stray}. Hence, the mean rate λ of electrons jumping between terminals due to thermal activity should be a huge number. Note that after an electron translocated in this way the voltage of C will shift by $\Delta V=\pm q/C$ volts depending on the sense of $\Delta E=(q/C)/d$, the step of electric field that is equivalent to, or that accounts for this monoelectronic translocation.

Since each of these jumps is the type of fluctuations of energy that we proposed as the origin of Johnson noise in resistors [6], it is not difficult to show the way this voltage noise is generated by these fluctuations occurring randomly in time and sign, with average rate λ. Thinking of a very cold resistor, Fig. 3 shows how its Johnson noise is built in time by these fluctuations that have been represented in the horizontal axis as Dirac's delta functions.

To give a number for λ let us use Eq. (4) for a single electron. This gives $\Delta U=q^2/(2C)$ Joules or $q/(2C)$ eV that we will take as the "mean energy per fluctuation that enters the resistor". The energy remaining in *C* after each fluctuation will continue its dissipation driven by the noise voltage *v(t)* resulting after the $\pm q/C$ volts increment due to such fluctuation and *v(t)* will continue decaying with time-constant

τ=RC. The null mean of the Johnson noise means that its causes (i. e. thermal fluctuations of electric field Δ*E*) will be 50 % positive (Δ*E* along *x*) and 50 % negative (Δ*E* along -*x*) on average. This balanced distribution agrees with the randomness of a thermal bath driving these sudden translocations of electrons between terminals.

Dividing by R the mean square voltage noise of C in TE that is kT/C V^2 due to thermal equipartition, we obtain P_d=kT/(RC) watts as the mean power dissipated in the resistor in TE. Equating P_d to λ times the mean energy per fluctuation ΔU=q^2/(2C) that enters the 2TD (that would be the mean power entering the resistor) the mean rate of fluctuations that results is [6]:

$$\lambda = \frac{2kT}{q^2 R} \qquad (7)$$

In a resistor of R=50 Ω at room T we have λ≈6×10^{15} fluctuations/second. Hence, the Johnson noise of 50 Ω resistors or that of transmission lines of Z_0=50 Ω looks like a continuous voltage. But thinking of a very cold resistor at T→0, its Johnson noise would be something like the spiky voltage shown in Fig. 3.

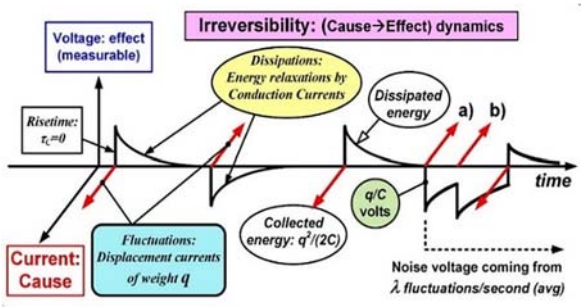


Fig. 3. 2-D model for Johnson noise generated from a random series of fluctuations in the volume V_Q of a resistor that are converted into voltage noise by its capacitance C.

Let us use the FBM for a 50 Ω resistor put in the test fixture of C_{stray}=0.5 pF of a noise analyzer to measure its Johnson noise at room T. Taking the material used for the resistor as a conductive one with a dielectric relaxation time of τ=ε/σ=1 ps, the capacitance due to this is: C_{mat}=τ/R=0.02 pF. Taking C=C_{stray}+C_{mat}≈0.5 pF for simplicity, the mean square voltage noise between terminals of this C at room T will be: kT/C=8.3×10^{-9} V^2 due to thermal equipartition. This voltage noise driving R=50 Ω means P_d=166 pW as the mean power dissipated in this resistor at room T. Following [3, 6] if you observe voltage noise there must be dissipation in the noisy device.

With regard the mean energy per fluctuation that enters this resistor ΔU=q^2/(2C)=0.16μeV, it is truly low, only is six parts per million the thermal energy kT≈26 meV at room T. The frequency of a photon of this energy only is: ΔU/*h*=39 MHz. From Eq. (7), the capacitance shunting this resistor would sense its 50 Ω resistance as a "dense rain" of λ=6.5×10^{15} fluctuations per second. If this was a 50 MΩ resistor its shunting capacitance only would undergo 6.5×10^9 fluctuations per second, thus 10^6 times lower as it would be the mean power dissipated in this case that would drop from 166 pW to P_d=166×10^{-18} W. This picture about how a capacitance "is sensing" its shunting resistance is very useful to envisage the effect of its voltage noise in the output signal of oscillators that use as resonator an L-C tank. We will show this later after a closer look at the way thermal energy pervading the volume V_Q of a 2TD is collected by its capacitance acting as an antenna for this purpose.

5. Charge Noise in C and Thermal Photons

From Maxwell equations, an electrical field is equivalent to positive and negative charge separated in space. The thermal bath that pervades the volume V_Q of a 2TD is electromagnetic energy (EME) traveling under the form of EM fields oscillating with different frequencies. When the EME travels in vacuum or through solid matter, it behaves as the continuous EM wave that was predicted by J. C. Maxwell. But when this EME interacts with charges existing in matter, it is absorbed or emitted in the form of packets of energy proportional to its frequency that are called "photons". The notion of EME as a set of individual photons that travel in space or through matter would be doubtful because its discrete nature is observed when it stops its travel and disappears as soon as it interacts with matter (to create an electron-hole pair for example).

Previous paragraph warns the reader about the doubtful meaning of sentences like "If a photon of energy U arrives in, or is absorbed by an electron…". Using the Plank constant *h*, it would be better to write: "If EME of frequency U/*h* arrives in, or is absorbed by an electron…". Once this warning has been given, we shall continue by saying that those oscillating fields of the thermal bath in the volume V_Q of a resistor can interact with some electrons of its inner material. This occurs with those conducting electrons of the material, often called "electrons in its conduction band (CB) or just free electrons", to mean that they can take part in conduction processes, a feature that electrons trapped in donor atoms embedded in this material, would not have. Neutral impurity atoms in n-doped silicon that are ionized by thermal energy produce free electrons in the CB of this material. This is the method we use to have free electrons occupying energy levels of its CB near its bottom level.

Free electrons that are already in the CB, in low energy states close to its bottom, can absorb thermal energy in order to occupy new states of higher energy within the CB and they continue being free electrons. In dual form, electrons in high energy states of the CB can emit EME to occupy states of lower energy in the CB. Concerning the inner material of our resistor with free electrons in energy levels near the bottom of its CB, the presence of its capacitance C sets a new degree of freedom for these free electrons that are able to gain

or to lose energy to jump between energy levels within the CB. Given the tiny amount of energy $\Delta U=q^2/(2C)$ shown in Section 4, the presence of the capacitance C means that there are new energy levels very close to the bottom of the CB, available to free electrons able to exchange EME with the thermal bath in the volume V_Q. This leads to consider the energy requirements of these exchanges and their effect.

If EME of frequency $\Delta U/h$ and with the proper polarization arrived in the V_Q of a resistor when its voltage in C is null, an electron of the CB could absorb it "to set a voltage of q/C V in C". If we prefer: "to occupy one of the new energy states that the existence of C creates near the bottom of the CB". In this way the energy ΔU of the photon thus absorbed would enter this resistor and would appear as a sudden step of electric field in its C with amplitude $\Delta E=(q/C)/d$ V/m and null rise time. Hence, the absorption of this photon of the thermal bath gives rise to a voltage step between terminals of $\Delta V=q/C$ volts amplitude and null rise time that would set $v(t)=+q/C$ volts in this resistor.

By contrast, if $v(t)$ already was $v_i(t)=+q/C$ V when this EME arrived in V_Q it could not be absorbed to rise $v_i(t)$ up to $v_f(t)=v_i(t)+q/C$ volts because the energy (and thus the frequency of the EME) required for it now is: $(2q)^2/(2C)-q^2/(2C)=3\Delta U$. However, the capacitance with a voltage $v_i(t)=+q/C$ V would have in its cathode an electron with an energy ΔU that could emit on its own by a spontaneous photon, or by a photon that was stimulated by EME of frequency $\Delta U/h$ that could exist in the thermal bath. In these two cases, an amount ΔU of EME would pass from the resistor to its thermal bath. And if $v(t)$ already was $v_i(t)=+q/C$ volts and an EME of frequency $3\Delta U/h$ arrived in V_Q with the proper polarization, a free electron with energy close to the bottom of the CB could absorb a photon of this EME to set in C a voltage $v(t)=+2q/C$ volts. In this case, an energy of 3ΔU coming from the absorbed photon would enter this resistor. It would appear as a sudden step of electric field of amplitude $\Delta E=(q/C)/d$ V/m in its C.

Given the broad energy band for the photons of a thermal bath at room T and their statistical arrival in the 2TD, the noise voltage in C could depart markedly from $v(t)=0$ from time to time. We refer to much larger departures than the $\Delta V=q/C\approx 0.3$ μV that results in this case where C≈0.5pF. Given the huge value of λ, a fast burst of fluctuations of the same sign that could vary $v(t)$ by 30 μV, for example, does not seem unlike. All the above and Fig. 3 show the way the Johnson noise of a resistor appears in time as this 2TD exchanges energy with its thermal bath due to its capacitance C between terminals. The sense of the flow of energy (inwards/outwards) at instant t_0 where a fluctuation takes place, will depend on the relative signs of the previous voltage $v(t)$ existing in C at t_0^- and that of the step $\Delta V=\pm q/C$ Volts generated by the fluctuation.

The energy that enters a resistor by its fluctuations and its unceasing dissipation due to the presence of $v(t)$ driving its resistance R, sustain in time a charge noise in its capacitance C whose measurable effect is the Johnson noise $v(t)$ at temperature T (see Fig. 3). It is worth noting that the charge noise due to fluctuations comes from the individual behavior of each electron, whereas the incessant dissipation of energy coming from the current $i_c(t)=v(t)$/R results from the collective behavior of electrons as they vary in the plates of the capacitance C. This last behavior is transferred to the charge noise of the displacement current $i_d(t)=-i_c(t)$ that also exists in V_Q as it is shown in Fig. 4.

In the dielectric relaxation taking place in the 2TD of Fig. 2 when its voltage $v(t)$ decays as $\exp(-t/RC)$, the conduction current $I_c(t)=v(t)/R$ is counterbalanced at each instant of time by the displacement current due to the electric field that decays in C. This must be so to keep null the current $i(t)$ that enters or leaves the 2TD through its terminals A or B because in TE, it is under open circuit conditions. Hence, the current associated to R (i. e. $i_c(t)=v(t)/R$) and the displacement current in C (i. e. $i_d(t)=C\times(\partial v(t)/\partial t)$ are both non-null but they cancel one another at each instant of time. From such exact cancellation it could be said that the voltage noise generated by these two currents giving rise to zero current at each instant of time should be null, at first sight. "At first sight" means that as we will show below, the mean square charge noise due solely to fluctuations is: 2kT/R coulombs/second, thus half the shot noise in the resistor.

This exact cancellation that discards any magnetic effect after each fluctuation, is considered in Fig. 4. It was found looking for magnetic effects that could produce the type of fluctuations we had proposed in our FBM for electrical noise [8]. Added to this, and from our attempts in the past century to measure the voltage noise of the shot noise assigned to a dc current I_{dc}=1 μA in a 1 MΩ resistor, we knew that "conduction current in resistors is noiseless" (see below). Hence, we realized that the null current of previous paragraph was 50 % noisy -$i_d(t)$- and 50% noiseless: -$i_c(t)$- (see Fig. 4). And since the mean square charge noise due solely to $i_d(t)$ is 2kT/R coulombs/second, we conclude that the total mean charge noise of the resistor was: 4kT/R C^2/s (or A^2/Hz) as it is written in the title of [7], where this 50 %-50 % partition of the Nyquist noise was proposed for the first time (to the best of our knowledge).

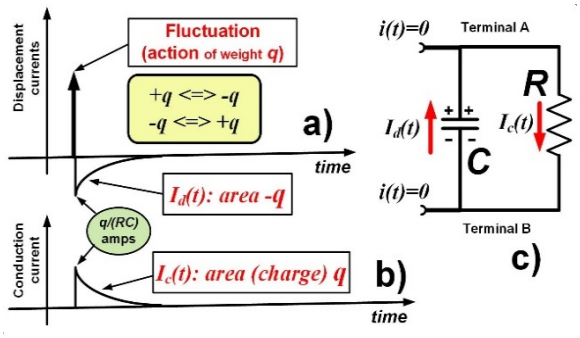


Fig. 4. Displacement and conduction currents triggered by a fluctuation causing a sudden voltage step of $\Delta V=q/C$ V in the 2TD of Fig. 2.

Considering that $i_c(t)$ and $i_d(t)$ are currents that are distributed in V_Q and go in opposed senses between plates of Fig. 2, we may use Fig. 4 to understand them better in a resistor with $v(t)$=0 between terminals up to the instant t_0, where the sudden jump of an electron from plate A to plate B takes place. This event would set $v(t_0^+)$=+q/C volts in C. As soon as $v(t_0^+)$ was set, a conduction current of [q/C]/R amps would start in R. It would be: $i_c(t)$=+[$q/(RC)$]×exp[-(t-t_0)/RC] (see Fig. 4-b). Integrating $i_c(t)$ from t=0 to $t\rightarrow\infty$ the charge that we obtain suggests that the charge passing from plate A to plate B "through R" is +q coulombs. This leads to the idea of a thermal action that displaces a charge +q C from plate B to plate A and the subsequent device reaction by the conduction current $i_c(t)$ to restore the 2TD to its initial state just before the fluctuation, thus to its state at the instant $t=t_0^-$. Note how the current $i_d(t)$ goes unnoticed in this action/reaction notion.

From the currents shown in Fig. 4 however, we see that the charge displaced in one sense by a fluctuation and that displaced back by its subsequent displacement current $i_d(t)$ would cancel one another in time, though at different instants. Given this cancellation, no other displacement of charge would be needed if the internal dipole due to the fluctuation in this 2TD disappears in this way. Then this intriguing question arises: what is the actual charge displaced by the current $i_c(t)$ that we use to consider as "the passage of electrical charges through R"? Since the presence of $i_d(t)$ is undeniable, we would say that the conduction current $i_c(t)$ "should not displace charge at all". This sounds heretic because $i_c(t)$ suggested that the charge "passing from plate A to plate B through R" is +q coulombs.

On the other hand, we knew long time ago that the voltage noise of a 1 MΩ resistor in TE and its voltage noise measured while a dc current I_{dc}=1 μA was set in it, were equal within experimental error. Years ago, we used to do this experiment as a calibrating routine of a home-made noise measuring setup. The 1 μW power dissipated by I_{dc} is low enough so as to not heat noticeably the big resistor of 1 MΩ/1 W that we used as a source of flat noise density S_{VJohn}=1.7×10^{-14} V^2/Hz. Viewing that the voltage noise density we measured for I_{dc}=1μA and for I_{dc}=0, was always S_{VJohn}, we have been considering that displacement currents are noisy, but conduction currents are noiseless for an unknown reason. Hence, currents like the two I_{sat} of a diode that occur in a space charge region (i. e. in a capacitive 2TD) should produce shot noise giving rise to voltage noise in the 2TD under study.

Regarding to this experiment, we shall say that if the current I_{dc}=1 μA was the flow of discrete electrons passing from the negative terminal to the positive one of this resistor, the number of electrons crossing it would be: 6.2×10^{12} electrons per second, whose shot noise converted into voltage noise by R=1 MΩ should give a density of voltage noise S_{Vcond}=2$q$$I_{dc}$R^2 V^2/Hz. This S_{Vcond}=3.2×10^{-13} V^2/Hz is much higher than S_{VJohn}. If I_{dc} was generating this noise, the experiment would reveal it unambiguously. Hence, conduction currents do not generate voltage noise.

From the null charge that should displace $i_c(t)$ when we consider that this task is done by $i_c(t)$ (see Fig. 4) and from the role of resistance "representing a series of dissipations of capacitive energy that are triggered by fluctuations" we could give a possible reason why conduction current is noiseless. If it would not involve the passage of discrete charges between two terminals, it would not produce voltage noise. Mimicking $i_c(t)$, conduction current should involve the occurrence of a series of dissipations of capacitive energy in the 2TD. If this was so, "dissipation current" would be a better name than conduction current for this type of current I_{dc} different from displacement ones. A requirement that I_{dc} should accomplish is to put the resistor out of TE, but "leaving unaltered the displacement currents that generate shot noise in the resistor in TE".

Searching for a model for conduction current from the above ideas we tried to keep the role of resistance in TE: "representing a series of discrete dissipations of capacitive energy that are triggered by fluctuations" but using an amount of energy for each dissipation that was proportional to $(V_{dc})^2$, the square of the dc voltage between terminals V_{dc}=R×I_{dc}. With this starting point and inspired by switched capacitor circuits emulating resistors, we envisaged a model for conduction current in resistors that would keep unaltered the displacement currents (fluctuations) that give rise to their Johnson noise [8]. However, this is a subject still under study that falls out of the scope of this work.

The action-reaction dynamics of Figures. 3 and 4 is an old notion we used concerning charge noise in a 2TD before reading [3]. From the noise viewpoint, there are λ sudden actions per unit time and λ slower reactions in opposed sense associated to dissipations as shown in Fig. 4 (a). Fig. 4 (c) is the lumped circuit that results if we separate the interpenetrating currents $i_c(t)$ and $i_d(t)$ that are distributed in the volume V_Q. Regarding the sense of $i_c(t)$ going from plate A to plate B, it is clear because the positive charge is in plate A. This makes the voltages V_A>V_B and applying Ohm's Law, $i_c(t)$ will go downwards "through R", as shown in Fig. 4.

Concerning the sense of $i_d(t)$, we must consider the electric field E set in C by the fluctuation and its time derivative. Since E points downwards (from plate A to plate B), but decreases as the voltage $v(t)$ decays in time, its time derivative is negative. Hence, $i_d(t)$ will go upwards, as shown. Since the external currents $i(t)$ of the 2TD of Fig. 2 are both null while $v(t)$ relaxes, we also obtain $i_d(t)$=-$i_c(t)$. This is the reason for the negative displacement current in Fig. 4 (a) starting at $t=t_0^+$, just when the fluctuation is already gone.

The notion of a mean square voltage is familiar for voltage signals. For charge signals, however, we must consider the charge in one plate of a capacitor and its mean square, since the net charge entering a 2TD like this one is null, as it should be known. From a charge noise point of view, the two opposed displacement currents of Fig. 4 (a) have a null overlap in time due to

their cause-effect link [6, 9]. This means that they are uncorrelated and their effects will add quadratically or "in power". Once stated this, we shall go to Fig. 4 (a) to consider the charge noise associated with each fluctuation. To do this, we must integrate current in time.

Since the charge brought to C (we mean: to its plate A for example) by the fluctuation of null dwell time (Dirac's delta) is $+q$ C, its square will be: q^2 C^2. With regard to the charge brought to the same plate A of C by the slow current $i_d(t)$ coming after, it is $-q$ C and its square will be: q^2 C^2. In this way, the mean square charge noise caused by each fluctuation in a 2TD like a resistor, a capacitor or an L-C tank circuit will be: $(q^2+q^2)=2q^2$ C^2. Mimicking quadratic voltages taken as "power over a resistance of 1Ω" in signal theory, let us call "charge noise power" P_Q to quadratic values of charge noise in a capacitor. Using this notion, we can say that given the huge rates λ for fluctuations giving Johnson noise and its random sign, they should give rise to a quite constant charge noise power in C.

Working with charge noise power we shall consider first the power due solely to fluctuations. Because of their null dwell-time, these impulsive currents of mean square charge q^2 C^2 each, do not overlap in time and their effects will add in power. The mean charge noise power of fluctuations taking place at the rate of Eq. (7) will be λ times q^2 C^2/s (i. e. coulombs^2/second). Multiplying Eq. (7) by q^2, we obtain P_{Qf}=2kT/R C^2/s, a numerical result that is only half the Nyquist density S_I=4kT/R A^2/Hz.

Since P_{Qf}=2kT/R C^2/s or S_I=2kT/R A^2/Hz are both magnitudes having to do with displacement currents causing shot noise, the lacking 50 % of S_I should be the shot noise due to displacement currents different from the fast fluctuations of null dwell-time. This leads to consider the noise of those decaying displacement currents $i_d(t)$ of opposed sign to their preceding fluctuation that were explained in regard to Fig. 4-a. This slow charge noise power is also P_{Qs}=2kT/R C^2/s. Thus, the total charge noise power P_Q that produces voltage noise in the resistor in TE at temperature T is:

$$P_Q = P_{Qf} + P_{Qs} = \frac{4kT}{R} \qquad (8)$$

To read properly Eq. (8), we must say that despite the equal charge noise power 2kT/R C^2/s of P_{Qf} and P_{Qs}, they have very different spectra. Whereas P_{Qf} comes from the individual behavior of the quantum of charge, P_{Qs} reflects the collective behavior of electric charges in the 2TD. The null dwell time of the fluctuations converts their displacement currents into an impulsive current noise whose spectral density is flat in frequency from the electrical viewpoint. This means that any electrical measurement in the 2TD should find voltage steps with null risetime reflecting the impulsive nature of these displacement currents producing shot noise of flat density S_I=2kT/R A^2/Hz. By contrast, the slower displacement currents tracking exponential decays of $v(t)$ will produce shot noise with Lorentzian spectrum of amplitude S_I=2kT/R A^2/Hz and cut-off frequency f_c=1/(2πRC) due to the time constant τ=RC of those exponential decays.

All in all, the Johnson noise of resistors will have two components. Its faster component a) will be a Lorentzian term of amplitude 2kTR V^2/Hz and cut-off frequency f_c=1/(2πRC) coming from the flat density of P_{Qf} (S_I=2kT/R A^2/Hz) being filtered by the circuit of Fig. 4 ©. Its second term, component b), will come from the Lorentzian shot noise of amplitude S_I=2kT/R A^2/Hz and f_c=1/(2πRC) being filtered by the circuit of Fig. 4 (c) whose time constant is τ=RC. This means that the -3 dB cut-off frequency f_{cs} of this slower term will be somewhat lower than f_c, and that its roll-off at high frequencies will be -40 dB/dec, two times faster than the -20 dB/dec of the component a). The sum in power of these two spectra of uncorrelated noises will show the well-known density S_V=4kTR V^2/Hz of Johnson noise at low frequencies, but its -3 dB bandwidth will be somewhat lower than f_c.

Concerning a value for frequencies close to f_c where the otherwise flat S_V will be decaying, we will consider again the noise measuring setup of C_{stray}= 0.5 pF for the Johnson noise of a resistor of R=50 Ω whose small but non null C_{mat} is included in C_{stray}. The time constant τ=RC=25 ps leads to f_c≈6.4 GHz. Hence, the Johnson noise of this resistor up to 2 GHz and even in the S band (2-4 GHz) would show the density S_V=4kTR V^2/Hz that everybody knows.

By contrast, this S_V density in the K band (8 GHz to 26 GHz) would be a neatly decreasing spectrum that would acquire a final roll-off approaching -20 dB/dec. Since measurements at these frequencies are not easy and most people use to work at "low f" (e. g. tens of MHz to avoid the filtering due to τ=RC) the limit f_c is barely reached and those measuring a flat S_V=4kTR V^2/Hz at $f<<f_c$, consider that the density S_I=4kT/R A^2/Hz that they infer from S_V is flat up to the limit given in [2] from quantum reasons.

6. Two Charge Noises: Empirical Evidence

The two different components of the phase noise of electronic oscillators (their carrier line of non-null width and a broad pedestal of much lower amplitude) is a striking feature that we are going to consider here. In frequency domain, the phase noise of an oscillator means that the spectral content of its output signal is not the line of null width $\delta(f-f_0)$ of a pure sinusoid of frequency f_0 (the expected "carrier" by design). Phase noise leads to find spectral content different from the above Dirac's delta.

The first noise term we refer to is "phase noise close to the carrier" because it is observed as a broadening of the otherwise monochromatic line at $f=f_0$. To keep this Section short, we will say that this part of the output spectrum is a Lorentzian line centered at f_0 that recalls the output spectrum of optical oscillators like lasers. In lasers, the broadening of their output line is already known as an effect of

spontaneous emissions of photons that disturb the coherent optical energy that is stored in their resonant cavity, usually a Fabry-Perot resonator.

In dual form, the L-C tank of an electronic oscillator (i. e. its resonator) should sustain in time a coherent or pure sinusoidal voltage signal, and the electronic loop of the oscillator is designed to do this task in the best way possible. Since the L-C tank has a capacitance C that is connected in parallel with an inductance L, it is evident that any sinusoidal voltage sustained in time by the feedback electronics of the oscillator will be disturbed by the spontaneous translocations of single electrons between the plates of C (i. e. fluctuations) that are the origin of voltage noise in 2TD.

Due to the null risetime of each voltage step of $\Delta V=\pm q/C$ volts that is added randomly (both in time and sign) to the otherwise sinusoidal oscillation of voltage in the capacitance of an L-C tank (see Fig. 5), the electronics of the oscillator loop (its Automatic Amplitude Control AAC), can do nothing to avoid the appearance of these steps. We mean nothing to remove the impulsive current generating them in C, hence in the output signal. To react against one of these voltage departures of the output signal from a perfect sinusoid, the AAC must wait for the voltage ΔV in C (effect) that indicates the occurrence of a fluctuation "that is already gone". From this picture about how Johnson noise or kT/C noise enters into the resonator, we can envisage the presence of a kind of "roughness" in the output signal (phase noise) that is unavoidable.

From Eq. (7) for λ and $\Delta V=\pm q/C$, it is quite easy to obtain the phase noise due solely to these voltage steps of null risetime. Fig. 5 taken from [10] shows the way each fluctuation disturbs the sinusoidal voltage (output signal) that would exist in the capacitance of the L-C resonator used by these oscillators.

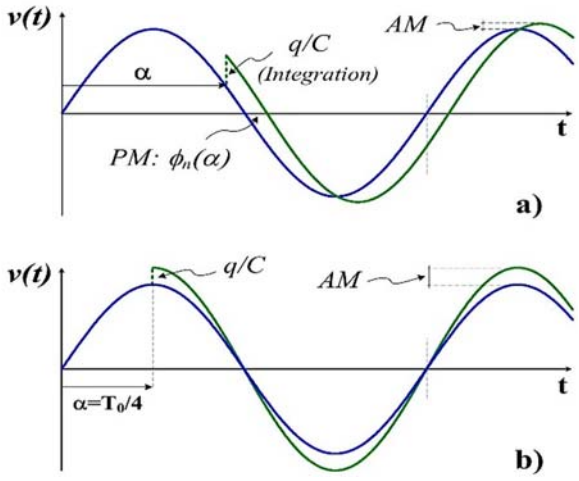


Fig. 5. Effects of a fluctuation taking place in the capacitance of an L-C oscillator when the damping of this change of energy is small during one period (see [10]).

Evaluating the phase noise due to disturbances like those shown in Fig. 5, taking place randomly both in time and sign with the average rate λ of Eq. (7), this phase noise spectrum is found [10]:

$$\mathcal{L}(\omega-\omega_0)=\frac{2D}{(\omega-\omega_0)^2+D^2} \qquad (9)$$

Eq. (9) is the Lorentzian line or "phase noise close to the carrier" that we commented at the beginning of this Section. This line agrees perfectly with a classical paper in the field published by D. B Leeson [11]. Thus, the Lorentzian line of the output spectra of electronic oscillators comes from the truly impulsive current that gives rise to P_{Qf}. This would explain the deep analogy between spontaneous emissions of photons in lasers and spontaneous translocations of single electrons in L-C resonators due to thermal activity and why only half the Nyquist noise ($P_{Qf}=2FkT/R$) gives phase noise close to the carrier. To include the total losses of the L-C resonator "loaded by its surrounding electronics" we have used its unloaded resistance R divided by the noise figure F of the electronics around the L-C tank (see below). The single-sided spectrum of Fig. 6 taken from [10] shows the narrow line around f_0 (the carrier frequency) coming from $P_{Qf}=2FkT/R$ C^2/s and the flat pedestal of $P_{Qs}=2FkTR$ V^2/Hz divided by $(V_O)^2/2$, the mean square of the sinusoidal voltage of amplitude V_O (thus $P_O=(V_O)^2/2R$) sustained in C by the loop.

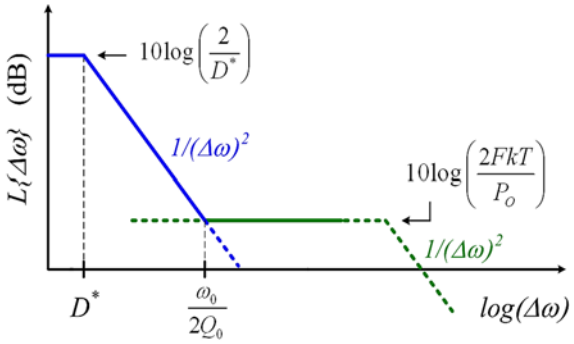


Fig. 6. Lorentzian Line of Phase Noise (e. g. "carrier" line) due to fluctuations in C, together with the Phase Noise Pedestal generated by the AAC (see [10]).

Concerning the slow 50 % of Nyquist noise we have called P_{Qs} we may say that gives rise to the pedestal of phase noise far from the carrier. This band-limited noise, being slower than P_{Qf}, is liable to be affected by the electronics of the loop and when this occurs, the band around f_0 of the aforementioned pedestal (i. e. its width) increases. The amplitude of this pedestal suggesting that only half the Nyquist noise $P_{Qs}=2kT/R$ C^2/s is involved in its generation, is another clue about the partition of the familiar density $S_I=4kT/R$ A^2/Hz of Nyquist noise into the fast and the slow shot noises that we have shown in [7]. Including the noise of the electronics by its noise figure F, the pedestal of phase noise results proportional to $P_{Qs}=2FkT/R$ C^2/s and the same occurs with the magnitude D (phase diffusivity) of Eq. (9) that

becomes proportional not to P_{Qf}=2kT/R C^2/s, but to P_{Qf}=2FkT/R C^2/s, as it must be from the meaning of noise figure F in the FBM [6, 10].

It is worth noting that the best fitting to the phase noise of electronic oscillators has been achieved by using instantaneous translocations of charge between nodes of the oscillator circuit [12]. Nevertheless, the packets of charge that these authors have used involve thousands of electrons (e. g. packets of fC, 10^{-15} C or more) and the physical meaning of those sets of many electrons jumping synchronously between two nodes of the circuit is unclear. Nevertheless, the merit of this work using "sudden voltage changes" is undeniable and really useful if the user of the program is not very concerned for the physical meaning of this proposal that "works well in simulations".

7. Conclusions

The model currently used for electrical noise in resistors is a dissipation-based model that needs to be improved to explain the noise of other two-terminal devices (2TD). With this notion in mind, we proposed a fluctuation-based model for electrical noise in 2TDs that fits in the Fluctuation-Dissipation framework that was proposed by Callen and Welton. It is a fluctuation-based model where the electrical capacitance between terminals of a 2TD is the reactive element that allows its thermal interaction. The voltage that appears in this capacitor whilst energy is being exchanged with its thermal bath, is Johnson noise in resistors, kT/C noise in capacitors and the voltage noise of 2TDs in general. Hence, the capacitance of each resistor between its terminals is the transducer that generates its density of Johnson noise S_V=4kTR V^2/Hz that we can measure.

Concerning the density S_I=4kT/R A^2/Hz that we can infer from S_V, this fluctuation-based model shows that S_I is formed by two different noises. The first term is impulsive shot noise of flat density S_{If}=2kT/R A^2/Hz until the quantum limit $f_Q \approx kT/h$ (f_Q≈6 THz at room T). The instantaneous translocations of electrons between the terminals of the device (i. e. fluctuations) give this ultrawideband noise. By contrast, the second term is a Lorentzian shot noise of amplitude $S_{Is}=2kT/R$ A^2/Hz whose cut-off frequency is equal to that of the 2TD (where the voltage noise is being measured). The phase noise of oscillators that use L-C resonators agrees with the impulsive origin of S_{If} and with this 50 %-50 % partition: $S_I = S_{If} + S_{Is}$=4kT/R C^2/s or A^2/Hz.

Acknowledgements

This work was supported by the Spanish Ministerio de Ciencia y Tecnología under the MCYT MAT2017-84817-C2-1-R and PID2020-118410RB-C22 projects.

References

[1]. J. B. Johnson, Thermal agitation of electricity in conductors, *Physical Review*, Vol. 32, 1928, pp. 97-109.

[2]. H. Nyquist, Thermal agitation of electric charge in conductors, *Physical Review*, Vol. 32, 1928, pp. 110-113.

[3]. H. B. Callen, T. A. Welton, Irreversibility and Generalized Noise, *Physical Review*, Vol. 83, 1951, pp. 34-40.

[4]. J. I. Izpura, On the electrical origin of flicker noise in vacuum devices, *IEEE Transactions on Instrumentation and Measurement*, Vol. 58, 2009, pp. 3592-3601.

[5]. J. I. Izpura, 1/f electrical noise in planar resistors: the joint effect of a backgating noise and an instrumental disturbance, *IEEE Transactions on Instrumentation and Measurement*, Vol. 57, 2008, pp. 509–517.

[6]. J. I. Izpura and J. Malo, A Fluctuation-dissipation model for electrical noise, *Circuits and Systems*, Vol. 2, 2011, pp. 112-120.

[7]. J. I. Izpura, Resistors as electromagnetic sensors: 4kT/R C^2/s (or A^2/Hz), in *Proceedings of the 8th International Conference on 'Sensors and Electronic Instrumentation Advances SEIA' 2022',* Corfu, Greece, 21-23 September 2022, pp. 84-90.

[8]. J. I. Izpura, Particle-wave duality in electrical noise, arXiv:1512.07149v2 [cond.mat.other] 2016.

[9]. J. I. Izpura, Electrical noise, Brownian motion and the arrow of time, *Journal of Modern Physics*, Vol. 6, 2015, pp. 131-140.

[10] J. Malo and J. I. Izpura, Thermodynamical phase noise in oscillators based on L-C resonators, *Circuits and Systems,* Vol. 3, 2012, pp. 61-71.

[11] D. B. Leeson, A simple model of feedback oscillator noise spectrum, *Proceedings of IEEE,* Vol. 54, 1966, pp. 329-330.

[12] A. Hajimiri, T. H. Lee, A general theory of phase noise in electrical oscillators, *IEEE Journal of Solid-State Circuits*, Vol. 33, 1998, pp. 179-194.

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 12-18

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Mechanical Modelling and Stress-Strain Prediction of Electroactive Polymer-Based Capacitive Sensor under Multiaxial Loads using Uniaxial Tensile Data

Nitin Kumar SINGH, Kazuto TAKASHIMA and Shyam S. PANDEY
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu campus, Kitakyushu, Fukuoka 808-0196, Japan,
Telefax: +81-93-695-6230
E-mail: nitinmjpruiitp@gmail.com, singh.nitin-kumar228@mail.kyutech.jp,
nitin.mt12@iitp.ac.in, shyam@life.kyutech.ac.jp

Received: 1 July 2022 /Accepted: 1August 2022 /Published: 31 October 2022

Abstract: In this paper, mechanical modelling and stress-strain prediction of the ionic type of electroactive polymer (EAP) based capacitive strain sensor under multiaxial loading has been performed using uniaxial tensile data. A highly stretchable sensor with > 400 % strain was fabricated using Styrene-ethylene-butylene-styrene (SEBS) rubber and Dodecyl benzene sulphonic acid (DBSA) doped polyaniline (PANI) composite material as a dielectric film, and carbon grease as top and bottom electrodes. From uniaxial tensile measurements, a three-dimensional constitutive model of the EAP strain sensor was validated. Experimental data obtained from the uniaxial tensile measurement was fitted according to the Ogden-based constitutive mathematical model. For predicting the stress-strain behaviour of the sensor under the multiaxial loads, 2nd term of the Ogden model was demonstrated to fit well with uniaxial tensile data. Such modelling can play a significant role in designing any type of EAP-based sensor for various operations under variable loading conditions.

Keywords: Ionic type of electroactive polymers, Strain sensors, Multiaxial loading, Constitutive modelling, Curve fitting, Uniaxial tensile testing.

1. Introduction

Electromechanical devices like actuators, sensors and generators have been widely used in robotics, healthcare, augmented reality and wearable devices [1- 5]. The last decade has witnessed a shift in the research trends from hard and brittle to soft, flexible and lightweight wearable devices. Highly elastic strain sensors that can detect large strains are highly in demand for deformable systems such as soft robotics and wearable devices. When mechanical stimuli are applied to the strain sensors then changes in capacitive and resistive values have been noticed, which makes the basis of this type of sensor technology [6-11].

Capacitive sensing has several advantages over resistive sensing like high linearity, reproducibility, and low hysteresis so currently researchers are, focusing the capacitive-based EAP sensors.

The total capacitance of EAP based capacitive sensor (C_{total}) is the sum of two capacitances, one is capacitance due to a change in its own dimension (C_1) and another is capacitance due to contact or electronic circuit, as given below:

$$C_{total} = C_1 \left(\frac{\varepsilon_r \varepsilon_0 A}{t}\right) + C_{parasitic},$$

where ε_r is the dielectric constant; ε_0 is the permittivity of free space; A is the active polymer area; t is the thickness of the dielectric film and $C_{parasitic}$ is the capacitance caused by contact. However parasitic capacitance is very small and can be neglected. During all types of practical applications, these EAP-based sensors undergo continuous loading and unloading conditions and multiaxial deformations.

For choosing a suitable mechanical sensor, it is very important to analyze stress-strain behaviour

under various conditions, while for elastomer-based soft sensors, it is very difficult and time-consuming to analyze the stress-strain behaviour of sensors in multiaxial loading conditions [13-14].

For designing a strain sensor for different types of loading comprehensive analytical modelling is needed, which could verify the sensor's performance experimentally. Existing papers do not provide complete modelling and neglect many crucial parameters which might affect the sensor's performance at large strain levels in the long run.

In this paper, the sensor was fabricated using Styrene-ethylene-butylene-styrene (SEBS) rubber and Dodecyl benzene sulphonic acid (DBSA) doped polyaniline (PANI) composite film (casting from the toluene solutions having 5 % concentration) sandwiched between upper and lower part with carbon grease-based electrodes. Tensile tests were performed, for analyzing the sensor's performance by using constitute equations, after post-treatment of experimental data global fitting parameters were evaluated.

These global parameters can predict sensor performance even at a large value of deformation in different types of loading. It was found that the stress-strain curve was well fit with the second order of the Ogden constitutive equation. By using fitting parameters in the Ogden model-based constitutive equation, stress can be predicted at any value of strain under multiaxial loads. This complete procedure is applicable to any type of elastomer-based capacitive strain sensor at different types of loadings.

2. Experimental

2.1. Materials

A conducting and stretchable freestanding film consisting of SEBS rubber (supplied from the Sigma-Aldrich co.) and DBSA doped polyaniline (synthesized in our laboratory at Kyutech, Japan) composite film sandwiched between carbon grease electrodes (purchased from MG chemicals) was used for fabricating Ionic type of EAP sensors. Carbon conductive tape for making the electrical connection was purchased from Nissin Co. Ltd. Our uniaxial tensile system developed at Kyushu institute of technology, Japan consisting a computer-controlled X-Y stage controller (Mark-102, Sigma Koki, Japan), load cell (LU-10 K, Kyowa), and digital Oscilloscope (TDS 2001C, Techtronic, USA).

2.2. Uniaxial Tensile Machine for Mechanical Characterization

Uniaxial tensile testing is most commonly used for characterizing the mechanical behavior of isotropic and elastic materials. Tensile measurement for the sensor used in this work was conducted using the uniaxial tensile testing system as shown in Fig. 1. It can be seen from this figure that our tensile measurement system has a 100 N load cell, mobile stage, X-Y stage controller, and an Oscilloscope. Multiple stretching and relaxation cycles under variable strains using a stage controller and mobile X-Y stages were controlled by a computer-controlled Lab-View program. Changes in the stress under variable strain.

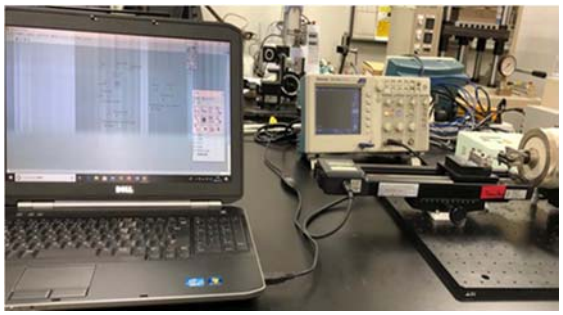


Fig. 1. Tensile measurement set-up.

2.3. Sensor Fabrication and Sample Preparation for Tensile Testing

The basic structure of the device has been shown in Fig. 2. This sensor is three layer device where DBSA doped Polyaniline SEBS rubber was used as a dielectric film and carbon grease was used as an electrode. The composite film was sandwiched between the upper and lower layer of carbon grease-based electrodes, as shown in the device structure in Fig. 2.

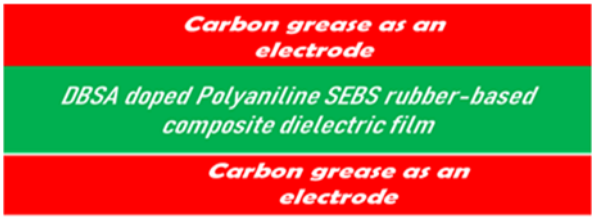


Fig. 2. Basic structure of stretchable carbon electrode-based ionic EAP strain sensor.

The detailed fabrication process has been described below [15]:

- Mix Polystyrene-block-poly(ethylene-ran-butylene)-block-polystyrene-graft-maleic anhydride (SEBS-g-MA) and toluene and stirred solution at 1500 rpm for 90 °C and 30 minutes at room temperature;
- After that mix DBSA-doped Polyaniline (5 %) in SEBS rubber solution;
- Then again stirred solution at 1500 rpm for 90 °C and 30 minutes at room temperature;
- Pour the solution into the desired shape of patriarch glass and heat at 90 °C for 120 minutes;
- Take dielectric film and apply carbon grease on both sides and fix carbon conducting tape on both

sides of electrodes for making the connection for measuring capacitance etc. Fig 3 shows SEBS rubber and DBSA doped polyaniline-based stretchable composite film works as a dielectric medium for ionic EAP-based sensors.

Fig. 3. DBSA doped dielectric film.

For conducting the tensile testing rectangular type of specimen having an active length of 2.8 cm, width of 1.5 mm, and thickness of 0.05 mm have been used as shown in Fig. 4. All tests were conducted at 1 mm/sec speed and fitting parameters were obtained at 100 % of strain value. Five samples were used for post-treatment of data (Among five samples, three samples were showing almost similar characteristics so one sample was chosen out of those three samples for post-treatment of data).

Fig. 4. Uniaxial tensile testing specimen.

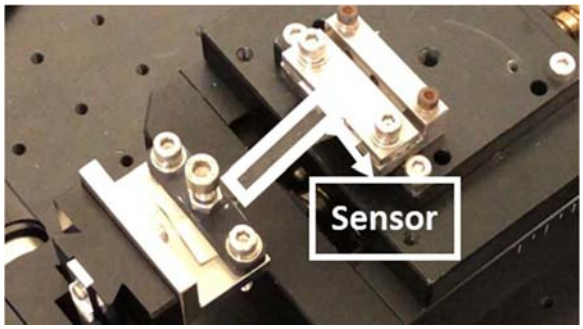


Fig. 5. Sensor fixed between two jaws of tensile machine.

Carbon grease was applied on both the upper and lower sides of the composite film and the non-conductive tape was used on both ends of the sensor for fixing the sensor between both jaws of the tensile machine. Fig .5 shows sensor is ready for uniaxial tensile testing for analyzing the mechanical behaviour of the sensor.

3. Results and Discussion

3.1. Basic Principle of Operation

When a deformable electroactive polymer film is sandwiched between highly deformable conducting electrodes on both sides of the EAP film then the whole system works like a variable parallel plate film capacitor and its capacitance changes on applying mechanical stress, which forms the basis of such type of capacitive sensing technology [15].

The capacitance of the EAP-based strain sensor under a relaxed state ($C_{relaxed}$) can be written as:

$$C_{relaxed} = \frac{\varepsilon_r \varepsilon_0 L_1 L_2}{L_3}, \quad (1)$$

where, ε_r, ε_0, are dielectric constant, the permittivity of free space, respectively and L_1, L_2, L_3 are the length, width and thickness of the sensor at relaxed state. While capacitance of EAP strain sensor in the stretched state can be written as:

$$C_{streched} = \frac{\varepsilon_r \varepsilon_0 \lambda_1 L_1 \lambda_2 L_2}{\lambda_3 L_3}, \quad (2)$$

where $\lambda_1, \lambda_2, \lambda_3$ are stretch ratio in X, Y and Z direction respectively, while l_1, l_2 and l_3 are lengths, width and thickness at stretched state, consequently:

$$\lambda_1 = \frac{l_1}{L_1}, \ \lambda_2 = \frac{l_2}{L_2}, \ \lambda_3 = \frac{l_3}{L_3}$$

Here we assume EAP strain sensor shows incompressible and isotropic behavior, consequently

$$\lambda_1 \lambda_2 \lambda_3 = 1 \quad (3)$$

$$l_3 = \frac{L_3}{\lambda_1 \lambda_2} \quad (4)$$

$$C_{deformed} = C_{relaxed} (\lambda_1 \lambda_2)^2 \quad (5)$$

In case of uniaxial deformation, $\lambda_1 = \lambda$, $\lambda_2 = \frac{1}{\sqrt{\lambda}} = \lambda_3$ consequently, we can write

$$C_{streched} = C_{\text{relaxed}} * \lambda \quad (6)$$

$$\Delta C = C_{relaxed} * \varepsilon, \quad (7)$$

where ε is the strain, while, λ is the extension coefficient and ΔC is the change in capacitance, respectively.

In Fig. 6 we can see the schematic diagram of basic principle of operation of electroactive polymer-based strain sensors (made from the elastomeric film as a dielectric medium).

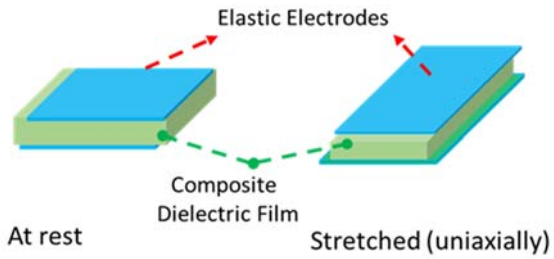


Fig. 6. Basic principle of operation of EAP -based strain sensors.

3.2. Mechanical Modelling of EAP Based Strain Sensors

The mechanical behavior of the EAP strain sensor does not follow Hook's law and nonlinear finite strain theory is used to explain EAP strain sensor behaviour as shown in Fig. 7.

Here we assumed that EAP strain sensor shows incompressible and isotropic behavior, consequently true stress T_i and strain energy potential *(W)* and stretch ratio λ_i can be linked by given equation [16-17].

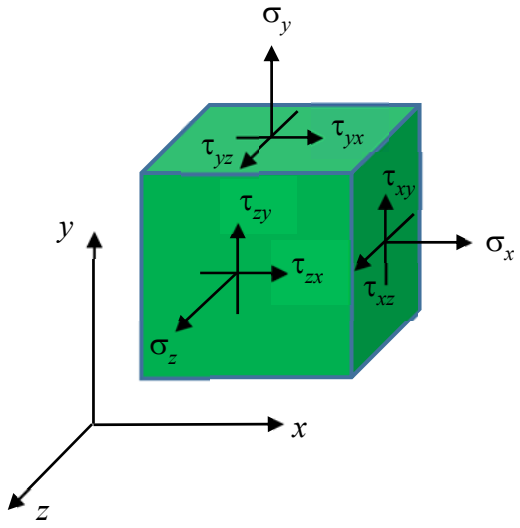


Fig. 7. Nonlinear finite strain theory for explaining EAP based strain sensor's behavior in 3D space.

$$T_i = \lambda_i \frac{\partial W}{\partial \lambda_i} - p \tag{8}$$

In above equation p is hydrostatic pressure.

Normal stress (perpendicular to a surface) can be written by equation given below:

$$\sigma_i = \frac{\partial W}{\partial \lambda_i} - \frac{1}{\lambda_i} p \tag{9}$$

The nine components of σ_{ij} can be given by the equation given as below.

$$\sigma_{ij} = \begin{pmatrix} \sigma_x & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_y & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z \end{pmatrix} \tag{10}$$

Above matrix can be written as below, where off diagonal elements represent shear terms.

$$\sigma_{ij} = \begin{bmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{bmatrix} \tag{11}$$

In the absence of normal shear stress τ_n (perpendicular to a surface), these three stresses perpendicular to these principal planes are called principal stresses. Fig. 8 shows the principal stresses and principal directions in 3D space.

$$\sigma_{ij} = \begin{pmatrix} \sigma_x & 0 & 0 \\ 0 & \sigma_y & 0 \\ 0 & 0 & \sigma_z \end{pmatrix} \tag{12}$$

According to the Cauchy's theorem:

$$T_i = \sigma_{ij} n_i \tag{13}$$

where T_i is the true stress vector at a point on a plane with normal vector n_i, λ is corresponded to the magnitude of normal stress vector (σ_n).

$$T_i = \lambda n_i = \sigma_{ji} n_i \tag{14}$$

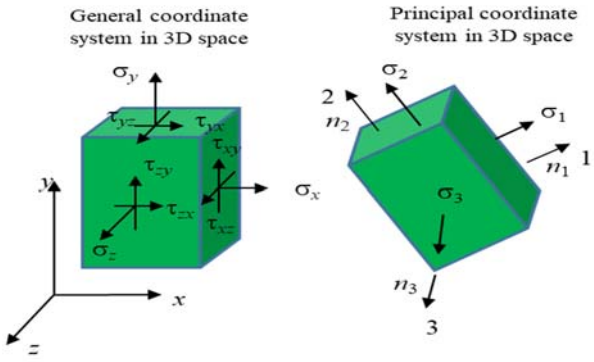


Fig. 8. Principal stresses and directions explanation using schematic diagrams.

$$\begin{aligned} &(\sigma_x - \lambda)n_1 + \tau_{xy}n_2 + \tau_{xz}n_2 = 0 \\ &\tau_{xy}n_1 + (\sigma_y - \lambda)n_2 + \tau_{yz}n_3 = 0 \\ &\tau_{xz}n_1 + \tau_{yz}n_2 + (\sigma_z - \lambda)n_3 = 0 \\ &\Rightarrow \begin{bmatrix} (\sigma_x - \lambda) & \tau_{xy} & \tau_{xz} \\ \tau_{xy} & (\sigma_y - \lambda) & \tau_{yz} \\ \tau_{xz} & \tau_{yz} & (\sigma_z - \lambda) \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \\ &= 0 \end{aligned} \tag{15}$$

$$\begin{aligned} &\begin{vmatrix} (\sigma_x - \lambda) & \tau_{xy} & \tau_{xz} \\ \tau_{xy} & (\sigma_y - \lambda) & \tau_{yz} \\ \tau_{xz} & \tau_{yz} & (\sigma_z - \lambda) \end{vmatrix} = 0 \\ &\Rightarrow -\lambda^3 + I_1\lambda^2 - I_2\lambda + I_3 \\ &= 0 \end{aligned} \tag{16}$$

This is for a homogeneous system of three linear equations, where n_i are the unknowns. Therefore, I_1, I_2 and I_3, called the first, second, and third stress invariants, respectively, can be given by following equations

$$I_1 = Tr|\sigma_{ij}| = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \quad (17)$$

$$I_2 = \frac{1}{2}\left[Tr(\sigma_{ij})^2 - Tr(\sigma_{ij}{}^2)\right] = \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2 \quad (18)$$

$$I_3 = det(\sigma_{ij}) = 1 \quad (19)$$

Different type of constitutive model for defining mechanical behavior of elastomers can be explained in terms of strain energy density (W), Invariants and coefficients as given below [18-21]:

$$W_{Mooney-Rivlin} = C_{10}(I_1 - 3) + C_{01}(I_2 - 3) \quad (20)$$

$$W_{Yeoh} = \sum_{i=1}^{N} C_{i0}\,(I_1 - 3)^i \quad (21)$$

$$W_{Neo-Hookean} = \frac{\mu}{2}(I_1 - 3) \quad (22)$$

$$W_{Ogden} = \sum_{i=1}^{N} \frac{\mu_i}{\alpha_i}(\lambda_1{}^{\alpha_i} + \lambda_2{}^{\alpha_i} + \lambda_3{}^{\alpha_i} - 3), \quad (23)$$

where λ (extension coefficient) =1+ε (strain)

Table 1. Matrix representation of strain vectors and formulas for calculating invariants in various types of loading.

Loading type	Uniaxial	Pure shear	Biaxial
Strain vector	$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \frac{1}{\sqrt{\lambda}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{\lambda}} \end{bmatrix}$	$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\lambda} \end{bmatrix}$	$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \frac{1}{\lambda^2} \end{bmatrix}$
Invariants (I_1 & I_2)	$I_1 = \lambda^2 + \frac{2}{\lambda}$ $I_2 = 2\lambda + \frac{1}{\lambda^2}$	$I_1 = I_2 = 1 + \lambda^2 + \frac{1}{\lambda^2}$	$I_1 = 2\lambda^2 + \frac{1}{\lambda^4}$ $I_2 = \lambda^4 + \frac{2}{\lambda^2}$

Table 1 shows the matrix representation of strain vectors for different kind of loads and equations for calculating invariants at different kind of loading conditions. With the help of Table 1 we can easily calculate the strain energy density for different kind of constitutive equations, as explained above (from equation 20 to 24) and further we can also establish relationship between stress and strain for any kind of constitutive equation by using equation 9.

3.3. Mechanical Characterization of EAP Strain Sensor

Force displacement characteristics as shown in Fig. 9, we can conclude that sensor shows purely elastic behavior as cycles are repetitive, following same trend and stress is almost constant in each cycle.

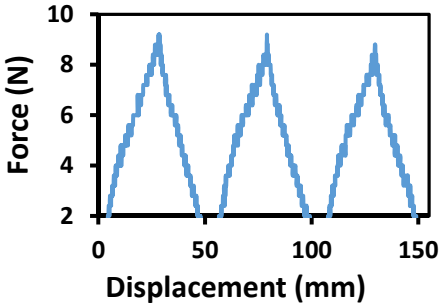


Fig. 9. Force displacement characteristics of Ionic EAP based strain sensor.

From Fig. 10 we can see that sensor is highly stretchable and can be elongated more than 400 % of strain value.

For elongation at break test sample's length, width, thickness was 2 cm, 1cm, 0.2mm respectively.

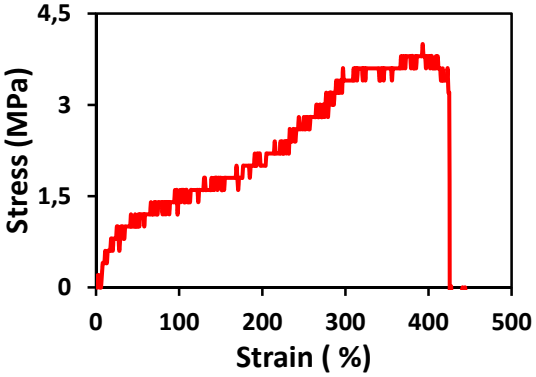


Fig. 10. Elongation at break test of EAP strain sensor.

3.4. Predicting Stress Strain behavior of EAP Strain Sensor under Multiaxial Loads

First stretching cycle up to 100 % applied strain was used for the mechanical modelling using Ogden based constitutive model and it can be seen from the Fig. 11 that there was good agreement between the experimental data and fitting of the data using the Ogden 2nd terms constitutive equation (in stress versus strain form). Fitting parameters (less than 10 % deviation for 5 samples) are summarized in the Table 2. For curve fitting cftool (in Matlab), hyperfit or abaqus can be used. Here curve fitting was done using Hyperfit software http://www.hyperfit.wz.cz(Ansys-compatible original form).

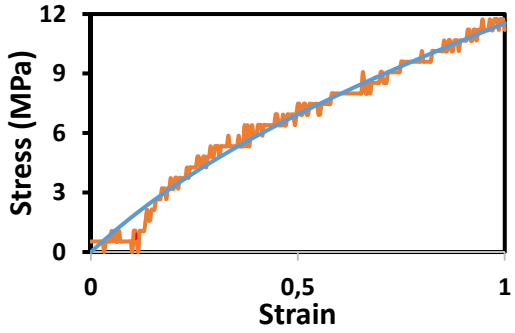


Fig. 11. Fitting of experimental data with Ogden Modelling for ionic type of EAP strain sensor.

Table 2. Ogden fitting parameters.

Fitting parameters	Ogden 2nd order
μ (MPa)	6.51,-0.02
α (MPa)	2.02, 3.18

With the help of fitting parameters, we can predict the stress-strain behavior (uniaxial loading) of the sensor at any value of strain by using equation 23, Table 1 and section 3.2 (as shown in figure Fig. 11 and Table 2). Similarly, we can also use these fitting parameters in predicting stress in biaxial and planar loading as well.

4. Conclusions

Sensor was highly stretchable and can be stretched up to 400 % of strain value. From uniaxial tensile experiments, a three-dimensional constitutive model of the EAP strain sensor was validated. Sensor was showing good agreement with Ogden's 2nd terms and these fitting parameters obtained from uniaxial tensile data can also be used for predicting sensor's behavior under multiaxial loading. This type of analytical modelling can be useful in designing elastomer-based strain sensor for various loading conditions.

Acknowledgement

One of the authors **Nitin Kumar Singh** would like to thank ministry of education, culture, sports and Science (MEXT), Japan for providing the scholarship.

References

[1]. N. K. Singh, K. Takashima, and S. S. Pandey, Tear strength estimation of electroactive polymer-based strain sensors, in *Electroactive Polymer Actuators and Devices (EAPAD) XXIV*, J. D. Madden, I. A. Anderson, and H. R. Shea, Long Beach, United States, Mar. 6–Apr. 11, 2022. SPIE, 2022.

[2]. N. K. Singh, K. Takashima, and T. Shibata, Dielectric elastomer based stretchable textile sensor for capturing motion, in *Electroactive Polymer Actuators and Devices (EAPAD) XXII*, Y. Bar-Cohen, I. A. Anderson, and H. R. Shea, Eds. Online Only, United States, Apr. 27–May 8, 2020. SPIE, 2020.

[3]. X. Liu et al., Dielectric elastomer sensor with high dielectric constant and capacitive strain sensing properties by designing polar-nonpolar fluorosilicone multiblock copolymers and introducing poly(dopamine) modified CNTs, *Composites Part B: Engineering,* Vol. 223, 2021, p. 109103.

[4]. N. K. Singh, K. Takashima, and S. S. Pandey, Electronic versus Ionic Electroactive Polymers (EAPs) Strain Sensors for Wearable Electronics: A Comparative Study, *Eng. Proc.,* 21, 1, 2022.

[5]. N. K. Singh, K. Takashima, and S. S. Pandey, Fatigue life prediction of electroactive polymer strain sensor, *Proceedings Volume 11587, Electroactive Polymer Actuators and Devices (EAPAD) XXIII*, 2021, 115871Z.

[6]. N. K. Singh, K. Patra., A. Kumar, Energy harvesting using Dielectric Elastomer, in *Proceedings of the International Conference on Advanced Materials for Power Engineering (ICAMPE-2015, Kerela, India),* 2015.

[7]. P. C. Binh, D. N. C. Nam, K. K. Ahn, Modeling and experimental investigation on dielectric electro-active polymer generator, *International Journal of Precision Engineering and Manufacturing*, Vol. 16, No. 5, May 2005, pp. 945–955,

[8]. N. K. Singh., K. Patra, Energy harvesting using dielectric elastomer, Master thesis, *Indian Institute of Technology Patna*, India, 2014.

[9]. K. Ozlem, O. Atalay, A. Atalay, and G. Ince, Textile based sensing system for lower limb motion monitoring, in Converging Clinical and Engineering Research on Neurorehabilitation III, *Springer International Publishing,* Cham, 2018, pp. 395–399.

[10]. Y.-D. Tao, G.-Y. Gu, and L.-M. Zhu, Design and performance testing of a dielectric elastomer strain sensor, *International Journal of Intelligent Robotics and Applications,* Vol. 1, No. 4, 2017, pp. 451–458.

[11]. J.-W. Zhang, Y. Zhang, Y.-Y. Li, and P. Wang, Textile-Based flexible pressure sensors: A review, *Polymer Reviews*, 2021, pp. 1–31.

[12]. K. Jung, K. J. Kim, and H. R. Choi, A self-sensing dielectric elastomer actuator, *Sensors and Actuators A: Physical,* Vol. 143, No. 2, 2008, pp. 343–351.

[13]. N. K. Singh, K. Takashima, and S. S. Pandey, Electrical characterization and analytical modeling for predicting stress-strain behavior of electroactive polymer-based sensors under multiaxial loads, in *Proceedings of the 8th International Conference on Sensors and Electronic Instrumentation Advances (SEIA' 2022),* Corfu, Greece, 21-23 September 2022, pp. 28-30 (accepted in June 2022).

[14]. A Qiu, Q Jia, H Yu, J. A. Oh, D Li, H. Y. Hsu, and J. Ma, Highly sensitive and flexible capacitive elastomeric sensors for compressive strain measurements, *Materials Today Communications*, 26, 2021, 102023.

[15]. N. Ni and L. Zhang, Dielectric Elastomer Sensors, in Elastomers, *InTech*, 2017.

[16]. X. Chen, N. Laforge, Q. Ji, H. Tan, J. liang, Introduction to mechanical metamaterials and their effective properties, *Computes Rendus. Physique*, November 2020, pp. 1–15.

[17]. Z. Suo, Theory of dielectric elastomers, *Acta Mechanica Solida Sinica,* Vol. 23, No. 6, 2010, pp. 549–578.
[18]. Y. Bai, C. Liu, G. Huang, W. Li, and S. Feng, "A Hyper-Viscoelastic Constitutive Model for Polyurea under Uniaxial Compressive Loading," *Polymers*, vol. 8, no. 4, p. 133, April 2016.
[19]. S. D. Barforooshi and A. K. Mohammadi, Study Neo-Hookean and Yeoh Hyper-Elastic Models in Dielectric Elastomer-Based Micro-Beam Resonators, *Latin American Journal of Solids and Structures*, Vol. 13, No. 10, 2016, pp. 1823–1837.
[20]. A. A. Abdelsalam, S. Araby, M. A. Hassan, and A. A. El-Moneim, "Constitutive modelling of elastomer/graphene platelet nanocomposites," *IOP Conference Series: Materials Science and Engineering*, Vol. 244, p. 012016, September 2017.
[21]. R. W. Ogden (1972). Large deformation isotropic elasticity–on the correlation of theory and experiment for incompressible rubberlike solids. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 326 (1567), pp. 565-584.

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 19-28

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Fibre Bragg Grating Based Optical Sensors Accuracy Improvement

[1, *] František URBAN jr., [2] Radek HELÁN and [3] František URBAN sr.

[1] Brno University of Technology, Faculty of Electrical Engineering and Communication, Dept. of Telecommunications, Technická 12, Brno, 61600, Czech Republic
[2] Network Group s.r.o., Dept. SFO, Olomoucká 91, Brno, 627 00, Czech Republic
[3] Brno University of Technology, Faculty of Electrical Engineering and Communication, Dept. of Microelectronics, Technická 10, Brno, 61600, Czech Republic
[1] Tel.: + 420736625804
E-mail: f.urban@nwg.cz

Received: 1 September 2022 /Accepted: 5 October 2022 /Published: 31 October 2022

Abstract: Fibre optic based sensors are becoming very successful tool for accurate and perturbation insensitive sensing of many non-electric and electric values. Favorite fibre sensoric structure capable of point and semi-distributed sensing of all effects influencing physical dimensions or refractive index of a fibre waveguide is the Fibre Bragg Grating. An interrogation of the grating reflectance spectrum is the important part of the fibre grating based sensorial system that can significantly affect the facilities of the measuring system and influence the applicability of the technology. The article analyses the spectrum scanning type of the grating sensor interrogation. We expose two main impacts that influence the accuracy of the measurement of the fibre grating spectral wavelength shift that bears the measured value. Both of them are connected with the quantity of the Grating Quality factor that we have defined for this purpose. The first impact refers to the way of scanning the wavelength spectrum of the sensorial grating and shows that the optimal measuring results are obtained when using the scanning filter wavelength spectrum width equal to the wavelength spectrum width of the sensorial fibre Bragg grating. The second impact lies in finding the grating design assuring the minimal size of the noise grating linewidth, also defined in this work. The work shows how the application of the specific apodization schemes to the fibre Bragg grating design can improve the noise grating linewidtg. Analysis of two suitable apodization profiles shows that the specific apodization of gratings can significantly improve the measuring accuracy of the scanning spectral interrogation of the grating based optical sensors.

Keywords: Fibre Bragg grating, Spectrum scanning, Spectrum linewidth, Apodization profile.

1. Introduction

Fibre Bragg Gratings (FBG) are the fibre longitudinal dispersive structures mostly formed as the periodical or quasi-periodical changes of the refractive index in the core of the singlemode optical fibre (SMF). The structure acts as the selective optical filter that reflects light in a narrow band of the wavelengths around the central wavelength λ_B given by

$$2\Lambda_G = s.\lambda_B/n_{eff}\ ; \quad (1)$$

where s is the integer, Braggs order; Λ_G is the Braggs period of the refractive index; n_{eff} is the effective refractive index of the waveguide with the grating.

Due to this reflection, the transmitted light lacks the part of the spectrum reflected from the grating. The reflected spectrum has the potential to be influenced by all the physical agents that can change either the

period Λ_G or the effective refractive index n_{eff} of the grating. The values that we speak of are namely temperature of the fibre, strains, bends, twists of the fibre, see Fig. 1. The induced spectrum changes can mostly be represented by the shift of the maximum of the reflexivity (minimum of the transmissivity) spectrum of the affected grating that is usually in a very accurately established relation to the measured agent.

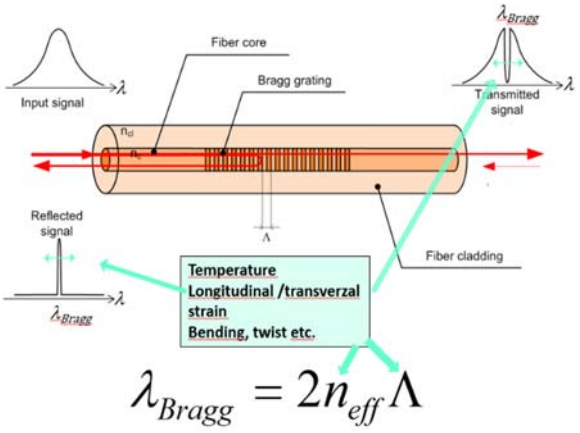


Fig. 1. Fibre Bragg Grating principle.

There is a number of methods for the evaluation of the grating spectral shift that can solve, in a compromising way, the tasks of the individual applications. One of the methods with the potential of a high resolution and a broad measuring range at the same time is the spectrum scanning.

In this way of interrogation, the source of a tuneable light is connected to the fibre isolator and then to the 50:50 symetrical coupler that sends half of the incident light power to the arm with the interrogated series of gratings and the other half of the power is sent to the arm with the reference detector. The light reflected from the FBGs is then forwarded back to the 50:50 symetrical coupler and through it to the measuring detector whose signal is then normalized using the reference detector received value. The maximum of the received power time funstion is then located and the time instant of the reception of the maximum is recalculated to the wavelength value. The structure of this interogation system is shown in the Fig. 2. In the application practice, two kinds of spectrum scanning are mostly used. The first one is based on the wavelength tuning of the narrow band optical source, often distributed feedback based laser diode (DFB), with its typical value of -20 dB drop spectral bandwidth and linewidth, $Df_{s\text{-}20} < 100$ MHz and $Dl_{s\text{-}20} \cong 0.8$ pm. The limitation of this approach is usually the tuning range and speed of the laser diode tuning. Standart telecom grade cost saving DFBs can give the modest tuning range of 10 nm in the O or C band. The second option is to scan the FBG spectrum with the light of the broadband source like superluniniscent light emiting diode (SLED) that is filtered by the tunable narrowband Fabry-Perot (FP) filter. That allows for increasing the scanning frequency to the kilohertz range and tuning range up to 100 nm, which then allows using more then 20 FBGs connected to the scanning interrogator in series. The tunable Fabry-Perot filters, however, give the one up to two orders of magnitude wider spectrum line, with the linewidth in half maximum $\Delta\lambda_{SFWHM}$ around: $\Delta\lambda_{SFWHM} \cong 60$ pm, i.e. $\Delta\lambda_{S\text{-}20} \cong \Delta\lambda_{SFWHM}/0.603 = 99.5$ pm. See Fig. 2.

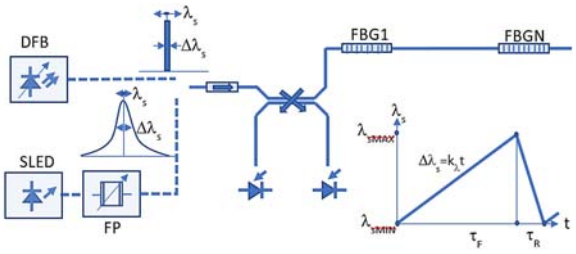


Fig. 2. Scaning interrogation scheme.

2. Spectrum Scanning and Noise

Scanning of the FBG spectrum produces the time domain photoelectric signal at the detector corresponding to the reflection spectrum of FBG. However, the time signal is not the fully linear copy of the FBG spectrum. The main peak of the time signal $I_m(t)$,

$$I_m(t) = \int_{min}^{max} Rg(\lambda).Is(\lambda - \Delta\lambda s(t)).d\lambda, \qquad (2)$$

where $\Delta\lambda_s(t) = k_\lambda.t$ is the wavelength shift of the scaning spectrum; $R_g(\lambda)$ is the spectral light reflectance of FBG; $I_s(\lambda)$ is the scanning light spectral intensity; received at the measuring detector, has its full width half maximum value different than the original FBG spectrum. It is altered by the scanning, depending on the linewidth of the scaning spectrum. The results of the scanning with the scanning spectrum FWHM equal to and wider than the FWHM linewidth of the original grating are shown in Fig. 3. The original grating design parameters of the grating represented in the Fig. 3 is L=11 mm, R_{max}= 85.6 %, $\Delta\lambda_{GFWHM}$ = 82.46 pm. Refractive index profile of this grating is the uniform sinusoid with the peak-to-peak amplitude of the modulation $\Delta n = 1.25\times10^{-4}$ and period of 450 nm superimposed to the constant level of the core refractive index of 1.468, mesured at the wavelength of 1310 nm. The graphs show that the wavelength position of the reflectance maximum stays untouched (because of the symetricity of both the grating and the scanning source spectra), but we see the change of the scanned spectra FWHM linewidths and also the change of the sharpness of the spectral maxima. That can be important for the noise influence in detecting the spectral maximum of the grating reflectance.

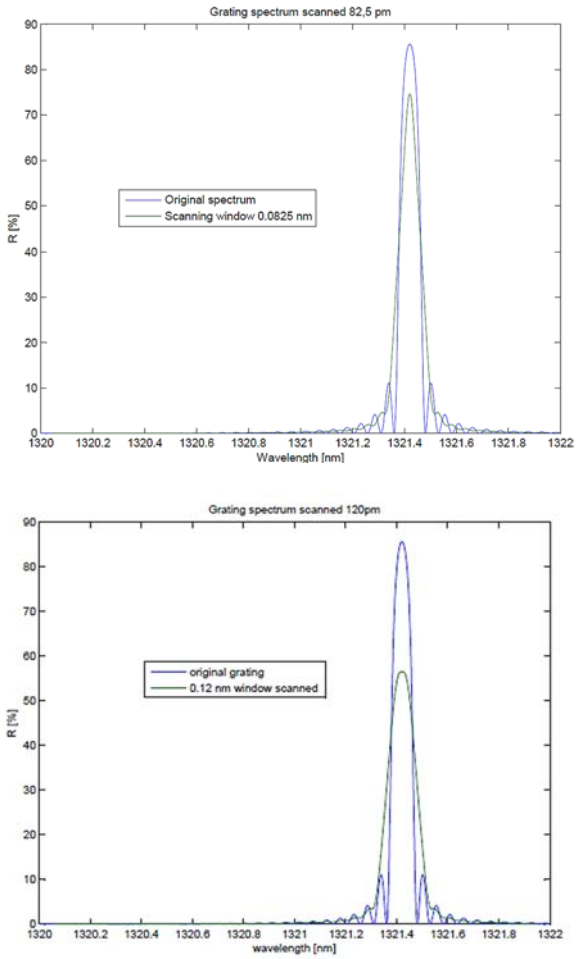


Fig. 3. Original and scanned spectrum of the uniform sinusoidal grating scanned with 82.5 pm (upper plot) and 120 pm (lower plot) scanning window.

The accuracy of the detection of the wavelength position of maximum grating reflexivity comes from the resolution of the wavelength measurement. The resolution of the wavelength measurement $\delta\lambda_G$ in this method is determined by the scanning period T_{SC}, period T_{SA} of sampling the time signal $I_m(t)$ and the scanning wavelength range λ_{Smax}-λ_{Smin}. It can be very small in case of fast sampling:

$$\delta\lambda_G=(\lambda_{Smax}-\lambda_{Smin}).\ T_{SA}\ /\ T_{SC} \le 0.01\ \text{pm} \qquad (3)$$

The time function $I_m(t)$ is influenced by the two fundamental factors. The first one is the signal to noise ratio seen at the measuring detector circuit, the second one is the linewidth of the central maximum of the reflectance spectrum of the grating. The noise that is superimposed on the time signal of the scanned grating influences the correct determination of the position of scanned grating reflexivity spectrum. There are few sources of noise in the scanning system, detector dark current noise I_D, thermal noise of the detector load circuit I_{TH} and the detector shot noise I_{SH}. The brief analysis shows that the dominating factor here is a relatively high level of received optical signal comparing to the usual levels seen in the fiber datacommunication systems. The practical example with the typical values of the receiver bandwidth BW = 1 MHz, the transimpedance load rezistor R_C = 50 kΩ, detector dark current I_D = 1 nA, detector photoelectric current for the 100 % grating reflectance I_0 = 50 μA (corresponding to the received powers in order of 50 μW) shows the RMS values of noise currents:

$$I_D=\sqrt{(2e.I_D.BW)} = 30\ \text{pA};\ I_{TH}=\sqrt{(4k.T.BW/R_C)} = 0.57\ \text{nA};\ I_{SH}=\sqrt{(2e.I_0.BW)} = 4\ \text{nA}. \qquad (4)$$

Therefore $I_N \cong I_{SH}$ = 4 nA. Signal to noise ratio then depends on the input optical power, represented by the detector photoelectric current I_0, on grating reflectance R_{max} and on receiver bandwidth BW:

$$I_S/I_N=(I_0.\ R_{max})/\sqrt{(\ 2e.I_0.R_{max}.BW)} = = K.\sqrt{R_{max}} \cong 10^4 \qquad (5)$$

From the point of view of the grating design, we see that the important factor to minimize noise is the square root of the grating reflectance $\sqrt{R_{max}}$. Further, our application example shows that the noise root mean square I_N is in the order of 10^{-4} of the signal current at the spectral top of the grating reflexivity. The superimposed noise on the detected optical power reflexed from the scanned grating gives the origin of the rising errors in determining the time and wavelength position of the reflexing spectral maximum, see Fig. 4.

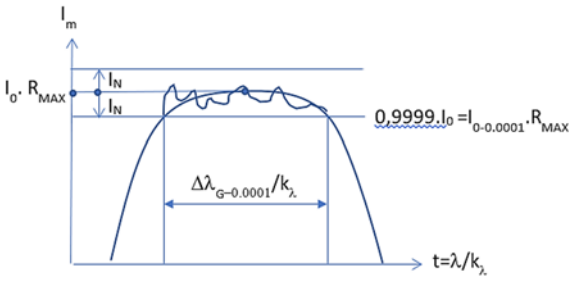


Fig. 4. Effect of superimposed noise to the location of reflectance maximum.

We can say, that the detected maximum in the current signal, composed of the light reflected from the scanned grating and superimposed noise current with the root mean square value $I_N=10^{-4}I_0$, will be most probably found in the time interval $\Delta\lambda_{G-0.0001}/k\lambda$ or in interval will be called the grating noise linewidth. It can be shown that the root mean square deviation σ of the location of detected maximum from its central position, in other words the mean error of the measurement is:

$$\sigma = \Delta\lambda_{G-0.0001}/(2.\sqrt{6}) \qquad (6)$$

The grating noise linewidth value $\Delta\lambda_{G-0.0001}$ is bound to the value of grating full width half maximum (FWHM) spectral linewidth $\Delta\lambda_{GFWHM}$. In the low reflectance uniform gratings we can work with the estimation $\Delta\lambda_{G-0.0001} \cong 0.022.\Delta\lambda_{GFWHM}$ which loses the fidelity for the cases of higher than 50 % reflectances and lower than 3 mm of lengths. Nevertheless, the next important factor in the grating design affecting the errors in the measurement of the position of maximum reflectance is the grating spectral linewidth in the half maximum $\Delta\lambda_{GFWHM}$. These two parameters, $\sqrt{R_{max}}$ and $\Delta\lambda_{GFWHM}$ are not independent in the grating design. As for that we define the grating quality factor Q_G:

$$Q_G = \sqrt{R_{max}}/\Delta\lambda_{GFWHM}; [-;\%,nm] \quad (7)$$

that depicts the suitability of the grating design for the accurate scanning interrogation, se Fig. 5. The plot here shows, that the best results in the accurate scanning interrogation of uniform gratings can be obtained for the grating lengths longer than 10 mm and the reflectance in spectral maximum around 60 %.

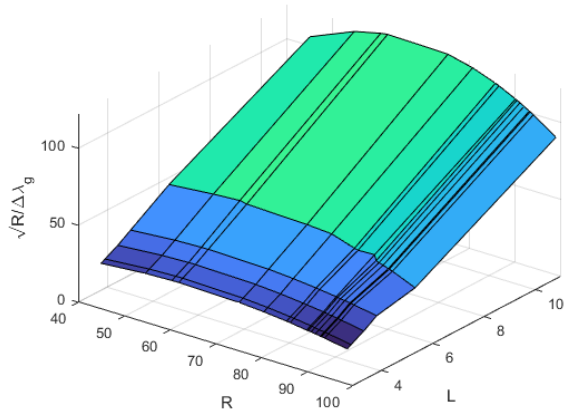


Fig. 5. Graph of the grating quality factor $Q_G = \sqrt{R_{max}}/\Delta\lambda_{GFWHM}$ on the grating reflectance R[%] and the grating length L[mm].

3. Grating Features

As we already stated, the maximum reflectance of the grating R_{max} and the reflectance spectral linewidth in the half maximum $\Delta\lambda_{GFWHM}$ of the grating are values mutually bonded. In the uniform gratings, the values of these two parameters relate to the physical length of the grating in the fibre. The graphs of Fig. 6 depict the relation of the grating half maximum linewidth $\Delta\lambda_{GFWHM}$ and the grating noise linewidth $\Delta\lambda_{G-0.0001}$to the typical lengths L of the gratings and to the reflectances R_{max}. We can see there that the value of FWHM linewidth of the uniform grating monotonically drops with increase of the grating length and with decrease of the maximal reflectance. To the contrary, the grating noise linewidth of the same uniform grating design shows the different behaviour. Whille the noise linewidth drops monotonically with the grating length, similarly like it is with FWHM linewidth, the shape of the dependence of the grating noise linewidth on the grating maximal reflectance shows the minimum close to the reflectivity of the 60%. Further, it can be seen, that the grating noise linewidth rises strongly for the grating lengths shorter than 5 mm and reflectances higher than 80 %.

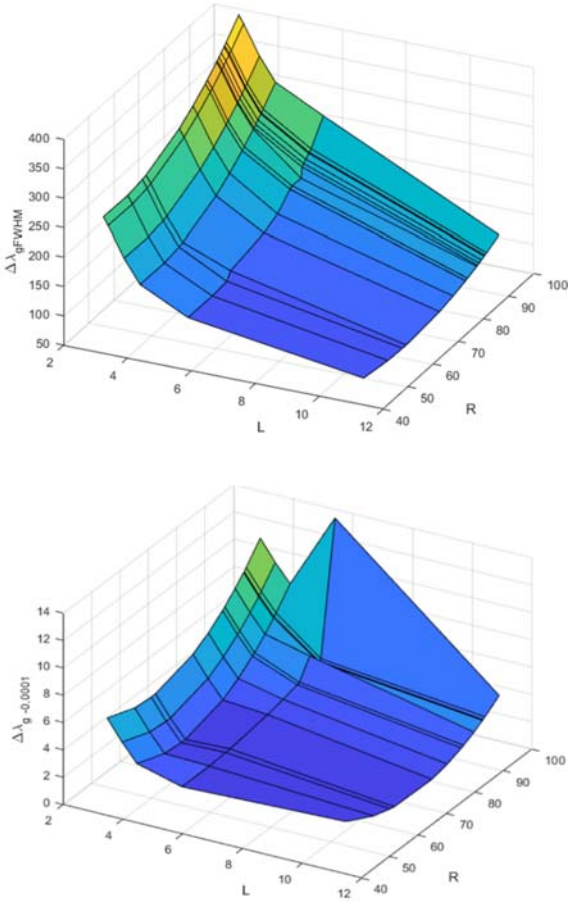


Fig. 6. Grating linewidth $\Delta\lambda_{GFWHM}$ [pm] and grating noise linewidth $\Delta\lambda_{G-0.0001}$ [pm] depending on the grating length L[mm] and the grating reflectance R[%].

When scanning the grating spectrum to locate the wavelength position of the reflectance maximum with use of an extra narowband tunable light source (DFB), $\Delta\lambda_{SFWHM} \leq 1$ pm, the time signal $I_m(t)$, eq. (2), is not distorted and linearly copies the spectral grating reflectance:

$$I_m(t) = I_{S0}.R_G(k_\lambda.t), \quad (8)$$

where I_{S0} is the photoelectric response to the light intensity of the light source; $R_G(k_\lambda.t) = R_G(\lambda)$ is the spectral reflectance of the sensorial grating.

The optimal values of the maximal grating reflectance to get the minimal noise linewidth derived

from Fig. 5 gives for the grating length L=11 mm, $R_{max} \cong$ 60.5 %. Then, for this way optimized uniform grating we get the FWHM grating linewidth $\Delta\lambda_{GFWHM}$= 63.8 pm, grating noise linewidth $\Delta\lambda_{G-0.0001}$= 1.3 pm and the mean error of establishing the wavelength of the maximum reflectance will be, according to the eq. (6), σ = 0.265 pm.

In the case of grating spectrum scanned by the tunable FP filter with the scanning spectrum FWHM linewidth similar to the spectrum linewidth of the grating itself, in spite of the shape of spectral intensity of the scanning light $I_S(\lambda)$ is similar to $R_G(\lambda)$, the linewidth and the shape of the signal $I_m(\lambda)$ is different from $R_G(\lambda)$. Analyzing this effect we find that the minimal value of the noise linewidth can be obtained when using the grating and the scanning filter with the equal value of $\Delta\lambda_{FWHM}$:

$$\Delta\lambda_{SFWHM} = \Delta\lambda_{GFWHM} \tag{9}$$

For the signal to noise ratio 10^4 we see the situation at the graph in Fig. 7, where the selected suboptimal sensual grating with L=11 mm, $R_{max} \cong$ 85.6 % and $\Delta\lambda_{GFWHM}$ = 82.46 pm gives the minimal value of scanned noise linewidth $\Delta\lambda_{GS-0,0001}$= 0.99 pm and the RMS error σ = 0.20 pm.

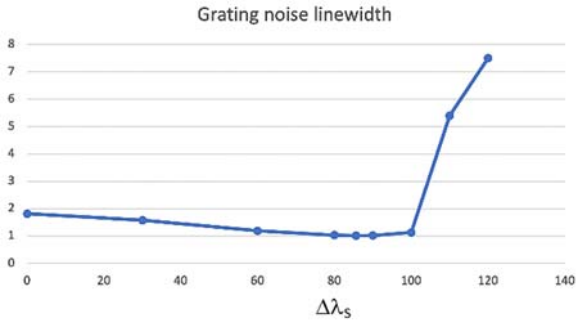


Fig. 7. Scanned noise linewidth $\Delta\lambda_{GS-0.0001}$ [pm] of the grating $\Delta\lambda_{GFWHM}$=82.5 pm scanned by the scanning spectrum FWHM linewidth $\Delta\lambda_S$ [pm].

Interesting remark can be made here, the obtained scanned noise linewidth of 0.99 pm is nearly half of the original grating noise linewidth $\Delta\lambda_{G-0.0001}$ = =1.82 pm. If the grating is scanned with the ultra narrowband tuneable source it gives the scanned noise linewidth $\Delta\lambda_{GS-0.0001}$= $\Delta\lambda_{G-0.0001}$= 1.82 pm and RMS error σ = 0.37 pm. Another feature of the scanning shows that the scanned noise linewidth rises abruptly when the sensual grating is scanned with the scanning spectrum FWHM linewidth higher that approx. 1.2 times the grating spectrum FWGM linewidth. Generalizing these findings, there are thre intervals for scanning source linewidth important from the point of the accuracy in the scanning interogation of the FBG based sensors:

$$\begin{aligned} &\Delta\lambda_{SFWHM} \ll \Delta\lambda_{GFWHM};\ \Delta\lambda_{GS-0,0001} = \Delta\lambda_{G-0.0001} \\ &\Delta\lambda_{SFWHM} \cong \Delta\lambda_{GFWHM};\ \Delta\lambda_{GS-0.0001} \cong 0.5.\ \Delta\lambda_{G-0.0001} \\ &\Delta\lambda_{SFWHM} > \Delta\lambda_{GFWHM};\ \Delta\lambda_{GS-0,0001} \gg \Delta\lambda_{G-0.0001} \end{aligned} \tag{10}$$

Progress in the grating production technology shows some other ways to further improve the accuracy of the measurement. The way lies in the modification of the grating structure by the apodization process. Simulation and the practical experiments have shown that we can decrease, by this way, the noise linewidth of the gratings noticeably below the limits achievable by the uniform gratings.

4. Apodized Gratings

Apodized grating is the grating structure where the amplitude of the periodic change of the refractive index along the fiber axis is modulated by the specific function. Well known are the apodized gratings with cos(x), $\cos^2$(x), sinc(x), Gaussian and other apodizing functions. Conventional use of the apodization is to suppress the side lobes of the grating spectrum and to increase the gradient of the sides of the main lobe of the reflectance spectrum.

Our goal with the apodization was narrowing the main peak of the grating reflectance spectrum with tolerance to the possible increase of the spectrum side lobes. We have investigated number of possible apodization schemes, but finally, we have limited the choice to the variation of cos(kx), sin(kx) and $\cos^2$(kx), $\sin^2$(kx) index amplitude apodization as these profiles showed up to be most suitable from the grating production point of view. FBG manufacturing process used for this work is based on the fibre side illumination through the grating phase mask, as seen in the Fig. 8(a). Fig. 8(b) shows the grating exposition workplace where the powerful excimer pulse UV KrF laser is used as the coherent light source, providing enough coherence length, that the two diffracted beams of +1 and -1 order generated at the phase mask can interfere and create the periodic interference pattern that spreads up to 0.5 mm below the phase mask periodically corrugated surface. This UV interference pattern is then projected to the stripped and hydrogenated fibre. Being absorbed in the fiber, the UV interference pattern maxima increase the refractive index in the core where the fiber contains small quantity of GeO_2 and H_2 as the photosensitive agents. This way the periodic refractive index variations along the fiber axis is formed. The refractive index modulation amplitude corresponds to the total UV exposition energy.

There are two principal ways to change the refractive index variations along the fiber axis and to provide the apodization for the created grating. The first one is based on the use of the intensity filter in the path of the UV beam irradiating the phase mask, as shown in the Fig. 8 a). The peak-to-peak refractive index variation induced in the fibre core exposed to the UV interference pattern under the phase mask is

proportional to the local intensity of the UV field above the phase mask and therefore proportional to the local transparency of the intensity filter. The example of the resulting index variation obtained by this method is called an impure apodization, see Fig. 8 c). It gives the index variation profile that keeps the amplitude of the variation according to the UV transparency of the filter but the mean value of the index variation and the effective refractive index n_{eff} of the fibre waveguide is not constant along the grating and changes also according to the filter transparency. That facility of the impure apodization leads to the effect of grating sections detuning and the grating partially behaves like the chirped grating. That method proved to give unsatisfactory results for our goal.

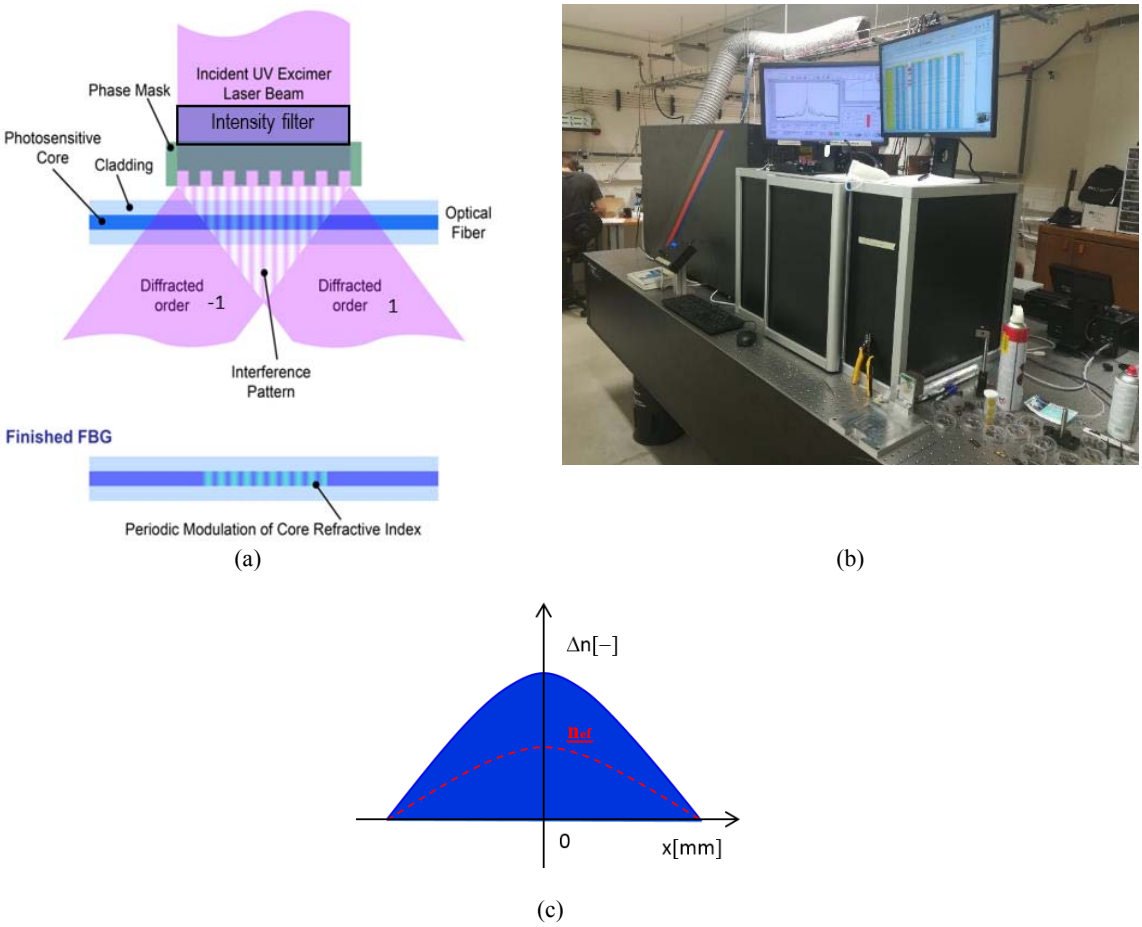


Fig. 8. a) Fibre side illumination through the grating phase mask; b) Grating exposition workplace; c) An example of index variation in the impure apodization.

The second option is to get the apodization profile by use of multiple exposition of the same interference pattern to the stretched and shifted fibre. The exposition setup is schematically shown in the Fig. 9. Here, after the first exposition of the grating component, the exposed fibre is stretched and axially shifted from its original position and subjected to the second exposition. That second step creates so called Moare effect in the overlapping of two periodic grating with slightly different grating period. This way the third and further steps can be made. Repeatability of the production limits that approach to the maximum of three consequent exposition steps. This method allows to achieve the apodization profiles based on the index amplitude functions cos(kx), sin(kx) and $\cos^2(kx)$, $\sin^2(kx)$. An important fact here is that the obtained apodization profiles can be made in the form of pure apodization. That means, the index variations always oscilate symmetrically around the constant mean value, as shown in the examples on Fig. 10, Fig. 11 and Fig. 12. With this design, the individual grating sections along the grating are not detuned from the central wavelength and the grating reflectance spectrum of the resulting structure is not widened.

Most promising apodization function from the scope of achieving the narrowest main spectral reflectance peak and the noise linewidth proved to be the index variation:

$$\Delta n(x) = n_c + \Delta n_{ph}(1 + (\sin^2(A.\pi x/L)+B).\sin(2\pi x/\Lambda_G)) \quad (11)$$

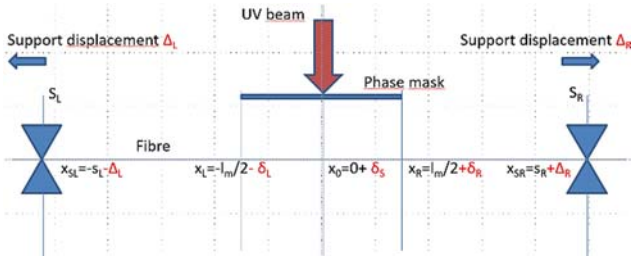


Fig. 9. Multiple exposition apodization scheme.

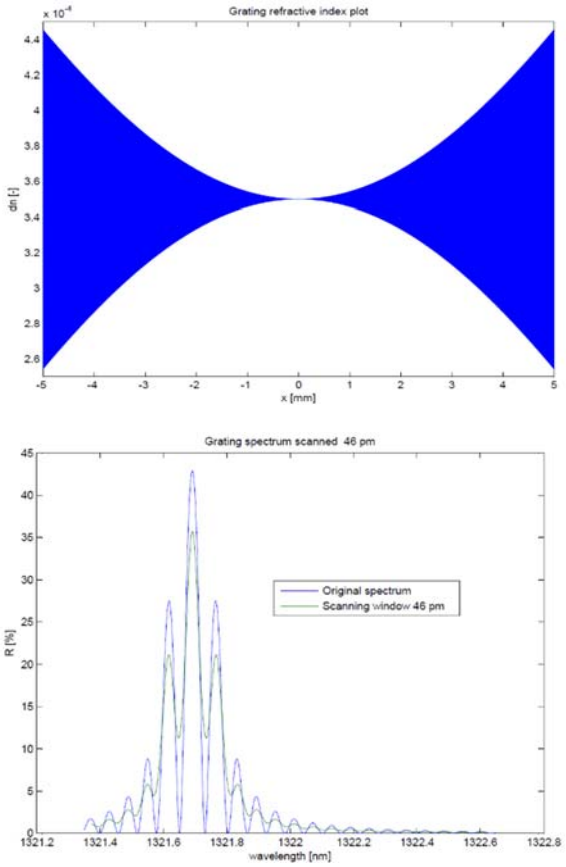


Fig. 10. Type 1 pure apodization grating design 2250-14.

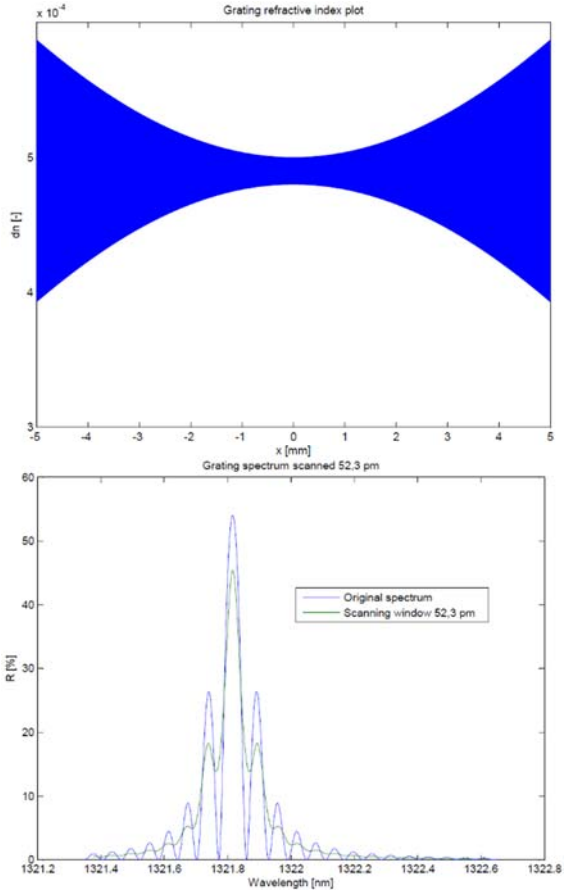


Fig. 11. Type 1 pure apodization grating design 2250-13.

The pure apodization index function from Eq. 11 can be formed by the three steps exposition process. In a complex proces of optimizing the shape coefficients Δn_{ph}, A and B, we found interesting and promissing results that have the potential for further refining. Partial results of the apodized grating design are shown in the Fig. 10 and Fig. 11. Here, both the presented gratings have the active length of 10 mm. The Fig. 10 represents the apodization grating design 2250-14 that achieved the narrowest noise linewidth of 0.912 pm, having though quite high side lobes that complicates the correct and errorless scanning process. The side lobes supression in this design is only 1.93 dB. After scanning with the optimized scanning spectrum $\Delta\lambda_{SFWHM} \cong \Delta\lambda_{GFWHM} = 46$ pm, the scanned noise linewidth is 0.495 pm and the scanned spectrum side lobe suppression 4 dB. The Fig. 11 represents the apodization grating design 2250-13 with side lobes supression (SLS) of 3.11 dB and 8.59 dB in the spectrum after the optimized scanning. That side lobe supression seems to be sufficient for reliable scannig proces. The noise linewidth achieved in this grating design is 1.05 pm which is slightly higher than in the previous design 2250-14. The scanned noise linewidth is still very good, 0.570 pm, about the half of the value for the uniform grating of the same length and reflectance.

Our second approach in using the apodization led us to the structures with the steep drop in the center of the reflectance spectrum main lobe. The apodization function that allows this feature is:

$$\Delta n(x) = \\ = n_c + \Delta n_{ph}(1 + \sin(A.\pi x/L).\sin(2\pi x/\Lambda_G)) \quad (12)$$

The representative of this design is shown at the picture in Fig. 12. The apodization profile shown there can be made by two subsequent partial expozitions with the same energy. The ($K.\Lambda_G/L$) stretch and $\Lambda_G/2$ shift has to be applied to the fibre for the second exposition. This way, two uniform gratings with equal index amplitudes, slightly different periods and the half period shift of one to the other are overlaping and creating resulting moare grating. Fig 12 shows the π - phase step in the center of the grating that is created by the superposition of the two original gratings. Analysing the reflectance spectrum of such the structure, we can observe the steep drop in the central reflectance maximum as can be seen in the Fig. 13.

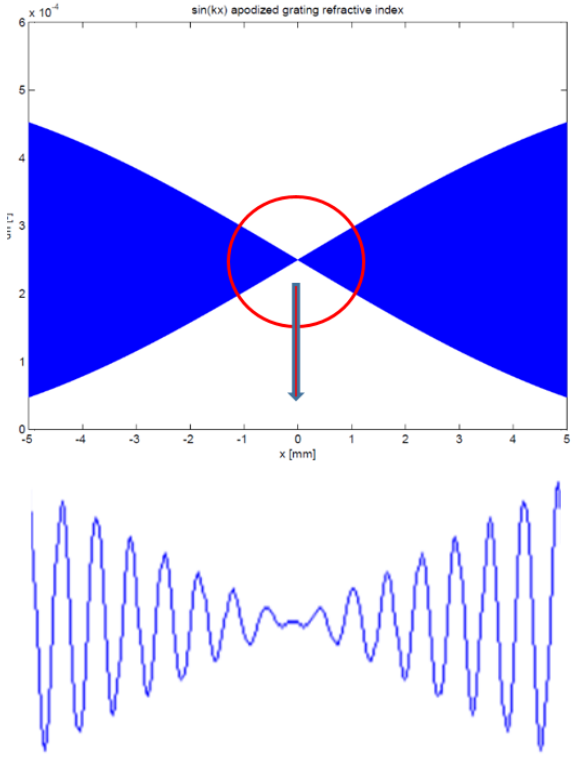


Fig. 12. Moare grating with central π -phase step.

This local reflectance minimum is very narrow and reaches the zero reflectance in the centre of the spectrum main lobe. Locating this central reflectance minimum can be therefore very accurate and the relevant noise linewidth very low. It is clearly seen that any scanning make the scanned spectrum noise linewidth only worse than the grating noise linewidth of this spectral minimum. Fig. 13 shows the original grating spectrum and scanned spectra for 50, 10 and 1 pm wide scanning spectrum. It demonstrates that the most suitable way of scanning interrogation in this case is using the ultranarrow tunable source. The scanned noise linewidth when scanned with 1 pm wide spectrum is 0,023 pm and corresponding RMS error σ = 0,005 pm, which is well below of the practically achievable scanning wavelength resolution. The results of the scanning with wider spectra are also in Table 1. Here we can see that the scanning with the spectrum linewidth wider than 10 pm rapidly deteriorates the scanned noise linewidth and measuring results. The Table 1 compares the sensorial grating features of the type 1 and type 2 apodized gratings with the uniform grating 10-60 of the same length.

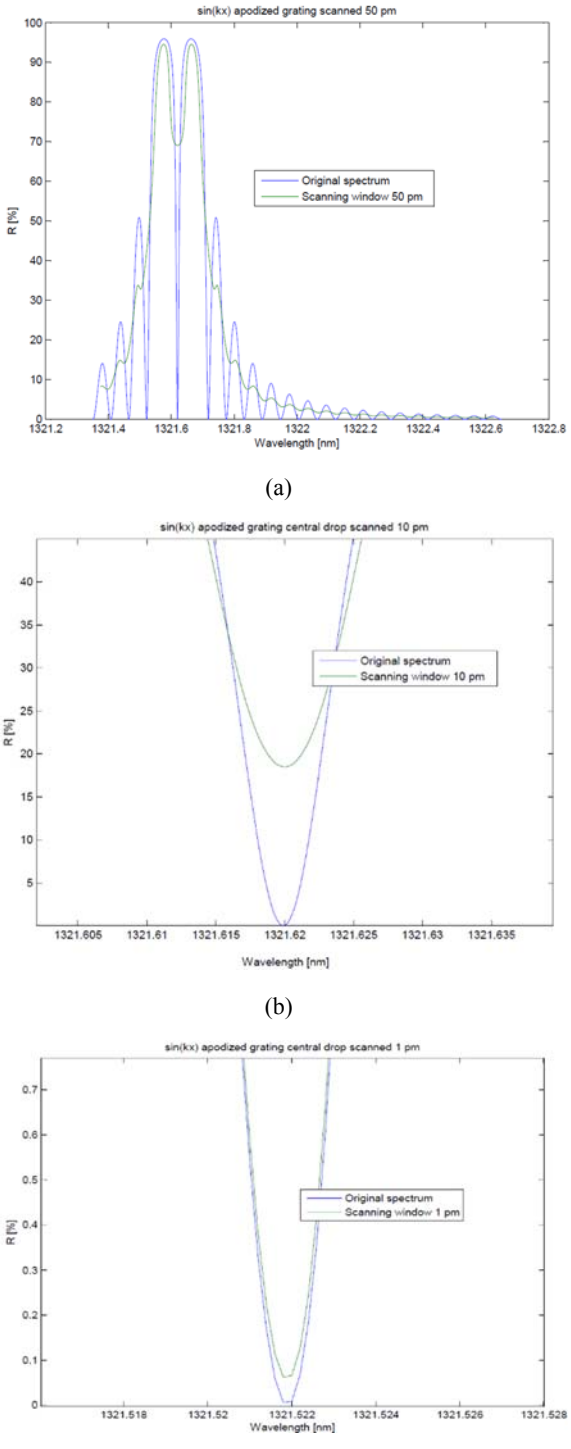


Fig. 13. Reflectance spectrum of the Moare grating and scanned spectra for a) 50 pm FWHM scanning linewidth, b) 10 pm FWHM scanning linewidth, c) 1 pm FWHM scanning linewidth.

Table 1. Comparison of uniform grating, type 1 pure apodized $\sin^2(kx)$ grating and type 2 pure apodized sin(kx) gratings.

	A	**B**	**Δn_{ph}**	**R_{max}**	**$\Delta\lambda_{GFWHM}$**	**$\Delta\lambda_{G\text{-}0.0001}$**	**$\Delta\lambda_{GS\text{-}0.0001}$ ($\Delta\lambda_{SFWHM}$)**	**σ**
	-	**-**	**x.10^{-4}**	**%**	**pm**	**pm**	**pm**	**pm**
10-60	0	1	0.8	60.5	63.8	1.3	1.125 (64 pm)	0.23
2250-13	0.28	0.1	5	54	52.35	1.05	0.570 (54 pm)	0.116
2250-14	0.35	0	3.5	43	46.0	0.912	0.495 (46 pm)	0.101
230-01	0.6	0	2.5	95.9	9.5	0.023	0.023 (1pm)	0.005
230-01	0.6	0	2.5	95.9	9.5	0.023	0.23 (10 pm)	0.046
230-01	0.6	0	2.5	95.9	9.5	0.023	3.10 (50 pm)	0.632

5. Conclusions

In this work, we have investigated the factors influencing the accuracy of the interrogation of the sensorial FBGs by scanning their spectra. The potentiality of the spectrum scanning interrogation method is the facility to achieve very fast reading of the grating reflectance spectrum maximum and so to follow even very fast changes of the measured values. Another potentiality of the method is a possible very low resolution of the measured values. To utilize that resolution, it was necessary to analyse the influence of the noise in the measurement. The analysis has showed the importance of some sensorial grating features to the measurement accuracy and errors. We have focussed the work to the evaluation of the grating spectral width and its impact to the measurement accuracy. We have found that the spectrum scanning gives the best results when the scanning spectrum width (FWHM) is equal to the spectral linewidth of the interrogated gratings. The scanning spectrum narrower than the grating spectrum linewidth gives still relatively good results with the RMS errors less than doubled with respect to the optimal scanning case. The increase of the scanning spectrum linewidth above the optimal value (equal to the grating linewidth) causes the rapid growth of the scanning noise linewidth and so the measuring errors.

The significance of the grating spectrum linewidth to the measuring errors minimization has led us to investigate the apodization of gratings, where we have found prospective ways to further increase the accuracy of the GBG interrogation without the need to increase the grating`s active length. We have found the apodization profiles $\sin^2(\pi x/L)$ and $\sin(\pi x/L)$ viable for grating fabrication and giving the encouraging results in decreasing the grating noise linewidth. Our results in this field (Figs. 10., 11., 12. and Table 1.) show that the proper design of the grating and the optimal scanning can noticeably increase the measuring accuracy and minimize the influence of the scanning process noise to the measurement. We therefore believe that the apodization of gratings has brought a novel look in the field of FBG sensing and more potential to further improvements of FBG sensor system design.

References

[1]. Luciano Mescia, Francesco Prudenzano, Advances on Optical Fiber Sensors, *Fibers,* 2, 1, 2013, pp. 1-23.

[2]. F. Sales, F. Mota, G. Guimarães, and A. Alexandria, Applications of Fiber Bragg Grating Sensors in the Industry, *International Journal of Advanced Engineering Research and Science*, 6, 2019, pp. 238-250.

[3]. J. K. Sahoa, N. Gupta, and D. Dhawan, Fiber Bragg gating sensors for monitoring of physical parameters: a comprehensive review, *Socity of Photo-Optical Instrumentation Engineers*, 59, 2020, pp. 1-35.

[4]. K. O. Hill and G. Meltz, Fiber Bragg Grating Technology Fundamentals and Overview, *Journal of Lightwave Technology*, 15, 1997, pp. 1263 - 1276.

[5]. J. M. L. Higuera, L. R. Cobo, A. Q. Incera, and A. Cobo, Fiber Optic Sensors in Structural Health Monitoring, *Journal of Lightwave Technology*, 29, 2011, pp. 587 - 608.

[6]. H. D. Lee, G. H. Kim, T. J. Eom, M. Y. Jeong, and C.-S. Kim, Linearized Wavelength Interrogation System of Fiber Bragg Grating Strain Sensor Based on Wavelength-Swept Active Mode Locking Fiber Laser, *Journal of Lightwave Technology*, 12, 33, 2015, pp. 2617 - 2622.

[7]. L. K. Cheng, W. Vliegenhart, and T. Habisreuther, Optical Fiber Grating based Technologies and Their Applications: from Nuclear Fusion to Medical, in *Proceedings of the Photonics Global Conference (PGC' 2012),* 2012, pp. 1-5.

[8]. J. P. Carvalho, Optical Sensing: Fiber Structures and Interrogation Techniques, PHD Thesis, *Universidade do Porto,* Porto, 2013.

[9]. T. K. Gangopadhyay, Prospects for Fibre Bragg Gratings and Fabry-Perot Interferometers in fibre-optic vibration sensing, *Sensors and Actuators A*, Vol. 113, Issue 1, 2004, pp. 20-38.

[10]. M. S. Ferreira, J. Bierlich, M. Becker, K. Schuster, J. L. Santos, and O. Frazão, Ultra-High Sensitive

Strain Sensor Based on Post-Processed Optical Fiber Bragg Grating, *Fibers*, pp. 142 - 149, 2014.
[11]. F. Urban jr, R. Helan, F. Urban sr., Sensoric Fiber Bragg Grating design for scanning interrogation in *Proceedings of the 8th International Conference on Sensors and Electronic Instrumental Advances (SEIA' 2022)*, Corfu, Greece, 2022, pp. 38-44.

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 29-36

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Possibility of Detecting Changes in Health Conditions using an Improved 2D Array Sensor System

[1, *] Yasutaka Uchida, [2] Tomoko Funayama, [1] Kazuyoshi Hori, [3] Misako. Yuge, [3] Nobuko Shinozuka and [4] Yoshiaki Kogure

[1] Faculty of Life & Environmental Sciences, Teikyo University of Science, Adachi-ku, Tokyo,120-0045, Japan
[2] Faculty of Medical Sciences, Teikyo University of Science, Uenohara-shi, Ymanashi, 409-0193, Japan
[3] Department of Nursing, Katoriomigawa Medical Center, Katori-shi, Chiba, 289-0332, Japan
[4] Professor Emeritus, Teikyo University of Science, Adachi-ku, Tokyo, 120-0045, Japan
Tel.: + 81369101010, fax: + 81369103800
E-mail: uchida@ntu.ac.jp

Received: 3 September 2022 /Accepted: 5 October 2022 /Published: 31 October 2022

Abstract: For improved detection of changes in body conditions, herein, we propose a two-dimensional system, wherein sensors are placed parallel and perpendicular to the direction of walking, based on an already proposed system that employs pressure sensors. The sensors placed parallel to the direction of walking identify the foot that steps on the sensors, and two others pairs of sensors are placed at positions corresponding to the inner and outer sides of the left and right feet, respectively, to accurately detect the foot that steps on the sensors during walking. This improved two-dimensional health monitoring system is applied to a hemodialysis patient known to have a wobble problem before and after treatment. This is performed to obtain effective values as features for machine learning. Using data obtained from the additional sensors, we develop an index that can evaluate a patient's sense of balance. The results indicate that classification is possible based on the walking speed obtained from the pressure sensors installed orthogonally to the direction of travel and the devised balance index. Using these values, the K-means method, which does not use supervised data, can be used to classify the subjects into three regions, and it is demonstrated that changes in gait before and after hemodialysis can be detected, although at an early stage.

Keywords: Health monitoring system, Hemodialysis, Gait balance, Pressure sensors, Machine learning, and K-means method.

1. Introduction

The increase in elderly population is turning into a global concern, particularly in Asia. In particular, Japan is rapidly converting into a super-aging society. According to the 2020 population statistics [1], the percentage of the total population aged 65 years and older (the aging rate) was as high as 28.8 %, which is higher than the proportion of population aged 75 years and older, which is in turn higher than the proportion of population aged between 65 and 74 years. As the elderly population continues to grow, aging of hemodialysis patients is also emerging as a major concern. The number of hemodialysis patients is increasing annually. As of 2020, the average age for dialysis induction was 70.88 years, and the average

age was 69.40 years [2]. Furthermore, the post-treatment fatigue experienced by aging hemodialysis patients has become a significant problem, as staggering and falls often result in bone fractures, leading to inhibition of the daily living activities of bedridden patients [3-5]. In addition, this is often accompanied by a risk of falls in partially or totally blind patients with retinopathy, owing to diabetic nephropathy. Patients with physical conditions that change before and after hemodialysis require additional monitoring compared to patients with comparatively constant disabilities, such as motor paralysis and sensory disturbances. Several recent reports have revealed that the maintenance of lower-limb muscle strength is an effective preventive measure against the aforementioned conditions, and renal rehabilitation is often performed in accordance with the physical condition of patients [6]. In addition, several studies have reported on the use of dialysis and amino acid supplementation to minimize amino acid loss [7].

A major cause of unsteadiness among hemodialysis patients is decreased blood pressure owing to inappropriate dehydration rates. The relationship between blood pressure and cerebral blood flow is known to correlate with a decrease in mean arterial pressure and cerebral blood flow below a certain mean arterial pressure [8]. Thus, attempts have been made to prevent reductions in the blood pressure by monitoring decreased cerebral blood flow based on the leg blood flow. However, such attempts are limited to the prevention of falls by maintaining a patient's muscle strength and detecting a drop in blood pressure during dialysis [9]. Therefore, these attempts have not yet accomplished gait support or fall prevention after dialysis treatment.

The increase in the number of elderly people living alone is likely to become another major problem with the increase in the number of patients requiring medical treatment in today's hyperaged society. The World Health Organization has reported on the problem of falls among the elderly since 2014 [10]. In particular, falls affect one in three adults aged 65 years and older and 50% of adults aged 80 years and older each year. Compared to the citizens of other countries, the citizens of Japan have relatively poor social interactions with their neighbours and are surrounded by fewer older adults who are willing to assist them. Although the government is implementing concerted programs to prevent the elderly from becoming isolated, the current efforts are still inadequate.

Hence, several studies have investigated fall detection [11- 16]. A typical example of a monitoring facility involves the use of a surveillance camera and a microphone connected to a network. However, current surveillance camera systems are expensive and require an external power supply. Nevertheless, because these systems are equipped with cameras and microphones, they can respond to unusual movements or sounds or to situations when motion is undetected for a certain period. However, there exists a trade off between accuracy and personal privacy when dealing with imaging and sound technologies. For images, complex processing such as using avatars is possible and is acceptable in terms of monitoring. Although the amount of information may be inadequate, a system that detects changes in physical conditions by detecting movement in a room or other locations as a quantity has been reported previously.

Recently, a method for analyzing gait trajectories indoors using a capacitive proximity-sensing device, called Smart Floor, was reported [17- 18]. The Smart Floor is capable of capturing trajectories over 24 h. This system is capable of accurately recognizing multiple activities within the same indoor area using a lightweight algorithm. This is effective as capturing the walking trajectory of an elderly person living alone can be helpful in detecting falls or immobility, which may result from changes in physical conditions. However, it is often difficult to identify individuals when multiple people walk in the same space; hence, new algorithms should be developed explicitly for elderly households.

So far, several studies on fall detection have used accelerometers or smartphones to detect changes in physical conditions by detecting body tilts and foot movements while walking. These sensors are connected to a network, and a server or other device employs deep-learning technology to detect differences from the norm. However, power supply problems are likely to arise when operating such systems for an extended period. Because these services are often event driven, they can provide rapid responses if used at institutions. If an incident occurs at home or in a remote area, the benefit of reporting the incident is clearly significant; however, this system lacks when an immediate response is required. Thus, current systems in Japan are lacking in terms of the foregoing.

We have proposed a simple monitoring system employing a flexible pressure sensor to enable long-term measurements. Further, we report the differences in data using machine learning analysis methods, such as support vector machines (SVM) and decision trees. Although this study involved activities that included rising from a bed and walking a few steps, it demonstrated the existence of a movement pseudo-limitation and the possibility of distinguishing between two people [19- 21]. In this study, we examined whether the proposed system could detect differences in the balance and walking speeds of patients undergoing dialysis treatments by obtaining these variables before and after dialysis treatments, which is assumed to be a time when changes in physical condition are more likely to occur [22].

2. Experiment

2.1. Subjects

The study subjects were six maintenance dialysis patients undergoing dialysis therapy at the Katori Omigawa Medical Center. Described next are certain

patient characteristics: age range of 64–91 years, average dialysis history of 102.5 ± 725.5 months, and the primary diseases were diabetic nephritis in five patients; however, one patient had nephrosclerosis. In addition, two of the subjects had visual impairment in both eyes owing to diabetic retinopathy and optical valves, and two patients were suspected to have lower extremity peripheral artery diseases based on the ankle brachial pressure index. Furthermore, a male subject in his 60s, who had not received hemodialysis treatment, was evaluated at a university laboratory as a reference.

The study design and protocol were reviewed and approved by the Ethics Committee for Human Subjects Research of Teikyo University of Science and the Ethics Committee of Katori Omigawa Medical Center. A written informed consent for participation in the study and publication of the study results was obtained from all study participants.

2.2. Overview of the New System

The improved 2D health monitoring system, as depicted in Fig. 1, had two pairs of position-detection sensors, one for each foot, which were placed parallel to the direction of walking; this arrangement is in contrast to the conventional perpendicular arrangement of sensors, represented by P. For example, eight sensors were placed to detect whether a weight was placed on the inner or outer side of the left foot, with even and odd numbers as subscripts, such as Q0 and Q1. These sensors were placed 1.5 cm apart so that the inner and outer sides of the foot could be detected. As the length of the sensors was only 62 cm, two sets of sensors were placed in series so that a length of approximately 120 cm could be analyzed. Considering a particular subject's physique, the distance between the left and right feet was assumed to be approximately 15 cm. For sensors placed perpendicular to the direction of travel, only the distance between P0 and P1 was assumed to be 10 cm at the starting point, and the distance thereafter was 15 cm. The sensors were affixed using a tape on a plastic sheet to highlight their positions, and the surface was protected using another plastic sheet. Figure 2 presents an image of the proposed system.

The configuration of the proposed system is presented in Fig. 3. A pressure sensor similar to a previously reported one was used in this study, and the resistance values varied from 100 Ω to 1 MΩ, depending on the pressure. A 1 kΩ resistor was connected in series with the sensor, and the voltage change of the resistor was employed as the input signal. Each sensor measured the voltage at a frequency of 1 kHz. The signals from 16 sensors were examined in pairs using a program created in Processing3 by connecting a personal computer (PC) and an Arduino Mega2560 R3. Considering the communication and write times, 50 pairs of data were acquired per second. The output signals obtained from each pressure sensor were imported to the PC as text data and processed using Microsoft Excel and Python. The definitions of the output pulses used in the analysis are depicted in the figure.

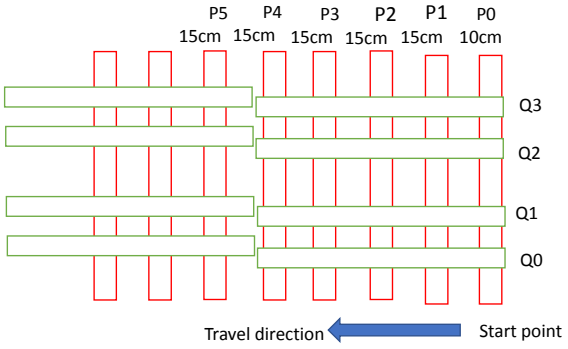


Fig. 1. Improved 2D health monitoring system.

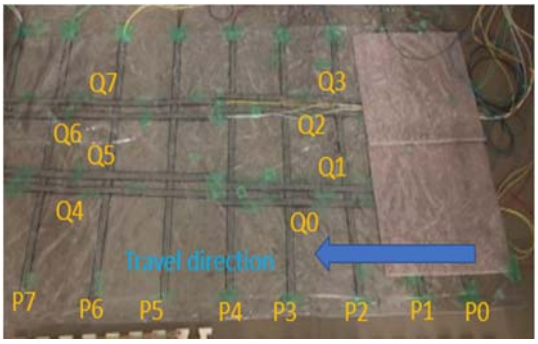


Fig. 2. Photograph of the system.

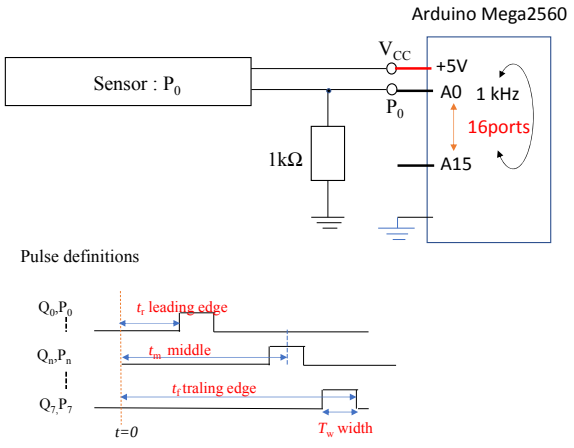


Fig. 3. Detection of output signals.

2.3. Measurement Method

The measurement interval between the two pairs of foot sensors had to be adjusted to match the motion interval between the subject's left and right feet; however, in this case, the measurement was not optimized for each individual patient as multiple

participants were considered. Therefore, the waveforms of the inner and outer feet of the left and right feet differed from subject to subject. Because focusing too much on the pressure sensor during walking could result in significant differences from the natural gait, we verbally instructed the subjects to walk only in the direction of the sensor parallel to the direction of gait.

3. Results

Fig. 4 presents the gait data of the reference subject obtained from a preliminary experiment. The upper figure depicts the positional relationship between the sensors and the feet, the middle figure presents the output signal obtained from the sensor positioned parallel to the direction of gait, and the lower figure presents the signal obtained from the sensor positioned perpendicular to the direction of gait.

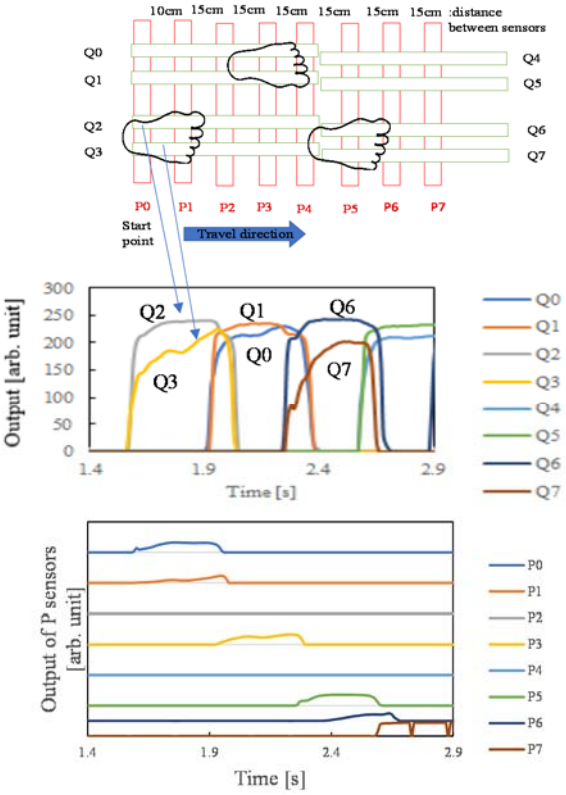


Fig. 4. Output waveforms of sensors Q and P corresponding to the sensor and foot positions, respectively; the upper figure shows the sensor and foot position, the middle figure shows the waveform of Q, and the lower figure shows the wave form of P.

The two sensors, labeled Q2 and Q3, detected the pressure applied to the outside and inside of the right foot, respectively. The waveforms illustrated in the middle panel of Fig. 4 represent signals that are produced in response to the motions of the outer and inner sides of the left and right feet, respectively. Differences in the shape of the rise in waveforms were also observed. The P-sensor signal corresponding to the walking direction, depicted in the lower part of Figure 4, presents the waveforms corresponding to the left and right feet; however, the amount of data is inadequate. Notably, machine learning, through the accumulation of data, can facilitate the identification of individuals.

An example of the data obtained from a subject undergoing hemodialysis is shown in Fig. 5. Compared to the subjects not undergoing hemodialysis, several cases with no signals after hemodialysis were obtained from the sensor represented by Q and the imbalance; however, a simple comparison is difficult owing to the ages of patients and differences in measurement locations.

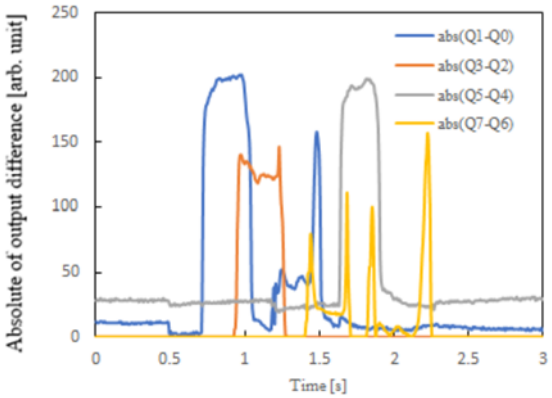


Fig. 5. An example of data from a subject undergoing hemodialysis.

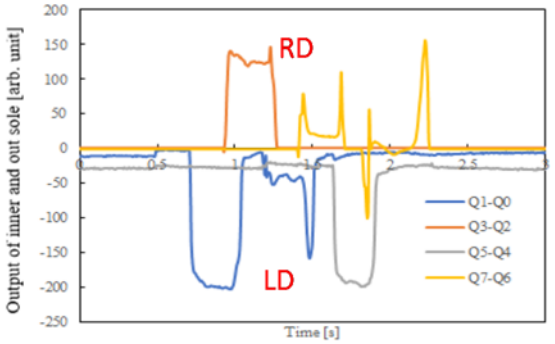


Fig. 6. The waveforms showing the difference in output between the inner and outer foot.

The waveforms highlighting the differences between the medial and lateral outputs obtained from the waveforms depicted in Fig. 5 are shown in Fig. 6. The amplitude of the outer signal was greater than that of the inner signal for both the left and right feet. Similarly, the waveforms corresponding to the first and second steps of the left foot had different amplitudes; however, the shapes were similar. In addition, the waveforms corresponding to the first and

second steps of the right foot had different amplitudes and similar shapes. These results indicate that it may be possible to distinguish between the right and left feet by identifying the signal difference between the inner and outer feet. However, the classification is difficult using the output difference waveform. We noticed that the left and right feet alternately stepped on the sensor, and a spike-like output was observed as the weight shifted from the left foot to the right foot and the left foot left the sensor. In addition, we discovered that the classification based on the output difference waveform was difficult.

Note that the balance of a subject is deemed to be good if the difference between the total time that the right foot is on the sensor and the total time that the left foot is on the sensor is equal, and total time is zero. From this figure, we attempted to evaluate the difference in the duration for which the right and left feet stepped on the line; however, we could not classify the subjects accordingly.

To classify the data presented in Fig. 6, the sum of the differences between each left and right feet was defined as follows. The sum of the medial and lateral differences of the left foot, Q1–Q0 and Q5–Q4, was expressed as LD, and the sum of Q3–Q2 and Q7–Q6 was expressed as RD. We attempted to plot the absolute values of LD and RD in two dimensions using the obtained values. The corresponding results are depicted in Fig. 7. If the difference between the times at which the right foot and left feet stepped on the sensor was similar, the balance of the subject was deemed to be good. Therefore, we attempted to evaluate the subjects based on this difference in times; however, we could not classify the subjects accordingly.

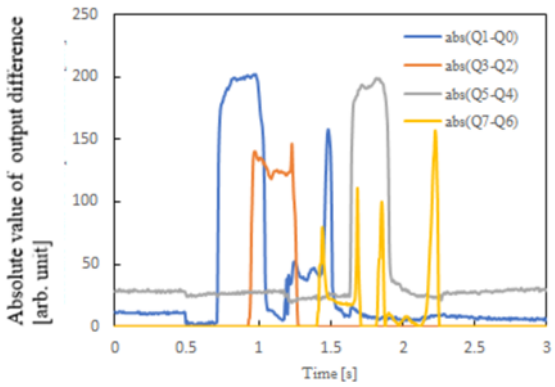


Fig. 7. Aggregate results of the absolute value of the left and right time difference.

Fig. 8 presents the results obtained by analyzing the data using a linear classification of SVMs, with LD and RD as the feature values; this was done because the waveforms corresponding to the movements of the left and right feet were repeated nearly in a similar manner. The support vector analysis classified nearly two areas; however, where LD was small and RD was large, an overlap was noted in certain areas circled by ovals.

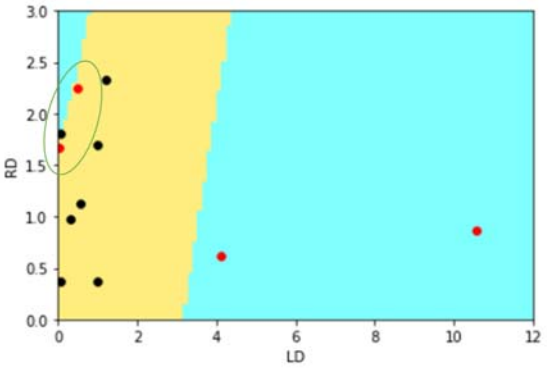


Fig. 8 SVM classification using LR and RD.

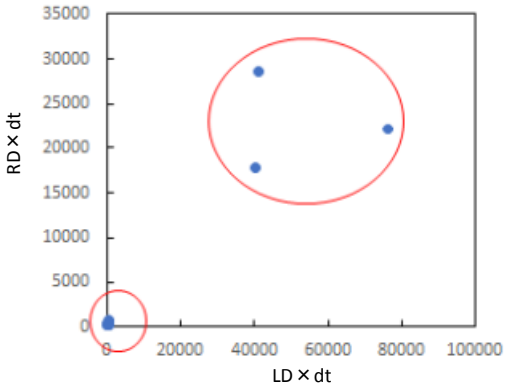


Fig. 9 The product of signal intensity and time dt from each of the left and right inner and outer feet.

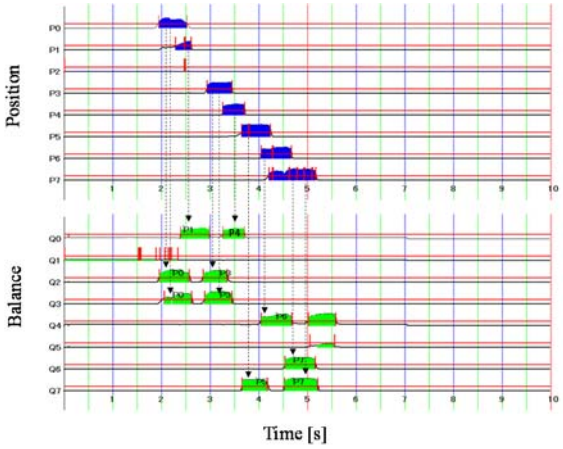


Fig. 10. Identification of the left and right feet that stepped on the sensors.

Next, the product of the signal strength and time dt measured from the inside and outside of each of the left and right feet, respectively, was obtained as an area and was plotted in two dimensions, as shown in Fig. 9. Eight points were clustered near the origin, and

the other three points were far apart and could be treated as features. However, detailed examination of the subjects at each point could not provide any evidence regarding the symptoms.

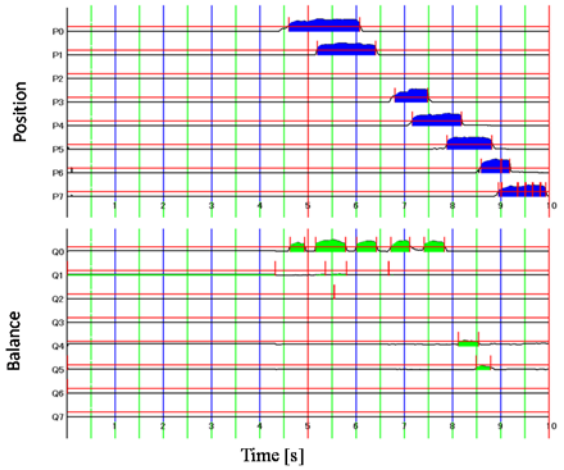


Fig. 11. Example of unbalanced measurement.

Fig. 10 presents the output signals obtained using the proposed system. The gait speed was obtained from the data collected from the P-sensor signals. The foot (left or right) that stepped on the sensor was identified by searching for the Q sensor that responded simultaneously to the signal depicted in Fig. 10, which corresponds to position P. The left and right feet alternated; however, some points could not be identified.

In the example presented in Fig. 11, the output signal progresses at almost constant intervals with respect to the direction of travel. In contrast, the Q signal, which indicates poor balance, can only be partially observed, indicating that it is moving to the left with respect to the direction of travel. This was confirmed based on the observations reported by the person who performed the measurement on that day, which indicates that the sensor was not stepped on properly.

Table 1 summarizes the measurement results depicted in Fig. 10, including the differences in the rise time and pulse width of the output signal obtained from the sensor and the results on the identification of the foot that stepped on the sensor. The results indicate that the right and left feet alternately stepped on the P sensor corresponding to the position. Because a variation was noted in the pulse width of the signal, we calculated the walking speed based on the rise time. The data collected from nearly all subjects indicated that the elapsed time and distance corresponding to each sensor position were nearly proportional to each other.

We observed that the gait speed before and after hemodialysis differed significantly among the subjects. Thus, using the walking speed as a new feature may be effective. We concluded that the reaction time of the sensor and the product of the time when the subject stepped on the sensor and the output voltage could not be used for effective classification of the left–right balance, and hence, is essential to consider another index.

Table 1. List of position sensor rise time and data on which foot is stepping on Pn, left or right.

Senor	Leading tr[s]	Tarling tf[s]	Middle tm[s]	Width tw[s]	L or R
P0	1.94	2.52	2.23	0.58	R
P1	2.29	2.61	2.45	0.32	L
P2	NA	NA	NA	NA	
P3	2.94	3.45	3.2	0.51	R
P4	3.26	3.71	3.49	0.45	L
P5	3.65	4.24	3.95	0.59	R
P6	4.05	4.67	4.36	0.62	L
P7	4.21	5.17	4.7	0.96	R

Based on these results, the numbers of steps nL for the left foot (L) and nR for the right foot (R) were obtained. The balance index is defined as follows:

$$|\Delta n|/n=|nL-nR|/(nL+nR). \quad (1)$$

The walking speed, number of times that the left foot stepped on the sensor, number of times that the right foot stepped on the sensor, and wobble coefficient defined in this study for all subjects are shown in Table 2.

Based on the relationship between the gait speed and balance index $|\Delta n|/n$, the data could be classified into two areas using an SVM method as shown in Fig. 12. However, the data with greater values of gait speed were separated from each other and could not be sufficiently classified; hence, further analysis is essential. We hypothesized that this could be attributed to the fact that the classification results obtained by the measurer could have been problematic and considered applying the K-means method, which does not provide supervised data.

Table 2. Wobble coefficient defined in this study for all subject.

Sample No	v (km/h)	nL	nR	|⊿n|/ n
1	1.44	3	4	0.143
2	1.245	3	0	1
3	3.112	1	2	0.333
4	1.859	5	2	0.429
5	1.286	3	2	0.333
6	0.873	6	0	1
7	1.507	3	4	0.143
8	2.878	2	1	0.333
9	1.34	3	5	0.286
10	0.487	1	7	0.857
11	1.294	4	0	1

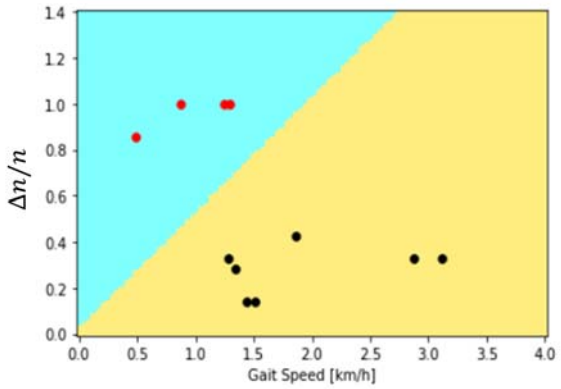


Fig. 12. Classified by the SVM methods.

As shown in Fig. 13, the pre- and post-hemodialysis balance indexes and gait speed data were classified into three clusters using the K-means method, reflecting the pre- and post-hemodialysis states of the subject. Here, triangular symbols represent the center of gravity for each cluster.

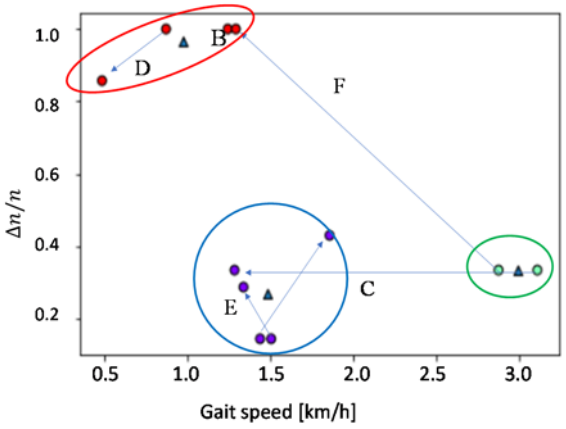


Fig. 13. Classified by the K-means methods.

4. Discussions

A flexible pressure sensor was used in the measurement considered in this study and enabled an objective evaluation of the severity of the changes in physical conditions by analyzing the measurement results of foot contact pressure during walking and the balance between the inner and outer contact pressures on the left and right sides, respectively. Further, we examined the possibility of clinical application of this system by examining several cases of individual subjects and selected subjects who were elderly (D), visually impaired (B, C), or had excessive water removal (C, F). The corresponding details are presented separately in this paper. We measured the blood flow in the lower limbs before and after dialysis as an indicator of post-dialysis staggering and related the results of the pressure sensor analysis to the effects of water removal and blood pressure decrease during dialysis. The pressure sensor results revealed that in all the subjects, a decrease in the walking speed and walking balance was noted immediately after dialysis. In addition, all subjects presented a decrease in lower-limb blood flow, indicating a relationship between the decrease in lower-limb blood flow and staggering to certain extent. These results suggest that the analysis data obtained using this system can detect changes in physical conditions based on the background and dialysis conditions of a patient and provide appropriate feedback to medical staff and family members on care appropriate to the condition of the patient.

5. Conclusions

Herein, we conducted an analysis based on an improved pressure sensor system on subjects undergoing hemodialysis treatment and adopted machine learning using the K-means method; further, we obtained the gait speed and balance index. Consequently, we proposed a system that could possibly detect changes in the physical conditions of patients after dialysis.

Acknowledgements

This work was supported by JAPS KANENHI Grant Number JP20K11924. We would like to thank the patients and the staff of Katori Omigawa Medical Center for their cooperation in this study.

References

[1]. Annual Report on the Ageing Society [Summary] FY2021, *Cabinet Office Japan*, July 2021.
[2]. Japanese Society for Dialysis Therapy, Aggregation on chronic dialysis patients at the end of 2020, (https://docs.jsdt.or.jp/overview/index.html)

[3]. T. Kutsuna, A. Matsunaga, T. Matsumoto, A. Ishii, K. Yamamoto, K. Hotta, N. Aiba, Y. Takagi, A. Yoshida, N. Takahira, and T. Matsuda, Physical Activity is Necessary to Prevent Deterioration of the Walking Ability of Patients Undergoing maintenance Hemodialysis, *Therapeutic Apheresis and Dialysis*, 14, 2, 2009, pp. 193-200.

[4]. H. Koyama, S. Fukuda, T. Shoji, M. Inaba, Y. Tsujimoto, T. Tabata, S. Okuno, T. Yamakawa, S. Okada, M. Okamura, H. Kuratsune, H. Fujii, Y. Hirayama, Y. Watanabe, and Y. Nishizawa, Fatigue Is Predictor for Cardiovascular Outcomes in Patients Undergoing Hemodialysis, *Clinical Journal of American Society of Nephrology*, 5, 2010, pp. 659-666.

[5]. M. Jhamb, F. Pike, S. Ramer, C. Argyropouos, J. Steel, M. A. Dew, S. D. Weisbord, L. Weissfield, and M. Unruh, Impact of Fatigue on Outcomes in the Hemodialysis (HEMO) Study, *American Journal of Nephrology*, 33, 2011, pp. 515-523.

[6]. J. Niwayama, T. Sanaka, Development of a new method for monitoring blood purification: The blood flow analysis of the head and foot by laser Doppler blood flowmeter during hemodialysis, *Hemodialysis International*, 2005, 9, pp. 56-62.

[7]. Verification of the Effect of Nutritional Status Improvement in Very Elderly Dialysis Patients Using a Stacked Dialyzer - From the Perspective of MIA Syndrome - (in Japanese), *Journal of Japan Association for Clinical Engineers*, Vol. 44, 2012, pp. 28-31.

[8]. Y. Abe, A. Matsunaga, R. Matsuzawa, T. Kutsuna, S. Yamamoto, K. Yoneki, M. Harada, R. Ishikawa, T. Watanabe, and A. Yoshida, Determinants of Slow Walking Speed in Ambulatory Patients Undergoing Maintenance Hemodialysis, *PLOS ONE*, 11, 2016, e0151037.

[9]. D. A. Drew, H. tighouart, S. Duncan, J. Rollins, A. Gupta, T. Scott, D. E. Weiner, and M. J. Sarnak, Blood Pressure and Cognitive Decline in Prevalent Hemodialysis Patients, *American Journal of Nephrology*, 49, 2019, pp. 460-469.

[10]. Kalache, A; Fu, D; Yoshida, S., et al., WHO global report on falls prevention in older age, *World Health Organization,* 2007.

[11]. M. M. Alam, and E. B. Hamida, Surveying Wearable Human assistive Technology for Life and Safety Critical applications: Standerds, Challenges and Opportunities, *Sensors*, 14, 2014, pp.9153-9209.

[12]. A. M. Herran, B. G. Zapirain, and A. M. Zorrilla, Gait analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications, *Sensors*, 14, 2014, pp. 3362-3394.

[13].S. Majurmder, T. Mondal, and M. J. Deen, Wearable Sensors for Remote health Monitoring, *Sensors*, 17, 2017, 130.

[14]. X. Qian, H. Cheng, D. Chen, Q. Liu, H. Chen, H. Jiang, and M-C. Huang, The Smart Insole: A Pilot Study of Fall Detection, *BODYNTES*, LNICST 297, 2019, pp. 37-49.

[15]. S. Subramaniam, S. Majumder, A. I. Faisal, and M. J. Deen, Insole-Based systems for Health Monitoring: current Solutions and Research Challenges, *Senosrs*, 22, 2022, 438.

[16].S. Subramanaiam, A. Faisal, and M. J. Deen, Wearable Sensor Systems for Fall Risk Assessment: A review, *Frontiers in Digital Health*, 4, 2022, 921506.

[17]. M. Sannier, S. Janaqi, V. Raducan, V. Barysheava, H. A. Haddou, S. Pla, G. Dray, and B. G. Bardy, Extracting Walking Trajectories at Home From a Capacitive Proximity Sensing Floor, *IEEE Sensors Journal*, Vol. 22, 2022, pp. 3695-3703.

[18]. S. Janaqi, M. Sannier, G. Dray, V. Raducanu, and B. Bardy, Walking on a Capacitive Sensing Floor, *in Proceedings of 8th International Conference on Sensors and Electronic Instrumentation Advances (SEIA' 2022)*, Corfu, Greece, 21-23 September, 2022, pp. 48-54.

[19]. Y. Uchida, T. Funayama, and Y. Kogure, Support Vector Machine Analysis to Detect Deviation in a Health Condition Monitoring system, *Sensors & Transducers*, Vol. 237, 2019, pp. 103-108.

[20]. Y. Uchida, T. Funayama, and Y. Kogure, Classification Accuracy of Support Vector Machine, Decision Tree and Random Forest Modules when Applied to a Health Monitoring with Flexible Sensors, *Sensors & Transducers*, Vol. 245, 2020, pp. 83-89.

[21]. Y. Uchida, T. Funayama, Y. Kogure, Bedside Monitoring System with Machine Learning to Independent Living People, in *Proceedings of the 24th Rehabilitation World Congress,* 7-9 September 2021, Aarhus, Denmark, Abstract Book, p.87.

[22]. Y. Uchida, T. Funayama, K. Hori, M. Yuge, N. Shinozuka, and Y. Kogure, Health monitoring system using a pressure sensor 2D array placed on floor mats, *in Proceedings of 8th International Conference on Sensors and Electronic Instrumentation Advances (SEIA' 2022)*, Corfu, Greece, 21-23 September, 2022, pp. 74-77.

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 37-44

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Development of a Virtual Reality Tool for Therapy and Diagnosis of Schizophrenia

[1] Andreia MENDES, [1,2] Duarte DUQUE and [1] Vítor CARVALHO
[1] 2Ai – School of Technology, IPCA, Campus do IPCA, 4750-810 Barcelos, Portugal
[2] ALGORITMI Research Centre / LASI, University of Minho, Guimarães, Portugal
Tel.: +351 253 802260
E-mail: amendes@ipca.pt, dduque@ipca.pt, vcarvalho@ipca.pt

Received: 2 September 2022 /Accepted: 3 October 2022 Published: 31 October 2022

Abstract: Schizophrenia is a serious mental disorder in which people interpret reality abnormally, often requiring lifelong treatment. In recent years new treatment methods have been tested, which include serious gaming, in particular, with virtual reality technologies. This project aims to develop a virtual reality tool capable of assisting in the therapy of schizophrenia patients. As a starting point, a systematic review was conducted in March 2021 using IEEE Xplore, Science Direct and dl.acm databases, from which 53 results were obtained, and 17 were filtered, revealing both, common aspects, as well as the differences, between healthy individuals and patients with schizophrenia when exposed to the technology and to different scenarios in virtual reality. Continued exposure to virtual reality exercises revealed improvements in the patient's quality of life, in some studies. This demonstrates that virtual reality can be used as a diagnosis and therapy tool for patients with schizophrenia. This project hopes to develop one such tool, and have it tested with schizophrenia patients while supervised by a medical team, to determine whether it is a viable alternative to traditional therapy, since it is easy and convenient to use.

Keywords: Virtual reality, Schizophrenia, Serious Games, Therapy, Unity.

1. Introduction

Schizophrenia is a serious mental disorder in which people interpret reality abnormally. Schizophrenia may result in some combination of hallucinations, delusions, and extremely disordered thinking and behavior that impairs daily functioning, and can be disabling, often requiring lifelong treatment.

Serious games are games designed to fulfil a serious goal. These games are already used in several ways, such as learning aid for children (e.g., mathematical games), or the practice of physical exercise (e.g., football).

Virtual reality (VR) technology is a type of technology that puts the user in a virtual world, with the use of VR glasses and controllers that detect the movements of the hands and head. Using VR grants the user as a feeling of immersion and presence on a" new level". VR has been evolving rapidly, being already used in several video games (e.g., Beat Saber, Minecraft, among others), as well as, in some cases, used for simulation of specific situations and environments for training (e.g., aircraft navigation simulation). However, the use of VR as a means of therapy is still a little studied topic, partly because the technology has a much faster evolution than the testing and approval period of the scientific community.

Studies comparing the use of VR in patients with schizophrenia and volunteers without any psychiatric disorder in various social, memory and spatial navigation exercises, demonstrated tendency-negative performance in patients with schizophrenia due to lack of motivation, when compared to volunteers without

psychiatric disorders [1]. On some studies, continued exposure to these exercises also revealed positive improvements in test performance and quality of life on patients with schizophrenia [2]. This demonstrates that it is possible to use VR as a diagnosis and therapy tool for schizophrenia. VR can provide easily accessible realistic exercises and environments for therapy and testing, as well as supervising and recording data about the patient, which can save a lot of resources when compared to creating and finetuning exercises to each patient in real life.

This paper is organized in six sections. Section 2 defines the research methods used for this paper; Section 3 analyses various studies made on the use of VR on patients with schizophrenia; Section 4 describes the project development; Section 5 explores the problems and changes during the development process and finally, Section 6 enunciates the current and future stages of the project.

2. Methods

For this research, the terms "Virtual Reality", and "Schizophrenia" or "Schizophrenic" were researched as keywords to be included on the abstract of all articles, or on all contents of the article for sources with few (-10) results. Sources included IEEE Xplore (yielding 13 results), Science Direct (yielding 31 results) and dl.acm (yielding 9 results), from which 53 papers in total were obtained.

Inclusion criteria were then applied as follows:

- The paper is written in English or Portuguese;
- The paper is clear and accessible;
- The paper mentions practical use of VR;
- The paper mentions solutions for therapy and/or diagnosing schizophrenia.

After inclusion criteria was applied, 38 paper remained and 15 were rejected.

Furthermore, papers were further selected if they could answer:

- How can VR help patients with Schizophrenia?
- How were VR solutions tested with schizophrenic patients?
- What were the test's results?
- Did the tests results favor or not the use of VR as a therapy/diagnostic tool for schizophrenia?

From these questions, 17 final papers were selected, and 21 more were rejected. Lastly, they were sorted into categories relevant to the kind of research performed, including spatial navigation, task simulation, facial expression recognition, social exercises, and memory exercises.

3. Literature Review

3.1. VR Task Simulation

This consists of challenging patients with common day-today tasks in a virtual world and analyze their performance. This exercise is particularly easy to translate into their daily tasks. Some studies use a virtual store, a virtual city, or even a virtual house where tasks are performed such as shopping with an item list, exploring and navigating an area, among others. It is remarkable that the use of VR for treatment and/or diagnosis of schizophrenia was a method generally accepted by patients on some studies, with them having noted that the virtual world was pleasant, learned to use the technology without difficulties, and kept focused on the world and the tasks requested, revealing even enthusiasm for them (Fig. 1) [3].

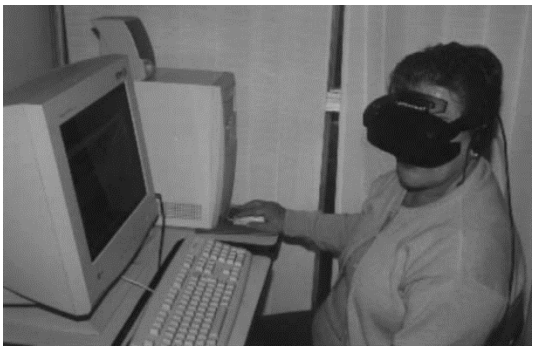


Fig. 1. Patient using the HMD (I-glasses) [3].

In several studies, the performance of patients with schizophrenia is compared to the performance of healthy volunteers in a virtual store. In each challenge, a shopping list is given, and after some time, it is withdrawn, and the challenge begins. The goal is to find all the items in the list, and "pay" for these items at the end. The results showed significant performance differences between volunteers and patients, possibly even finding different severities of schizophrenia depending on performance (Number of correct items, time spent...). In a similar study with the same task of purchasing multiple items in a store, performance in a virtual environment was compared to performance in reality, on patients with schizophrenia. The results were similar in terms of correct items collected, and very similar in terms of efficiency (distance traveled), this given that the task in the real world included a more intense social variable compared to the virtual task (existence of other customers, it was permitted to request help from the shop workers) [4].

3.2. VR Spatial Navigation

This consists of virtual world navigation, and performance analysis. This translates into the spatial cognition capacity of the patient. One of the studies carried out consists of the appropriation of the "Virtual Morris water maze" test (traditionally used for the testing of spatial cognition in animals), using VR (which has the advantage of easily creating and using environments difficult to experience in the real world). This test consists of patients repeatedly finding an

invisible point in an arena using landmarks, changing their starting position in each iteration. A second test was performed, the only difference being the arena, which rotated slowly. The time spent on the task and the used route are analyzed. Patients with schizophrenia showed worse results when compared to a control group (time spent, distance traveled, among others), which demonstrates spatial cognition deficits. This was especially accentuated in the rotating arena [5, 6].

Another study consists of navigating a virtual housing area and maze. Patients received a 3D map of their virtual environment with a marked goal, and after a period of study, they navigated the virtual environment starting on a random initial position. The time used and the success or failure of the tasks are analyzed here. The results reveal that schizophrenia and schizoaffective patients have a lower performance in general when compared to a healthy group. There was also a noticeable difference in the performance of schizoaffective and schizophrenic patients, the latter having a worse performance overall [7].

In a study using an eight-arm radial maze (consisting of a center and eight possible "arms" Fig. 2) in VR, one group of patients with schizophrenia and another group of healthy volunteers were observed. Four of the eight arms contain rewards, and in each iteration the rewards are placed on the same arms. These tests working memory and operational memory, with re-entering an arm already visited correlating to working memory errors, and entering an arm that has never had a reward correlating to operational memory errors. The results reveal a significant difference between the two groups, being the schizophrenia patients group the one with the most errors of both types. Interestingly, after some iterations of the test, the number of errors in the operating memory on the schizophrenia patients group decreases, while the number of errors of working memory remained [8].

Fig. 2. View of the eight-arm radial maze [8].

Another study compared two navigation tests, one in a virtual park and the other in a virtual maze, the virtual park being rich in references, and the maze void of any references, with the goal of finding a pot of gold. The objective of this test is to observe the difference in the processing capacity of object-to-object navigation (virtual park with references) and individual-to-object navigation (virtual maze without references). Two groups participated in the tests, a control group with healthy individuals and a group of patients with schizophrenia. The results revealed significant differences in performance in the park test, with patients with schizophrenia having significantly lower performance than the control group. However, the virtual maze test revealed no significant performance differences between the two groups [9].

A similar study evaluates navigation in a virtual labyrinth without references, with the objective of comparing the processing capacity of the individual-to-object navigation in patients with schizophrenia to healthy volunteers. This test had several iterations. In general, in the first and second iteration there was not much difference between the groups, then from the third iteration the performance of the group with schizophrenia was significantly worse. Magnetic resonance imaging of the brain in each iteration revealed that the area dedicated to individual-to-object navigation was significantly decreasing over iterations in the healthy group, and the area dedicated to object-to-object navigation was increasing significantly. In patients with schizophrenia, the object-to-object navigation area remained inactive, and the individual-to-object navigation area kept similar levels of activity at each iteration. This may be because, in each iteration, the healthy group was building a mental map of the maze, effectively transitioning to object-to-object navigation, while the group with schizophrenia used only individual-to-object navigation, since schizophrenia particularly affects object-to-object navigation [10].

One study compared the performance of patients with schizophrenia with a group of healthy volunteers in a series of tasks in a virtual city, including a planning phase where you can analyze the city map and the tasks to be performed. This test analyzes the navigation phase in particular and provides the map and to-do list during the test to avoid overloading working memory, thus considering only the efficiency of the route (distance traveled when compared to a pathing algorithm A*), the amount of errors and successes achieved, and their association with demotivation in schizophrenia. The study showed that only patients with schizophrenia had strong associations in demotivation and navigation efficiency, revealing a deficiency in goal-oriented motivation. This can be caused by the inability to execute a plan efficiently, which causes demotivation [1].

3.3. VR Facial Expression Recognition Exercises

In one study, software was developed that contained human models capable of producing various expressions linked to certain emotions (angry, fear,

sad, contented, satisfied, neutral), delivering context linked to this emotion in a story format told by the avatars (with lip movement). The reactions of a group of patients with schizophrenia were compared to a group of healthy volunteers. It was observed that the group of patients with schizophrenia had an abnormal fixation at sites like the avatar's forehead, or completely off the face area. This was quite different from the control group, which settled mostly on the eyes and mouth of the avatar (Fig. 3). These results have not changed even without the context told by the avatars (facial expressions only) [11].

A similar study challenged a group of patients with schizophrenia and a group of healthy volunteers to detect emotions on different pictures containing faces, including real and virtual faces, as soon as possible. The objective of this test is to understand if there is a difference between the sense of virtual and real faces in individuals with schizophrenia and/or healthy individuals.

The group of patients with schizophrenia performed generally worse in the correct identification of emotions, and there were no significant differences between the identification of emotions of real and virtual faces, except for revulsion, which was significantly more mistaken with other emotions by both groups on the virtual faces. This can be explained by the low number of polygons on the virtual face, which caused wrinkles in the nasal area to not be so pronounced (which is a major feature of revulsion) [12].

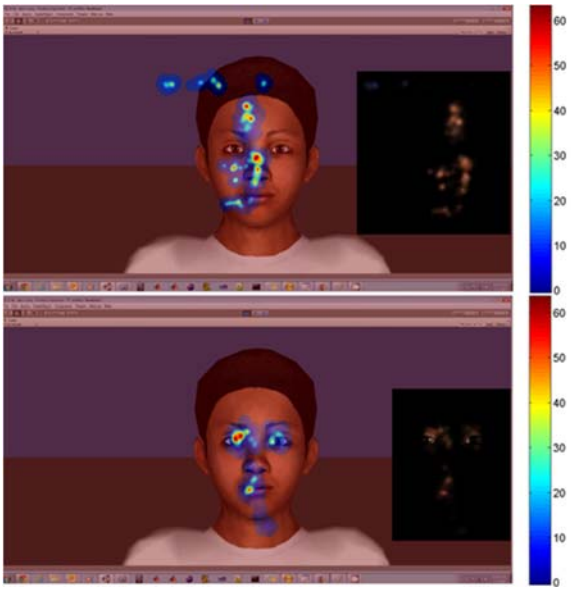


Fig. 3. Masked maps overlaid on at the side of heat map visualizations. Top figure shows the SZ group gaze while the lower figure depicts the control group gaze aggregated across subjects and trials [11].

3.4. VR Memory Exercises

In one study, it was observed the performance of a group of patients with schizophrenia and another group of healthy volunteers in an exercise that consisted of picking up objects in various locations in a virtual dwelling, and then returning them to the same place, in a given order and after a certain time, with a number of variable objects. This test revealed that the group of patients with schizophrenia generally performed worse, in terms of errors and time spent in the challenge [13].

Another study analyzed the associations of memory between objects, places, and people. The test consisted of traveling through a virtual city, to a specific place where a person (virtual avatar) and an object were waiting. After some time, users traveled to another place, with another person and object, and repeated this process several times (Fig. 4). In the end, the user tried to remember which parameter was associated to which another parameter (person, object and place).

Fig. 4. (A) Aerial view showing the layout of the virtual city, with green dashes indicating the path taken; (B) Encounter with a character paired with an object at a location; (C) An example of a recognition trial from the "Person" condition [14].

Two groups, one with patients with schizophrenia, and the other with healthy volunteers performed the test. The brain activity of the participants was analyzed, which revealed that the group of patients with schizophrenia generally had less activity around object-oriented memory, but still managed to complete the tasks, which can be explained by the main processing area of object-oriented memory not being very affected, but being the related peripheral areas affected, which make the processing faster and more efficient [14].

3.5. VR Social Exercises

In one study, a virtual boutique was created, where patients with schizophrenia performed various levels

of shop worker training (organize inventory, deal with sales). The store includes a virtual manager, as well as virtual customers who interact with each other, and with the patient, in the case of the manager. Before and after these tests, the cognitive abilities of the patient were analyzed, and compared with that of patients who performed a similar exercise in a real store with a therapist as a manager. The performance study revealed that the VR group generally had better performance and more confidence in its capabilities, which can be explained by the VR environment being a safe medium when compared to real life, while allowing for an immersive experience and presence in the requested challenge [15].

In one study, the behavior of a group of patients and a group of healthy volunteers was compared in an exercise that consisted of asking all avatars in a virtual office if they could help using a new software, and later asking for money to buy a gift for the head of that office. Avatars could respond negatively or neutrally, and the dialogue was triggered manually by an examiner. The purpose of this test was to observe whether a virtual environment can create stress in patients with schizophrenia. After the tests it was remarkable that a number of patients with schizophrenia gave up on the tests due to the VR system, the task, or private issues. However, the sample of patients is not sufficient to ascertain whether the side effects of VR use caused this situation. The group of patients with schizophrenia reported, in general, having suffered more stress than the healthy group. This reveals that virtual worlds can simulate difficult situations to challenge patients with schizophrenia [16].

One study used realistic virtual avatars to simulate VR conversations with patients with schizophrenia. This was part of several therapy sessions, accompanied by a therapist. After the sessions, social tests revealed changes and improvement in social performance, with fewer negative symptoms and less social anxiety [2].

Another study challenged patients with schizophrenia to join a social context presented in the virtual world, considering the situation presented (e.g. joining a table in a canteen Fig. 5). This exercise increased in difficulty, between easy, medium and difficult. After the tests there were observed decreases in the negative symptoms of patients with schizophrenia, as well as patients reporting their satisfaction with the technology used, and the intention to continue with the tests [17].

3.6. Analysis of the Literature Review

The tests analyzed between SP and HI reveal a significant difference in performance [1, 5-7, 9-15], which demonstrates the feasibility of using VR as a diagnosis tool, since the difference in performance allows the identification of SP when compared to HI. On some studies, it was also possible to observe improvements in the quality of life of SP when continually exposed to exercises in VR [2, 15, 17], which demonstrates this could be used as a means of therapy for schizophrenia.

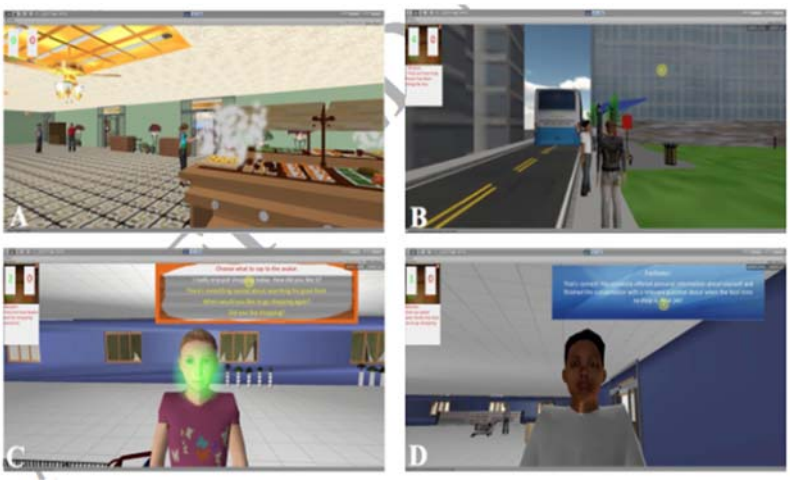


Fig. 5. Sample scenes from the virtual tool [17].

Schizophrenia is usually treated with an individually tailored combination of therapy and medicine. Personalizing therapy exercises for each SP is resource and time consuming, whereas VR therapy benefits of being easily customizable, as well as offers a safe and pleasant virtual environment for SP to try exercises they wouldn't otherwise be comfortable with, in a more exciting and engaging way [3, 15, 17]. It's notable that in some studies the VR environment was so realistic as to induce stress in virtual social environments, as it would in the real world [16], as well as produced similar results to non-VR methods [4].

Therefore, a VR therapy tool could be a valuable asset for the treatment of schizophrenia, providing safe, yet realistic environments, while being accessible and convenient to use.

The use of a VR tool in schizophrenia therapy is a novelty project, even more so considering the lack of practical tests performed in the area. Even among the few researched tests, some couldn't provide reliability to draw accurate conclusions [4, 12, 16, 18].

4. Project Development

This project aims to develop a VR tool, in the form of a serious game, capable of assisting medical professionals on the therapy sessions of patients with schizophrenia, in a convenient and interactive way.

The development of this project will receive guidance from a medical team, which will later validate the game's capability as a therapy tool with schizophrenia patients. The specific requirements that will specify the game are currently being evaluated. Patient interaction will be recorded to allow the game a continuous development and patient engagement.

The game will make use of the Oculus Quest 2 device, since it is wireless and easy to use, as well as it supports hand tracking. The game will consist of two main environments: A home environment and a city environment, where the player will be tasked with performing day-to-day tasks, of customizable and adaptive difficulty, and have its performance evaluated [18].

4.1. VR Hardware

Firstly, a choice had to be made regarding which VR system would be chosen to play the game on.

Initially it was intended for the game to be able to be played on as many VR systems as possible, however, this was later abandoned since there wasn't a universal software that would work for all platforms. Each VR brand has their own software, and with the VR systems being a relatively new technology, no universal software had yet been fully developed. This then led to focusing on the most important aspect the game required: It was relevant that the game be as immersive as possible, so the Oculus 2 Quest was chosen, which is wireless and supports *handtracking*, as well as being among the most accessible VR systems.

4.2. VR Software

For the development of the game, the Unity platform was chosen, since it supports development for the Oculus Quest 2, as well as it has ready to use frameworks with practical examples available. Three frameworks were explored during the development process: XR Toolkit (XR), OVR toolkit (OVR) and Mixed reality toolkit (MRTK). Initially, XR and OVR were both being used, with XR being the one most generally used, since XR allows integration with many VR systems, while OVR was specifically used on situations where *handtracking* was required, since XR does not support it. Later, XR was abandoned since it wasn't compatible with HTTC Vive (another VR system by steam), and replaced with MRTK, which had very similar capabilities, while being compatible with that system. It should be noted that at this point, the game was being developed to be played with either controllers or *handtracking*, so it was important to be compatible with many systems. This was later abandoned to focus solely on the Oculus Quest 2 system and *handtracking* only.

4.3. Game Specifications

The game will consist of going through a normal day and performing all kinds of daily tasks throughout. For this reason, the game will take place in a home environment (including a bedroom, bathroom, kitchen, living area and garden) as well as a city environment (including a marketplace). The tasks are to be done reflecting a day-to-day order, which means some areas or tasks may be repeated (for example, having lunch in the living area, and later returning for dinner). They are also divided into different time frames, such as morning, afternoon, and night, and they must be performed in a specific order in their specific area (teeth must be brushed in the bathroom) to be completed.

When all tasks are completed in an area in the current timeframe, the player may move on to the next area and/or timeframe.

The garden area has no tasks associated with it, being instead a general area where the player first starts the game, and calibrates the game as needed. It also serves as a general relaxing area that the player may visit to take any necessary breaks.

When all tasks in all areas and timeframes are completed, the day is completed, and the game ends.

The medical team will have access to an application in a separate computer where they can supervise and manipulate the game in real time (in the form of task difficulty). The game will also provide information about itself and the player to this application in real time (such has current progress, performance, among others.).

5. Discussion

Some changes had to be made as the development process evolved.

As stated before, initially it was intended for the game to be playable on as many VR systems as possible, which was not doable due to the lack of a common development software, so one specific VR system had to be chosen.

There were, however, other problems and details which had to be changed within the game plan itself.

Some assets within the environment had to be changed or removed, since it could be dangerous for schizophrenia patients to interact with these elements while forgetting they are in a virtual environment (for example, attempting to sit in a chair that exists in the game, but not in reality).

A calibration process had to be created to keep the player within the play area in the real world. During the game, the player can move in real space to move in game as part of the normal gameplay. However, when the area changes, the place the player is standing in, although it could be recorded for the next area, could be intersecting, or fully immersed in that area's assets (for example, appearing inside a table). To combat this, the player starts a new area in a specific position, in the centre of the area. However, this now means that there is a disparity between where the player is standing in real life (where they were before the area changed), and in game (the centre of the area). For this reason, a calibration process is required every time there is an area change, asking the player to return to the middle of the play area in real life, synchronizing the positions in real life and in game.

Some tasks also had to be entirely removed due to system limitations. *Handtracking* is still a fairly recent technology, which means it doesn't properly work in all situations. In particular, the Oculus Quest 2 heavily struggles to detect the hands when they are close to the cameras, which are in the headset, and they must be always visible in front of the cameras in order to be tracked. This meant that tasks that required the hands to be close to the head had to be discarded (for example, brushing teeth).

6. Project Current Stage and Future Work

The game is currently undergoing in its final development stages and will later be sent to a medical team for validation.

The medical team will test the game with schizophrenia patients, and later share the results.

Results are expected to include patient acceptance towards the game, any improvements to daily life provided by continued exposure to the game, comparison to traditional therapy results, among others.

Acknowledgements

This work was funded by the project "NORTE-01-0145-FEDER-000042", supported by Northern Portugal Regional Operational Program (Norte2020), under the Portugal 2020 Partnership Agreement, through the European Regional Development Fund (FEDER). It was also funded by national funds, through the FCT – Fundação para a Ciência e Tecnologia and FCT/MCTES in the scope of the project UIDB/05549/2020.

References

[1]. I. Siddiqui, S. Saperia, G. Fervaha, S. Da Silva, E. Jeffay, K. K. Zakzanis, O. Agid, G. Remington, and G. Foussias, Goal-directed planning and action impairments in schizophrenia evaluated in a virtual environment, *Schizophrenia Research,* Vol. 206, 2019, pp. 400–406.

[2]. M. Rus-Calafell, J. Gutierrez-Maldonado, and J. Ribas-Sabaté, A virtual reality-integrated program for improving social skills in patients with schizophrenia: A pilot study, *Journal of Behavior Therapy and Experimental Psychiatry,* Vol. 45, Issue 1, 2014, pp. 81–89.

[3]. R. M. E. M. da Costa and L. A. V. de Carvalho, The acceptance of virtual reality devices for cognitive rehabilitation: a report of positive results with schizophrenia, *Computer Methods and Programs in Biomedicine,* Vol. 73, Issue 3, 2004, pp. 173–182.

[4]. K. E. Greenwood, R. Morris, V. Smith, A.-M. Jones, D. Pearman, and T. Wykes, Virtual shopping: A viable alternative to direct assessment of real life function? *Schizophrenia Research*, Vol. 172, Issue 1, 2016, pp. 206–210.

[5]. I. Fajnerová, M. Rodriguez, F. Spaniel, J. Horáček, K. Vlcek, D. Levcik, A. Stuchlík, and C. Brom, Spatial navigation in virtual reality — from animal models towards schizophrenia: Spatial cognition tests based on animal research, in *Proceedings of the International Conference on Virtual Rehabilitation (ICVR),* 2015, pp. 44–50.

[6]. I. Fajnerová, M. Rodriguez, L. Konrádová, P. Mikolás, K. Dvorská, M. Ungrmanova, J. Horáček, K. Vlcek, D. Levcik, A. Stuchlík, and C. Brom, Spatial memory in a virtual arena: Human virtual analogue of the Morris water maze, in *Proceedings of the International Conference on Virtual Rehabilitation (ICVR),* 2013, pp. 186–187.

[7]. M. Kargar, S. Askari, A. Khoshaman, and A. Mohammadi, Differential diagnosis of schizophrenia and schizoaffective disorder from normal subjects using virtual reality, *Psychiatry Research,* Vol. 273, 2019, pp. 378–386.

[8]. E. A. Spieker, R. S. Astur, J. T. West, J. A. Griego, and L. M. Rowland, Spatial memory deficits in a virtual reality eight-arm radial maze in schizophrenia, *Schizophrenia Research,* Vol. 135, Issue 1, 2012, pp. 84–89.

[9]. G. Weniger and E. Irle, Allocentric memory impaired and egocentric memory intact as assessed by virtual reality in recent-onset schizophrenia, *Schizophrenia Research,* Vol. 101, Issue 1, 2008, pp. 201–209.

[10]. J. Siemerkus, E. Irle, C. Schmidt-Samoa, P. Dechent, and G. Weniger, Egocentric spatial learning in schizophrenia investigated with functional magnetic resonance imaging, *NeuroImage: Clinical*, Vol. 1, No. 1, 2012, pp. 153–163.

[11]. E. Bekele, D. Bian, J. Peterman, S. Park, and N. Sarkar, Design of a virtual reality system for affect analysis in facial expressions (vr-saafe); application to schizophrenia, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 25, Issue 6, 2017, pp. 739–749.

[12]. M. Dyck, M. Winbeck, S. Leiberg, Y. Chen, and K. Mathiak, Virtual faces as a tool to study emotion recognition deficits in schizophrenia, *Psychiatry Research,* Vol. 179, Issue 3, 2010, pp. 247–252.

[13]. I. Fajnerová, I. Oravcová, A. Plechatá, L. Hejtmánek, V. Sahula, K. Vlček, and T. Nekovárová, The virtual episodic memory task: Towards remediation in neuropsychiatric disorders, in *Proceedings of the International Conference on Virtual Rehabilitation (ICVR),* 2017, pp. 1–2.

[14]. C. Hawco, L. Buchy, M. Bodnar, S. Izadi, J. Dell'Elce, K. Messina, R. Joober, A. Malla, and M. Lepage, Source retrieval is not properly differentiated from object retrieval in early schizophrenia: An fmri study

using virtual reality, *NeuroImage: Clinical*, Vol. 7, 2015, pp. 336–346.
[15]. M. M. Tsang and D. W. Man, A virtual reality based vocational training system (vrvts) for people with schizophrenia in vocational rehabilitation, *Schizophrenia Research*, Vol. 144, Issue 1, 2013, pp. 51–62.
[16]. K. Hesse, P. A. Schroeder, J. Scheeff, S. Klingberg, and C. Plewnia, Experimental variation of social stress in virtual reality – feasibility and first results in patients with psychotic disorders, *Journal of Behavior Therapy and Experimental Psychiatry*, Vol. 56, 2017, pp. 129–136.
[17]. L. H. Adery, M. Ichinose, L. J. Torregrossa, J. Wade, H. Nichols, E. Bekele, D. Bian, A. Gizdic, E. Granholm, N. Sarkar, and S. Park, The acceptability and feasibility of a novel virtual reality based social skills training game for schizophrenia: Preliminary findings, *Psychiatry Research,* Vol. 270, 2018, pp. 496–502.
[18]. A. Mendes, D. Duque and V. Carvalho, A review on the use of virtual reality on therapy and diagnosis of schizophrenia, in *Proceedings of the 8th International Conference on Sensors and Electronic Instrumentation Advances (SEIA 2022),* Corfu, Greece, 21-23 September 2022, pp. 120-124.

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 45-51

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Measurement Site Selection and Validation in Extracorporeal Circuit for Predictive Maintenance of Blood Coagulation using Photoacoustic Imaging of LED Light Source

[1] Takahiro Wabe, [2,3] Ryo Suzuki, [4] Akimitsu Fujii, [5] Yohsuke Uchida, [2,3] Kazuo Maruyama and [1] Yasutaka Uchida

[1] Teikyo University of Science, Department of Life & Health Sciences, Faculty of Life & Environmental Sciences, 2-2-1 Senjyusakuragi, Adachi-ku, Tokyo, Japan
[2] Faculty of Pharma-Science, Teikyo University, 2-11-1 Kaga, Itabashi-ku, Tokyo, Japan
[3] Advanced Comprehensive Research Organization (ACRO),Teikyo University, 2-21-1 Kaga, Itabashi-ku, Tokyo, Japan
[4] Tokyo college of Medico-Pharmaco Technology, 2-11-1 Higashikasai, Edogawa-ku, Tokyo, Japan
[5] Shuto iko, 1-7-3 Nishisinjuku, Sinnjuku-ku, Tokyo, Japan
Tel.: + 8169101010, fax: + 8169103800
E-mail: takahiro091@gmail.com

Received: 5 September 2022 /Accepted: 7 October 2022 /Published: 31 October 2022

Abstract: Blood purification therapy that removes pathogenic substances and unnecessary blood components in blood is performed by extracorporeal circulation therapy that draws blood out of the body. In extracorporeal circulation therapy, blood is filled in the circuit and circulated outside the body by a pump, but the blood coagulates due to a foreign body reaction in the circuit. Anticoagulants such as heparin are used as blood coagulation measures, but they are not absolute. There is a need for a measurement method that can be predicted before blood coagulation. Photoacoustic imaging was used for predictive maintenance measurements. In photoacoustic imaging used for extracorporeal circulation therapy, the amount of data used must be reduced because it is constantly measured. Accordingly, in this study, we identified areas where photoacoustic imaging changes significantly over time due to blood coagulation, and reduced the volume of data by limiting the area to be measured to 233 pixels where 218240 pixels are normally required for measurement. In addition, because we were able to identify the measurement site, based on that, we develop a predictive maintenance system based on changes in blood coagulation over time in an extracorporeal blood circulation circuit that simulates the environment of extracorporeal blood circulation therapy.

Keywords: Blood coagulation, Photoacoustic imaging, LED, Extracorporeal circulation therapy, Predictive maintenance.

1. Introduction

Blood purification therapy is a treatment method that removes toxic substances and unnecessary blood components that cause diseases accumulated in the blood. Hemodialysis is a typical example, and it covers a wide range of areas such as apheresis and blood cell adsorption therapy [1-3]. The blood

purification therapy introduced here is a treatment method that utilizes extracorporeal circulation. Blood purification therapy uses a pump to pump blood out of the body to provide treatment tailored to each symptom. Blood drawn out of the body for treatment causes blood coagulation when it comes into contact with an artificial foreign substance in the circuit. When blood coagulation occurs, it causes blood clots, so it is customary to stop blood purification therapy when signs appear and then replace it with a new extracorporeal circulation circuit. Since the signs of blood coagulation cannot be detected in advance with a margin, the medical staff involved in the treatment is forced to take immediate action, which puts a heavy load on them. In addition, when replacing the extracorporeal circulation circuit, there is a problem that the patient's blood must be discarded together with the circuit. As a measure against blood coagulation that causes these problems, an anticoagulant typified by heparin is generally used [4-7]. However, anticoagulants are not absolute and it is difficult to completely prevent blood coagulation during extracorporeal circulation therapy. As a safety measure, blood coagulation is detected using a pressure sensor in the extracorporeal circulation circuit. The pressure sensor reacts ex post facto only after the circuit is clogged by blood coagulation or when the viscosity of the blood becomes extremely high and the pressure rises. Therefore, a highly sensitive measurement method (predictive maintenance) that can observe changes in the state of blood coagulation in extracorporeal circulation therapy in advance is required. Therefore, we proposed photoacoustic imaging as a new measurement method that continuously captures the progress of blood coagulation in the circuit and leads to predictive maintenance. The principle of photoacoustic imaging is as follows. The object to be measured irradiated with light energy causes volume expansion due to heat and generates elastic waves (photoacoustic waves). Photoacoustic waves are received and imaged by an ultrasonic probe. Photoacoustic imaging is a promising imaging technology for a wide range of biomedical applications with the potential for high contrast and high spatial resolution. By applying diagnostic imaging to photoacoustic imaging, blood coagulation in the extracorporeal circulation circuit we thought that predictive and conservative measurement was possible. In addition, since it is assumed that the measurement will be performed by photoacoustic imaging in the hospital, an LED light source was used as the light source instead of the laser light source used for conventional photoacoustic imaging. In the previous paper, we reported the basic experimental results showing the possibility of blood coagulation measurement by small photoacoustic imaging of measuring instruments [8-13]. When the photoacoustic imaging image is analyzed using the capacity as it is from the image obtained by the above measurement, the amount of data is large and the processing is burdensome. Therefore, instead of measuring the entire extracorporeal circulation circuit, we thought that we could try to reduce the volume of data within the range that does not affect predictive maintenance by identifying the sites where changes are significant before and after blood coagulation. In addition, the selection of the measurement target site will lead to the miniaturization of the light source and acoustic probe in the future. In selecting the measurement target site, the threshold value of the change in brightness over time, which is a feature amount for the photoacoustic imaging measurement data, was analyzed and examined using MATLAB.

In addition, using MATLAB (Diagnostic Feature Designer), we explore the possibility of developing a predictive maintenance system based on changes in blood coagulation over time in an extracorporeal blood circulation circuit that simulates the environment of extracorporeal blood circulation therapy.

2. Method

2.1. Photoacustic Imaging of the Air Trap Chamber (Bmode)

For the photoacoustic imaging device, AcousticX (CYBERDYNE, INC.), which uses an LED as a light source, was used. As a preliminary experiment, as a result of measuring the extracorporeal circulation circuit by photoacoustic imaging, the reaction of the air trap chamber was the strongest, and almost no reaction was seen in other parts. Therefore, the measurement site was further narrowed down using the air trap chamber as the measurement target. At the time of measurement, an extracorporeal circulatory device and circuit actually used in the medical field were used to flow blood. Fig. 1 shows a schematic diagram of the experimental equipment used in this study.

In addition, a transparent gel spacer (CYBERDYNE, INC.) With an acoustic impedance equivalent to that in water was directly sandwiched between the LED light source and the ultrasonic probe for measurement. The transparent gel spacer does not affect the application of light energy.

As shown in Fig. 1, two high-intensity high-speed pulse-driven LED light source arrays were attached to both sides of the ultrasonic probe using a jig. The angle was about 40°. The distance between the ultrasonic probe and the air trap chamber via the transparent gel spacer was about 1 cm. The LED array used a wavelength of 850 nm, and the light energy was about 200 μJ / pulse, the pulse width was 70 ns, and the repetition frequency was 4 kHz. An ultrasonic probe with a center frequency of 7 MHz was used. The addition average was performed 64 times.

When conducting blood purification therapy for extracorporeal circulation, a large amount of blood is used. It costs money to use a large amount of blood, and it takes time and money to dispose of it after the experiment. In the early stages of research, it is

difficult to use large amounts of blood within the budget, so we decided to use cheap, commercially available blood (sheep blood) that does not affect the reproducibility of blood coagulation in the circuit. The details of the sheep blood used are as follows. It is shown in Table 1.

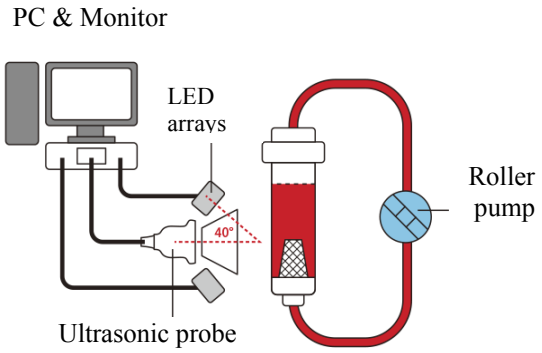


Fig. 1. Configuration diagram of the experiment.

Table 1. Blood used.

Aseptic storage blood of sheep	100ml/container
KOHJIN BIO	No.12070210
Anticoagulant	Alsever liquid

The commercially available blood used is anticoagulated with the anticoagulant ALSEVER'S SOLUTION for transport. Of course, blood does not coagulate as it is, so calcium, which is a coagulation factor, is required to promote blood coagulation. This time, we used calcium gluconate as a coagulation promoter. The upper limit solubility of calcium gluconate is 3.3 g per 100 ml. 1.1 g of calcium gluconate dissolved in 33 ml of saline was injected into 50 ml of blood to promote blood coagulation.

2.2. Selection of Measurement Points

ImageJ, a free software, was used to analyze the luminance of the B-mode image (grayscale image) obtained in Section 2.1, and MATLAB was used to extract the feature amount, and the change in luminance due to blood coagulation was analyzed.

For the input image, 4 × 4 pixels were set as one cell, and 1 × 10 cells were set as one block. Each time the analysis sites using ImageJ, an image processing software, were analyzed so as to overlap each other, the luminance distribution was investigated by shifting them one block at a time in the vertical direction. In the horizontal direction, 4 × 4 pixels were set as one cell. The analysis was performed in the range of 1 × 10 cells as one block. Fig. 2 shows the relative positional relationship between the allocated block numbers and the air trap chamber.

The brightness 126 - 173 used in the measurement was measured in the vertical and horizontal directions of the air trap chamber based on the experimental results reported previously. Vertical and horizontal measurements are described later in Tables 2 and 3, respectively. Based on these results, sites where the photoacoustic wave changes significantly before and after blood coagulation in the extracorporeal blood circulation circuit were identified. In order to show the relative positional relationship between the air trap chamber and the measurement site, the measurement results in Tables 3 and 4 were superimposed on the diagram of the air trap chamber to form Fig. 2.

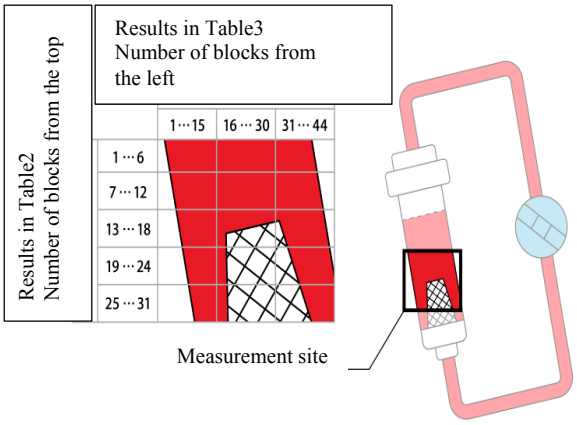


Fig. 2. Relative view of measurement results and air trap chamber.

2.3. Prototype of Predictive Maintenance Program for Changes in Blood Coagulation Over Time in the Extracorporeal Blood Circulation Circuit

Measurements were performed 11 times each at (a) 0 minutes after calcium gluconate injection, (b) 2 minutes, (c) 5 minutes, and (d) 10 minutes. The measurement results had to be biaxial for the convenience of analysis in MATLAB. Therefore, the luminance range was divided into two, 126 - 149 and 150 - 173. The number of cells showing brightness levels of 126 - 149 and 150 - 173 at the measurement site was counted. Furthermore, the time to blood coagulation was expressed as (a) Long, (b) Medium (c) Short, and the state of blood coagulation. In addition, 10 minutes after calcium gluconate injection, blood coagulation is completed, so it was defined as (d) Urgent. The measurement results are shown in Table 2.

Table 2. Measurement result by photoacoustic imaging to the selected luminance range and measurement site.

		Luminance		condition
		126～149	150～173	
min.	1	288	32	Long
	2	306	43	Long
0	3	397	2	Long
	4	424	0	Long
	5	437	0	Long
	6	423	0	Long
	7	432	0	Long
	8	431	0	Long
	9	459	0	Long
	10	482	0	Long
	11	512	0	Long
2	1	170	84	Medium
	2	174	81	Medium
	3	205	87	Medium
	4	237	78	Medium
	5	241	95	Medium
	6	283	96	Medium
	7	283	105	Medium
	8	328	81	Medium
	9	318	89	Medium
	10	330	80	Medium
	11	330	85	Medium
5	1	260	81	Short
	2	215	102	Short
	3	208	101	Short
	4	175	115	Short
	5	161	71	Short
	6	131	86	Short
	7	135	84	Short
	8	137	84	Short
	9	134	94	Short
	10	165	114	Short
	11	198	106	Short
10	1	311	56	Urgent
	2	304	60	Urgent
	3	308	62	Urgent
	4	288	69	Urgent
	5	288	63	Urgent
	6	292	59	Urgent
	7	260	77	Urgent
	8	238	71	Urgent
	9	223	65	Urgent
	10	227	57	Urgent
	11	212	67	Urgent

Since the number of measurements was small, we explored the temporal changes in blood coagulation in the extracorporeal blood circulation circuit by classification learning in order to explore the possibility of regression (supervised learning). Analysis results by MATLAB are shown in Fig. 3. The vertical axis represents the number of cells showing luminance levels 126-149 within the measurement range, and the horizontal axis represents the number of cells showing luminance levels 150-173. The blood coagulation states shown in Table 2 are expressed as (a) Long: blue, (b) Medium: red, (c) Short: yellow, and (d) Urgent: purple. In addition, ○: Correct and ×: False are ○: Correct if they match the prediction model created from the existing data in Table 2, and ×: False if they do not match. Since the number of measurements was small, cross-validation (5 divisions) was performed to prevent overfitting.

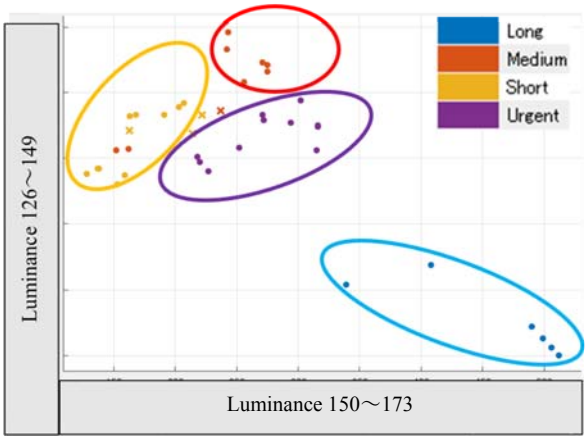


Fig. 3. Results of classification learning with MATLAB.

Long at 0 minutes after injecting calcium gluconate is concentrated in the lower right of the graph, and Medium, Short, and Urgent are gathered in the upper left. It can be seen that the distribution moves counterclockwise from Long as it becomes Urgent.

A suitable algorithm was examined for the data obtained in Tables 2. Fine KNN and fine Gaussian SVM had the highest accuracy at 90.96 %.

As an example, we confirmed the occurrence of prediction correctness for the true value of fine KNN. The confusion matrix obtained from fine KNN is shown in Fig. 4. The vertical line represents the actual class, and the horizontal line represents the predicted class. A mixture matrix is a classification table that expresses the number of measurements and the ratio of correctly predicted Urgent data to those whose actual data is Urgent when evaluating the performance after creating a learning system. It is possible to extract the feature amount of correctness judgment of prediction.

The positive predictive value (PPV) was 100 % for Long and Urgent, 88.9 % for Medium and 76.9 % for Short. The precision of fine KNN is also high, and the bias of prediction from the positive predictive value (PPV) is not high. We believe that a prototype for predictive maintenance could be made under the conditions in this section.

3. Results

3.1. Photoacoustic Imaging (B mode) Results of Air Trap Chamber

Fig. 5 shows the changes over time in blood coagulation that occurred in the air trap chamber using photoacoustic imaging (B mode). The circled area

indicates the wall surface of the air trap chamber, the left side of the wall surface is the transparent gel spacer, and the right side is the blood in the circuit. White streaks like ripples of blood in the circuit are photoacoustic waves due to blood coagulation.

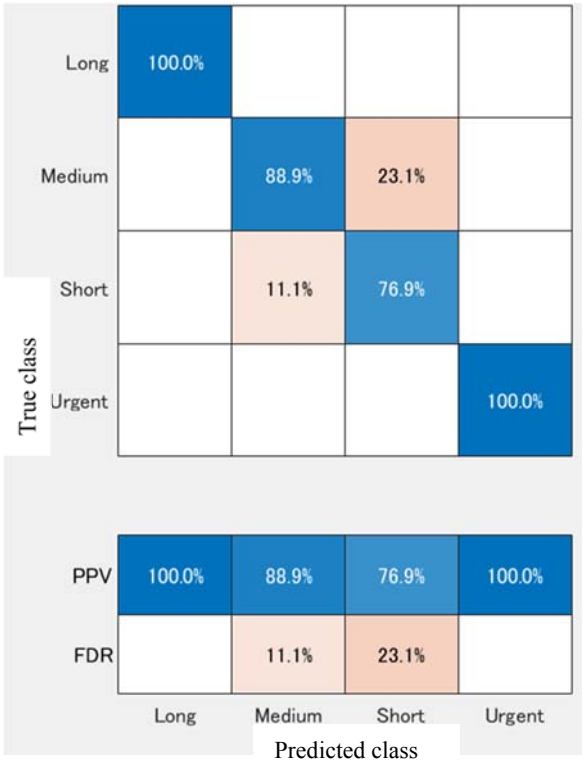


Fig. 4. Confusion matrix obtained from fine KNN.

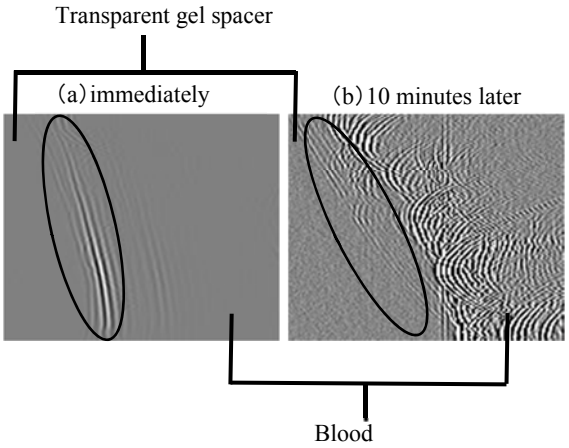


Fig. 5. B mode image by photoacoustic imaging.

3.2. Results of Selection of Measurement Points

The figure (grayscale image) obtained by photoacoustic imaging before and after blood coagulation was subjected to luminance analysis with ImageJ, and the feature quantity was extracted with Rank Features of MATLAB Diagnostic Feature Explorer. Before and after blood coagulation in the extracorporeal circulation circuit, we identified a site where changes in photoacoustic waves were remarkable. Feature extraction was performed by T-test in consideration of increasing the number of measurements in the future. Table 3 shows the measurement results in the vertical direction, and Table 4 shows the measurement results in the horizontal direction.

Table 3. Feature extraction of changes in brightness over time in grayscale images (Vertical direction).

Blocks from the top	T-test	Blocks from the	T-test
1	0.5706	17	0.7727
2	0.6335	18	0.8397
3	0.5937	19	0.7737
4	0.597	20	0.6836
5	0.6816	21	0.6737
6	0.7041	22	0.7307
7	0.7527	23	0.6914
8	0.7233	24	0.7023
9	0.6575	25	0.6722
10	0.7435	26	0.7083
11	0.6734	27	0.7075
12	0.7509	28	0.7581
13	0.6977	29	0.6901
14	0.7829	30	0.6655
15	0.7692	31	0.6762
16	0.8103		

Table 4. Feature extraction of changes in brightness over time in grayscale images (Horizontal direction).

Blocks from the left	T-test	Blocks from the left	T-test
1	0.2867	23	1.4659
2	0.2881	24	1.6364
3	0.2807	25	1.6005
4	0.3019	26	1.5406
5	0.3172	27	1.4555
6	0.3211	28	1.2525
7	0.3567	29	1.2027
8	0.4333	30	1.0604
9	0.4822	31	0.8732
10	0.4895	32	0.8315
11	0.5309	33	0.6781
12	0.6778	34	0.6012
13	0.8276	35	0.6029
14	0.7841	36	0.6345
15	0.7822	37	0.6445
16	0.798	38	0.5969
17	0.5462	39	0.6622
18	0.7156	40	0.6299
19	1.0942	41	0.6708
20	1.5675	42	0.7122
21	1.3059	43	0.6332
22	1.2722	44	0.6036

Table 5 shows a part of the two-dimensional mapping of each component of Table 3 and Table 4. The grayscale image obtained by the measurement was divided into 44 in the horizontal direction and 31 in the vertical direction. The parts showing remarkable changes (0.9 or more) are shown in gray. The shaded areas in the rows and columns are the areas where the T-test changed significantly before and after blood coagulation in the horizontal and vertical directions, respectively. All blocks 19-29 in Table 4 and blocks 13-31 in Table 3 correspond to the mesh filter in the air trap chamber.

From Table 5, it was found that it is possible to measure significant changes over time in photoacoustic waves before and after blood coagulation near and directly above the mesh filter inside the air trap chamber. Blood coagulation appears prominently near the mesh filter because it is a filter that receives blood flow from the front, so it is presumed that the foreign body reaction is stronger than in other parts. It is also possible that blood coagulation (blood clot) generated in other parts of the circuit has peeled off and is caught in the mesh filter. Furthermore, due to the blood flow directly above the mesh filter, blood clots and the like stay and the reaction is thought to be high.

4. Conclusions

In order to realize predictive and conservative measurement of blood coagulation that occurs in the extracorporeal circulation circuit used in the treatment method using extracorporeal circulation represented by blood purification therapy, the previously proposed photoacoustic imaging raises the following problems. (1) The amount of obtained image data is large, and (2) the light source and ultrasonic probe for measurement are large. In order to solve these problems, in this paper, we measured the changes over time in blood coagulation in the extracorporeal circulation circuit, identified the sites where photoacoustic waves were prominent, and selected the measurement sites. From Table 5, by setting the threshold value of 0.9, which is the product of the results of the horizontal and vertical photoacoustic waves, it was possible to select the measurement site where the change is remarkable before and after blood coagulation (gray part). As a result, the measurement range is reduced, and the possibility of reducing the size of the photoacoustic imaging measurement unit (light source, ultrasonic probe) is shown. In addition, when measuring the entire air trap chamber, 218240 pixels are required, but the measurement target site is narrowed down to 233 pixexel to reduce the dose of data by about 1/1000, and the extracorporeal blood circulation circuit using photoacoustic imaging. We were able to reduce the data handled for predictive maintenance of blood coagulation in the body. Therefore, it is considered that the problem of photoacoustic imaging operation for predictive maintenance measurement of blood coagulation occurring in the extracorporeal circulation circuit has been solved.

Machine learning was performed using MATLAB from grayscale images obtained from photoacoustic imaging of the time course of blood coagulation in the extracorporeal blood circulate learning, we made a prototype with supervised learning classification, and found that fine KNN and fine Gaussian SVM were the most accurate prediction models with an accuracy of 90.9 %. From the positive predictive value (PPV) of the mixed matrix, the prediction of Long and Urgent before and after clear blood coagulation was 100 %, with on circuit. In machine the data easily distinguishable, while Medium was 88.9 % and Short was 76.9 %.

Therefore, a predictive maintenance system based on the temporal change of blood coagulation in an extracorporeal circulation circuit simulating the environment of extracorporeal circulation therapy was investigated.

Acknowledgments

T. Wabe and Y. Uchida would like to thank Mr. N. Sato of CYBERDYNE, INC Research and Development Dept, for his cooperation in this experiment.

References

[1]. J. Utley, Pathophysiology of cardiopulmonary bypass: current issues, *J. Card. Surg.*, Vol. 5, Issue 3, 1990, pp. 177-189.

[2]. Y. Mori, Blood compatible material, Kobunshi Ronbunshu, Vol. 42, No. 42, October 1985, pp. 601－615 (in Japanese).

[3]. J. K. Kirklin, A. D. Pacifico, Complement and the damaging effects of cardiopulmonary bypass, *J. Thorac. Cardiovasc. Surg.*, Vol. 86, Issue 6, 1983, pp. 845-857.

[4]. L. Gott, J. D. Whiffen, R. C. Dutton, Heparin bonding on colloidal graphite surfaces, *Science*, Vol. 142, Issue 3597, 1963, pp. 1297-1298.

[5]. I. O. Salyer, Medical Application of Plastic, in Biomedical Material Symposium No. 1, *Interscience*, New York, Vol. 105, 1971.

[6]. M. Murase, A. Usui, M. Maeda, Y. Tomita, F. Murakami, K. Teranishi, T. Koyama, T. Ito, Ot. Abe, Nafamostat mesilate reduces blood loss during open heart surgery, *Circulation*, Vol. 88, Issue 5 Pt2, November 1993, pp. 11432-11436.

[7]. M. Hiroura, A. Usui, M. Kawamura, M. Hibi, K. Yoshida, F. Murakami, J. Iwase, Nafamostat mesilate reduces bloodcell adhesion to cardiopulmonary bypass circuit: an invitro study, *J. Extra Corpor. Technol.*, Vol. 26, Issue 3, September 1994, pp. 121- 125.

[8]. T. Wabe, R. Suzuki, K. Maruyama, Y. Uchida, Possibility for temporal observationof thrombus generated in extracorporeal circulator circuit by photoacoustic imaging using LED, in *Proceedings of the 5th International Conference on Sensors Engineering and Electronics Instrumentation*

Advances (SEIA'19), Canary Islands (Tenerife), Spain, 25-27 September 2019, pp. 157-160.

[9]. T. Wabe, R. Suzuki, K. M*aruyama, Y. Uchida,* The Measurement of Blood Coagulation Process in Extracorporeal Circuit Using LED Photoacoustic Imaging, *Sensors & Transducers*, Vol. 237, Issue 9-10, September-October 2019, pp. 88-94.

[10]. T. Wabe, R. Suzuki, K. Maruyama, Y. Uchida, Possibility for temporal observation of thrombus generated in extracorporeal circulator circuit by photoacoustic imaging using LED, in *Proceedings of the 5th International Conference on Sensors Engineering and Electronics Instrumentation Advances (SEIA' 2019),* Canary Islands (Tenerife). Spain.25-27 September 2019. pp. 157-160.

[11]. T. Wabe, R. Suzuki, K. Maruyama, Y. Uchida, Possibility of Extracting Feature Value from the Changes in Brightness over Time of Blood Coagulation in the Extracorporeal Circuit, *Sensors & Transducers,* Vol. 246, Issue 7, November 2020, pp. 64-70.

[12]. T. Wabe, R. Suzuki, A. Fujii, Y. Uchida, K. Maruyama, Yasutaka Uchida, Basic Research on Blood Coagulation Measurement of Extracorporeal Circulation Circuit Using Photoacoustic Imaging of LED Light Source, *Sensors & Transducers,* Vol. 253, Issue 6, November 2021, pp. 1-8.

[13]. T. Wabe, R. Suzuki, A. Fujii, Y. Uchida, K. Maruyama, Y. Uchida, Selection of measurement site for predictive maintenance of blood coagulation in an extracorporeal circulation circuit using LED photoacoustic imaging and an extracorporeal circulation device, *Proceedings of the 8th International Conference on Sensors Engineering and Electronics Instrumentation Advances (SEIA' 2022)*, 21-23 September 2022, pp. 55-59.

Table 5. Selection of measurement points.

Table3		19	20	21	22	23	24	25	26	27	28	29
Table4	T-test	1.0942	1.5675	1.3059	1.2722	1.4659	1.6364	1.6005	1.5406	1.4555	1.2525	1.2027
1	0.5706	0.624351	0.894416	0.745147	0.725917	0.836443	0.93373	0.913245	0.879066	0.830508	0.714677	0.686261
2	0.6335	0.693176	0.993011	0.827288	0.805939	0.928648	1.036659	1.013917	0.97597	0.922059	0.793459	0.76191
3	0.5937	0.649627	0.930625	0.775313	0.755305	0.870305	0.971531	0.950217	0.914654	0.86413	0.743609	0.714043
4	0.597	0.653237	0.935798	0.779622	0.759503	0.875142	0.976931	0.955499	0.919738	0.868934	0.747743	0.718012
5	0.6816	0.745807	1.068408	0.890101	0.867132	0.999157	1.11537	1.090901	1.050073	0.992069	0.853704	0.81976
6	0.7041	0.770426	1.103677	0.919484	0.895756	1.03214	1.152189	1.126912	1.084736	1.024818	0.881885	0.846821
7	0.7527	0.823604	1.179857	0.982951	0.957585	1.103383	1.231718	1.204696	1.15961	1.095555	0.942757	0.905272
8	0.7233	0.791435	1.133773	0.944557	0.920182	1.060285	1.183608	1.157642	1.114316	1.052763	0.905933	0.869913
9	0.6575	0.719437	1.030631	0.858629	0.836472	0.963829	1.075933	1.052329	1.012945	0.956991	0.823519	0.790775
10	0.7435	0.813538	1.165436	0.970937	0.945881	1.089897	1.216663	1.189972	1.145436	1.082164	0.931234	0.894207
11	0.6734	0.736834	1.055555	0.879393	0.856699	0.987137	1.101952	1.077777	1.03744	0.980134	0.843434	0.809898
12	0.7509	0.821635	1.177036	0.9806	0.955295	1.100744	1.228773	1.201815	1.156837	1.092935	0.940502	0.903107
13	0.6977	0.763423	1.093645	0.911126	0.887614	1.022758	1.141716	1.116669	1.074877	1.015502	0.873869	0.839124
14	0.7829	0.856649	1.227196	1.022389	0.996005	1.147653	1.281138	1.253031	1.206136	1.139511	0.980582	0.941594
15	0.7692	0.841659	1.205721	1.004498	0.978576	1.12757	1.258719	1.231105	1.18503	1.119571	0.963423	0.925117
16	0.8103	0.88663	1.270145	1.058171	1.030864	1.187819	1.325975	1.296885	1.248348	1.179392	1.014901	0.974548
17	0.7727	0.845488	1.211207	1.009069	0.983029	1.132701	1.264446	1.236706	1.190422	1.124665	0.967807	0.929326
18	0.8397	0.9188	1.31623	1.096564	1.068266	1.230916	1.374085	1.34394	1.293642	1.222183	1.051724	1.009907
19	0.7737	0.846583	1.212775	1.010375	0.984301	1.134167	1.266083	1.238307	1.191962	1.12612	0.969059	0.930529
20	0.6836	0.747995	1.071543	0.892713	0.869676	1.002089	1.118643	1.094102	1.053154	0.99498	0.856209	0.822166
21	0.6737	0.737163	1.056025	0.879785	0.857081	0.987577	1.102443	1.078257	1.037902	0.98057	0.843809	0.810259
22	0.7307	0.799532	1.145372	0.954221	0.929597	1.071133	1.195717	1.169485	1.125716	1.063534	0.915202	0.878813
23	0.6914	0.75653	1.08377	0.902899	0.879599	1.013523	1.131407	1.106586	1.065171	1.006333	0.865979	0.831547
24	0.7023	0.768457	1.100855	0.917134	0.893466	1.029502	1.149244	1.124031	1.081963	1.022198	0.879631	0.844656
25	0.6722	0.735521	1.053674	0.877826	0.855173	0.985378	1.099988	1.075856	1.035591	0.978387	0.841931	0.808455
26	0.7083	0.775022	1.11026	0.924969	0.901099	1.038297	1.159062	1.133634	1.091207	1.030931	0.887146	0.851872
27	0.7075	0.774147	1.109006	0.923924	0.900082	1.037124	1.157753	1.132354	1.089975	1.029766	0.886144	0.85091
28	0.7581	0.829513	1.188322	0.990003	0.964455	1.111299	1.240555	1.213339	1.167929	1.103415	0.94952	0.911767
29	0.6901	0.755107	1.081732	0.901202	0.877945	1.011618	1.12928	1.104505	1.063168	1.004441	0.86435	0.829983
30	0.6655	0.72819	1.043171	0.869076	0.846649	0.975556	1.089024	1.065133	1.025269	0.968635	0.833539	0.800397
31	0.6762	0.739898	1.059944	0.88305	0.860262	0.991242	1.106534	1.082258	1.041754	0.984209	0.846941	0.813266

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 52-60

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Advantages and Limits of Weighing Pharmaceutical Gelatin Capsules with Dual Energy X-ray Transmission

[*] Christine BAUER, Rebecca WAGNER and Alexander ENNEN
Fraunhofer IIS, Fraunhofer Institute for Integrated Circuits IIS, Division Development Center X-Ray Technology, Flugplatzstr. 75, 90768 Fürth, Germany
[1] Tel.: +49 911 58061-7663, fax: +49 911 58061-7299
E-mail: christine.bauer@iis.fraunhofer.de

Received: 2 September 2022 /Accepted: 10 October 2022 /Published: 31 October 2022

Abstract: Quality control is an important step during the manufacturing of products and of crucial interest in the field of pharmaceuticals. These products need to be clean, in a stable condition and efficient. Pharmaceutical capsules should not contain any foreign items, be precisely filled and in an immaculate state. Unfortunately, the weight of the capsule shells alone can vary due to production variations. Given that mainly the filled capsules are weighed to check the correct filling, the amount of filling itself may vary and lead to unwanted side effects or decreasing effectiveness. Dual energy X-ray transmission may be a powerful tool in non-destructive testing and can be used for weighing materials. In this approach, five types of dietary supplements in pharmaceutical gelatin capsules with varying weight were examined and weighed to determine the limitation and advantages of this method. The samples were investigated in movement both in their blister packaging and without.

Keywords: Dual energy, X-ray transmission, Basis material decomposition, Drugs, Drug quality, Drug weighing, Capsule.

1. Introduction

Gelatin hard capsules are widely used in the field of pharmaceuticals. They may be filled with different kinds of powder or granulate and carry highly diverse contents, such as dietary supplements or active pharmaceutical ingredients. Depending on the filling, slight variances in the amount of pharmaceutical ingredients may have a tremendous impact on the drug's efficacy and its adverse effects for the patient.

In the production process of gelatin capsules it is possible to measure the weight of filled capsules on a conveyor belt or in the packaging machine [1] for example by measuring the capacitance profile of the capsule [2]. Unfortunately, the weight of the filling alone cannot be determined. This would be of great interest as the mass of the capsule shells are varying due to production variations or differences between the various manufacturers.

X-ray transmission (XRT) may offer a solution for the above described problem. With this technique, the attenuation, which is proportional to the mass of the transmitted object, can be obtained. Unfortunately, this includes not only the filling but also the shell and (if existing) the packaging. If those two vary, errors in the estimation of the mass of the filling or the filled capsule occur.

In contrast to standard XRT, the dual energy approach (DE-XRT) may offer a more robust solution. The method is well known since 1976 [3] and proved to be a powerful tool in non-destructive testing. The most prominent usage are luggage scanners at airports, but DE-XRT also shows promising results in the sectors of recycling [4], mining [5], food security [6]

and sorting applications [7]. The method in combination with an analysis algorithm provides information about the material of the examined objects. The results are images showing the areal density of two chosen materials [8]. With this approach it may be possible to separate the signals of filling and shell. Thus to weigh the filling inside the capsule or to weigh filled capsules inside a package.

As X-rays have no effect on the quality of the product, DE-XRT is save to use in pharmaceutical environments [9].

Two approaches were used to test the idea of weighing objects by DE-XRT. The first is to determine the weight of capsules (shell and filling) in their blister packaging, the second is to obtain the weight of the filling alone for capsules outside of the package. The principle is the same from a technical point of view.

This paper gives an overview over the examined samples and provides an introduction to DE-XRT and the weighing of objects by this method. Results for capsules with five different fillings of dietary supplements are presented for measurements with and without blister packaging. The benefits and limitations of the method are illustrated.

2. Experimental

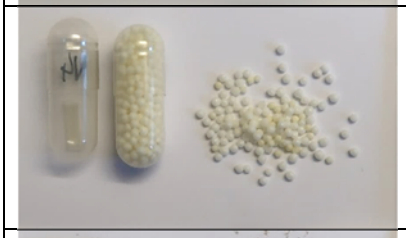
2.1. Sample Materials

After conducting a proof-of-principle study with one capsule type [10], five different types of filled pharmaceutical gelatin capsules were examined for this study. All samples are over the counter dietary supplements and were bought in a drugstore. There is no relation between the authors of this paper and the drugstore or the manufacturer. The samples were randomly chosen, not based on a prior estimation of their X-ray properties. Two of the samples contain granulated material while the other three are filled with powder. The gelatin capsules themselves were transparent or dyed and could be opened by hand without being destroyed. The studied samples and their ingredients can be seen in Table 1 and Table 2, respectively.

The capsules were opened and loaded with different amounts of filling to cover a wide range of weights and to prove that the weight is proportional to the areal density (see Table 3). For every sample type, five empty capsule shells were measured as well as three to seven capsules with their original amount of filling. In total, ten to twenty-five filled capsules of each type were examined (for more details see Table 1).

All samples were weighed previously with an analytical balance (Kern ACJ 220-4M, weighing range (*max*) = 220 g, minimum load (*min*) = 10 mg, verification value (*e*) = 1 mg, readability (*d*) = 0.1 mg) before performing the DE-XRT measurements. The weighing was conducted three times for the capsules with filling, as well as the filling and the capsule shell alone. The maximum difference that occurred between the three measurements was in the region of the last digit in agreement to the verification value of 1 mg.

Table 1. Studied samples.

	**Vitamin B complex depot** 25 samples (7 originally filled)
	**Zinc + vitamin C depot** 20 samples (5 originally filled)
	**Devil's claw** 10 samples (3 originally filled)
	**Immune complex** 10 samples (5 originally filled)
	**L-carnitin** 10 samples (4 originally filled)

The weights of the filled capsules, as well as of shells and fillings alone are given in Table 3. The weight of the original filling of the vitamin B capsules was determined to (455 ± 5) mg. Fig. 1 shows a plot of the weight of the shells of these capsules. Their mean weight is 96.6 mg with a standard deviation of 2.2 mg and 10 mg difference between the lightest and the heaviest shell. These two values amount to approximately 0.5 % and 2 % of the weight of the original filling. If only the total mass of the filled capsules would be measured during quality control, these uncertainties in the capsule shell would propagate to the uncertainty of the mass of the filling. Thus, separate weighing of capsule shell and filling makes sense.

Table 2. Ingredients of the five studied samples.

Sample	Ingredients
Vitamin B complex	Sugar, gelatin, nicotinamide (niacin), maltodextrin, starch, calcium D-pantothenate (pantothenic acid), coating agent: shellac, pyridoxine hydrochloride (vitamin B6), riboflavin (vitamin B2), thiamine mononitrate (vitamin B1), anti-caking agents: silicon dioxide and talc, glucose syrup, vegetable fat (coconut), pteroylmonoglutamic acid (folic acid), sodium chloride, colorant: iron oxide, D-biotin, cyanocobalamin (vitamin B12).
Zinc + C	L-ascorbic acid (vitamin C), sugar, gelatin, coating agent: shellac, cornstarch, zinc sulfate, glucose syrup, vegetable oil (coconut), release agent: talc.
Devil's claw	Devil's claw root powder (43 %, Harpagophytum Procumbens), maltodex-trin, gelatin (bovine), L-ascorbic acid (vitamin C), anti-caking agent: magnesium salts of fatty acids, cholecalciferol (vitamin D)
Immune complex	L-ascorbic acid (vitamin C), gelatin, bulking agent: calcium phosphate, ferrous sulfate, anti-caking agents: magnesium salts of fatty acids and silicon dioxide, zinc oxide, copper sulfate, pyridoxine hydrochloride (vitamin B6), retinyl acetate (vitamin A), Pteroyl monoglutamic acid (folic acid), sodium selenate, cholecalciferol (vitamin D3), cyanocobalamin (vitamin B12)
L-carnitin	L-carnitine L-tartrate, magnesium oxide, gelatin (bovine), anti-caking agent: magnesium salts of fatty acids, pyridoxine hydrochloride (vitamin B6), cyanocobalamin (vitamin B12)

Table 3. Minimum, maximum, mean and standard deviation (SD) of the weight of the capsule shell alone, the filling and the refilled samples in mg. The weight of the originally filled capsules is the mean value over the number of capsules with original filling given in Table 1.

[mg]		Vitamin B complex	Zinc + C	Devil's claw	Immune complex	L-carnitin
Shell	Min	91.0	93.1	92.3	93.5	113.2
	Max	101.1	101.7	97.3	101.4	120.3
	Mean	96.6	97.2	95.2	96.6	115.5
	SD	2.2	2.2	1.5	2.1	1.8
Orig. Filling	Min	449.5	511.2	489.9	438.4	845.8
	Max	463.8	547.4	505.5	452.3	925.8
	Mean	454.6	525.7	497.1	446.2	879.8
	SD	4.9	13.3	7.9	5.7	36.5
Filling	Min	54.9	47.0	183.8	155.3	244.7
	Max	600.9	641.7	505.5	535.4	925.8
	Mean	338.3	391.2	382.8	387.9	660.3
	SD	149.4	190.3	109.5	119.1	237.2
Capsule + filling	Min	151.9	146.6	277.0	250.9	355.9
	Max	694.6	737.2	590.6	629.7	1040.2
	Mean	434.6	488.7	470.3	481.2	774.0
	SD	149.7	190.0	105.3	118.3	237.4

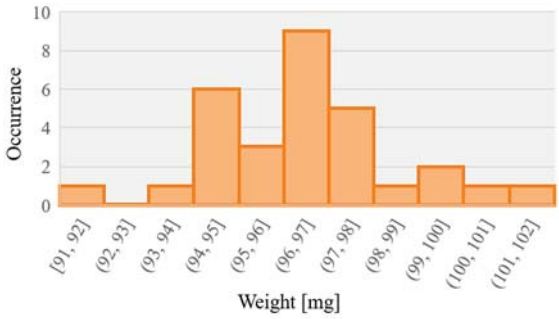


Fig. 1. Distribution of weight of vitamin B capsule shells measured by an analytical balance.

2.2. Dual Energy X-ray Transmission Measurements

The DE-XRT measurements were carried out with a Comet MXR-225HP/11 X-ray source in combination with a Hamamatsu C11800-08U detector. This dual energy line detector has a pixel pitch of 0.4 mm. It contains two detection layers. The first records the low energy (LE) photons. Thereby it acts as a filter for the second layer, which collects the high energy (HE) photons. Thus, two images are recorded at the same time. The measurements were performed with a tube voltage of 70 kV and a tube current of 6.8 mA. The samples were placed in a drawer system (Fig. 2) and moved between source and detector with a velocity of approx. 170 mm/s, which corresponds to an exposure time per recorded line of 2.67 ms.

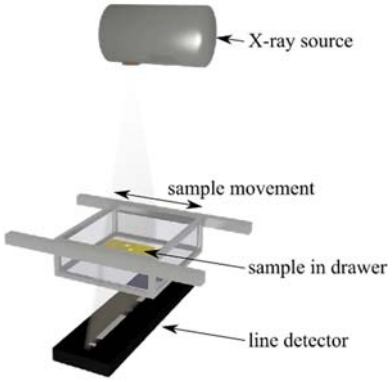


Fig. 2. Schematic of the used DE-XRT drawer system (from [11], licensed under CC BY 4.0).

2.3. Basis Material Decomposition

For analyzing the two images obtained by the dual energy detector, a method called basis material decomposition (BMD) was used. This method provides a new set of images showing the areal density (mass per area) of two chosen basis materials for every pixel (Fig. 3) [11, 12].

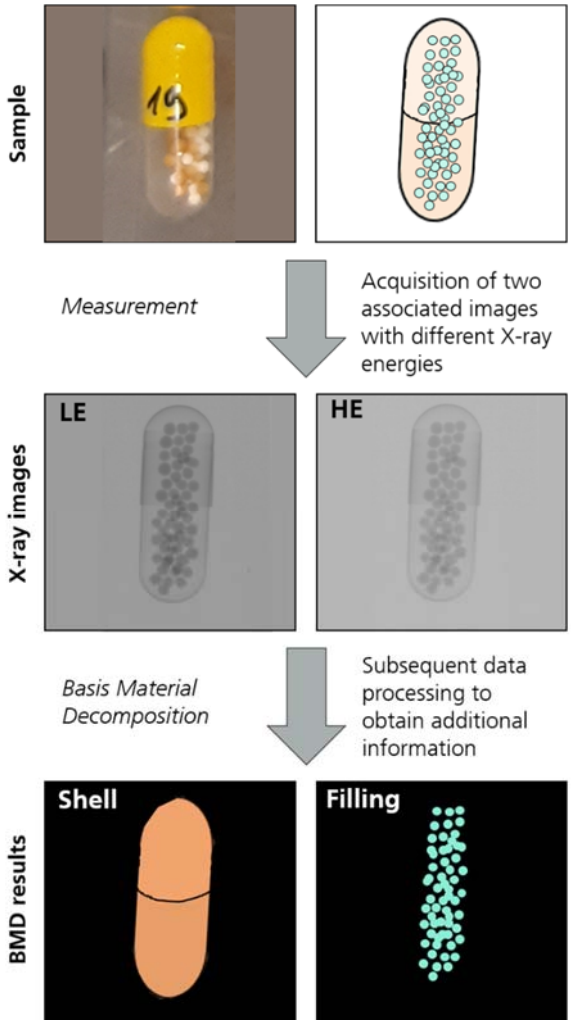


Fig. 3. Scheme of the workflow showing a sample, DE-XRT images and result obtained after data processing.

The foundation of BMD is the dependence of the intensity of X-rays I transmitted through an object on this objects' mass attenuation coefficient μ' and its areal density p, as well as on the X-ray energy E. This dependence is described by Lambert-Beer's law:

$$I(E) = I_0(E)\exp(-\mu'(E)p). \quad (1)$$

I_0 is the unattenuated intensity to which the object is exposed. In case of laboratory X-ray sources, $I_0(E)$ is a polychromatic spectrum. It can be found by simulations or measurements [13].

For X-rays passing through more than one material, equation (1) transforms into

$$I(E) = I_0(E)\exp(-\sum_i \mu'_i(E)p_i), \quad (2)$$

where i counts the materials.

The mass attenuation coefficients are published in tables for chemical elements [14]. For complex materials, they can be calculated from the weight fractions w of their chemical elements using [14]

$$\mu' = \sum_j w_j \mu'_j. \quad (3)$$

The intensity detected behind an object can be calculated by integrating over the energy in equation (2). When the samples examined by DE-XRT contain two materials with known mass attenuation coefficients, equation (2) contains two unknowns: the areal densities of the two materials. Thus, when two measurements are executed with different X-ray energies, a set of two equations with two unknowns is created. Solving this set of equations results in the areal densities of the two basis materials for every pixel (for more details see [5]). Thus, two basis material images are created. Each of them shows the areal density of only one of the basic materials as illustrated schematically in Fig. 3. If a third material is present in the examined object, it appears as a superposition of the two chosen basis materials.

In the application presented in this paper, the blister packaging, the capsule shells and the filling were not pure elements but chemically complex objects. Even though the ingredients are known from the folding box (Table 2), their relative amount is unknown. Thus, the mass attenuation coefficients cannot be calculated using equation (3). A suitable description had to be determined experimentally. To this end, DE-XRT images of blister packaging, empty capsules and fillings were recorded. For each of these components, the effective atomic numbers Z_{eff} of suitable basis materials were varied in the subsequent BMD analysis, until the currently considered component was visible in only one of the basis material images. Here, non-integer values of Z_{eff} indicate that the mass attenuation coefficient is a weighted average of the coefficients of the neighboring chemical elements. For instance, Z_{eff} = 7.4 means that the X-ray attenuation is between that of N (Z = 7) and O (Z = 8), the values of Z_{eff} found in this way are given in Table 4.

Table 4. Effective atomic numbers of the package, the capsules and its filling used for the BMD analysis. The zinc + C depot filling has two components.

Z_{eff}	Package	Capsule	Filling
Vitamin B complex	11.0	6.4	7.4
Zinc + C	11.0	6.4	7.2 20.0
Devil's claw	11.0	7.2	7.4
Immune complex	11.0	7.0	10.5
L-carnitin	11.0	7.5	8.3

In case of the blister packaging, this value comes about as a superposition with the tape used to fix it to the drawer during measurements.

2.4. Weighing by DE-XRT

The areal density obtained by BMD for every pixel is proportional to the weight. Therefore, to obtain an estimate of the mass of a particular capsule, the areal densities of all pixels belonging to it have to be

summed up. Multiplication by the area covered by a pixel gives the mass.

The correlation between capsules and their related pixels was obtained using a binary image. The first step to this end was only done for measurements of capsules in the blister packaging. Here, an X-ray projection of an empty package was subtracted from the projection of the filled pack to increase the contrast between capsule and environment. The X-ray projections were then binarized by thresholding. Morphological operations were used to remove single bright and dark pixels. The resulting binary images showed only the filled capsules. Then, connected pixels were assigned to one object, that is, to one capsule. The binary image was then used as a mask for the basis material image.

3. Results and Discussion

For this publication, two approaches were tested. At first, the filled capsules were weighed by DE-XRT inside of their blister packaging (Fig. 4, left). In a second experiment, the samples were measured outside of the packaging (Fig. 4, right).

Fig. 4. Left: Photograph of vitamin B complex capsules in their blister package. Right: Photograph of loose vitamin B complex capsules.

3.1. Packaging vs. Filled Capsule

Fig. 4 shows a photograph of vitamin B complex capsules in their packaging. The DE-XRT images (Fig. 5) show capsules with different amounts of filling in their blister package. Although the gray value is proportional to the amount of filling (a darker gray value means more attenuation due to more mass or higher μ'), it is also influenced by the local thickness of the package.

Fig. 6 shows the basis material images obtained by BMD. While one of them shows only the blister packaging, the other shows the capsules. The effective atomic numbers used to create these images were Z_{eff} = 11.0 (packaging) and Z_{eff} = 7.4 (filling). As a consequence, the capsule shell (Z_{eff} = 6.4) is visible in the basis material image of the filling, too. However, its signal in this image is slightly too high, as the chosen basis materials do not match the shell's effective atomic number.

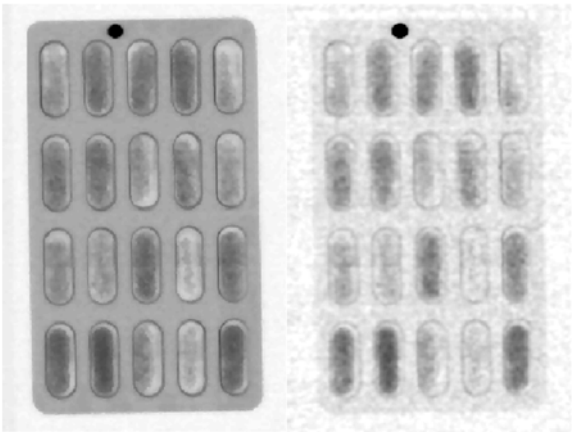


Fig. 5. Left: Low energy XRT image of the vitamin B complex capsules. Right: High energy XRT image of the vitamin B complex capsules.

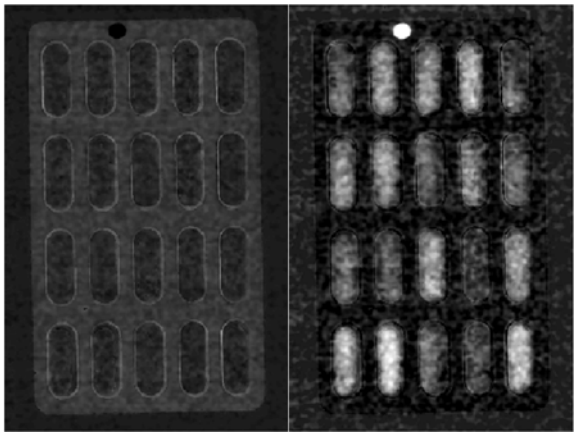


Fig. 6. Left: BMD image of material 1 (packaging) with Z_{eff} = 11.0. Right: BMD image of material 2 (vitamin B complex capsule and filling) with Z_{eff} = 7.4.

The weight of the capsules obtained from the basis material images is proportional to the values obtained with the analytical balance (Fig. 7). A line of the form

$$y = mx + n \tag{4}$$

was fitted to the data points. The slope m is slightly higher than expected, mostly due to inaccuracies in the determination of Z_{eff}. This systematic error would not cause problems in an application as it is possible to calibrate the weight determined by DE-XRT with an analytical balance. The y-intercept n is close to zero, as expected. The scattering of the data points around the fit is caused by random errors in the weight determined by DE-XRT. These are mainly due to the limited signal to noise ratio (SNR) of the X-ray

projection. The noise is not only propagated to the basis material images, but also limits the accuracy of the binarization (section 2.4). Possibilities to increase the SNR are discussed in section 4.

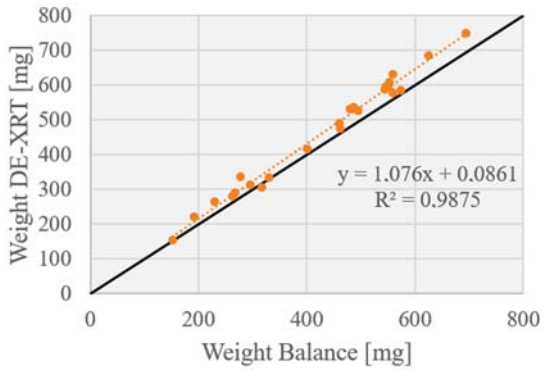


Fig. 7. Weight estimated by BMD versus weight of capsule shell and filling determined by an analytical balance for each vitamin B complex sample. The black line indicates the ideal fit for the weight determined by DE-XRT and the analytical balance.

The parameters for the line fits of the five different kinds of capsules are given in Table 5. In all cases, the first basis material was chosen to represent the packaging, the second to represent the filling (for the respective values of Z_{eff}, see Table 4).

Table 5. Parameters for linear fits of weight determined by DE-XRT versus weight measured with an analytical balance for filled capsules in their blister package.

Sample	**Slope *m***	***n* [mg]**	***R*²**
Vitamin B complex	1.08	0.09	0.9875
Zinc + C	0.99	8.59	0.9902
Devil's claw	1.07	-27.55	0.9781
Immune complex	0.93	273.37	0.4898
L-carnitin	1.04	22.87	0.9881
Ideal fit	1	0	1

For Devil's claw, the weight found by DE-XRT agrees well with that measured by the analytical balance.

The zinc + C depot filling consist of two components with different X-ray attenuation (Fig. 8(a)). Only the component with the lower effective atomic number, which constitutes the higher fraction of the filling, is used for the weight estimation by BMD. The other component is visible in the basis material image of the blister packaging. Still, the parameters of the line fit are closer to the ideal case than for any of the other samples. The scattering is low, showing that the presence of two components does not deteriorate the result. Even if this was the case, as long as the ratio between both components is constant, a calibration with an analytical balance would be possible.

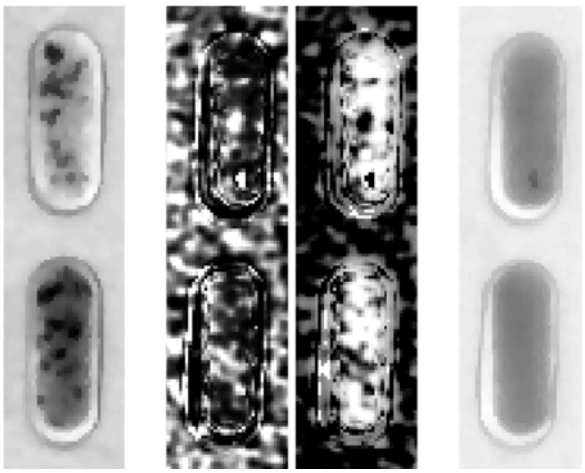


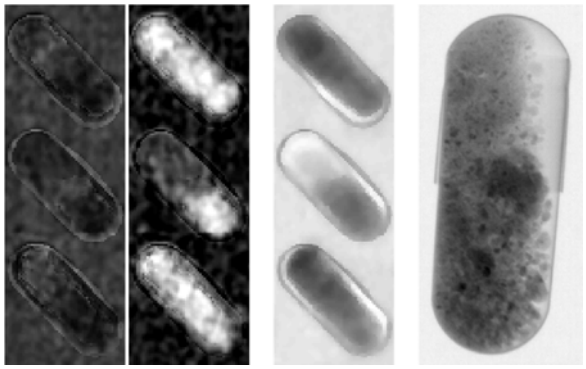
Fig. 8. (a) LE projection of zinc + C depot capsules showing two components with different attenuation, (b) basis material images of packaging (left) and capsule (right) for immune complex, (c) LE projection of immune complex.

The case of the immune complex samples was challenging due to the low contrast of only 0.5 between Z_{eff} of capsule shells and filling. This low contrast leads to a low SNR in the basis material images (Fig. 8(b)). Consequently, there is a large random error in the estimated weight, leading to low values of R^2. The low contrast in Z_{eff} further leads to stronger effects of small errors in Z_{eff}. Thus, the weight estimated by BMD is approximately 240 mg higher than expected, which leads to a large value of the *y*-intercept of 273 mg.

The X-ray projections of immune complex also reveal an inclusion of higher attenuation in one of the capsules (Fig. 8(c)). Whether it is some kind of contamination or due to an uneven mixture of the ingredients, is unknown. In any case, this example shows that X-ray imaging can reveal such imperfections, too.

The estimation of Z_{eff} for L-carnitin was also challenging. It was not possible to find a Z_{eff} that completely removed the filling from the basis material image of the blister package (Fig. 9(a)). The filling taken from the capsules, directly after opening them for the first time, was no fine powder but contained agglomerations. These may be the reason for areas of varying X-ray attenuation visible in the X-ray projections (Fig. 9(b)). In addition, in measurements with a detector (Teledyne Dalsa Shad-o-Scan 8K) enabling a better spatial resolution (but no DE-XRT measurements), unevenly distributed components of higher X-ray attenuation are visible (Fig. 9(c)). These components and agglomerations might impede the finding of a suitable Z_{eff}. Nevertheless, even though

the weight determined by DE-XRT is about 50 mg higher than the expected values, the scattering is low.



(a) (b) (c)

Fig. 9. (a) Basis material images of blister packaging (left) and capsules (right) for L-carnitin showing imperfect material separation, (b) LE projection of L-carnitin showing areas of varying attenuation, (c) high resolution projection of a L-carnitin capsule with pixel size of 27 µm.

3.2. Capsule vs. Filling

The separation of two materials into different basis material images requires some contrast in these materials' effective atomic numbers. Table 4 shows that the highest contrast between capsule shell and filling exists for L-carnitin and immune complex. As explained in the previous section, the estimation of Z_{eff} for L-carnitin was problematic due to the presence of agglomerations in the filling. Thus immune complex was chosen for attempting to estimate the weight of the filling by DE-XRT.

Fig. 10 shows the basis material images for these capsules.

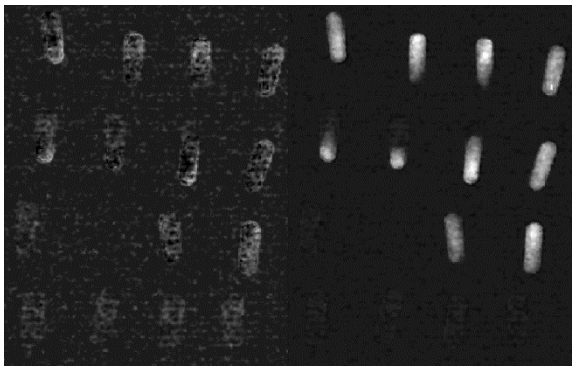


Fig. 10. Areal density image of capsule shells (left) and filling (right) for immune complex capsules without package. Five empty capsule shells are visible on the left image.

The image of the capsule shells contains some traces of the filling. These are not visible in the empty shells towards the bottom of the image. This indicates an inaccurate estimation of Z_{eff} values. Indeed, as with L-carnitin (Fig. 9), it was not possible to determine an effective atomic number that perfectly separates capsule shell and filling into the two basis material images. Thus, the slope and R^2 are higher and lower respectively, than expected for the ideal case (see Table 6).

Table 6. Parameters for linear fits of weight determined by DE-XRT versus weight measured with an analytical balance for filled capsules in their blister package.

Sample	**Slope *m***	***n* [mg]**	***R*²**
Vitamin B complex	1.12	111.03	0.9898
Immune complex	1.08	4.17	0.9663
Ideal fit	1	0	1

As a second example, Fig. 11 shows the basis material images for vitamin B. The difference in the effective atomic numbers between capsule shell and filling is only $\Delta Z_{eff} = 1$. Thus the weight is about 150 mg higher than the expected value, which is more than in the example of the same capsules in their package (Fig. 6 and Fig. 7). Accordingly, the value of *n* deviates considerably from zero (Table 6).

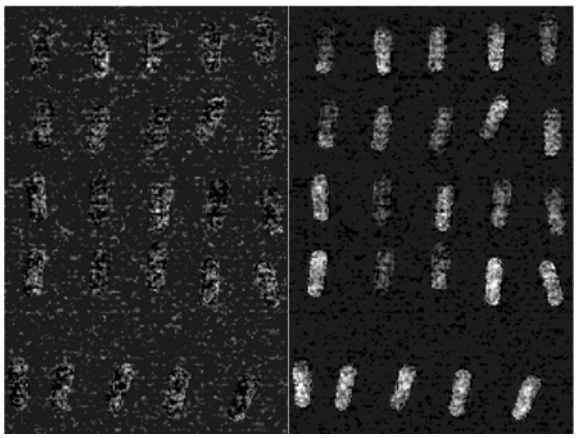


Fig. 11. Areal density image of capsule shells (left) and filling (right) for vitamin B without package.

3.3. Influence of Spatial Resolution on Estimated Weight

The accuracy of the separation between filling and capsule shell is influenced by the spatial resolution of the measurement system. Due to the two consecutive detection layers in the dual energy detector, the HE image has a slightly higher magnification than the LE image. This is compensated by image registration, which requires interpolation between the pixels in at least one of the images. Abrupt changes of the object thickness, like at the transition between filling and shell or between shell and air, rarely coincide with

borders between pixels. Thus, edge pixels are typically incompletely covered by the object. In particular, the coverage will be different for pixels in the LE and the HE detection layers. For these edge pixels, the interpolation required by image registration creates inconsistencies in the gray values of LE and HE projections. These would be further increased by an imperfect registration. The result of the BMD analysis is inaccurate for edge pixels as a consequence and may show artifacts.

The capsules have an approximately cylindrical shape. X-ray imaging projects the attenuation properties of this three dimensional form into two dimensions. Consequently, a considerable portion of the mass of the capsule shell can be found around the edge pixels (Fig. 9(c)). The wall thickness of capsule shells is typically around 100 µm, which is less than the pixel size of 400 µm used for DE-XRT imaging. Thus, the edge pixels of the generated basis material images, which are subject to artifacts, include not only a part of the capsule shell which contains a considerable portion of the weight, but a part of the filling, too.

As a result, the low resolution prevents the determination of the weight of the capsule shells. This would require pixels smaller than the wall thickness. For the filling, the influence is smaller because most of the mass is close to the center and thus not influenced by the artifacts. Still, the smaller the pixels, the smaller is the ratio of edge pixels to non-edge pixels. A better spatial resolution would thus improve the accuracy of the weight determination.

As an example, in Fig. 11 a capsule contains (1350 ± 60) pixel. Roughly 10 % of these pixels (136 ± 4) are considered edge pixels. When creating an image with one half of the pixel size by interpolation, the total number of pixels of a capsule increases by a factor of four to (5410 ± 240) pixels. The number of edge pixels only increases by a factor of approximately two to (324 ± 7) pixels. Unfortunately, a high resolution detector capable of dual energy was not available for this study.

4. Conclusion and Outlook

The presented examples show the potential of DE-XRT: the determination of weight is possible for filled capsules in blister packaging and for the filling of capsules outside of the package. Weighing capsules with different, known amounts of filling as presented in this paper can be used as a calibration. Then, the weight of other capsules or their filling can be determined from the calibration curve. This can be done in real-time, so that DE-XRT can be integrated into a production line.

Although the bought samples were randomly chosen, they cover different properties. They contain powder or granulated material, the fillings consists of one or more detectable components and also agglomerations were visible. The variability of the effective atomic numbers Z_{eff} found for capsule shells and fillings lead to pairs of low and of high Z_{eff} contrast. This spectrum of properties allows to see also the limits of the method.

One general limitation to the accuracy of the method is given by the difference in the effective atomic numbers of blister package, capsule and filling, the smaller these are, the lower is the SNR of the basis material images and the higher is the random error of the weight determined by DE-XRT. A small contrast in the effective atomic numbers also increases the effect of inaccuracies in the values of Z_{eff} on the estimated weight and increases the systematic error.

Another restriction of the presented method is the limitation to two material differentiation. A separation of blister packaging, shell und filling into three basis material images is not possible by this method. This limitation also causes a requirement for the uniformity of the filling. For instance, a separation of the shell of the capsule and two components of the filling into three basis material images cannot be done. This may lead to restrictions for the kinds of capsules that can be investigated. However, as shown in the example of the zinc + c depot, as long as the mixing ratio is constant and suitable values of Z_{eff} can be found for one filling component, a calibration with an analytical balance is possible.

Even for the presented examples that fulfill above requirements, the scattering in plots like Fig. 7 has to be reduced to achieve an accuracy acceptable for industrial application. This requires a higher SNR of the projection data. However, the power of X-ray sources is limited and longer exposure times are contradicted by the need for speed in industrial production. Another option are detectors using time delayed integration (TDI). These increase SNR by averaging the detected signal over several detector lines. This technique also allows to use smaller pixels that would otherwise decrease the SNR, because a smaller area collects less photons. Smaller pixels would also improve the spatial resolution and thus increase the accuracy of the weight determination.

Apart from weighing filled capsules or the filling alone, DE-XRT may also detect quality issues like agglomerations and contaminations. Of cause, it also includes the possibilities provided by standard XRT, like detecting deformations or capsules missing in a blister package.

Acknowledgements

The authors would like to thank Hamamatsu Photonics Deutschland GmbH for providing the measurement setup.

References

[1]. F. Bessler, W. Bauer, X-Ray Net Weight Control of Pharmaceutical Products, in *Proceedings of the 7th World Conference on Nondestructive Testing*, Shanghai, China, 25-28 October 2008.

[2]. M. Bssrmen, F. Pernus and B. Likar, High speed precision weighing of pharmaceutical capsules, in *Proceedings of the IEEE International Conference on Industrial Technology (ICIT)*, Churchill, Victoria, Australia, 10-13 February 2009, pp. 1-6.

[3]. R. E. Alvarez, A. Macovsk, Energy-selective reconstructions in X-ray computerized tomography, *Physics in Medicine and Biology*, Vol. 21, Issue 5, 1976, pp. 733-744.

[4]. M. Firsching, M. Ottenweller, S. Rüger, J. Leisner, Comparison of X-ray and RGB-based Component Detection on Waste PCBs for Recycling, in *Proceedings of the 8th International Conference on Sensors and Electronics Instrumentation Advances (SEIA'2022)*, Corfu, Greece, 21-23 September 2022, pp. 133-134.

[5]. C. Bauer, R. Wagner, B. Orberger, M. Firsching, A. Ennen, C. Garcia Pina, C. Wagner, M. Honarmand, G. Nabatian, I. Monsef. Potential of Dual and Multi Energy XRT and CT Analyses on Iron Formations, *MDPI Sensors*, Vol. 21, Issue 7, 2021, pp. 2455.

[6]. C. Bauer, R. Wagner, J. Leisner, Foreign body detection in frozen food by dual energy x-ray transmission, *Sensors & Transducers*, Vol. 253, Issue 6, 2021, pp. 23-30.

[7]. J. Leisner, C. Bauer, R. Wagner, Circular Economy: Material Sorting in Waste Streams Using Dual Energy X-ray Transmission, in *Proceedings of the 8th International Conference on Sensors and Electronics Instrumentation Advances (SEIA'2022)*, Corfu, Greece, 21-23 September 2022, pp. 93-96.

[8]. M. Firsching, J. Mühlbauer, A. Mäurer, F. Nachtrab, N. Uhlmann, Quantitative sorting using dual energy X-ray transmission imaging, in *Proceedings of the Optical Characterization of Materials Conference (OCM'13)*, Karlsruhe, Germany, 6-7 March 2013, pp. 259-264.

[9]. K. Uehara, T. Tagami, I. Miyazaki, N. Murata, Y. Takahashi, H. Ohkubo, T. Ozeki, Effect of X-ray exposure on the pharmaceutical quality of drug tablets using X-ray inspection equipment, *Drug Development and Industrial Pharmacy*, Vol. 41, Issue 6, 2015, pp. 953–958.

[10]. C. Bauer, R. Wagner, J. Leisner, Determination of the Weight of Pharmaceutical Capsules with Dual Energy X-ray Transmission, in *Proceedings of the 8th International Conference on Sensors and Electronics Instrumentation Advances (SEIA'2022)*, Corfu, Greece, 21-23 September 2022, pp. 97-99.

[11]. R. Wagner, C. Bauer, M. Firsching, J. Leisner, From Food to Mining: Dual Energy XRT Applications, *Advances in Sensors: Reviews*, Vol. 8, 2022, pp. 153-171.

[12]. M. Firsching, F. Nachtrab, N. Uhlmann, R. Hanke, Multi-Energy X-ray Imaging as a Quantitative Method for Materials Characterization, *Advanced Materials*, Vol. 23, Issue 22-23, 2011, pp. 2655–2656.

[13]. E. Y. Sidky, L. Yu, X. Pan, Y. Zou, M. Vannier, A robust method of X-ray source spectrum estimation from transmission measurements: Demonstrated in computer simulated, scatter-free transmission data, *Journal of Applied Physics*, Vol. 97, Issue 12, 2005, pp. 124701.

[14]. Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients from 1 kev to 20 mev for elements z = 1 to 92 and 48 additional substances of dosimetric interest (https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients).

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 61-68

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Detection of Motion Restriction with Smart Insoles

[1, *] Tomoko Funayama, [2] Yasutaka Uchida and [3] Yoshiaki Kogure

[1] Faculty of Medical Sciences, Teikyo University of Science,
Uenohara-shi, Ymanashi, 409-0193, Japan
[2] Faculty of Life & Environmental Sciences, Teikyo University of Science,
Adachi-ku, Tokyo, 120-0045, Japan
[3] Professor Emeritus, Teikyo university of Science, Adachi-ku, 120-0045, Japan
[1] Tel.: + 81554634411, fax: + 81554636944
E-mail: funayama@ntu.ac.jp

Received: 9 September 2022 /Accepted: 11 October 2022 /Published: 31 October 2022

Abstract: In this study, a wireless smart insole was used to measure walking with joint motion restriction. This smart insole outputs data for four parts—toe, heel, inside, and outside—and the color of these parts changes according to the degree of weighting. Motion restrictions involving the ankle joint were performed to assess changes in the physical condition. Normal walking without motion restriction was also assessed. Raw data from the smart insole were graphed and visually predicted. Subsequently, we used support vector machines and K-means methods to detect the classifications. Analysis of the data with and without ankle joint restrictions showed a trend toward higher classification accuracy for those with restrictions. The peak value of each waveform during the gait cycle, rate of decrease in the value after each peak, and data inside the insole were identified as potential detection possibilities. The use of smart insoles may facilitate the determination of changes in physical conditions. This will lead to an assessment of the physical condition based on objective data.

Keywords: Smart insole, Physical condition assessment, Wearable devices, Walking monitoring, Wireless data.

1. Introduction

Assessing health conditions in daily life is beneficial for older people and those with health problems, and wearable devices are becoming increasingly popular as a form of healthcare. The feet are one of the most sensory-sensitive parts of the body, and as they are the only body part in contact with the floor when standing, they are responsible for correcting posture. Owing to their distance from the heart, blood circulation tends to be impaired during sitting and standing positions. The sole is the body part through which health conditions are likely to be indicated.

The indications of insoles for orthopedic diseases, including osteoarthritis, diabetes, and other ailments related to old age, are being investigated in the rehabilitation and health promotion fields [1–5]. The world is becoming an ageing society. As the population ages, the number of people facing health challenges will increase [6]. Therefore, having a healthy body and being active is important as long as possible. For this purpose, the assessment of activities will be useful. Research and development of smart insoles that can acquire digital data to assess the effects of aging and disease and for purposes such as fall prevention has also progressed in recent years [7–15]. The sensors used in these devices include acceleration, temperature, and pressure sensors [16–[24]. Considerable amount of data can be acquired from various sensors. Therefore, focusing on which data are valid and how they are needed is important to

determine the health conditions of an individual [25]. We have previously conducted a study using smart insoles with occupational and physical therapists [26]. Occupational and physical therapists reviewed the response to smart insoles and studied their motor conditions, making observational judgments regarding the sensor responses of smart insoles during simulated motion restrictions. The response of the pressure sensor on the smart insole serves as objective and important data for rehabilitation. Herein, we examined the pressure sensor response of a wireless smart insole when walking with restricted ankle joints by simulating restricted motion of the upper and lower limbs to assess the changes in physical condition. The subjects' ankle joint range of motion was measured beforehand [27]. This study was approved by the Ethics Committee on Research of the Teikyo University of Science, with humans as participants.

2. Experimental Method

2.1. Device and Measurement System

The smart insoles used were a wireless type (FEELSOLE®) that measures four parts per foot and eight parts in total on both sides. Before the insoles can be used, they must be calibrated. Calibration was performed four times: no pressure with no feet in the shoes, standing on both feet, and standing on one foot on each side. A 10-s operation was possible. The colors of the four parts (toe, heel, inside, and outside) changed according to the weight applied. Therefore, the user can visually determine the differences in the application of weights to the four parts. The data were saved on an iPad Air (Apple) and could be viewed on the screen. They were transferred from the tablet to the PC via email and made available for analysis. The sampling frequency was 50 Hz, and the data were output in the CSV format.

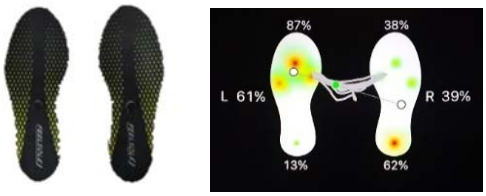


Fig. 1. Exterior of FEELSOLE and tablet screen.

2.2. Measurement Method

In this study, to assess changes in physical condition, simulated walking with restricted upper and lower limb motion was performed, and the pressure sensor response of the smart insoles was examined with and without restricted ankle joint motion focused. Ankle joint motion was simulated using a supporter. The responses of the sensors to the four regions of one leg and eight regions of both feet were examined. The ankle joint range of motion was measured in advance. The range of motion (ROM) of the ankle joint is shown in Table 1, where R signifies the right side, and L represents the left side. There were two subjects: a female in her 50s (Case A) and a male in his 70s (Case B).

Table 1. Ankle joint range of motion.

		Case A	Case B
Plantar flexion	**R**	50	45
	L	55	50
Dorsi flexion	**R**	20	20
	L	20	-5

2.3. Analysis Method

The characteristics were visually predicted by graphing the raw data from the smart insoles. Subsequently, with respect to the features, we studied how they could be classified as teaching data for machine learning, support vector machines (SVM), and we used the k-means method to check the degree of classification. SVM was performed with the Python library scikit-learn.

3. Results

3.1. Insole Data in Normal Condition

Figs. 2 and 3 show the data at the four insole regions when walking normally without restrictions in Case A. The four regions were the heel, toe, and inside and outside of the insole. The horizontal axis represents the data-measurement period. With respect to time, the number 500 on the graph represents 10 s. This corresponds to everything after Fig. 2. Figs. 2 and 3 show the right and left insoles, respectively.

The data values corresponding to the inside region have a left–right difference, with the right inside being larger than the left inside.

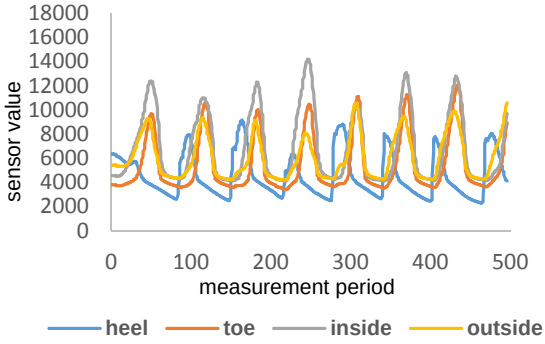


Fig. 2. Right insole with normal walking in Case A.

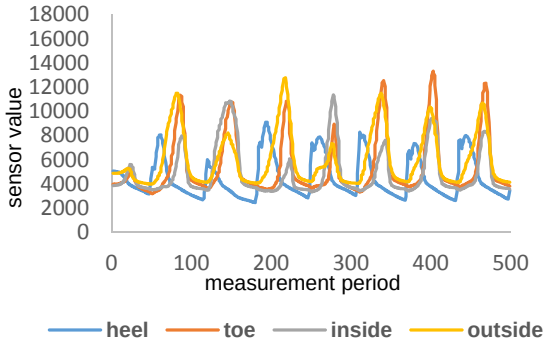


Fig. 3. Left insole with normal walking in Case A.

Figs. 4 and 5 show the data at the four insole regions when walking normally without restrictions in Case B. Figs. 4 and 5 show the right and left insoles, respectively.

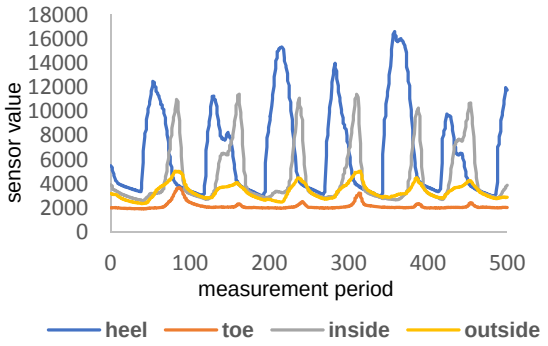


Fig. 4. Right insole with normal walking in Case B.

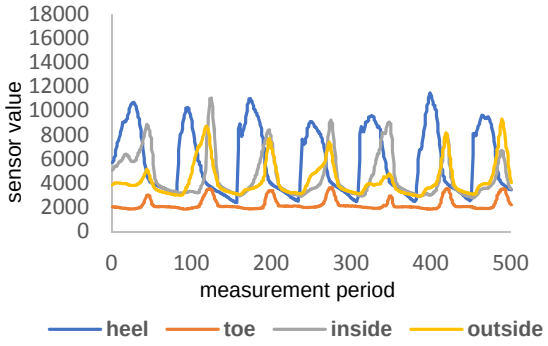


Fig. 5. Left insole with normal walking in Case B.

There were left–right differences in the data for all four regions, with particularly large left–right differences in the heel and outside region.

Even without the simulated motion restriction, there were differences from left to right and from condition to condition, even in the same case. For example, Case B showed larger heel data and a left–right difference, whereas in Case A, the heel data was small. The range of motion of the ankle joints is listed in Table 1. Case B has a larger left–right difference than Case A.

3.2. Insole Data in Ankle Joint Restriction

Figs. 6 and 7 show data from the right insole when the right ankle joint was restricted. Fig. 6 shows Case A, and Fig. 7 shows Case B.

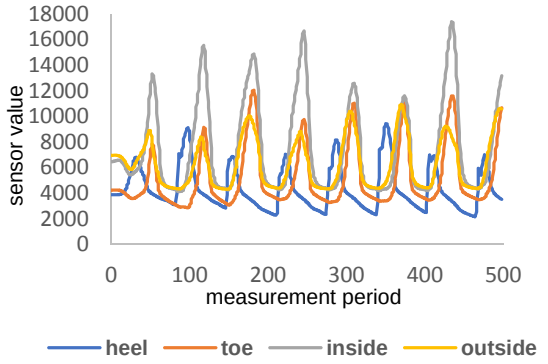


Fig. 6. Right insole with right ankle restriction in Case A.

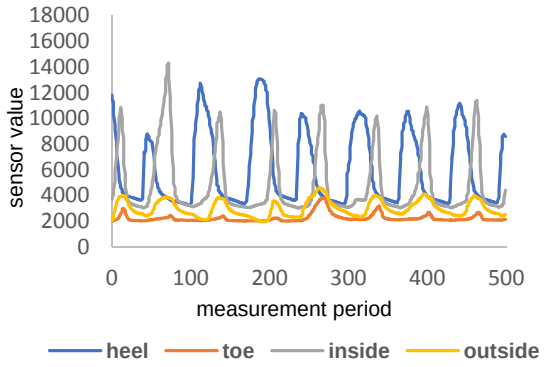


Fig. 7. Right insole with right ankle restriction in Case B.

Both cases changed from normal walking when the ankle joint was restricted. The differences with and without restrictions were greater at the heel in Case B and at the inside in Case A. Even for the inside in Case B, differences with and without restrictions were showed. In Case B, the waveform peak values for the outside and toe are lower and those for the heel and inside are higher, with and without restrictions. Decrease in value after waveform peak differed depending on ankle joint restrictions and insole regions.

3.3. Peak Value and Rate of Decrease

The peak value and rate of decrease of After the peak value in each waveform were analyzed. The rate of decrease was calculated using the least-squares method (LSM). The mean was calculated for each

region of the right insole with respect to the percentage decrease after the waveform peak with and without right ankle joint restriction, as calculated using LSM. The rate of decrease corresponds to the gradient of the waveform. The absolute values of the mean for each region in Cases A and B are shown in Figs. 8 and 9, respectively. The vertical axis of the graph is the average of the Smart Insole data for each region. Walking without joint restrictions was considered normal.

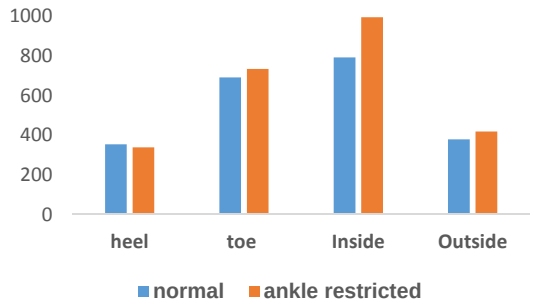


Fig. 8. Mean rate of decrease after waveform peak in Case A with and without restrictions.

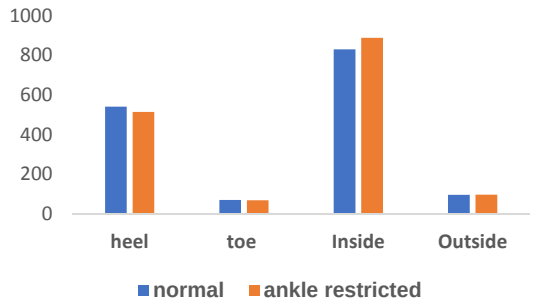


Fig. 9. Mean rate of decrease after waveform peak in Case B with and without restrictions

The data corresponding to the inside region were higher when there were restrictions in both cases, namely Cases A and B.

3.4. SVM Analysis Using Peak Values and Decreasing Rates

The possibility of classification by SVM was examined with respect to the peak and decrease rates of the right insole data with and without right ankle joint restrictions. Since this study was in its initial stages, we primarily focused on how the measured data can be classified. We will add new tests to confirm the accuracy of the classification in the future, however, in this study, we confirmed the possibility that SVM can be used to classify the data as supervised data. We conducted comparisons using LSM and Subtraction of two weighted average (SWA) in SVM. The rate of decrease in the waveform with SWA was calculated using a subtraction of the values before and after successive weighted averages to eliminate and smooth out abrupt changes. Among the subtracted values, the value with the largest decrease in value before and after, i.e., the rate of decrease, was extracted. The weighted average was obtained using five values (target value, two each before and after) calculated as follows. Owing to the large influence of the central region, the percentages were calculated by assigning 50% to the target value of the waveform, 20% to the values before and after one peak, and 10 % to the values before and after two peaks. SWA was calculated using the peak value and rate of decrease for each waveform. Since one dataset was used per wave, there were 6–7 data sets for each in the SVM. The SVM parameters were C = 10.0, the kernel was rbf, and gamma was 1×10^{-6}. This value was selected to the extent that it does not result in overtraining, although not fully optimized. All the subsequent SVMs used the same parameters.

A. Detection using inside and outside data

SVM analysis was performed using the peak value and rate of decrease data for each waveform as learning data and four types of teaching data: with and without right ankle joint restrictions, inside and outside of the right insole, and combinations of them. The results of SVM in Case A are shown in Figs. 10 and 11. Figs. 10 and 11 show LSM and SWA, respectively. The dots in the figure are black on the inside and red on the outside of the insole when the insole is normal without ankle joint restrictions and blue on the inside and white on the outside when the insole has ankle joint restrictions.

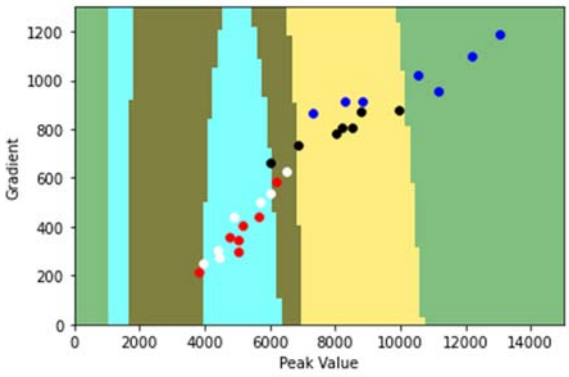


Fig. 10. SVM using LSM: inside and outside in Case A with and without restrictions.

In Case A, 60.7 % accuracy was obtained for LSM, as shown in Fig. 10, and 63.0 % for SWA, as shown in Fig. 11. Case B was also implemented using the same method as the SVM for Case A. Figs. 12 and 13 show LSM and SWA in Case B.

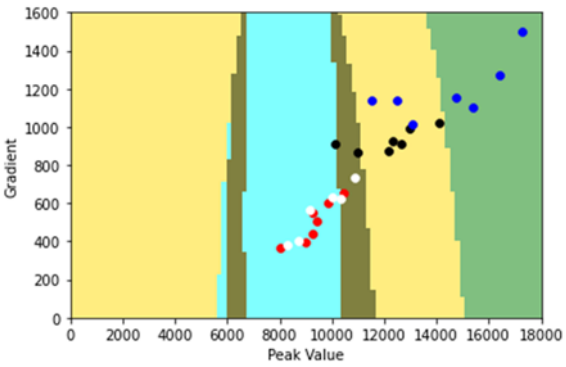


Fig. 11. SVM using SWA: inside and outside in Case A with and without restrictions.

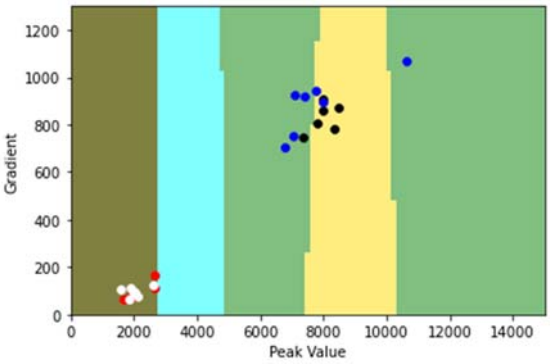


Fig. 12. SVM using LSM: inside and outside in Case B with and without restrictions.

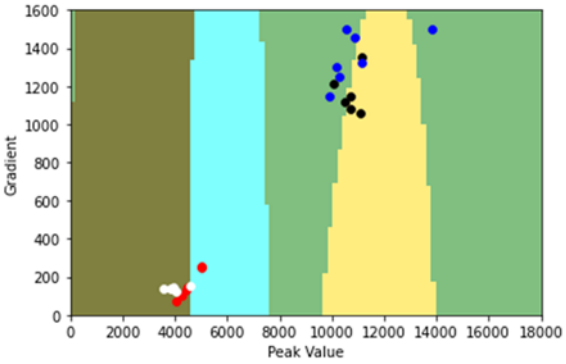


Fig. 13. SVM using SWA: inside and outside in Case B with and without restrictions.

In Case B, 65.4 % accuracy was obtained for LSM, as shown in Fig. 12, and 80.8 % for SWA, as shown in Fig. 13.

The accuracy was 60.7 % for LSM and 63.0 % for SWA in Case A and 65.4 % for LSM and 80.8 % for SWA in Case B. The results tended to be slightly more accurate for SWA than for LSM. The color of the dots indicates that more differences appear inside than outside when the ankle joint is restricted. This is shown in both Cases A and B.

B. Detection using heel and toe data

The heel and toe, which are considered points of attention in gait analysis, were analyzed. The peak value and decrease rate of the waveform were set as learning data, and SVM was performed with four types of data for teaching: distinction between the heel and toe of the right insole and the presence or absence of right ankle joint restrictions. Figs. 14 and 15 present Cases A and B. Results for the heel and toe are shown only for waveform peaks and the rate of decrease as calculated by SWA. The dots in the figure are black on the heel and red on the toe of the insole when the insole is normal without ankle joint restrictions and blue on the heel and white on the toe when the insole has ankle joint restrictions.

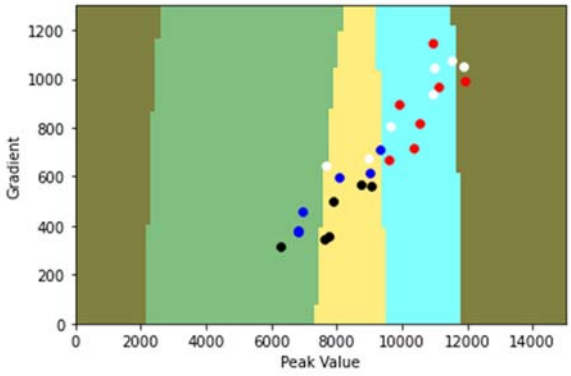


Fig. 14. SVM using SWA: heel and toe in Case A with and without restrictions.

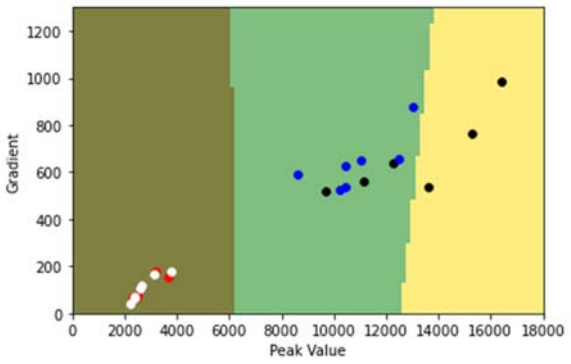


Fig. 15. SVM using SWA: heel and toe in Case B with and without restrictions.

Accuracy was 57.7 % in Case A and 65.4 % in Case B. These results were obtained using pairs of heel and toe data. While, regarding inside and outside pairs, they were 63.0% and 80.8%, respectively. Accuracy was higher for the inside and outside pairs than for the heel and toe pairs. Furthermore, the color of the dots in the figure indicates that although there is a difference between the heel and toe, detecting a difference owing to the presence or absence of restrictions is difficult.

C. Left and right insole data without restrictions

SVM analysis was performed on four types of insoles: left and right and inside and outside, when walking normally without joint restrictions. The waveform peaks and decrease rates calculated using the SWA are shown below. The results for Cases A and B are presented in Figs. 16 and 17, respectively. The dots in the figure are black for the right inside, red for the right outside, blue for the left inside, and white for the left outside.

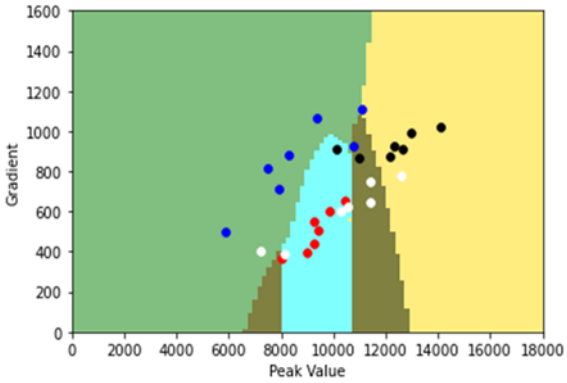


Fig. 16. SVM using SWA: inside and outside in Case A without restriction.

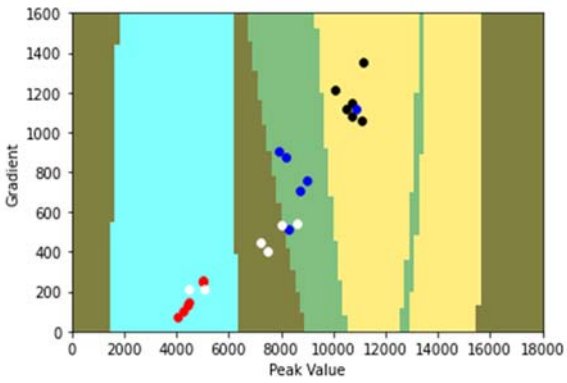


Fig. 17. SVM using SWA: inside and outside in Case B without restriction.

The accuracies were 67.9% for Case A and 79.2% for Case B. The colors of the dots in the figure also indicate that both Cases A and B tend to be classified into four categories.

SVM analysis was performed on four types of insoles—left and right, heel, and toe—when walking normally without joint restrictions. The waveform peaks and decrease rates calculated using the SWA are shown below. The results for Cases A and B are presented in Figs. 18 and 19, respectively. The dots in the figure are black, red, blue, and white for the right heel, right toe, left heel, and left toe, respectively.

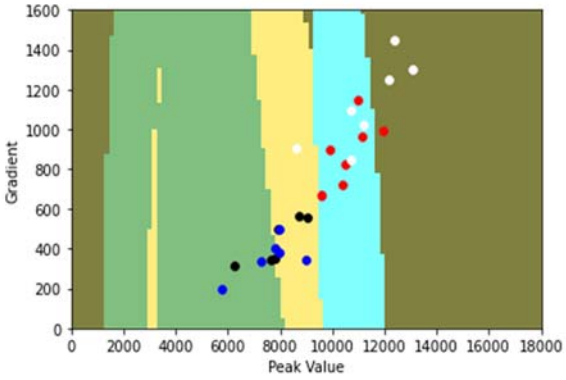


Fig. 18. SVM using SWA: heel and toe in Case A without restriction.

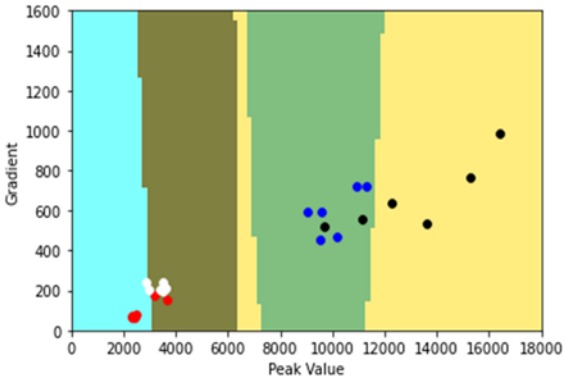


Fig. 19. SVM using SWA: heel and toe in Case B without restriction.

The accuracies were 57.7 % for Case A and 79.2 % for Case B. According to the colored dots in the figure, the trend is different between Cases A and B. Case A has higher values for the toes and Case B has higher values for the heels. Accuracy for the inside and outside of the insole was 67.9 % and 79.2 % for Cases A and B, respectively, which is equal to or better than that for the heel and toe.

3.5. Detection using K-means Clustering Method

The K-means clustering method was performed by inputting data for four patterns of combinations of inside and outside insoles, with and without ankle joint restrictions. The data were parsed using weighted calculated peaks and percentage reductions. This is unsupervised learning, which does not use information data on the insole parts and restrictions. Cases A and B were classified into four categories. However, for Case B, the centers of gravity of the two clusters are close. Figs. 20 and 21 show the results of the K-means method using SWA data for Cases A and B.

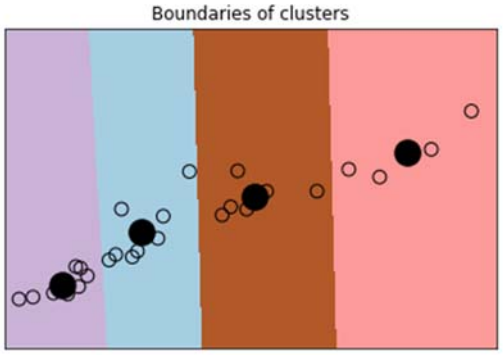


Fig. 20. K-means using SWA: inside and outside in Case A with and without restrictions

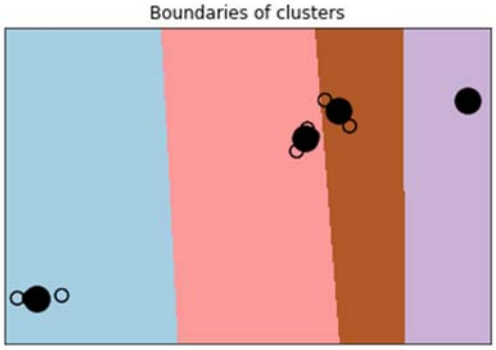


Fig. 21. K-means using SWA: inside and outside in Case B with and without restrictions

4. Discussion

The response to the smart insoles varied from left to right and condition to condition, even in the same case. However, the SVM classification, which simulates ankle joint restrictions, was more detectable for the inside and outside insole pairs than for the heel and toe pairs. This tendency was similar when discriminating between the left and right sides during normal walking, without restrictions. In addition, Case B, wherein the left–right difference in ankle joint range of motion was greater than that in Case A, had a higher accuracy by SVM analysis. With and without simulated ankle joint restriction, the rate of decrease inside the insole was higher for those with restriction from the colored points in the figure, that is, the inside of the unrestricted foot tended to change more slowly. However, whether these results apply to normal walking without restrictions remains to be investigated. The possibility that supporters or other factors may influence responses to smart insoles should also be considered.

According to studies already performed by occupational and physical therapists using the same smart insoles, they agree that the toes are considered important in terms of predicting motion restrictions and useful sensor parts. However, other aspects differed depending on the therapist. Smart insoles can be objectively applied and can complement therapists' assessments. In this study, we investigated the peak of each waveform in the insole data and the rate of decrease after the peak. The SVM analysis figure shows that the difference is particularly noticeable inside the insole. The inside and outside of the insoles are considered to be more difficult to assess by therapist observation than the heel and toes. In particular, the soles of the feet can be used not only to detect the subject's gait pattern but also to assess fatigue and other health conditions. Falls among older adults can also cause serious injuries such as broken bones and head trauma and may result in the need for nursing care. The use of smart insoles potentially makes the assessment of changes in physical conditions easier than ever before.

Therefore, we suggest that when assessing walking with smart insoles, three points should be considered in addition to examining the peak value and the rate of decrease after the peak: (1) the differences from the norm based on the characteristics of each individual case, (2) the response of the pressure sensor of the smart insole to common physical characteristics such as ankle joint range of motion and gait, and (3) the relation between the sensor and the sole of the foot.

5. Conclusions

In this study, we investigated two subjects with a walking time of 10 s using a small amount of data. We did not use machine-learning test data for the SVM. Therefore, the results of this study cannot be generalized. However, we did find the possibility that the data of the inside waveform peak and the rate of decrease can be useful. We plan to increase the number of subjects and extend the measurement time of the insoles.

Acknowledgements

This work was supported by JSPS KANENHI Grant Number JP20K11924.

References

[1]. D. Lin, E. Papi, A. H. McGregor, Exploring the clinical context of adopting an instrumented insole: a qualitative study of clinicians' preferences in England, *BMJ Open*, Vol. 9, Issue 4, 2019, e023656.

[2]. B. C. Chang, J. Y. Wang, B. S. Huang, H. Y. Lin, W. C. C. Lee, Dynamic impression insole in rheumatoid foot with metatarsal pain, *Clinical Biomechanics*, Vol. 28, Issue 2, 2012, pp. 196-201.

[3]. Lorenzo Brognara, Emmanuel Navarro-Flores, Lorenzo Iachemet, Nuria Serra-Catalá, Omar Cauli, Beneficial Effect of Foot Plantar Stimulation in Gait Parameters in Individuals with Parkinson's Disease, *Brain Sciences*, Vol. 10, Issue 2, 2020, 69.

[4]. T. W. Seo, J. Y. Lee, B. H. Lee, The reliability test of a smart insole for gait analysis in stroke patients, *Korean Physical Therapy Science*, Vol. 29, No.1, 2021, pp. 30-40.

[5]. K. J. Kelleher, W. D. Spence, S. Solomonidis, D. Apatsidis, The effect of textured insoles on gait patterns of people with multiple sclerosis, *Gait & Posture*, Vol. 32, Issue 1, 2010, pp. 67-71.

[6]. World Health Organization 2022, World health statistics 2022 monitoring health for the SDGs, *Sustainable Development Goals*, 2022.

[7]. S. Saidani, R. Haddad, R. Bouallegue, R. Shubair, A New Proposal of a Smart Insole for the Monitoring of Elderly Patients, in *Proceedings of the 35th International Conference on Advanced Information Networking and Applications*, Toronto, Canada, 12 - 14 May 2021, Vol. 2, pp. 273-284.

[8]. B. Najafi, E. Ron, A. Enriquez, I. Marin, J. Razjouyan, D. G. Armstrong, Smarter Sole Survival: Will Neuropathic Patients at High Risk for Ulceration Use a Smart Insole-Based Foot Protection System?, *Diabetes Science and Technology*, Vol. 11, Issue 4, 2017, pp. 702-713.

[9]. E. M. Macdonald, B. M. Perrin, L. Cleel, M. I. C. Kingsley, Podiatrist-Delivered Health Coaching to Facilitate the Use of a Smart Insole to Support Foot Health Monitoring in People with Diabetes-Related Peripheral Neuropathy, *Sensors*, Vol. 21, Issue 12, 2021, 3984.

[10]. H. Nagano, R. K. Begg, Shoe-Insole Technology for Injury Prevention in Walking, *Sensors*, Vol. 18, Issue 5, 2018, 1468.

[11]. V. Bucinskas, A. Dzedzickis, J. Rozene, J. S. Zemaitiene, I. Satkauskas, V. Uvarovas, R. Bobina, I. Vilkonciene, Wearable Feet Pressure Sensor for Human Gait and Falling Diagnosis, *Sensors*, Vol. 21, Issue 15, 2022, 5240.

[12]. M. Munoz-Organero, J. Parker, L. Powell, S. Mawson, Assessing Walking Strategies Using Insole Pressure Sensors for Stroke Survivors, *Sensors*, Vol. 16, Issue 10, 2016, 1631.

[13]. J. K. Kim, M. Nam, K. B. Lee, S. G. Hong, Gait event detection algorithm based on smart insoles, *Electronics and Telecommunications Research Institute (ETRI)*, Vol. 42, Issue 1, 2020, pp. 46–53.

[14]. D. D. L. Rodriguez, J. Assal, Biofeedback can reduce foot pressure to a safe level and without causing new at-risk zones in patients with diabetes and peripheral neuropathy, *Diabetes/Metabolism Research and Reviews*, Vol. 29, Issue 2, 2013, pp. 139–144.

[15]. E. M. Macdonald, B. M. Perrin, N. Hyett, M. I. C. Kingsley, Factors influencing behavioural intention to use a smart shoe insole in regionally based adults with diabetes: a mixed methods study, *Foot and Ankle Research*, Vol. 12, 2019, 29.

[16]. N. Hegde, E. Sazonov, SmartStep: A Fully Integrated, Low-Power Insole Monitor, *Electronics*, Vol. 3, Issue 2, 2014, pp. 381-397.

[17]. S. Subramaniam, S. Majumder, A. I. Faisal, M. J. Deen, Insole-Based Systems for Health Monitoring: Current Solutions and Research Challenges, *Sensors*, Vol. 22, Issue 2, 2022, 438.

[18]. S. Saidani, R Haddad, N. Mezghani, S. Saidani, R. Haddad, N. Mezghani, R. Bouallegue, A survey on smart shoe insole systems, in *Proceedings of the International Conference on Smart Communications and Networking (SmartNets)*, Hammamet, Tunisia, 16-17 November 2018, pp. 1-6.

[19]. A. K. S. Mahmud, M. E. H. Chowdhury, M. B. Reaz, S. Kiranyaz, Z. B. Mahbub, S. H. Md Ali, A. A. A. Bakar, Design and Implementation of a Smart Insole System to Measure Plantar Pressure and Temperature, *Sensors*, Vol. 22, Issue 19, 2022, 7599.

[20]. S. S. Lee, S. T. Choi, S. I. Choi, Classification of Gait Type Based on Deep Learning Using Various Sensors with Smart Insole, *Sensors*, Vol. 19, Issue 8, 2019, 1757.

[21]. A M. Tana, F. K. Fussa, Y. Weizmana, Y. Woudstra, O. Troynikovb, Design of Low Cost Smart Insole for Real Time Measurement of Plantar Pressure, *Procedia Technology*, Vol. 20, Issue , 2015, pp. 117–122.

[22]. T. E. Roden, R. L. Grand, R. Fernandez, J. Brown, J. (Ed) Deaton, J. Ross, Development of a smart insole tracking system for physical therapy and athletics, in *Proceedings of the 7th International Conference on Pervasive Technologies Related to Assistive Environments*, New York, United States, 27 - 30 May 2014, No. 40.

[23]. A. M. Cristiani, G. M. Bertolotti, E. Marenzi, S. Ramat, An Instrumented Insole for Long Term Monitoring Movement, Comfort, and Ergonomics, *IEEE Sensors*, Vol. 14, Issue 5, 2014, pp. 1564-1572.

[24]. F. Lin, A. Wang, C. Song, W. Xu, Z. Li, Qin Li, A comparative study of smart insole on real-world step count, in *Proceedings of the IEEE Signal Processing in Medicine and Biology Symposium (SPMB' 2015)*, Pennsylvania, USA, 12 December 2015, 15789440.

[25]. S. Yoo, H. Gil, J. Kim, J. Ryu, S. Yoon, and S. K. Park, The Optimization of the Number and Positions of Foot Pressure Sensors to Develop Smart Shoes, *Ergonomics Society of Korea*, Vol. 36, Issue 5, 2017, pp. 395-409.

[26]. Y. Uchida, T. Funayama, Y. Kogure, R. Kimura, D. Souma, A study on the number of pressure sensors in smart insoles and the detection of body condition changes, *Proceedings of the Human Interface Society of Japan*, Kyoto, Japan, Vol. 24, No. 1, 2022, pp. 85-86.

[27]. T. Funayama, Y. Uchida, Y. Kogure, Assessment of Motion Restrictions Using Smart Insoles, in *Proceedings of 8th International Conference on Sensors and Electronic Instrumentation Advances (SEIA' 2022)*, Corfu, Greece, 21-23 September, 2022, pp. 129-132.

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 69-74

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Effect of a Rauch-Tung-Striebel Algorithm on Different Global Navigation Satellite System Time Transfer

[1,2] Mengshi CHEN, [1,2,3,*] Haibo YUAN, [1,3] JiHai ZHANG, [1,3] Hong ZHANG, [1,2] Zongyuan LI and [1,2] Yiheng WANG

[1] National Time Service Center, Chinese Academy of Sciences, Xi'an 710600, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China,
[3] Key Laboratory of Time and Frequency Primary Standards, Chinese Academy of Sciences, Xi'an 710600, China
[1] Tel.: + 8618220533997
E-mail: chenmengshi@ntsc.ac.cn, yuanhb@ntsc.ac.cn

Received: 5 September 2022 /Accepted: 5 October 2022 /Published: 31 October 2022

Abstract: Global Navigation Satellite System (GNSS) Precise point positioning (GNSS PPP) is widely applied in high precision international time transfer. With the development of GNSS precision orbit and clock products, GNSS time transfer technology has also attracted the attention of researchers. Traditional PPP has a long convergence time. In this paper, the improved algorithm based on the RTS method is used to calculate the time transfer results of different navigation systems to evaluate the improvement of this method on the time transfer results of different systems. The results show that the time transfer link based on Global Positioning System (GPS) and BeiDou Navigation Satellite System (BDS) can both achieve sub-nanosecond precision. The frequency stability of the time transfer results can reach 2E-14 with the average time of 1E6 seconds. The improved method can effectively reduce the influence of the convergence of the forward filter on the time transfer results.

Keywords: GNSS PPP, Kalman filter, Time transfer, Frequency and time, Time stability.

1. Introduction

With the development of GNSS, satellite navigation technology has been widely used in many fields of our life. In recent decades, many high-precision time transfer technologies have also been achieved through satellites. GNSS PPP time comparison technology is one of important technology for BIPM to conduct international time transfer [1]. GNSS PPP time transfer technology has been widely applied for low cost, high precision, and easy maintenance [2]. Though the traditional static PPP has high accuracy, it`s convergence time is long [3, 4]. It takes approximately 20 min required for 95 % of solutions to reach a horizontal accuracy of 20 cm or better for static PPP [5, 6]. PPP time transfer has the same result. For different satellite navigation systems, the number of satellites, satellite elevation angles, and signal quality observed by the user at the same time are different. These factors can also cause PPP to have different convergence times.

Traditional static PPP only uses forward filer. The first series of positioning and time transfer data will have poor accuracy. For post-processing calculation, the fixed interval smoother can improve the accuracy. The fixed interval smoother combines forward and backward filter result by the noise matrix. This method can reduce the influence of the one-way filter convergence problem on the calculation of the initial part of the data. But introducing the inverse filtering result directly introduces additional noise, especially in the last part of the result. In order to solve the above

problems, based on previous work a post-processing RTS method is investigated in this paper [7].

2. Method

The traditional static PPP model has multiple model corrections, and it can be written as:

$$\begin{aligned} P_{r,j}^S &= \rho_r^S + c\cdot(dt_r - dt_{s,j}) + T_{r,j}^S + I_{r,j}^S + \\ &\quad + (d_{r,j} - d_j^S) + \varsigma_j + \varepsilon_{r,j}^S(P_{r,j}^S) \\ L_{r,j}^S &= \rho_r^S + c\cdot(dt_r - dt_{s,j}) + T_{r,j}^S - I_{r,j}^S + \\ &\quad + \lambda_j^S(w_{r,j}^S + N_{r,j}^S) + \varsigma_j + \xi_{r,j}^S(L_{r,j}^S) \end{aligned} \quad (1)$$

where $P_{r,j}^S$ and $L_{r,j}^S$ are the code and carrier phase observations, ρ_r^S is the geometric distance, c is the speed of light in vacuum, dt_r and $dt_{s,j}$ are the receiver and satellite clock offset, $T_{r,j}^S$ is the tropospheric delay, $I_{r,j}^S$ is the ionospheric delay, $(d_{r,j} - d_j^S)$ is the hardware latency, λ_j^S is the wavelength of carry phase, $(w_{r,j}^S + N_{r,j}^S)$ is the ambiguity parameter, ς_j is the correction not listed, $\varepsilon_{r,j}^S(P_{r,j}^S)$ and $\xi_{r,j}^S(L_{r,j}^S)$ are the observation noise.

GNSS PPP uses a Kalman filter for parameter estimation. The estimation parameters include antenna coordinates, receiver clock, tropospheric parameters and ambiguity. The Kalman filter model is shown in Equation 2:

$$\begin{aligned} K &= Q*H*(H'*Q*H+R)^{-1} \\ Xp &= X + K*l \\ Qp &= (I - K*H')*Q \end{aligned} \quad (2)$$

where K is the Kalman gain, X and Xp are the state vector, Q and Qp are the covariance matrix of states, l is the innovation (measurement - model), R is the covariance matrix of measurement error, H is the transpose of the design matrix, I is an identity matrix.

For forward filtering, the state parameters and the covariance matrix of the current epoch are calculated from the state parameters and the covariance matrix of the previous epoch. For post processed calculations, the fixed interval smoothed filter is used to combine the result of forward and backward filtering. For a calculation process with a total of N epoch, for epoch k, the state parameters and covariance matrix of epoch k can be calculated from epoch 1 to epoch k by forward filtering, and from epoch n to epoch k by backward filtering. The fixed interval smoothed filter combines the forward and backward results in a smoother, then get the final result. In this way, the results of the unconverged part of the forward filter are improved. However, this calculation introduces additional noise at the end of the data. Thus, a new method based on Rauch-Tung-Striebel (RTS) algorithm is applied in this paper.

First, for a total of N epoch, a series of data from the epoch 1 is calculated by backward filtering. For post processed calculation, sufficient data are used to calculate the covariance matrix and state parameters for epoch 1 with high precision. The station geodetic coordinates, clock offsets, troposphere, ambiguity and corresponding covariance and noise matrices are used as initial information for forward filtering [8]. The result of this forward filtering from epoch 1 to epoch N is the final result. In this way, the influence of convergence at the beginning and the end of the calculation results can be reduced.

3. Calculation and Analysis

The PPP time comparison data UTC(TP)-UTC(ORB) of ORB and TP stations from MJD 59580 to 59587 are shown in follow section. The precise orbit and clock files are from GeoForschungsZentrum (GFZ). The calculation results of GPS and BDS are shown in following part. The time comparison results between the two stations published by the International Bureau of Weights and Measures (BIPM) are used as reference.

3.1. Time Transfer Results

Time transfer results using the above seven-day data processing are presented. The three methods mentioned above are all used for the calculation, and the time transfer results between the two stations through different two satellite navigation systems are different.

As can be seen from the Fig. 1, benefit from the high accuracy of the calculation results of the traditional static PPP, the results calculated by the above three methods are very close. Only the results of the forward filtering have an obviously difference from the reference value at the beginning of each day.

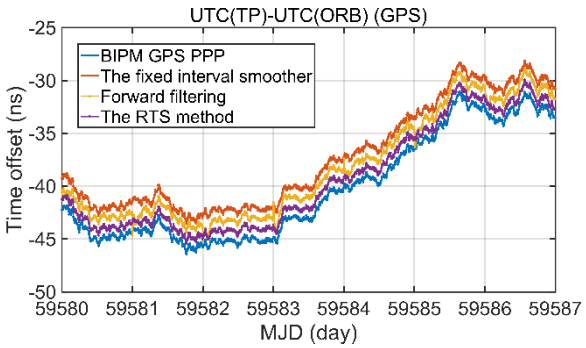


Fig. 1. The GPS PPP time transfer result of UTC(TP)-UTC(ORB) calculated by the above three methods and the results published by the BIPM. The observation data are from TP and ORB laboratories during MJD59580 to MJD59587.

From Fig. 2, it can be seen that the calculation results of Beidou are similar to GPS. The results calculated by the RTS method and the fixed interval smoother are very close to the reference values, and the results calculated by the forward filtering are slightly fluctuating.

In order to perform a detailed comparison of the impact of various methods on the results of different systems, the calculation results are shown as the difference between the clock offset between the two stations calculated by the above three methods and the reference.

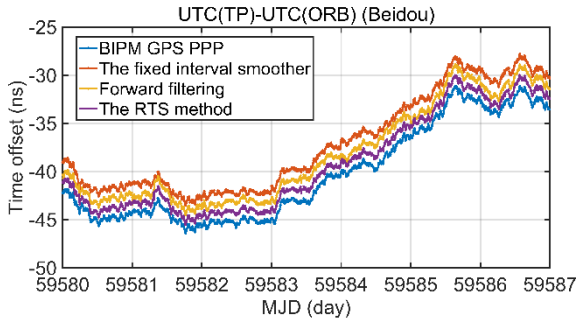


Fig. 2. The BDS PPP time transfer result of UTC(TP)-UTC(ORB) calculated by the above three methods and the results published by the BIPM. The observation data are from TP and ORB laboratories during MJD59580 to MJD59587.

3.2. Analysis of Time Transfer Results

The differences of the results of the different methods are shown below. The difference between the forward filtering method and the RTS method are shown in Fig. 3 and Fig. 4.

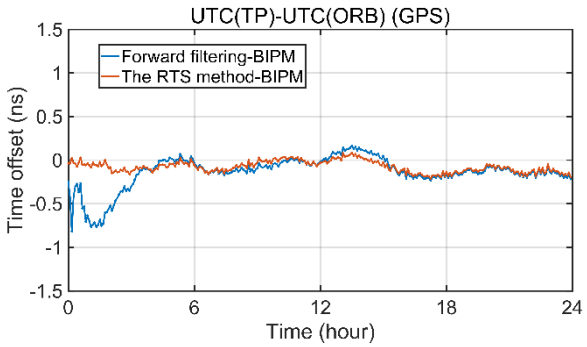


Fig. 3. Difference between time comparison results calculated by RTS algorithm and Forward filter GPS PPP. The observation data are from ORB and TP station, during MJD 59580.

As can be seen from Fig. 3, the difference between the GPS PPP results calculated by the two methods and the reference value is on the order of nanoseconds. When compared with the RTS method, the result with forward filter has an obvious convergence process. In the first 6 hours of each day's data, the result of forward filter has large fluctuations. The root mean square (RMS) of the calculated results for the seven days data are used to quantitatively compare the two methods. The RMS of the results using the forward filtering for seven consecutive days is 319 ps, and the RMS of the first 6 hours of each day is 427 ps. The RMS of the results using the RTS method for seven consecutive days is 115 ps, and the RMS of the first 6 hours of each day is 88 ps, much better than the forward filter.

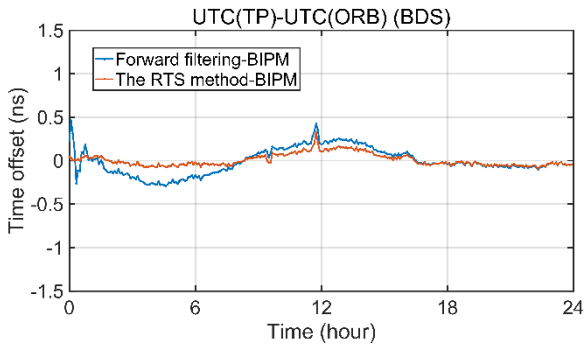


Fig. 4. Difference between time comparison results calculated by RTS algorithm and Forward filter BDS PPP. The observation data are from ORB and TP station, during MJD 59582.

The BDS data shows the same characteristics to GPS data. In the first 12 hours, there is a difference of several hundred picoseconds between the forward filter and the RTS method. But in the last 12 hours the two are very close. The RMS of the results using the forward filtering for seven consecutive days is 351 ps, and the RMS of the first 6 hours of each day is 471 ps. The RMS of the results using the RTS method for seven consecutive days is 125 ps, and the RMS of the first 6 hours of each day is 98 ps.

It can be seen that the RTS method improves the calculation results of the forward filtering of BDS and GPS, and it is more significant in the initial part of the calculation, which reduces the offset between the calculation results and the standard value. At the same time, it is similar to the result of forward filtering after convergence, and the characteristics of forward filtering are maintained. In contrast, GPS converges faster than BDS. After the forward filter has converged, and when using the RTS method, the time transfer accuracy of the two is comparable.

The difference between the fixed interval smoother method and the RTS method are shown in Fig. 5 and Fig. 6.

Fig. 5 shows that the results calculated with these two methods by GPS PPP are similar, the difference from reference value is within 0.5 ns. The results calculated by the two methods are obviously different in the last half series data. The calculation result of the RTS method is closer to the result calculated by the

forward filtering. The RMS of the results with the fixed interval smoother for seven consecutive days is 134 ps, and the RMS of the last 6 hours of each day is 151 ps. The RMS of the results using the RTS method for of the last 6 hours of each day is 138 ps. The result with RTS method has higher accuracy in the last 6 hours of daily data.

It can be seen from Fig. 6 that the difference between the time transfer results of the two stations through BDS PPP and the reference value is maintained within 0.5 nanoseconds. The calculation result of the RTS method is closer to the result calculated by the forward filtering. The RMS of the results with the fixed interval smoother for seven consecutive days is 122 ps, and the RMS of the last 6 hours of each day is 130 ps. The RMS of the results using the RTS method for of the last 6 hours of each day is 127 ps. The result with RTS method has higher accuracy in the last 6 hours of daily data.

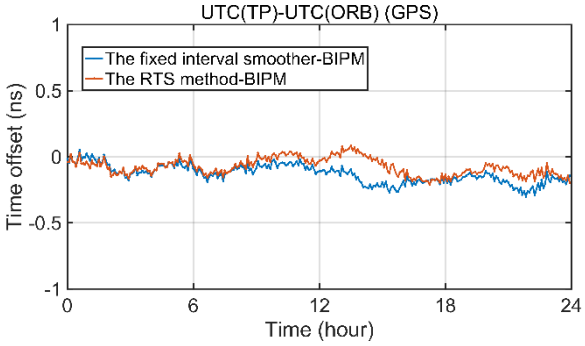


Fig. 5. Difference between time comparison results calculated by RTS algorithm and Fixed interval smoother GPS PPP. The observation data are from ORB and TP station, during MJD 59580.

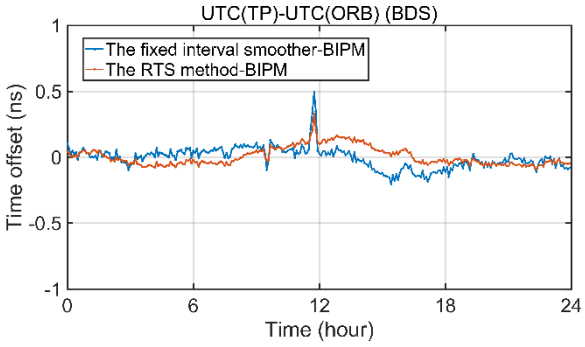


Fig. 6. Difference between time comparison results calculated by RTS algorithm and Fixed interval smoother BDS PPP. The observation data are from ORB and TP station, during MJD 59582.

The time transfer results calculated by the two navigation systems using these two methods are considerable precise, and the offset between the result and the reference value is within 0.5 nanoseconds. During the last 6 hours of data, the time-transfer results calculated using the RTS algorithm are slightly closer to the reference values.

3.3. Frequency Stability Analysis

The stability analysis is shown in following parts. Overlapping Allen deviation is used to evaluate the frequency stability of the alignment results [9].

Fig. 7 and Fig. 8 show the overlapping Allan variance of UTC(TP)-UTC(ORB) calculated by the three methods via GPS and BDS PPP and published by BIPM. It can be seen, for both GPS and BDS, the results of the RTS method and the fixed interval smoother have little difference compared to the value published by BIPM. However, the frequency stability of the results of the forward filtering is an order of magnitude lower than the others at an average time of 1000 seconds. As the average time increases to 100,000 seconds, the frequency stability of the results of the forward filtering raised to the same level as the others. It can be seen that both methods improve the short-term stability and reduce the impact of non-convergence compared with forward filtering for different navigation systems.

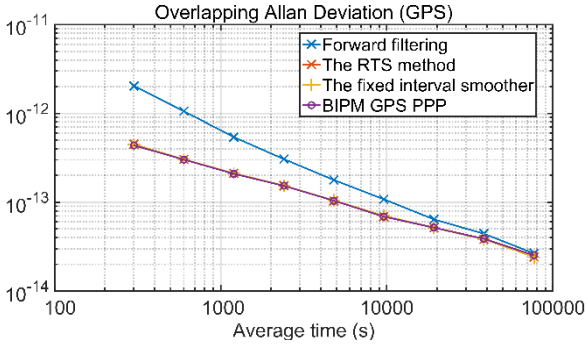


Fig. 7. Overlapping Allan deviation of UTC(TP)-UTC(ORB) calculated by different methods from MJD 58590 to MJD 59587.

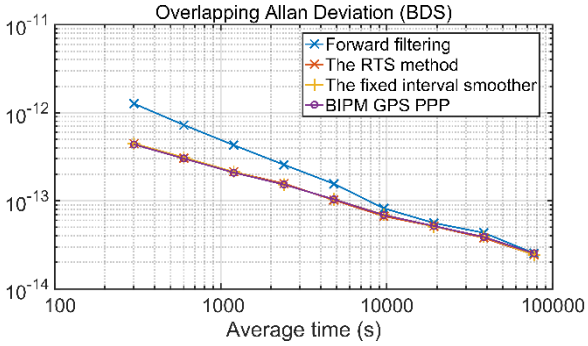


Fig. 8. Overlapping Allan deviation of UTC(TP)-UTC(ORB) calculated by different methods from MJD 58590 to MJD 59587.

Fig. 9 and Fig. 10 depicts the overlapping Allan variance calculated from 0:00 to 12:00 on MJD 59580 using the three methods by GPS and BDS PPP. The frequency stability of the results calculated with forward filtering is obviously worse than the other two

methods. It can be seen that in the previous part of the calculated data, both methods effectively improve the end effect with forward filtering.

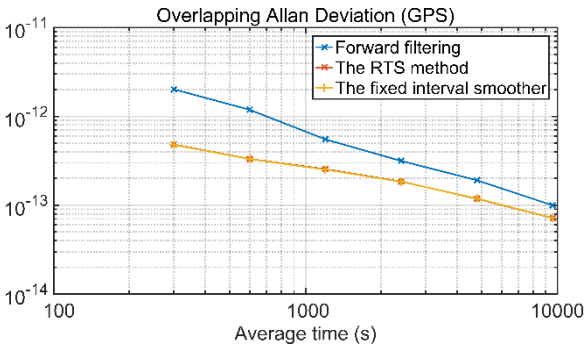


Fig. 9. Overlapping Allan deviation of UTC(TP)-UTC(ORB) calculated by different methods from 0:00 to 12:00 on MJD 59580.

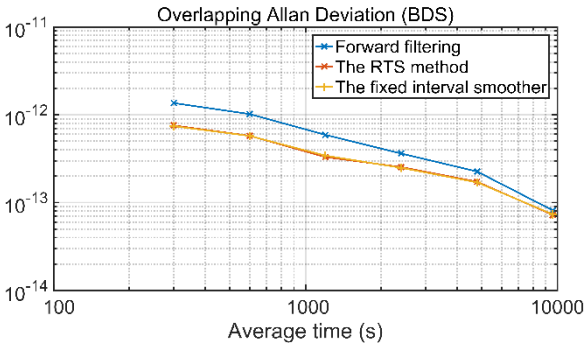


Fig. 10. Overlapping Allan deviation of UTC(TP)-UTC(ORB) calculated by different methods from 0:00 to 12:00 on MJD 59580.

The overlapping Allan deviation from 12:00 to 24:00 on MJD 59580 calculated by the three methods differs little from each other. Therefore, this part of the results is presented in Table 1.

Table 1 depicts the overlapping Allan variance calculated from 12:00 to 24:00 on MJD 59580 using the three methods. In the last half part of the calculated data, the results of the three methods have the same order of magnitude of frequency stability. The result calculated by the fixed interval smoother shows slightly worse frequency stability when the average time reach 10000 seconds. This is caused by the direct introduction of backward filtering results.

4. Conclusion

The results show that the PPP time transfer results based on GPS and BDS have high precision. The RMS of the offset between the reference values and PPP time transfer results using forward filtering, the fixed interval smoother and RTS methods are within 0.5 ns. Long-term frequency stability of time transfer results can reach 2E-14 with the average time of 1E6 seconds.

Table 1. Overlapping Allan deviation of UTC(TP)-UTC(ORB) calculated by different methods from 12:00 to 24:00 on MJD 59580.

Average time (s)	Overlapping Allan deviation		
	Forward filtering	The RTS method	The fixed interval smoother
GPS PPP			
3.00E+02	4.01E-13	4.03E-13	4.04E-13
6.00E+02	2.97E-13	2.98E-13	2.99E-13
1.20E+03	2.03E-13	2.02E-13	2.00E-13
2.40E+03	1.48E-13	1.47E-13	1.46E-13
4.80E+03	1.03E-13	9.87E-14	9.76E-14
9.59E+03	3.07E-14	2.99E-14	3.26E-14
BDS PPP			
3.00E+02	4.19E-13	4.17E-13	4.35E-13
6.00E+02	2.70E-13	2.70E-13	2.73E-13
1.20E+03	1.46E-13	1.47E-13	1.46E-13
2.40E+03	9.68E-14	9.82E-14	1.00E-13
4.80E+03	7.28E-14	7.35E-14	7.34E-14
9.59E+03	3.73E-14	3.17E-14	3.28E-14

For different navigation systems, the calculation results of the convergence time are different. Both the fixed interval smoother and the RTS methods can considerably improve the influence of the forward filter non-convergence on the accuracy of time transfer results. In the first few hours of the calculation, the error is reduced by a factor of several, and the frequency stability is improved by about an order of magnitude. But in the last hours of the results, the frequency stability of results calculated by the RTS method and the forward filter method are both slightly better than the fixed interval smoother.

In conclusion, the improved RTS method can effectively reduce the influence of the convergence of the forward filter on the time transfer results for GNSS PPP, and do not introduce large noise. This method can achieve sub-nanosecond PPP time transfer.

References

[1]. Kouba, J., Héroux, P., Precise Point Positioning Using IGS Orbit and Clock Products, *GPS Solutions*, 5, 2001, pp. 12–28.

[2]. Zumberge J. F., Heftin M. B., Jefferson D. C., et al., Precise point positioning for the efficient and robust analysis of GPS data from large networks, *Journal of Geophysical Research*, 102, B3, 1997, pp. 5005-5017.

[3]. Choy, S., Bisnath, S. & Rizos, C. Uncovering common misconceptions in GNSS Precise Point Positioning and its future prospect, *GPS Solutions*, 21, 2017, pp. 13–22.

[4]. Seepersad G., Bisnath S., Challenges in assessing PPP performance, *J Appl Geodesy*, 8, 3, 2014, pp. 205–222.

[5]. Bisnath S., Gao Y., Current state of precise point positioning and future prospects and limitations.

Observing Our Changing Earth, *International Association of Geodesy Symposia*, Springer, Berlin, Heidelberg, 133, 2009, pp. 615–623.

[6]. Hèroux P. et al., Products and applications for precise point positioning—moving towards real-time, in *Proceedings of ION GNSS 2004 Meeting*, Institute of Navigation, Long Beach, California, USA, 21–24 September 2004, pp 1832–1843.

[7]. Mengshi C et al, An Rauch-Tung-Striebel Algorithm to Improve the End Effect of Static GNSS PPP, in *Proceedings of the 4th IFSA Frequency & Time Conference*, Corfu, Greece, 21-23 September 2022, pp. 10-13.

[8]. Springer Handbook of Global Navigation Satellite Systems. Springer Handbooks, *Springer*, Cham.

[9]. Allan D. W., Barnes J. A., A Modified "Allan Variance" with Increased Oscillator Characterization Ability, in *Proceedings of the Thirty 5th IEEE Annual Frequency Control Symposium*., 1981, pp. 470-475.

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 75-81

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Evaluation on Negative Voltage Analysis Model for Gate Driving of MOSFET Application

[1, *] Ching-Guo Chen, [2] Shiu-Hui Lee, [1] Chih-Ming Yu, [1] Wen-Nan Huang, [2] Jin-Shyan Lee, [1] Hsing-Chi Meng and Tung-Ming Lai

[1] Potens Semiconductor Corp., 6F-3, 32, Gaotie 2nd Rd., Zhubei City, Hsinchu County, 30274, Taiwan
[2] Nation Taipei University of Technology, 1, Sec. 3, Zhongxiao E. Rd., Taipei, 10608, Taiwan
[1] Tel.: +886-3-6682068, fax: +886-3-6682279
[*] E-mail: tim@potens-semi.com

Received: 9 September 2022 /Accepted: 15 October 2022 /Published: 31 October 2022

Abstract: In power converter or inverter applications, MOSFETs are used as switches to control the value of output current or voltage. The MOSFET turn on/off are controlled by the PWM control IC or MCU with gate driver to achieve high frequency switching. In general, the absolute maximum negative voltage rating of the driving pin at the control IC or gate driver is not sufficient and the driving pin could be damaged by negative voltage which could be induced by the high frequency switching during the MOSFET turn on/off. The main purpose of this paper is to derive and adopt the negative voltage analysis model to evaluate and find out the MOSFET parameter relation which leads to the control IC damaged by negative voltage. The experimental results are demonstrated on the LLC converter with 400 V input voltage and 12 V output voltage. The negative voltage waveforms are measured from the LLC half bridge and synchronous rectifier circuits. Several experimental results are presented to validate this analysis model.

Keywords: Parameters of MOSFET, PWM control IC, Gate driver, Negative voltage, LLC converter.

1. Introduction

In many of switching power supply applications, the gate drive pin of the control IC could be damaged by the negative voltage over the maximum rating. The driving of the power MOSFET is related to the switching speed of the turn-on/off in the switching power supply design [1-3]. The characteristic of the MOSFET and the capability of the gate driver need to be checked at the same time [4- 8]. Due to the parasitic inductance of the MOSFET lead and PCB circuit trace, the driving circuit loop could generated the negative voltage by high current deviation at the reverse recovery time of the MOSFET body diode [9-11]. This paper introduces negative voltage analysis model for gate driving. The next section presents operation principle of gate driving and derivation of analysis model. Experimental results are presented in Section 3 and the conclusion is given in Section 4.

2. Implementation of Circuit

Among circuitry topologies for the switching power supply, LLC converter is chosen for negative voltage measurement and its main DC/DC and synchronous rectifier circuits are shown as Fig. 1. The S_1 and S_2 are the high voltage MOSFETs of half bridge and configured to output the square wave voltage. The low voltage MOSFETs S_3 and S_4 are the power switches of the secondary side synchronous rectifier to minimize power loss.

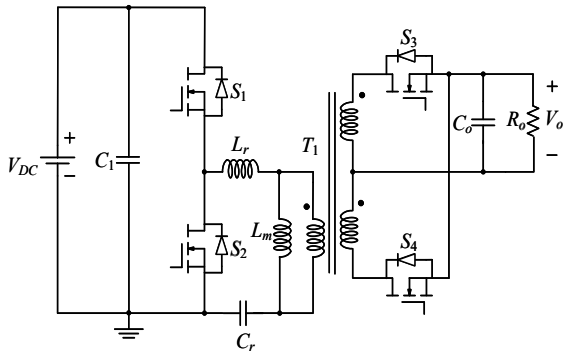


Fig. 1. The LLC converter for application.

2.1. Gate Resistor

The switching speed of the turn-on/off is related the parasitic capacitance of the MOSFET and gate driving circuit [12-14]. The basic driving circuit is shown as Fig. 2. The resistor R_{gs} is to make the gate-source voltage down to zero volts while the gate-source voltage is open. Therefore, we recommend placing 10 kΩ~100 kΩ resistor for reducing malfunction of the switch. The gate resistors R_{g_ext} and R_g and the capacitance C_{iss} would affect the switching speed and the switching loss. For gate resistor selection to reduce switching loss, the following equation is recommended for the setting of gate to source voltage rise/fall time by 5 time constants:

$$t_{rise/fall} = 5 \times (R_{g\_ext} + R_g) \times C_{iss} \ , \quad (1)$$

where C_{iss} is the input capacitance. R_g is the internal gate resistor. R_{g_ext} is the external resistor to change the switching speed for efficiency, thermal or EMI optimization in the gate drive circuit [15, 16]. From Potens' experience, we choose:

$$\frac{t_{period}}{t_{rise/fall}} \geq 50 \ , \quad (2)$$

where t_{period} is the period (cycle duration). The relation of the period and the switching frequency f_s is:

$$t_{period} = \frac{1}{f_s} \quad (3)$$

From equation (1) to (3), R_{g_ext} can be expressed as below:

$$R_{g\_ext} \leq \frac{1}{250 \times f_s \times C_{iss}} - R_g \ , \quad (4)$$

We can use equation (4) to determine the suitable external resistor for the gate drive circuit.

2.2. Negative Voltage Analysis Model

The gate drive circuit loop with parasitic inductance L_{P1}, the gate drive voltage V_{GS}, the gate drive resistor R_g, the gate drive resistor voltage V_R, the gate drive current I_g, and gate to source voltage V_{gs} are shown in Fig. 3.

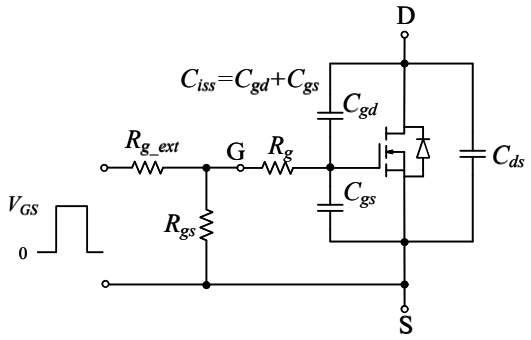


Fig. 2. The basic gate drive circuit.

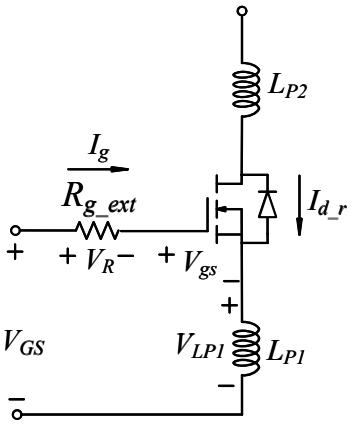


Fig. 3. The gate drive circuit with parasitic inductance.

In reverse recovery related period, I_{d_r} is defined as the current of body diode. Fig. 4 is the relation between the reverse current of body diode and the voltage of parasitic inductance. The I_{RM} is the maximum reverse current of body diode, t_{rr_P} is the period of the positive induced voltage and t_{rr_N} is the period of the negative induced voltage. Therefore, the magnitude of the negative voltage can be derived as

$$V_{LP1} = L_{P1} \times \frac{dI_{d\_r}}{dt} = L_{P1} \times \frac{0 - I_{RM}}{t_{rr\_N}} = -L_{P1} \times \frac{I_{RM}}{t_{rr_N}} \ , \quad (5)$$

The relation of the reverse recovery charge, the reverse recovery time and the maximum reverse current is shown as

$$Q_{rr} = \frac{1}{2} I_{RM} \times t_{rr} \ , \quad (6)$$

The duration of the negative voltage means the energy stress on the gate drive pin of the control IC. The more duration the negative voltage sustains; the more control IC will be damaged. Therefore, we can drive the negative voltage energy as

$$A = |V_{LP1}| \times t_{rr\_N} \tag{7}$$

Substitute equation (5) and (6) into (7), we can derive the area of the negative voltage energy as

$$A = 2 \times L_{P1} \times \frac{Q_{rr}}{t_{rr}} \tag{8}$$

The equation (8) means that the energy might damage the IC is related to the ratio of the reverse recovery charge and the reverse recovery time.

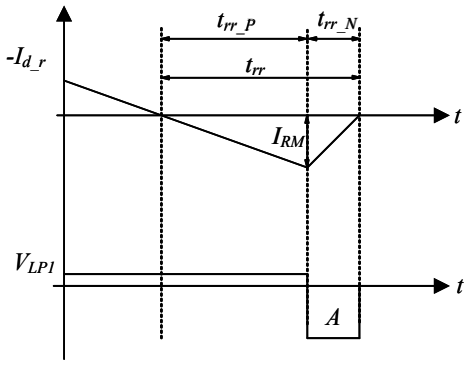


Fig. 4. Relation between Id current and negative voltage.

From the above relation, we can established the negative voltage analysis model considers the influence on the gate drive circuit and MOSFET equivalent circuit. The MOSFET parameters are the inputs of model and the parasitic parameter is measured from circuit loop inductance. By this analysis model, we can obtain the negative voltage as the output and estimate the influence of the negative voltage on the gate drive pin of the control IC from this model. Fig. 5 is the negative voltage analysis model. In gate drive circuit, the gate drive voltage can be derived as

$$\begin{aligned} V_{GS} &= V_{L_{P1}} + V_R + V_{gs} \\ &= L_{P1} \times \frac{dI_{d\_r}}{dt} + V_{gs} + I_g \times R_g \end{aligned} \tag{9}$$

When the gate drive resistor is increased, the magnitude of the negative voltage is decreased shown as the following relationship:

$$R_g \uparrow \Rightarrow |V_{GS}| \downarrow \tag{10}$$

From above relation, the smaller the negative voltage is, the smaller the energy is. Therefore, the possibility of IC damaged is being low by increasing the gate drive resistor.

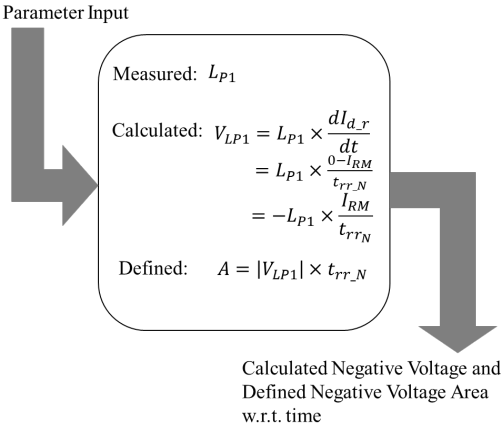


Fig. 5. Negative voltage analysis model.

3. Verification Based on Experimental Result

To demonstrate the effectiveness of the proposed analysis model, a 300 W LLC converter platform is chosen for demonstration. The negative voltage is measured and verified through this LLC converter platform with primary side half bridge and secondary side synchronous rectifier. Fig. 6 is the LLC circuit evaluation board and its main component selection and circuit parameters are given in Table 1.

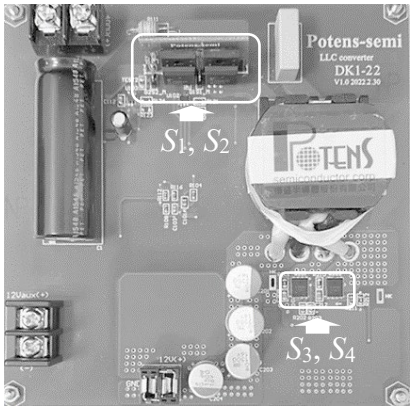


Fig. 6. The LLC circuit evaluation board.

Table 1. Parameters of main circuit.

Parameter	Value	Description
C_1	220 μF	450 V electrolytic capacitor
S_1, S_2	< 600 mΩ	650V Super Junction MOSFET
T_1	L_m=600 μH, L_r=100 μH	CC33, N_p:N_{s1}: N_{s2} = 34:2:2
C_r	68 nF	1 kV film capacitor
S_3, S_4	< 20 mΩ	65 V MOSFET
C_o	1500 μF × 4	16 V electrolytic capacitor

To compare more example, we take four different Super Junction MOSFETs for negative voltage analysis of S_1/S_2 pair [17 - 20]. Fig. 7 - 10 are the reverse current waveforms of PJF20N65N, PJF14N65N, PJF11N65N and PJF08N65N, which are measured by the reverse recovery tester. Then the measured data (the maximum reverse current of body diode, the period of the negative induced voltage, the reverse recovery charge, the reverse recovery time) and the calculation data (the negative voltage, the negative energy and the ratio of the reverse recovery charge and the reverse recovery time) are shown in Table 2 and the parasitic inductance is measured and it is about 5 nH. Table 2 shows that the negative voltage energy is proportional to the ratio of the reverse recovery charge and the reverse recovery time. It means the negative voltage energy in MOSFET application can be judged by the ratio of the reverse recovery charge and the reverse recovery time directly while choosing a new MOSFET. Fig. 11- 14 are the negative voltage of V_{GS} waveforms of PJF20N65N, PJF14N65N, PJF11N65N and PJF08N65N. These figures show that the period of the negative induced voltage and the negative voltage are close to the calculated data in Table 2. The comparison of calculated and measured negative voltage is shown in Table 3. There is above 85 % accuracy for the calculated data from the negative voltage analysis model.

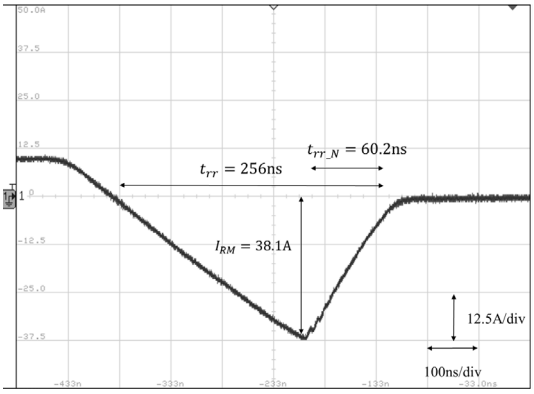


Fig. 7. The reverse current of PJF20N65N.

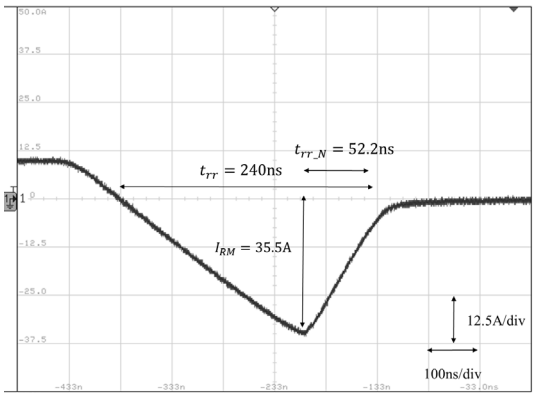


Fig. 8. The reverse current of PJF14N65N.

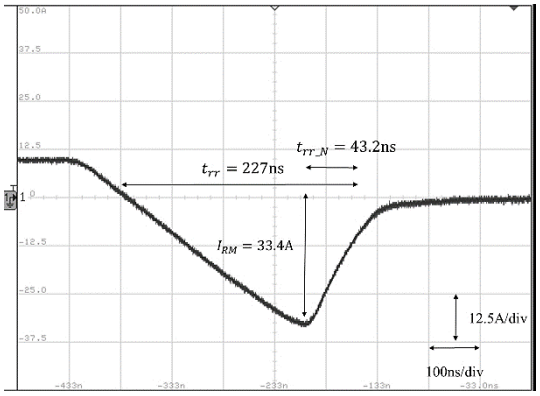


Fig. 9. The reverse current of PJF11N65N.

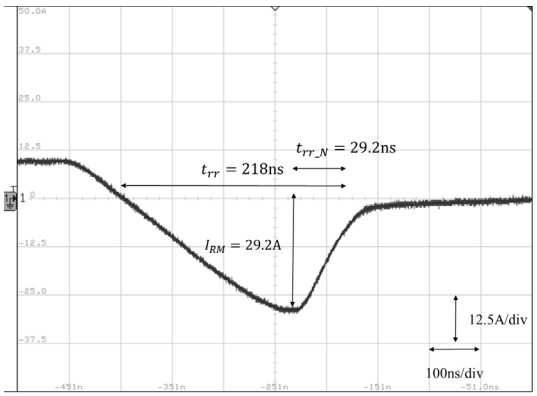


Fig. 10. The reverse current of PJF08N65N.

Table 2. MOSFET parameters and calculated negative voltage.

Part No.	t_{rr} (ns)	Q_{rr} (nC)	I_{RM} (A)	t_{rr_N} (nS)	V_{LP1} (V)	A (V•ns)	$\frac{Q_{rr}}{t_{rr}}$
PJF20N65N	256	5233	38.1	60.2	-3.2	193	20.4
PJF14N65N	240	4576	35.5	52.2	-3.4	177	19.1
PJF11N65N	227	4025	33.4	43.2	-3.85	168	17.7
PJF08N65N	218	3888	29.2	42.7	-3.4	145	15.5

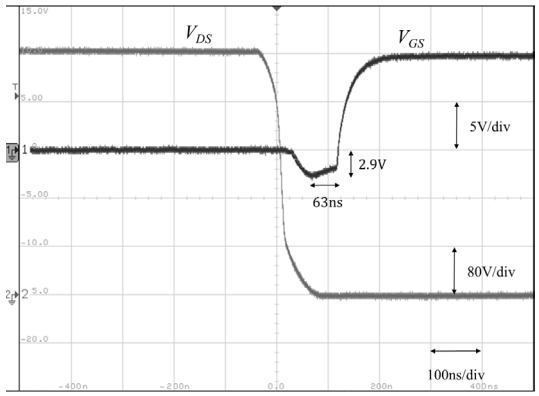


Fig. 11. The negative voltage of PJF20N65N.

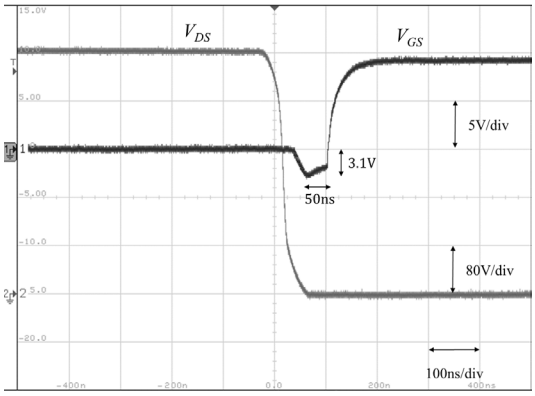


Fig. 12. The negative voltage of PJF14N65N.

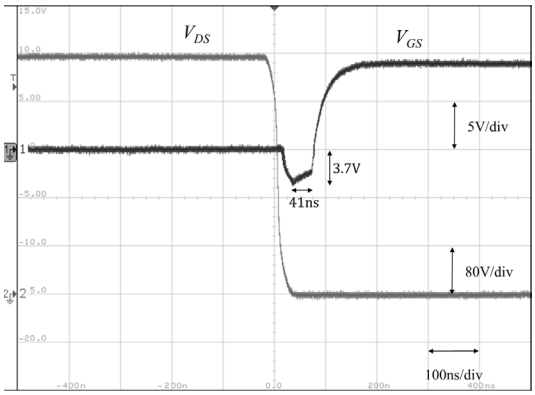


Fig. 13. The negative voltage of PJF11N65N.

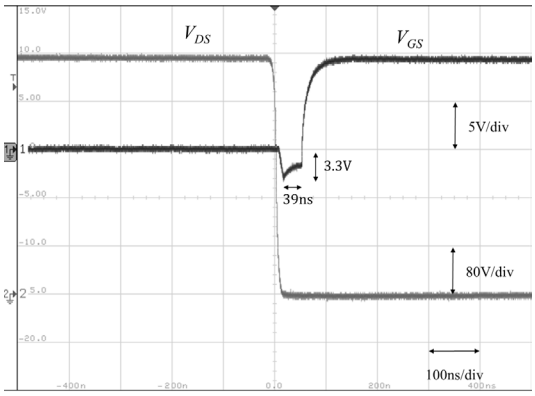


Fig. 14. The negative voltage of PJF08N65N.

Table 3. Calculated and measured negative voltage.

Part No.	Calculated data		Measured data		Error	
	V_{LP1} (V)	A (V·ns)	V_{LP1} (V)	A (V·ns)	V_{LP1} (%)	A (%)
PJF20N65N	-3.2	193	2.9	183	10.3%	5.5%
PJF14N65N	-3.4	177	3.1	155	9.7%	14.2%
PJF11N65N	-3.85	168	3.7	152	5.4%	10.5%
PJF08N65N	-3.4	145	3.3	129	3.0%	12.4%

For synchronous rectifier, we also take four different low voltage MOSFETs for negative voltage and R_g impact analysis of S_2/S_4 pair [21 - 24]. Fig. 15 shows the turn on waveforms of the PDEC69F0BX-5 which R_g is 26 Ω, and Fig. 16 shows the turn on waveforms of the PDC6974X-5 which R_g is 1.8 Ω. Fig. 17 shows the turn on waveforms of the PDC6986BX-5 which R_g is 0.9 Ω. Fig. 18 shows the turn on waveforms of the PDC6988X-5 which R_g is 0.3 Ω. These results shows that a larger R_g will have a smaller negative voltage. From above analysis, we could also conclude that the large ratio of the reverse recovery charge and the reverse recovery time is the highest possibility to damage the gate drive pin of the control IC.

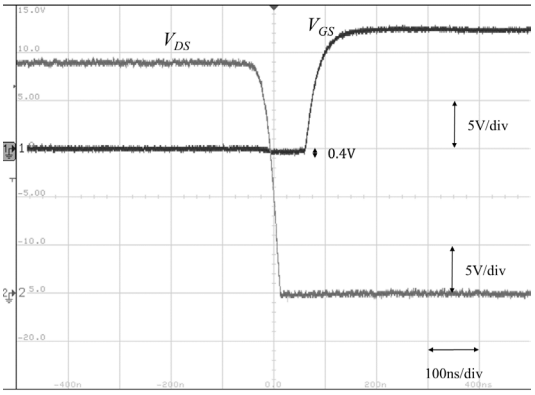


Fig. 15. The negative voltage of PDEC69F0BX-5.

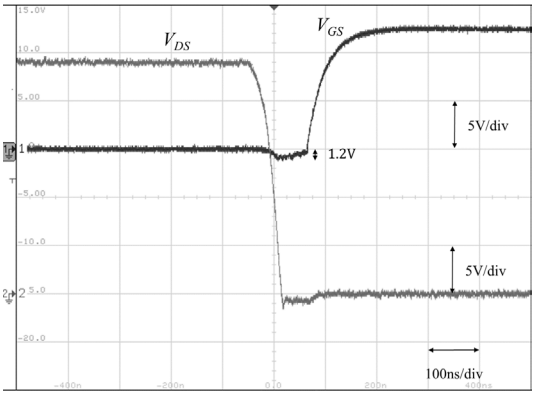


Fig. 16. The negative voltage of PDC6974X-5.

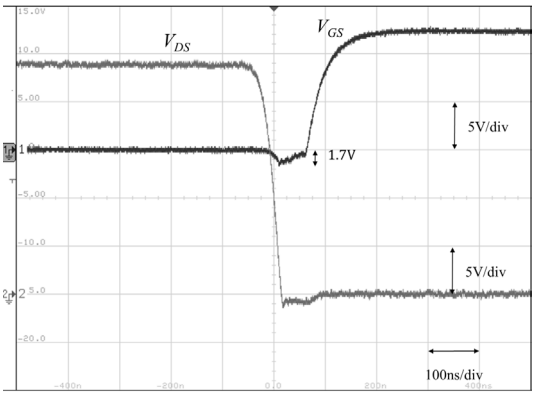


Fig. 17. The negative voltage of PDC6986BX-5.

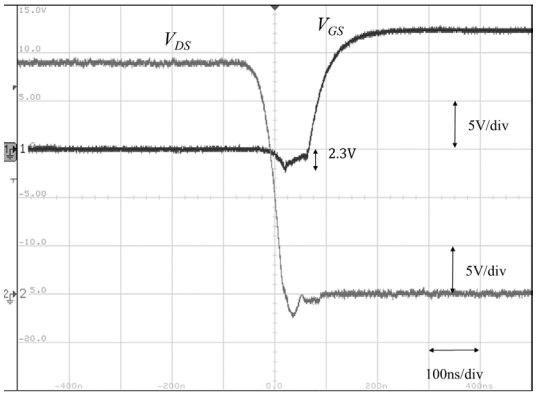


Fig. 18. The negative voltage of PDC6988X-5.

4. Conclusions

This paper shows an analysis method to evaluate the influence of the negative voltage caused by the parasitic inductance of the MOSFET lead and PCB circuit trace and an impact on the drive pin of the control IC. The reverse recovery time is also the highest possibility to damage the gate drive pin of the control IC. The more duration the negative voltage sustains, the more possibility that the control IC will be damaged. However, we could reduce the negative voltage drop at the gate drive pin by increasing the gate drive resistor to avoid IC damaged.

References

[1]. C. G. Chen, S. H. Lee, C. M. Yu, W. N. Huang, J. S. Lee, H. C. Meng and T. M. Lai, Negative Voltage Analysis Model for Evaluation on Control IC Driving of MOSFET Application, in *Proceedings of the 4th International Conference on Microelectronic Devices and Technologies (MicDAT '2022)*, Corfu, Greece, 21-23 September 2022, pp. 5-8.

[2]. B. Sun and R. Burgo, Ultra-low input-output capacitance PCB-embedded dual-output gate-drive power supply for 650 V GaN-Based half-bridges, *IEEE Trans. Power Electron*., Vol. 34, No. 2, Feb. 2019, pp. 1382-1393.

[3]. R. Mitova, J. C. Crebier, L. Aubard and C. Schaeffer, Fully integrated gate drive supply around power switches, *IEEE Trans. Power Electron*., Vol. 20, No. 3, May 2005, pp. 650-659.

[4]. STMicroelectronics, AN2626: MOSFET body diode recovery mechanism in a phase-shifted ZVS full bridge DC/DC converter, *STMicroelectronics*, Application Note, 2007.

[5]. Infineon Technologies, AN_201609_PL52_031: Benefits of low side MOSFET drivers in SMPS, *Infineon Technologies*, Application Note, 2016.

[6]. M. Miura-Mattausch, S. Ooshiro and M. Suetake, Circuit simulation models for coming MOSFET generations, in *Proceedings of the International Conference on Simulation Semiconductor Processes and Devices*, Seattle, USA, 6-8 September 2000, pp. 106-111.

[7]. W. S. Choi, S. M. Young and D. W. Kim, Analysis of MOSFET failure modes in LLC, in *Proceedings of the 31st International Telecommunications Energy Conference (INTELEC' 2009)*, Incheon, Korea, 18-22 October 2009, pp. 1-6.

[8]. C. D. G. Vladimir, T. J. Manuel, B. B. Hector, Z. T. Erwin, Q. M. Felix and L. A. M. Fernanda, Modeling and simulation of a MOSFET transistor in Verilog-A considering parasite elements, in *Proceedings of the IEEE Workshop on Power Electronics and Power Quality Applications (PEPQA' 2015)*, Bogota, Colombia, 2-4 June 2015.

[9]. J. Victory, S. Pearson, S. Benczkowski, T. Sarkar, H. Jang, M. B. Yazdi and K. Mao, A physically based scalable SPICE model for shielded-gate trench power MOSFETs, in *Proceedings of the 28th International Symposium on Power Semiconductor Devices and ICs (ISPSD' 2016)*, Prague, Czech, 12-16 June 2016, pp. 219-222.

[10]. P. Hower, C. Kaya, S. Pendharkar and C. Jones, Reverse-recovery safe operating area of diodes in power integrated circuits, in *Proceedings of the 24th International Symposium on Power Semiconductor Devices and ICs*, Bruges, Belgium, 3-7 June 2012, pp. 65-68.

[11]. R. Bonyadi, O. Alatise, J. Hu, J. A. Ortiz-Gonzalez, L. Ran and P. A. Mawby, Compact electro-thermal reliability modelling and experimental characterisation of bipolar latch up in SiC and CoolMOS power MOSFETs, *IEEE Trans. Power Electron*., Vol. 30, No. 12, January 2015, pp. 6978-6992.

[12]. V. Dimitrov, P. Goranov, and D. Hvarchilkov, An analytical approach to model the switching losses of a power MOSFET, in *Proceedings of the IEEE International Power Electronics and Motion Control Conference (PEMC' 2016)*, Varna, Bulgaria, 25-28 September 2016, pp. 928-933.

[13]. Toni Lopez and Eduard Alarcon, Power MOSFET technology roadmap toward high power density voltage regulators for next-generation computer processors, *IEEE Trans. Power Electron*., Vol. 27, No. 4, April 2012, pp. 2193-2203.

[14]. B. Nguyen, X. Zhang, A. Ferencz, T. Takken, R. Senger and P. Coteus, Power MOSFET technology roadmap toward high power density voltage regulators for next-generation computer processors, *IEEE Trans. Power Electron*., Vol. 27, No. 4, April 2012, pp. 2193-2203.

[15]. Y. C. Son, K. Y. Jang and B. S. Suh, Integrated MOSFET inverter module for low-power drive system, *IEEE Trans. Ind. Appl*., Vol. 44, No. 3, April 2008, pp. 878-886.

[16]. X. Yu and P. Yeaman, Temperature-related MOSFET power loss modeling and optimization for DC-DC converter, in *Proceedings of the 28th Annual IEEE Applied Power Electronics Conference and Exposition (APEC' 2013)*, Long Beach, USA, 17-21 March 2013, pp. 2788-2792.

[17]. PJF20N65N datasheet, Potens, Hsinchu County, Taiwan, 2022. Available: https://potens-semi.com/upload/product/PJF20N65N.pdf

[18]. PJF14N65N datasheet, Potens, Hsinchu County, Taiwan, 2022. Available: https://potens-semi.com/upload/product/PJF14N65N.pdf

[19]. PJF11N65N datasheet, Potens, Hsinchu County, Taiwan, 2022. Available: https://potens-semi.com/upload/product/PJF11N65N.pdf

[20]. PJF08N65N datasheet, Potens, Hsinchu County, Taiwan, 2022. Available: https://potens-semi.com/upload/product/PJF08N65N.pdf

[21]. PDEC69F0BX-5 datasheet, Potens, Hsinchu County, Taiwan, 2022. Available: https://potens-

semi.com/upload/product/PDEC69F0BX-5.pdf
[22]. PDC6974X-5 datasheet, Potens, Hsinchu County, Taiwan, 2022. Available: https://potens-semi.com/upload/product/PDC6974X-5.pdf
[23]. PDC6986BX-5 datasheet, Potens, Hsinchu County, Taiwan, 2022. Available: https://potens-semi.com/upload/product/PDC6986BX-5.pdf
[24]. PDC6988X-5 datasheet, Potens, Hsinchu County, Taiwan, 2022. Available: https://potens-semi.com/upload/product/PDC6988X-5.pdf

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 82-92

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Low-cost Integrated Photonic Platform Developed via a Sol-gel Dip-coating Method: A Brief Review

[1] Łukasz KOZŁOWSKI, [1] Muhammad SHAHBAZ, [1, *] Muhammad Ali BUTT, [2] Cuma TYSZKIEWICZ, [2] Paweł KARASIŃSKI, [1] Andrzej KAŹMIERCZAK and [1] Ryszard PIRAMIDOWICZ

[1] Warsaw University of Technology, Institute of Microelectronics and Optoelectronics, Koszykowa 75, 00-662 Warszawa, Poland
[2] Silesian University of Technology, Department of Optoelectronics, ul. B. Krzywoustego 2, 44-100 Gliwice, Poland
E-mail: ali.butt@pw.edu.pl

Received: 31 August 2022 /Accepted: 3 October 2022 /Published: 31 October 2022

Abstract: Optical devices and circuits are now key elements in a variety of applications, including biotechnology, automotive, food quality control, chemistry, etc. It is highly desired to find low-cost solutions to the photonics integrated circuits (PICs). In this paper, a comprehensive review of silica-titania (SiO_2:TiO_2) materials obtained by a sol-gel method come together with a dip-coating technique, and the numerical analysis of different optical devices based on the SiO_2:TiO_2 optical platform are discussed. The sol-gel method is highly effective and does not require the use of expensive high-tech equipment. We discuss recent progress in our consortium on a cost-effective SiO_2:TiO_2 optical platform for integrated photonics applications, with a concentration on materials, devices, and optical sensing principles investigated via numerical simulations.

Keywords: Integrated photonics, Photonic integrated chip, Photonic sensor, Ring resonator, Grating, Photonic crystal, Silica titania, Sol-gel dip-coating method.

1. Introduction

Optical devices and circuits are now essential components in various applications, including biotechnology, automotive, food quality control, and chemistry, to mention a few. The attraction in optical sensing is supported by the exceptional benefits made possible by photonic technologies, such as high sensitivity, low cost, compactness, integration with electronic devices, and metal-free operation. Since technology has advanced so quickly in the last decade, silicon (Si) photonics has become one of the most efficient technology platforms for producing a wide range of functional optical components [1-5].

On the other hand, silica (SiO_2), titania (TiO_2), and silica-titania (SiO_2:TiO_2) materials obtained by the sol-gel technique have received much attention due to their potential optical applications [6-9]. The substantial performance of integrated photonic sensors is supported by technological characteristics and advances, among which ring resonators, subwavelength gratings (SWG), and one-dimensional photonic crystals (PhC) have fascinated the attention of scientists in recent years [6, 10-12]. In the following

sections, we discuss the recent developments on a low-cost silica-titania optical platform for integrated photonics applications, with special consideration to materials, devices, and optical sensing principles based on the numerical analysis performed in our group. Building planar (1D) and ridge (2D) waveguides for photonic integrated circuits (PICs) with acceptable optical characteristics and low optical transmission losses has been processed using SiO_2, TiO_2, and SiO_2:TiO_2 compounds obtained by the sol-gel method [13-15]. The sol-gel method and dip-coating technique are simple and inexpensive [16]. They allow the coating of large surfaces and do not require extremely high-temperature processing. Fabrication of high-quality thin films with exceptional thermal and mechanical stability using them is possible [17, 18]. A sufficient refractive index (R.I) contrast (Δn) between the substrate and the guiding layer is required for a functional WG [19].

The refractive indices and thickness of sol-gel films can be modified by fine-tuning the set-off precursors and solvents, in addition to the compound molar ratio and reaction temperature. Rare-earth elements [20, 21], laser dyes [22], and other organic compounds [23] can be doped into these WGs, permitting SiO_2 and TiO_2 sol-gel materials to be utilized in optical amplifiers, laser-active media, and sensing applications [24, 25]. The sol-gel process is highly effective and does not necessitate using pricey high-tech equipment. The sol-gel technique can adjust the R.I of the waveguide films. The optical losses sustained by these waveguides are equivalent to waveguides obtained using the chemical vapor deposition (CVD)/low-pressure chemical vapor deposition (LPCVD) method [26, 27]. Several research groups have developed optical waveguides based on a silica-titanium platform for different applications in recent years [28-30].

A recent study [9] demonstrated a desirable waveguide platform based on SiO_2:TiO_2 that is simple to develop and inexpensive. The material is an excellent choice for the visible and near-IR wavelength range due to its broad transparency range. Another study [10] discussed developing an optical waveguide system utilizing the sol-gel method and dip-coating technique. The optical characterization of the waveguide system is investigated at a visible wavelength. However, the system can operate from visible to near-IR wavelength ranges.

The same approach involving the sol-gel method [11] is used to develop high-quality SiO_2:TiO_2 thin films on a glass substrate. The coating process produces thin films with reduced surface roughness and losses, making them an excellent choice for the waveguiding purpose. A recent work [6] proposes numerical modeling of extremely attractive SWG waveguides for filtering and sensing applications using the finite element method for modelling and nano-imprint lithography (NIL) process to develop low-cost integrated photonic devices. The experimental analyses of these sensor designs are yet to be performed. However, the development of the silica-titania platform and its characterization has been performed at the Silesian University of Technology, Poland.

2. Materials and Fabrication Process

The coating process involves depositing a thin layer of materials onto a substrate in either the solid or liquid phase (solution). According to production needs for coated layer thickness, coated surface roughness, rate, and coating product size—which can be determined by coating velocity, coated film width, and designing capabilities—coating strategies may be used. Silica titania (SiO_2: TiO_2) has been widely researched because of its potential optical properties and applications [31]. It is a material that constitutes the new low-cost technological platform for PICs via the aforementioned combination of the sol-gel method and dip-coating technique [32]. SiO_2:TiO_2 itself is a very attractive material regarding integrated photonics because its R.I may vary between 1.6-2.2 and its spectral range stands from visible to NIR [12]. Moreover, the material combined with this fabrication technique gives the possibility to create optical interconnects with minimal transmission losses [33-36].

First of all, the fabrication process of a single SiO_2:TiO_2 waveguide film will be discussed. This process is relatively simple [25, 37] as depicted in Fig. 1.

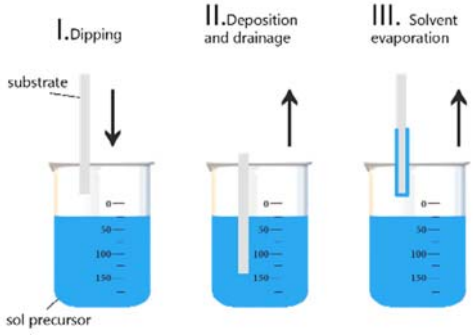


Fig. 1. Sol-gel dip-coating fabrication method.

To carry out the fabrication process the following components are required:

1. Substrate – which is the material on which SiO_2:TiO_2 will be deposited will be fully covered by it, creating a full SiO_2:TiO_2 waveguide film, e.g., soda-lime glass or BK7 glass. BK7 glass is preferred due to its lower thermal expansion coefficient and lower surface roughness.
2. Sol precursor – a previously prepared colloidal solution that contains the covering material, e.g., SiO_2:TiO_2.

The preparation of the sol precursor for SiO_2:TiO_2 consists of adding SiO_2 and TiO_2 precursors

respectively being tetraethyl ortosilicate (TEOS) and tetraethyl ortotitanate (TET) with a homogenizing factor which in this case was ethyl alcohol (C_2H_5OH). Further to carry out the reaction of hydrolysis and condensation, hydrochloric acid (HCl) needs to be used [13].

After having the BK7 glass covered by SiO_2:TiO_2 to harden the material it is necessary to carry out heat treatment. Numerous factors, such as pH and solution concentration, might have a significant impact on the efficacy of the dip-coating process by altering, for instance, its viscosity. Nevertheless, because coating techniques used in modern studies have been well-documented in earlier works, pH and solution viscosity are rarely discussed by writers. Using more viscous solutions causes more clumping and thicker coating layers, so fewer dip-coating cycles are necessary to achieve a specific thickness or amount of deposited mass. The resulting coating, meanwhile, may shatter and agglomerate particles in unfavorable locations. To get around these problems, the dip-coating technique might offer a more homogeneous coating by combining a less viscous solution with numerous repetitions.

As for the dependence of R.I on the variables of the process, it should be noted that it depends on the stoichiometric ratio between precursor components as shown in Fig. 2.

The withdrawal speed of the substrate from the sol affects predominately the thickness of the deposited film as presented in Fig. 3.

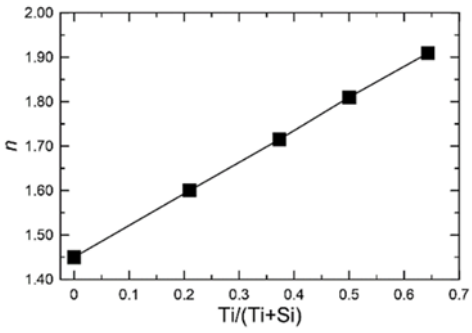


Fig. 2. Sol-gel dip-coating fabrication method [10].

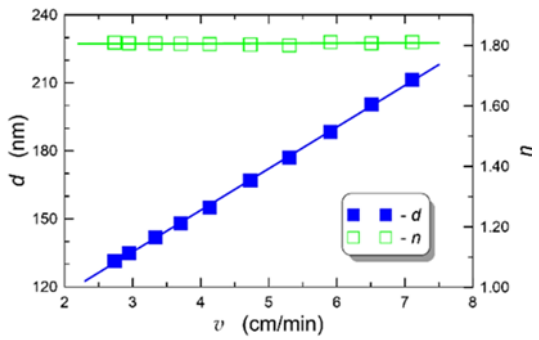


Fig. 3. Sol-gel dip-coating fabrication method [10].

The main benefit of using this method over other conventional methods of thin film deposition as plasma-enhanced chemical vapor deposition (PECVD), low-pressure chemical vapor deposition (LPCVD), or physical vapor deposition (PVD) is that it is much easier to conduct, extremely cheaper and does not require foundries or advanced equipment and laboratories [38].

Once the platform is prepared, the next step is the fabrication of waveguide interconnects and other waveguide structures. There are several methods of fabrication of such structures. Some approaches are shown in Fig. 4. All the following three methods are currently being explored in our consortium.

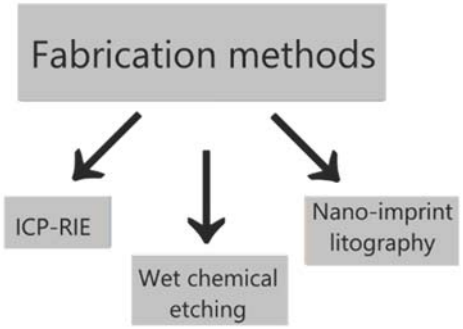


Fig. 4. WG structure fabrication processes.

An exemplary rib WG fabrication process via reactive ion etching (RIE) [39] is shown in Fig. 5.

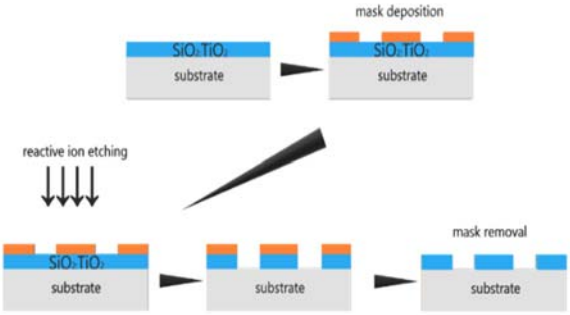


Fig. 5. SiO_2:TiO_2 ridge WG fabrication on a glass substrate.

The main design variables of a basic WG are its width w and its height h. Nevertheless, the crucial parameters for a WG to operate properly are the refractive indices of the WG, substrate, and ambient medium. In Fig. 4 the refractive indices are $n_{wg}=1.7$, $n_{sub}=1.5$, and $n_{amb}=1$.

Another interesting WG fabrication process [40] is shown in Fig. 6 which involves the wet-chemical etching of substrate and later on depositing a thin-film of SiO_2:TiO_2 material via the dip-coating method. This

WG structure can provide a low-cost solution to the PICs.

Both WGs shown in Fig. 5 and Fig. 6 are layouts that can be fabricated on SiO_2:TiO_2 with the use of the previously mentioned fabrication method.

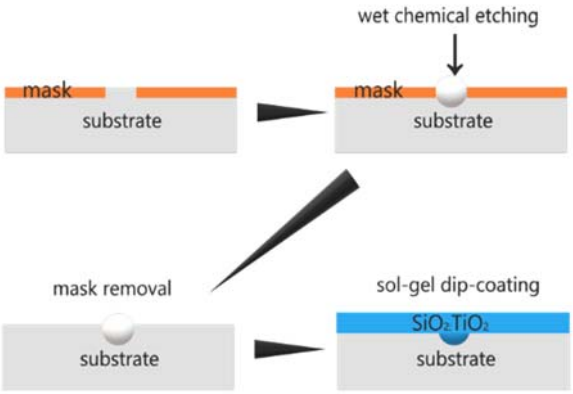


Fig. 6. Inverted rib waveguide fabrication on a glass substrate.

The most convenient method of fabricating waveguide interconnects and structures consist of lithography and etching. The mainstream technologies are electron beam lithography (EBL) or conventional UV lithography combined with RIE or inductively coupled plasma (ICP) RIE.

Those technologies are well-researched and applied for a long time which gives them the advantage of being reliable. However, those are technologies that require advanced and expensive laboratory equipment making them inaccessible and not cost-effective. To fabricate an integrated photonic chip, it is necessary to use the services of foundries.

The approach of using different methods that are not as expensive and do not require sophisticated equipment could be a game changer regarding integrated photonics. One of the technologies that could open a lot of new fabrication possibilities for integrated photonics could be nano-imprint lithography (NIL). The simplicity and cost-effectiveness of NIL are its biggest advantages over the previously mentioned conventional methods [11]. Nevertheless, this technology is not as well developed and well researched concerning other methods so it needs more time and attention to match other more well-known technologies.

The simplicity of NIL lies in the fact that it is easily reproducible and the fabrication of waveguide structures is etch-less. However, there are several variables such as force applied during imprinting, amount of sol-gel, and post-annealing temperature that should be properly optimized to avoid cracks and guarantee reproducibility. It is shown in Fig. 7. If some automatization would be added to the process, then it could revolutionize the whole integrated photonics industry.

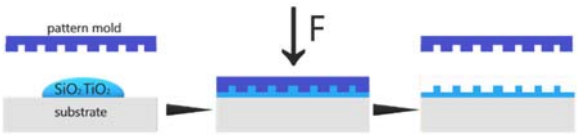


Fig. 7. NIL process.

3. Devices and Applications

The main and most interesting applications for PIC platforms based on SiO_2:TiO_2 materials would be sensing and biosensing. Of course, it can find many other functions same as other conventional materials used for integrated photonics [26, 41].

Three investigated structures for sensing involving: ring resonators, subwavelength gratings (SWG), and 1-D photonic crystals (1-D PhC) are discussed below.

3.1. Ring Resonators

The SiO_2:TiO_2 optical waveguide's R.I sensing capabilities were investigated using the ring resonator device [10]. Sensitivity (*S*), a figure of merit (*FOM*), and the *Q-factor* are three crucial factors that should be wisely considered while creating sensing devices. The following expression is used to determine the sensitivity of the ring resonator.

$$S = \frac{\Delta\lambda_{res}}{\Delta n}, \quad (1)$$

where $\Delta\lambda_{res}$ is the change in resonance wavelength and Δn represents the shift in ambient R.I.

The fraction of the sensitivity and full width at half maximum (FWHM) of the resonance dip is known as a *FOM*. *FOM* is calculated by using the following expression:

$$FOM = \frac{S}{FWHM} \quad (2)$$

For a variety of functions, integrated resonators with superior Q-factors are mostly needed. The Q-factor is calculated as follows:

$$Q - Factor = \frac{\lambda_{res}}{FWHM} \quad (3)$$

The ring resonators are particularly sensitive to changes in the ambient medium, resulting in the redshift of the resonance wavelength. The value of shift in the resonance wavelength depends on the device's geometric variables [10]. A schematic diagram of a ring resonator based on a SiO_2:TiO_2 optical waveguide is shown in Fig. 8.

The radius of the ring (*R*) was selected to be 15 μm. The distance of coupling between the bus waveguide and ring is symbolized as *g*. The surrounding medium

was filled with a dielectric material with an R.I of 1.33. The height of the device was set at 200 nm [10].

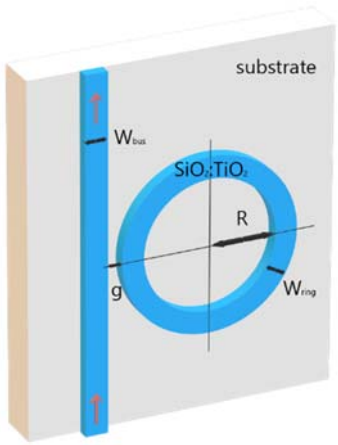


Fig. 8. A schematic view of a ring resonator based on SiO_2:TiO_2 optical waveguide [10].

3.1.1. Geometric Variables of the Ring Structure

Fig. 9a shows the ring resonator's transmission spectrum. It is worth mentioning that six resonance dips (λ_{dip}) with an FSR of ~6.4 nm and the FWHM of λ_{dip} ~0.44 nm were achieved, as illustrated in Fig. 9b. The coupling efficiency (CE) was calculated for each λ_{dip}, as illustrated in Fig. 9c. CE was calculated by estimating the ER of the bus waveguide and obtaining the ER for g in the 50-350 nm range. The max. value ER of ~2.5-2.6 dB was achieved for g = 250 nm to 300 nm [10].

3.1.2. Sensing Analysis and Device Performance

The R.I of the ambient medium was varied from 1.33 to 1.365 with 0.005 step size to test the sensing abilities of the sensor [10].

As the R.I of the ambient medium increased, the λ_{dip} executed a redshift, as shown in Fig. 10a. For this analysis, W_{bus} = W_{ring} = 800 nm was used, which resulted in substantial light confinement in the ring structure and a low evanescent field. As a result, light-matter interaction is low, with S = ~90 nm/RIU, FOM = ~204.5 RIU1, and Q-factor = 2239. Fig. 10b and Fig. 10c shows normalized electric field distributions at λ = 985.3 nm and = 987 nm, which correspond to the on-resonance and off-resonance states, respectively [10].

As we know, boosting the evanescent field ratio (EFR) can improve sensitivity. It is possible to get higher EFR by shrinking the width of the ring waveguide (W_{ring}). Reducing W_{ring} = 500 nm improved the *S*, *FOM*, and *Q-factor* of the device, as shown in Fig. 11.

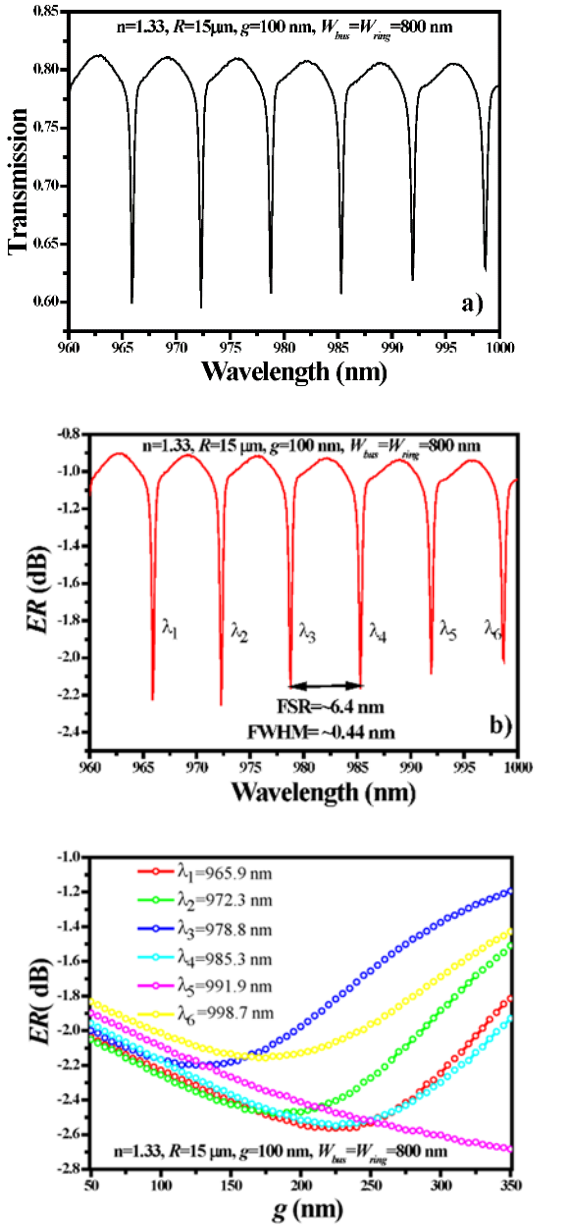


Fig. 9. Spectral properties of the ring resonator configuration: (a) transmission spectrum, (b) FSR vs ER, (c) optimization of g [10].

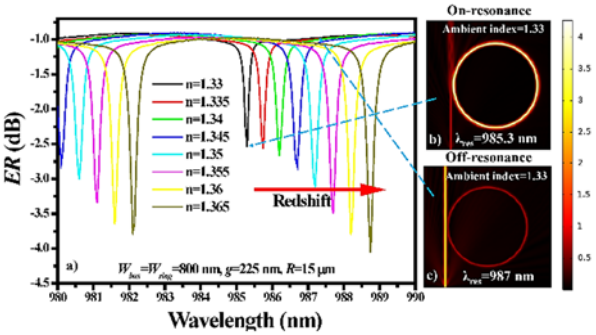


Fig. 10. (a) Transmission spectrum of the ring resonator with various ambient refractive indices. Distribution of norm. The electric field in the ring resonator structure in the (b) on-resonance state and (c) off-resonance state [10].

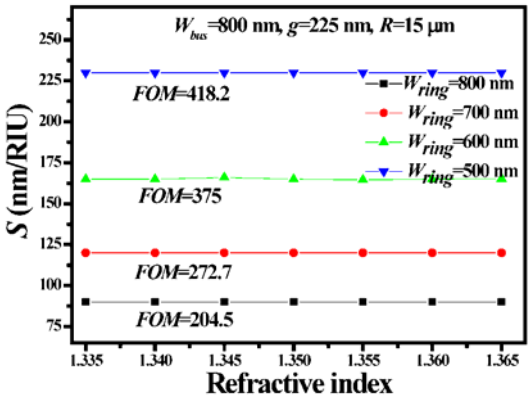


Fig. 11. Analysis of *S* and *FOM* depending on W_{ring}. The other geometric variables for instance W_{bus}, R , and g were set at 800 nm, 15 μm, and 225 nm, respectively [10].

3.1.3. Summary of the Performance of the Device at Different Values of W_{ring}

The device performance at various Wring values is summarized in Table 1. We can observe that by lowering Wring from 800 nm to 500 nm, the sensitivity (S), a Figure of Merit, and Q-factor of the device were improved to 230 nm/RIU, 418.2 RIU-1, and 2247.5, respectively.

Table 1. Ring resonator performance for different values of W_{ring} [10].

W_{ring} (nm)	*S* (nm/RIU)	*FOM* (RIU^{-1})	*Q-Factor*
500	~230	~418.2	2247.5
600	~165	~375	2240
700	~120	~272.7	2227.5
800	~90	~204.5	2239

The findings in this study are incredibly attractive and competitive with those of Si photonics and optical fiber-based sensors. A semiconductor waveguide R.I sensor with S = 235 nm/RIU was tested for a bio-sensing application [42]. According to Carlborg et al., who achieved an experimental sensitivity of 246 nm/RIU with a far more complicated ring resonator layout structure [43], the resulting sensitivity value is similar to silicon nitride slot waveguide ring resonator's sensors.

3.2. Subwavelength Grating Waveguide (SWG Waveguide)

Another investigated structure is an SWG waveguide. It is a structure that may be used to implement a waveguide NIR filter [12, 44-45] or a waveguide Fabry–Pérot (FP)-sensor [12] as shown in Fig. 12 (top) and Fig. 12 (bottom), respectively.

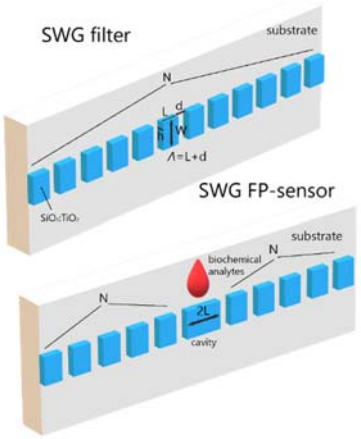


Fig. 12. SWG NIR-filter and FP-sensor structures design [12].

As shown in the schematic representation, an SWG is a structure that relies on many different variables, which if changed affect the response of the whole structure. The main variables of an SWG are W – width of the waveguide, H – the height of the WG, Λ – period of the grating being the sum of L – length of a WG segment, and d – the distance between the segments, N – number of periods. There is also a parameter named *duty cycle* which is the ratio of $\frac{L}{\Lambda}$.

3.2.1. NIR-filter

A NIR-filter is one of the SWG applications. It works based on cutting a certain wavelength window. Light after passing through an SWG in the transmission spectrum is observed as the input spectrum minus a specific range of wavelengths [44]. As a result, SWG WGs can work as bandstop filters. In Fig. 13 the response of the SWG depending on the duty cycle is shown. The change in the duty cycle affects the range of the wavelengths that are blocked by the SWG. The lower the duty cycle, the wider the stopband gets, and the wavelength range shifts towards slightly longer wavelengths.

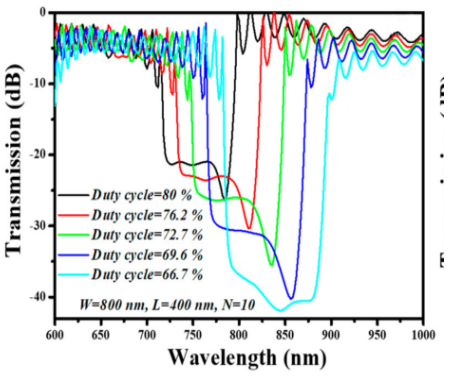


Fig. 13. Transmission spectrum of an SWG NIR filter depending on duty cycle [12].

In Fig. 14 the distribution of the electric field for two different wavelengths is shown, which depicts the operation of an SWG in a very understandable way. For the wavelength in the stopband, the electric field stops propagating through the SWG because of the Bragg reflection [46], and for a wavelength not included in the stopband, the electric field mapping is ordinary.

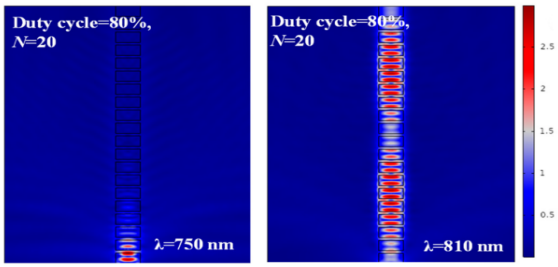


Fig. 14. Electric field mapping in the stopband region (left) and above the stopband region (right) [12].

3.2.2. FP-sensor

A slightly more interesting and more complex photonic structure containing an SWG is an FP-sensor [47]. The sensor is shown in Fig. 12. It is very similar to a regular SWG but instead of being uniform, it has a cavity of 2L size from the waveguide material in the middle of the structure. This results in the formation of a cavity sandwiched between two DBRs. Thus, it creates an FP-sensor. The response of an FP-sensor is shown in Fig. 15.

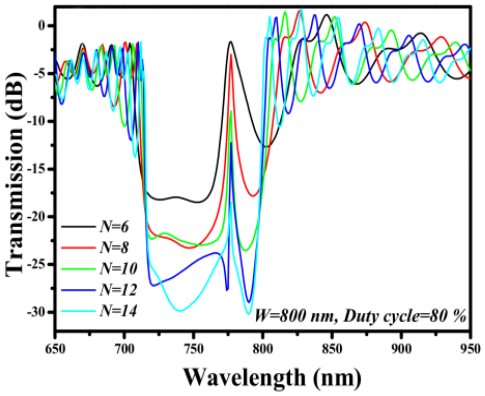


Fig. 15. Transmission spectrum of an SWG FP-sensor [12].

As can be observed, the response also contains a filtered band of light, but it has a specific wavelength within the stopband that is transmitted. It can be called the FP-wavelength. The FP-wavelength changes with the change of the duty cycle as shown in Fig. 16.

It is also affected by the change in the width of the waveguide *W*. The transmission characteristic depending on the change of width is shown in Fig. 17.

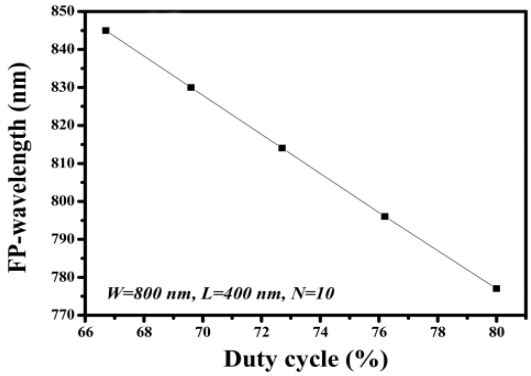


Fig. 16. FP-wavelength on duty cycle dependence [12].

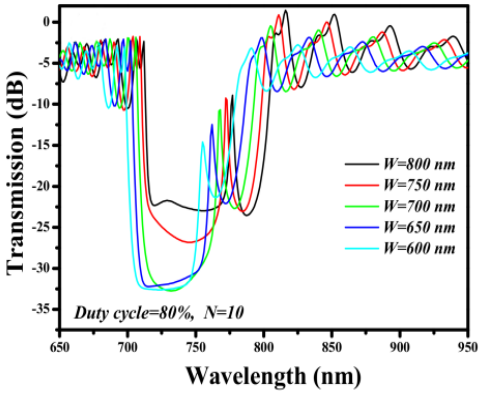


Fig. 17. FP-sensor operation dependance on the width of the waveguide [12].

A shift of the FP-wavelength is observed in the transmitted spectrum. Reading and comparing the shift of the FP-wavelength is the basic principle of operation of an SWG FP-sensor.

By changing the n_{amb} it is also possible to observe the shift of the FP-wavelength. Therefore, by disposing of a biochemical sample on the SWG, the sample can be characterized [48-50].

3.3. 1-D Photonic Crystals

A sensing structure that is somehow similar to an SWG is a 1-D photonic crystal (1-D PhC). The fundamental design structure of a 1-D PhC [51] is shown in Fig. 18. It is a simple waveguide that has a set of identical air holes evenly distributed along the waveguide.

The structure looks similar to an SWG and also has some basic variables that define it, being *a* – lattice constant, *r* – radius of air holes, *W* – width of the waveguide, and *N* – total number of air holes.

This basic 1-D PhC structure performs as a waveguide filter and by changing its variables, for example, the n_{wg}, *r*, or *a*, slight changes in the transmission spectrum are achievable which enables the use of this type of filter for different specific

applications [52]. Some transmission spectra for different variables are shown in Fig. 19. The most important visible effect of the manipulation of n_{wg}, r, or a is the shift of the stopband region. For lower n_{wg} and a the stopband appears for shorter wavelengths. However, for lower r the stopbands shifts towards longer wavelengths.

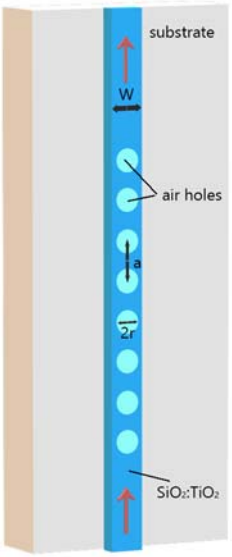


Fig. 18. Schematic 1-D PhC design with basic variables [11].

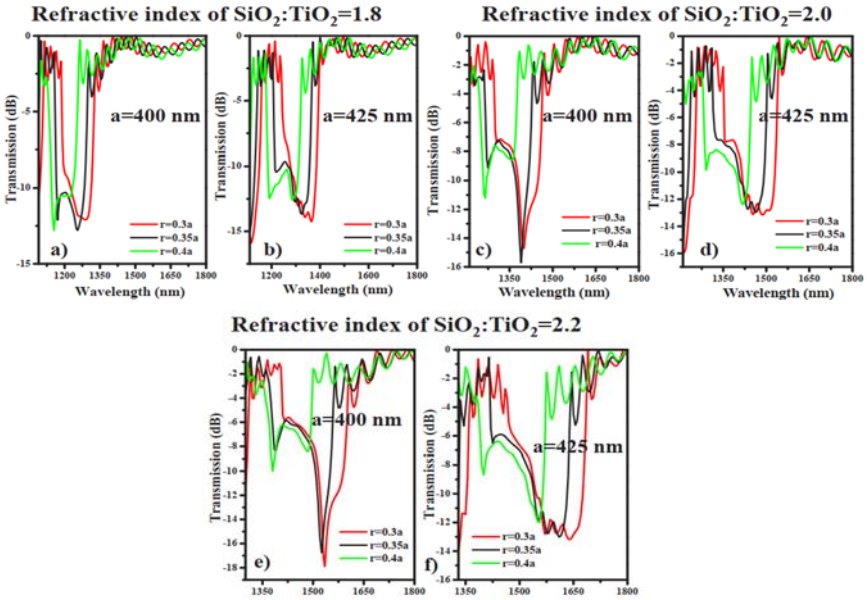


Fig. 19. 1-D PhC filter operation for different variables (a-f) [11].

Another approach to the application of a 1-D PhC would be to design an FP-sensor by adding a cavity in the middle of the periodic structure of the 1-D PhC. The presence of the cavity adds to the transmission spectrum a certain wavelength within the stopband region.

Fig. 20 shows the operation of a 1-D PhC. In a) the schematic diagram of the structure can be observed, showing the cavity and the air holes, b) the transmission spectrum with the visible peak wavelength in the stopband region and c) the Electric field pattern for the resonance peak.

The principle of operation of a 1-D PhC lies in the fact of changing the n_{amb} just as in the SWG FP-sensor. After changing ambient R.I, a shift in the transmission spectrum is visible, therefore the structure can operate as a sensor [53]. An exemplary response is shown in Fig. 21.

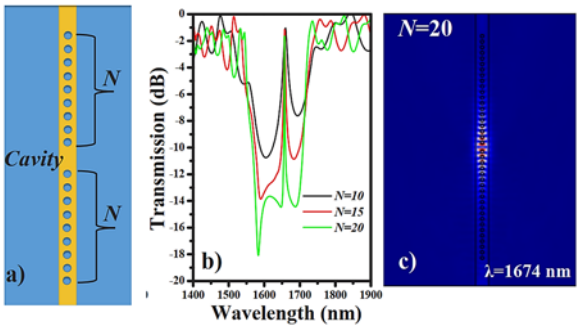


Fig. 20. a) Schematic FP-filter design, b) transmission spectrum, c) Electric field pattern in the resonance peak [11].

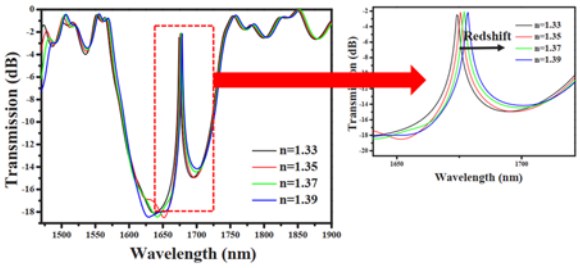


Fig. 21. Sensor operation of a 1-D PhC depending on the change of n_{amb} [11].

The sensitivity of the device can be easily calculated by the formula (1) previously mentioned.

The sensitivity of the device is also affected by the width of the waveguide as in Table 2. The device becomes more sensitive if narrower waveguides are used.

Table 2. Sensitivity dependence of the structure on waveguide width [11].

W (nm)	*N*	*a* (nm)	*r*	*S* (nm/RIU)
1000	20	425	0.3a	75
800	20	425	0.3a	100-108
600	20	425	0.3a	175

4. Conclusions

The numerical analyses and performances of different devices such as ring resonators, subwavelength grating, and photonic crystal waveguides show promising results and the SiO_2:TiO_2 platform can be utilized in several applications due to its exceptional physical, chemical, and optical properties. This platform has great potential when paired with nano-imprint lithography for the fabrication of integrated photonic devices in a single step. We believe that the studies presented in this mini-review will be useful for the researchers working on the topic of silica-titania sol-gel deposited via the dip-coating method and the realization of photonic devices based on this interesting platform.

Acknowledgements

This work was funded with the project "Hybrid sensor platforms of integrated photonic systems based on ceramic and polymer materials" carried out within the TEAM-NET program of the Foundation for Polish Science financed by the European Union under the European Regional Development Fund, POIR.04.04.00-00-14D6/18.

References

[1]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy. Optical elements silicon photonics, *Comput. Opt.,* 43, 2019, pp. 1079–1083.

[2]. M. A. Butt, N. L. Kazanskiy, Two-dimensional photonic crystal heterostructure for light steering and TM-polarization maintaining applications, *Laser Phys.*, 31, 2021, 036201.

[3]. D. X. Dai, S. P. Wang, Asymmetric directional couplers based on silicon nanophotonic waveguides and applications, *Front. Optoelectron*, 9, 2016, pp. 450–465.

[4]. L. Dan, D. S. Citrin, S. Hu, Compact high-performance polarization beam splitter based on a silicon photonic crystal heterojunction, *Opt. Mater.,* 109, 2020, 110256.

[5]. G. R. Chen, J. W. Choi, E. Sahin, D. T. Ng, D. H. Tan, On-chip 1 by 8 coarse wavelength division multiplexer and multi-wavelength source on ultra-silicon-rich nitride, *Opt. Express*, 27, 2019, pp. 23549–23557.

[6]. M. A. Butt, A. Kaźmierczak, C. Tyszkiewicz, P. Karasinski, R. Piramidowicz, Recent Advances in the Realization of a Low-Cost Integrated Photonic Platform Developed via a Sol-Gel Dip-Coating Method, in *Proceedings of the 8th International conference on Sensors and Electronics Instrumentation Advances (SEIA'2022)*, 21-23 September 2022, Greece, 2022, pp. 45-47.

[7]. P. Karasinski, C. Tyszkiewicz, R. Piramidowicz, and A. Kazmierczak, Development of integrated photonics based on SiO_2:TiO_2 sol-gel derived waveguide layers: state of the art, perspectives, prospective applications, *Proc. SPIE*, 11364, 2020, 1136414.

[8]. P. Karasinski, C. Tyszkiewicz, A. Domanowska, A. Michalewicz, and J. Mazur, Low loss, long time stable sol-gel derived silica-titania waveguide films, *Mater. Lett.,* 143, 2015, pp. 5–7.

[9]. A. Lukowiak, R. Dylewicz, S. Patela, W. Strek, and K. Maruszewski, Optical properties of SiO_2-TiO_2 thin film waveguides obtained by the sol-gel method and their applications for sensing purposes, *Opt. Mater.*, 27, 9, 2005, pp. 1501–1505.

[10]. Butt, M. A., Kazmierczak, A., Tyszkiewicz, C., Karasinski, P., Piramidowicz, R. Mode Sensitivity Exploration of Silica–Titania Waveguide for Refractive Index Sensing Applications, *Sensors*, 21, 2021, 7452.

[11]. Butt, M. A., Tyszkiewicz, C., Karasiński, P., Zięba, M., Hlushchenko, D., Baraniecki, T., Kaźmierczak, A., Piramidowicz, R., Guzik, M. and Bachmatiuk, A.,

Development of a low-cost silica-titania optical platform for integrated photonics applications, *Optics Express,* 30, 13, 2022, pp. 23678-23694.

[12]. Butt, M. A., Tyszkiewicz, C., Wojtasik, K., Karasinski, P., Ka ´zmierczak, A., Piramidowicz, R. Subwavelength Grating Waveguide Structures Proposed on the Low-Cost Silica–Titania Platform for Optical Filtering and Refractive Index Sensing Applications, *Int. J. Mol. Sci.*, 23, 2022, 6614.

[13]. Karasinski, P., Jaglarz, J., Mazur, J., Low loss silica-titania waveguide films, *Photonics Lett. Pol.,* Vol. 2, No. 1, 2010, pp. 37-39.

[14]. Kazmierczak, A., Slowikowski, M., Pavlov, K., Filipiak, M., Vervaeke, M., Tyszkiewicz, C., Ottevaere, H., Piramidowicz, R., Karasinski, P. Efficient, low-cost optical coupling mechanism for TiO_2-SiO_2 sol-gel derived slab waveguide surface grating coupler sensors, *Opt. Appl.* 4, 2020, pp. 539–549.

[15]. Innocenzi, P., Martucci, A., Guglielmi, M., Armelao, L., Pelli, S., Righini, G., Battaglin, G. Optical and surface properties of inorganic and hybrid organic–inorganic silica–titania sol–gel planar waveguides, *J. Non-Cryst. Solids,* 259, 1999, pp. 182–190.

[16]. Karasinski, P., Tyszkiewicz, C., Piramidowicz, R., Kazmierczak, A., Development of integrated photonics based on SiO_2:TiO_2 sol-gel derived waveguide layers:state of the art, perspectives, prospective applications, *Proceedings of the SPIE 11364, Integrated Photonics Platforms: Fundamental Research, Manufacturing and Applications, SPIE Photonics Europe,* Strasbourg, France, 29 March–2 April 2020, 1136414.

[17]. Dezfuli, S. M., Sabzi, M. Deposition of self-healing thin films by the sol–gel method: A review of layer-deposition mechanisms and activation of self-healing mechanisms, *Appl. Phys*. *A*, 125, 2019, p. 557.

[18]. Inoue, H., Iwamoto, T., Horie, K., Makishima, A., Ikemoto, M. Preparation and properties of sol-gel thin films with porphins, *J. Opt. Soc. Am.,* B, 9, 1992, pp. 816–818.

[19]. M. Butt, E. S. Kozlova, S. N. Khonina, and R. V. Skidanov, Optical planar waveguide sensor based on (Yb,Nb): RTP/RTP(001) system for the estimation of metal coated cells, *CEUR Workshop Proc.,* 1638, 2016, pp. 16–23.

[20]. X. Orignac, D. Barbier, X. M. Du, R. M. Almeida, O. McCarthy, and E. Yeatman, Sol-gel silica/titania-on-silicon Er/Yb-doped waveguides for optical amplification at 1.5 μm, *Opt. Mater.,* 12, 1, 1999, pp. 1–18.

[21]. A. I. Gomez-Varela, Y. Castro, A. Duran, P. A. Beule, M. T. Flores-Arias, and C. Bao-Varela, Synthesis and characterization of erbium-doped SiO_2-TiO_2 thin films prepared by sol-gel and dip-coating techniques onto commercial glass substrates as a route for obtaining active Gradient-index materials, *Thin Solid Films,* 583, 2015, pp. 115–121.

[22]. Y. Sorek, R. Reisfeld, and A. M. Weiss, Effect of composition and morphology on the spectral properties and stability of dyes doped in a sol-gel glass waveguide, *Chem. Phys. Lett.*, 244, 5-6, 1995, pp. 371–378.

[23]. A. Yimit, K. Itoh, and M. Murabayashi, Detection of ammonia in the ppt range based on a composite optical waveguide pH sensor, *Sens. Actuators, B,* 88, 3, 2003, pp. 239–245.

[24]. A. D'Orazio, M. D. Sario, L. Mescia, V. Petruzzelli, and F. Prudenzano, Design of Er3+Yb3+ doped silica/titania planar waveguide amplifier, in *Proceedings of 2003 5th International Conference on Transparent Optical Networks*, Warsaw, Poland, 2003.

[25]. P. Karasinski, Sol-gel derived optical waveguide films for planar sensors with phase modulation, *Opt. Appl.*, 34, 4, 2004, pp. 467–475.

[26]. Karasinski, P., Tyszkiewicz, C., Rogozinski, R., Jaglarz, J., Mazur, J. Optical rib waveguides based on sol-gel derived silica-titania films, *Thin. Solid Films,* 519, 2011, pp. 5544–5551.

[27]. Khlyustova, A., Cheng, Y., Yang, R. Vapor-deposited functional polymer thin films in biological applications, *J. Mater. Chem. B*, 8, 2020, pp. 6588–6609.

[28]. Royer, F., Jamon, D., Rousseau, J. J., Roux, H., Zins, D., Cabuil, V. Magneto-optical nanoparticle-doped silica-titania planar waveguides, *Appl. Phys. Lett.,* 86, 2005, 011107.

[29]. Roy, R. D., Sil, D., Jana, S., Biswas, P. K., Bhadra, S. K., Experimental study of perfectly patterned silica-titania optical waveguide, *Photonic Sensors*, 2, 2012, pp. 81–91.

[30]. Almeida, R., Marques, A., Pelli, S., Righini, G., Chiasera, A., Mattarelli, M., Spectroscopic assessment of silica-titania and silica-hafnia planar waveguides, *Philos. Mag.,* 84, 2006, pp. 1659–1666.

[31]. G. Brusatin, M. Guglielmi, P. Innocenzi, A. Martucci, G. Battaglin, S. Pelli, G. Righini, Microstructural and optical properties of sol-gel silica-titania waveguides, *Journal of Non-Crystalline Solids*, Vol. 220, Issues 2–3, 1997, pp. 202-209.

[32]. Leszek Golonka, Pawel Bembnowicz, Dominik Jurków, Karol Malecha, Henryk Roguszczak, Rafał Tadaszak, Ceramic Microsystems, in *Proceedings of the International Microelectronics Assembly and Packaging Society Conference,* Poland Chapter, 2010.

[33]. Giancarlo Righini, Stefano Pelli, Sol-Gel Glass Waveguides. *Journal of Sol-Gel Science and Technology*, 8. 1997, pp. 991-997.

[34]. Vasconcelos, Helena, Optical Waveguides Based on Sol-Gel Coatings, in Electromagnetic Propagation and Waveguides in Photonics and Microwave Engineering (Patrick Steglich, Ed.), *Intechopen*, 2020.

[35]. Rafał Tadaszak, Anna Lukowiak, Leszek Golonka, Sergiusz Patela, Hybrid sol-gel-glaze planar optical waveguides on LTCC substrate - Preliminary works, *Optica Applicata*, 41, 2011.

[36]. K. Nowak et al., Sol-gel-based optical waveguides on LTCC substrates, in *Proceedings of the 31st International Spring Seminar on Electronics Technology,* 2008, pp. 518-522.

[37]. Muhammad A. Butt, Thin-Film Coating Methods: A Successful Marriage of High-Quality and Cost-Effectiveness - A Brief Exploration, *Coatings*, 12, 2022, 1115.

[38]. N. Golshani et al., Low-loss, low-temperature PVD SiN waveguides, in *Proceedings of the 17th IEEE International Conference on Group IV Photonics (GFP),* 2021, pp. 1-2.

[39]. Mark P. Hiscocks, Christopher J. Kaalund, François Ladouceur, Shane T. Huntington, Brant C. Gibson, Steven Trpkovski, David Simpson, Eric Ampem-Lassen, Steven Prawer, James E. Butler, Reactive ion etching of waveguide structures in diamond, *Diamond and Related Materials,* Vol. 17, Issue 11, 2008, pp. 1831-1834.

[40]. Axel Grosse et al, Deep wet etching of fused silica glass for hollow capillary optical leaky waveguides in

microfluidic devices, *J. Micromech. Microeng*, 11, 2001, 257.
[41]. Nikolay L. Kazanskiy, Svetlana N. Khonina, Muhammad A. Butt, Advancement in Silicon Integrated Photonics Technologies for Sensing Applications in Near-Infrared and Mid-Infrared Region: A Review, *Photonics,* 9, 2022, 331.
[42]. Wang, Y., Gao, B., Zhang, K., Yuan, K., Wan, Y., Xie, Z., Xu, X., Zhang, H., Song, Q., Yao, L., et al. Refractive Index Sensor Based on Leaky Resonant Scattering of Single Semiconductor Nanowire, *ACS Photon.*, 4, 2017, pp. 688–694.
[43]. Carlborg, C. F., Gylfason, K. B., Kaźmierczak, A., Dortu, F., Polo, M. J. B., Catala, A. M., Kresbach, G., Sohlström, H., Moh, T., Vivien, L., et al., A packaged optical slot-waveguide ring resonator sensor array for multiplex label-free assays in labs-on-chips, *Lab Chip,* 10, 2010, pp. 281–290.
[44]. Junjia Wang, Ivan Glesk, and Lawrence R. Chen, Subwavelength grating filtering devices, *Opt. Express*, 22, 2014, pp. 15335-15345.
[45]. Wang, J., Glesk, I. and Chen, L., Subwavelength grating Bragg grating filters in silicon-on-insulator, *Electron. Lett.,* 51, 2015, pp. 712-714.
[46]. Junjia Wang, Ivan Glesk, Lawrence R. Chen, Subwavelength grating devices in silicon photonics, *Science Bulletin*, Vol. 61, Issue 11, 2016, pp. 879-888.
[47]. N. L. Kazanskiy; M. A. Butt; S. N. Khonina, Silicon photonic devices realized on refractive index engineered subwavelength grating waveguides-A review, *Optics & Laser Technology*, 138, 2021, 106863.
[48]. N. Kazanskiy, S. Khonina, M. Butt, A. Kaźmierczak, R. Piramidowicz, State-of-the-Art Optical Devices for Biomedical Sensing Applications—A Review, *Electronics,* 10, 2021, 973.
[49]. M. A. Butt, S. N. Khonina & N. L. Kazanskiy, Numerical analysis of a miniaturized design of a Fabry–Perot resonator based on silicon strip and slot waveguides for bio-sensing applications, *Journal of Modern Optics,* 66, 11, 2019, pp. 1172-1178.
[50]. Wei Liang, Yanyi Huang, Yong Xu, Reginald K. Lee, and Amnon Yariv, Highly sensitive fiber Bragg grating refractive index sensors, *Appl. Phys. Lett.*, 86, 2005, 151122.
[51]. Nikolai Lvovich Kazanskiy, Muhammad Ali Butt, One-dimensional photonic crystal waveguide based on SOI platform for transverse magnetic polarization-maintaining devices, *Phot. Lett. Pol.,* Vol. 12, No. 3, 2020, pp. 85-87.
[52]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy, Recent advances in photonic crystal optical devices: A review, *Optics & Laser Technology*, 142, 2021, 107265.
[53]. Muhammad A. Butt, Nikolay L. Kazanskiy, Svetlana N. Khonina, Advances in Waveguide Bragg Grating Structures, Platforms, and Applications: An Up-to-Date Appraisal, *Biosensors,* 12, 2022, 497.

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 93-98

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2022
http://www.sensorsportal.com

Modeling Branched Electrical Circuits with Different Types of Conductivity

[1] Nikolai A. PARFENTEV and [2] Natalia A. PARFENTYEVA

[1] All-Russian Research Institute for Optical and Physical Measurements (VNIIOFI), Moscow, Russia
[2] Moscow State University of Civil Engineering (MGSU) National Research University, Moscow, Russia
Tel.: +7 916 858 8010, + 7 925 0544560
E-mail: nikolaiparfentiev41@gmail.com, nparfentyeva@gmail.com

Received: 31 August 2022 /Accepted: 3 October 2022 /Published: 31 October 2022

Abstract: As a model of a developed circuit with different types of conductivity, the principles of operation of bipolar transistors are considered n-p-n and p-n-p. It is shown that an abrupt increase in the input voltage upon the appearance of a collector voltage is associated with disturbances in the energy balance, when the electron energy spent on the metal-emitter barrier transition cannot be compensated for in the base-metal transition. The balance is disturbed due to the different energy spectra of electrons in base and collector.

Fundamental difference in the mechanisms of increasing (in modulus) the base-emitter input voltage with the appearance of a collector potential for transistors of various types are revealed. In is shown that this effect is associated with the sorting of electrons having different average energies in the n-p-n transistor and the mixing of electrons in the emitter of the p-n-p transistor. The concept adopted by us can be useful in the creation of semiconductor devices. It includes not only the properties of semiconductors, but also the properties of metals, with the help of which it is connected to the circuit.

Based on the formulated conditions for electrical circuits with compensated contact potential differences [11], an analysis of the phenomena in a bipolar transistor was carried out.

The contact differences of potentials in a branched chain can be neglected.

1) A relatively small current value at which the Peltier and Zeeman effects can be neglected (intense temperature exchange with the environment).

2) Constant current in all elements of loop current.

3) The conservation of the value of the average energy of charge carriers. As an example, detailed analysis of contact phenomena in a bipolar transistor is carried out. The fundamental difference between the modes of operation of the base-emitter junction in the n-p-n transistor is shown: at zero and non-zero values of the collector voltage.

Keywords: Branched chain, Contact phenomena, Bipolar transistor, Semiconductor.

1. Introduction

A real electrical circuit usually contains components made from different materials. In classical electrical engineering, this circumstance is usually not given much attention [1-10]. In real conditions, contact potential differences are inevitably present in a closed circuit, but they completely compensate each other, allowing the physical laws, discovered and formulated several centuries ago, to be used quite reasonably for their calculation. On the other hand, the Peltier

and Zeeman effects (associated with contact phenomena) successfully work in technology, allowing both to measure the temperature and change its value in the required direction. The physical nature of the contact potential difference provides an opportunity to both increase the energy of carriers falling into the zone of its action, and to lower it. Naturally, specific changes in energy depend on the polarity of the potential difference, the direction of the carrier velocity, and the polarity of the carriers themselves. These phenomena are of particular importance in semiconductor technology, which has penetrated everywhere and has changed practical life, in which they occupy a large and important place. However, contact phenomena are not always correctly explained, since the laws that determine their influence are still not clearly defined.

The purpose of this work is to analyze contact phenomena in classical circuits with different types of conductivity – bipolar transistors of type n-p-n and p-n-p. At the same time, the mechanisms of formation of input characteristics, fundamentally different for transistors of different types, are analyzed. The in-depth understanding, we offer the physical processes in transistors can enter the textbook and reference books on electronics and electrical engineering. It may be necessary and useful material for future engineers and researchers.

2. An Example of a Branched Circuit with Contact Potential Differences – Bipolar Transistor n-p-n

A classic example of where the contact difference manifests itself in a paradoxical way is the bipolar transistor, about which everything from school textbooks to serious fundamental books on electronics seems to be written [1]. The first thing we learn from these sources is that a transistor is a device whose main function is to amplify or convert a weak signal from a measuring sensor into a voltage or current supplied to the device.

The most common way to turn on a transistor is with a common emitter circuit for maximum power gain. In this case, both the input current and the input voltage increase at the output of the circuit. For practical calculations, equivalent circuits of real transistors are used, in which there are models of the input base circuit and the output-collector circuit. The main element of this circuit is the current source in the collector circuit, and the magnitude of this current is linearly related to the magnitude of the current in this circuit, and the coefficient determining this relationship is much greater than unity. However, some unexpected phenomenon occurs in the circuit – with the help of the collector current in bipolar transistors of the n-p-n type, it is necessary to increase the voltage supplied to the base. Moreover, this phenomenon, noticeable on real input (basic) characteristics, is of an abrupt nature – it manifests itself immediately after a slight increase in the collector voltage (current). In classical equivalent circuits, this effect is described by the feedback coefficient between the collector voltage and the base voltage.

Fig. 1 shows the equivalent transistor circuitry for small signals. In the input circuit is placed a source of voltage proportional to the voltage of the output circuit.

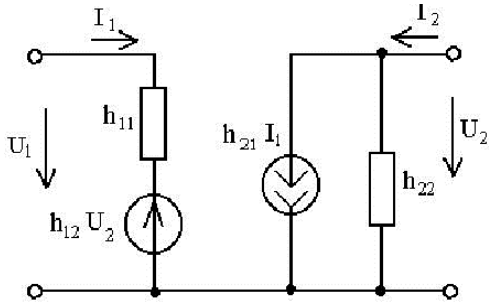


Fig. 1. Equivalent transistor circuitry.

The historically equivalent circuitry of a transistor (a linear model of a nonlinear device) emerged at a time when computer simulations had not yet gained momentum. Physical models developed on the basis of solid-state physics were used to describe the operation of a transistor. The band theory contributed to a deep understanding of the processes that determine the mechanisms of the transistor.

However, the linear model of the transistor substitution scheme, created for practical purposes, turned out to be imperfect and did not take into account the peculiarities of the movement of electrons in chains with different types of conductivity.

The interpretation of the increase in the input voltage of transistors as a consequence of the influence of the collector voltage is an obvious logical fallacy. Indeed, both the base and collector voltages are voltages of the same sign. An increase in the base voltage causes an increase in current, the part of the collector voltage appears can only increase the amount of current. In fact, in order to achieve the same current value, it is necessary to increase the supplied base voltage.

The purpose of this article is a physical interpretation of this phenomenon, showing the obvious fallacy of the accepted interpretation. The longevity of this misconception is probably due to the fact that the effect itself manifests itself in the "inoperative" region of the transistor, when the required current gain (the main concern of the transistor designer) is far from the maximum value. LTSpice, an excellent electronic component simulator created to advertise products from the famous Analog Device, fails when trying to describe the operation of a transistor in this area.

Let's analysis of contact phenomena in a bipolar transistor n-p-n.

A fruitful approach to the analysis of a bipolar transistor is the n-p (p-n) junction model in form of a charged capacitor with movable plates, the distance

between which can be controlled by changing the voltage applied to the junction. When a reverse voltage is applied, the distance between the plates increases – this is how a varistor works (a capacitor with a variable electric capacitance value). When a direct voltage is applied, the distance between the charged layers decreases so much that a current begins to flow between them and the actually controlled capacitor turn into a controlled resistor (which a variable resistance value).

Unlike a capacitor, where charge carriers are concentrated within the rigid boundaries of metal linings, the charge layers of the semiconductor junction can change their thickness. In particular, the thickness of the charge layer increases with temperature, which explains the dependence of the voltage transition characteristics on temperature.

Now consider in detail what happens in a real transistor switching circuit. In practice, a device whose basic principles of operation are determined by semiconductors is inevitably connected to the measuring circuit using metal contacts. Naturally, in each such connection, a contact potential difference arises. In the diagram shown in Fig. 2, they are represented by voltage sources in the form of circles with a polarity designation – the method adopted in the LTSpice program.

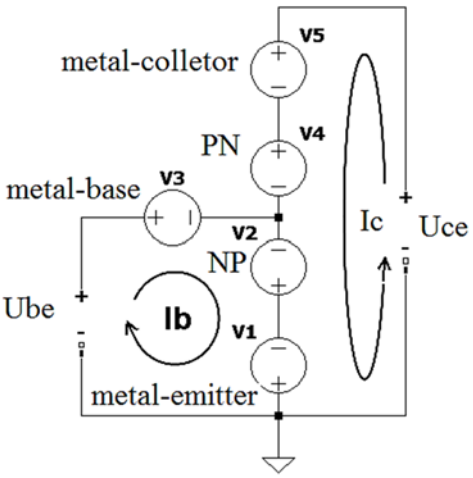


Fig. 2. Connection diagram with a common emitter of a n-p-n bipolar transistor. V1 – contact potential difference metal-emitter; V2 – np transition emitter-base; V3 – contact potential difference metal-base; V4 – p-n base-collector junction; V5 – potential difference collector metal.

Blocking and anti-blocking layers may appear at the metal-semiconductor contact. Then such a contact, when an electric current passes through it, will be nonlinear, since its properties change depending on the applied voltage. The existence of a barrier layer at the metal-semiconductor interface and the possibility of controlling its thickness with the help of an external voltage determines the rectifying properties of the contact.

This property, as is well known, is used for point diodes, used, for example, in integrated circuits. The input characteristics of the transistor necessarily include the original curve of the dependence of the current on the input voltage in the absence of the collector voltage, and, consequently, the collector current. At zero collector current, the contact differences of the metal connections of the transistor fully compensate each other, without exerting any influence on the formation of the input voltage, which is completely dependent on the junction voltage np. Naturally, this characteristic is not linear and corresponds to the current voltage characteristic of a semiconductor diode. At first glance, the appearance of the collector current does not change the situation in any way: the contact differences remain the same, and the weak base current, being part of the total emitter current flowing to the base, should not be influenced by the contact differences on the transistor electrodes. In this case, in both cases, the contact potential difference V1, which arises during the metal-semiconductor connection due to the different electron concentration in the metal and semiconductor, is the first barrier for electrons.

As a result, electrons penetrate into the emitter, the average energy of which is less than the average energy of electrons in the metal contact. Let us consider in more detail what happens in the transitions of the bipolar transistor itself in Fig. 3.

Electrons entering the base layer through the emitter move in crossed electric fields. The relatively weak base voltage (Eb) forces them to move along the thin base layer, while the high collector voltage can drag them into the collector junction zone.

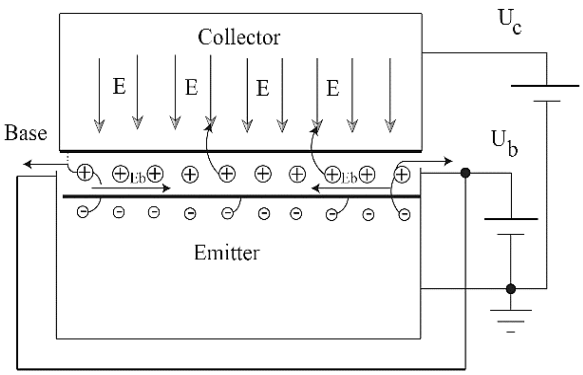


Fig. 3. Scheme of a bipolar transistor n-p-n (circuit with a common emitter).

For the most part, electrons with a relatively large kinetic energy get there. Such electrons have a high mobility and are likely are captured by the collector, avoiding the possibility of getting into the base contact. As a result, out of the total number of electrons that passed into the collector through the metal-emitter contact junction, most of them have a higher energy than the electrons entering the base. As a result, the energy spectrum of electrons going to the base differs from the energy spectrum of electrons reaching the collector. The average energy of these electrons is less than the average energy of the electrons entering the

emitter. The contact difference V3, which increases their energy, is nevertheless not able to restore its value to the value that they had in the metal in contact with the emitter. Thus, the base-collector branching circuit is a device that sorts electrons by the amount of their energy. The base current does not change, it coincides with the fraction of the current flowing from the metal to the emitter, but the energy spectrum of electrons changes. Thus, when the collector current appears in the base circuit, irreplaceable energy losses occur. The kinetic energy of electrons that have lost energy during the transition of the metal-emitter contact cannot be restored to their original value when passing through the base-metal contact. Thus, an abrupt change in the input voltage with the appearance of even a small collector voltage, or rather the collector current, turns out to be associated with the contact potential difference of the contact base. Based on the above, it is possible to formulate the conditions under which the contact potential differences do not affect the current in the branched circuit:

1) A relatively small current value at which the Peltier and Zeeman effects can be neglected.

2) Equal current value in all circuit elements.

3) The invariability of the mean value of energy spectrum of carriers.

In the base circuit of the transistor, to make up for energy losses when a collector current appears and a concomitant change in the energy spectrum of electrons, it is necessary to increase the input voltage of the transistor.

Fig. 4 shows the dependence of the ratio of the average energy Em of electrons in a metal to the minimum energy of free electrons E0, depending on the value of the minimum energy of the conduction band, expressed in meV, calculated at normal temperature. As follows from the graph, the presence of a contact potential difference that decreases the average energy of electrons is equivalent to an increase in their minimum energy and, consequently, a greater inhomogeneity of the energy spectrum of electrons.

A small increase in temperature will not significantly affect the nature of this dependence, although the conductivity of semiconductors increases.

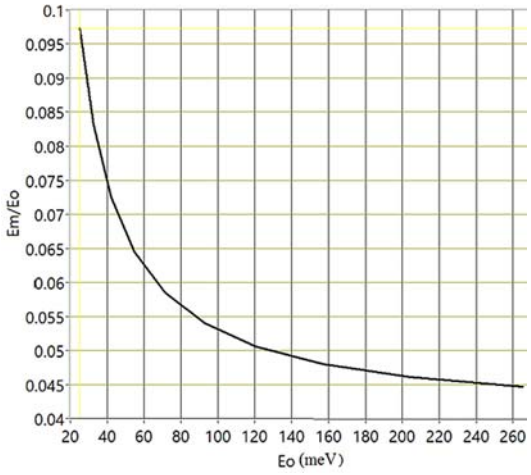


Fig. 4. Dependence of the ratio of the average energy of electrons in a metal to the minimum energy on the minimum energy at normal temperature.

A feature of the type of input characteristics of the n-p-n transistor is the high constancy of the amplitude of the input voltage jump. Fig. 5 shows a sample of a typical input characteristic (n-p-n) of an Analog Device MAT-02 transistor.

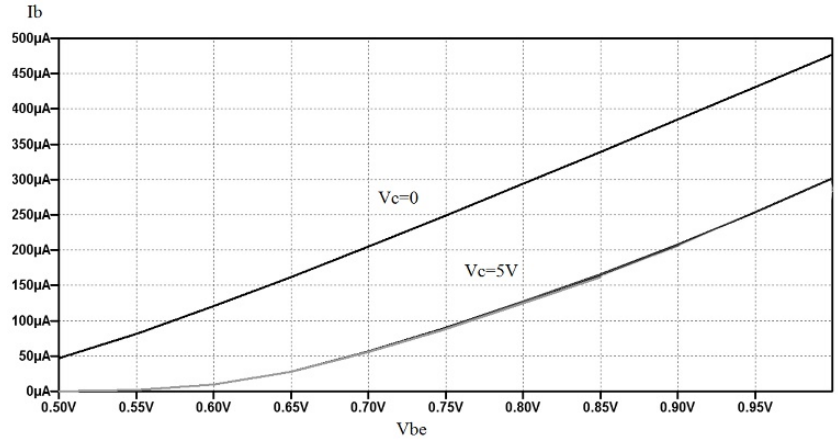


Fig. 5. Input characteristics of transistor MAT-02.

It is easy to see that the same value of the input current is achieved by increasing the input voltage by an almost constant increase in the input voltage by an amount equal to 0.2 volts. The constancy of this value is an additional argument in the erroneous interpretation of the input voltage jump from the appearance of the collector voltage. Of course, the accepted interpretation of the effect creates additional difficulties for calculations and design of transistors, since apart from the properties of the transistor itself, the properties of the metal used for contact with the terminals of the transistor must be taken into account. Studying the

characteristics of p-n-p transistors, one can notice manifestations of the opposite effect – when the collector current appears, the absolute value of the input voltage decreases. Let's move on to the analysis of the principle of operation of p-n-p transistors, which differ in a different mechanism, causing an increase (in absolute value) of the input voltage when the collector current appears.

3. The Mechanism of Formation of Input Characteristics of the p-n-p Transistor

The given characteristics of p-n-p transistors are similar (up to sign) with the characteristics of n-p-n transistors. The input characteristics of this type transistors exhibit the effect of increasing the input voltage (modulo) as collector voltage increases. This phenomenon cannot be interpreted within the collector and base circuit feedback model.

Consider in detail the p-n-p transistor circuit (Fig. 6).

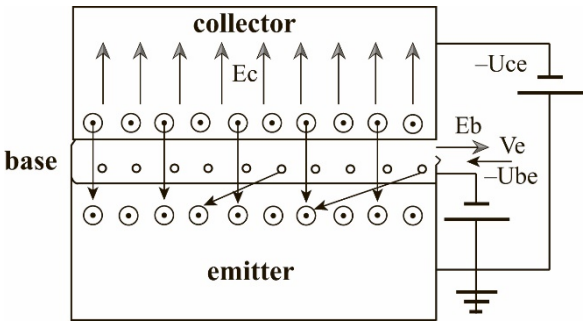


Fig. 6. Scheme of a bipolar transistor p-n-p (circuit with a common emitter).

Similar to the n-p-n type transistor, this transistor is a device where crossed electric fields are present. When a sufficient negative voltage is applied, the layers of electrons in the base and holes in the emitter converge so much that the electrons of the base begin to pass into the emitter. Note that the energy of electrons coming from the base has a relatively high value, but in the absence of collector voltage, they come to the plus contact of the base voltage source and all contact potential differences compensate for each other.

When a collector voltage appears, in addition to the base electrons, electrons torn out of the collector begin to flow into the emitter with a relatively small energy characteristic of the process of moving holes in the collector. The energy of the base electrons is higher than the energy of the collector electrons because only the potential difference between the base and the emitter leads to the appearance of current in the transistor, and the voltage on the collector is not enough for it to occur. Then, both streams of electrons are mixed in the emitter and through the common ground wire move to the positive contacts of voltage sources. Due to the significant predominance of the number of electrons from the collector over the number of electrons from the base, the average energy of the electrons is lower than the energy of the electrons of the base.

As a result, the energy balance in the closed base circuit is disturbed and to achieve the same value of the current in the base, it is necessary to increase the voltage of the base source.

The considered mechanisms for increasing the base voltage are radically different in bipolar transistors of different types: in the n-p-n transistor, the cause of the phenomenon is the sorting of electrons of different energies between the base and the collector, in the p-n-p transistor the cause is the mixing of electrons with different average energies coming from the base and collector.

4. Analyze a Schema with a Common Base

A common emitter transistor switching circuit was chosen for analysis because the input characteristics of the common-base circuit do not contain a base current measurement.

The input characteristics of a common base transistor n-p-n circuit show a decrease in voltage between the emitter and the base when a collector voltage is applied. At first glance, this circumstance contradicts the provisions of the proposed model.

However, in this case, only the emitter current is measured, which includes the sum of the base and collector currents. With a high current gain, the base current is significantly less than the collector current. Therefore, to obtain the same value of the emitter current, it is sufficient to apply a smaller value of the voltage between the emitter and the base. Naturally, the input characteristics of a p-n-p transistor look similar.

The general emitter circuit provides base current measurements that reflect the physical processes occurring in circuits with different types of conductivity. In the future, we plan to develop a complete model based on experimental data on the study of transistors of different types and inclusion schemes.

This article mainly contains qualitative descriptions of the processes in bipolar transistors. Its results were reported at the MICDAT 2022 conference [11]. In the future, it is planned to develop a model with correct proof of the relationship between changes in the input voltage and contact differences in metal-semiconductor potentials that occur when connecting a transistor.

5. Conclusions

1. Conditions have been formulated under which contact potential differences in branched electrical circuits can be neglected.

2. The error of the classical interpretation of the input voltage jump when the collector current appears is shown.

3. A physical interpretation of the effect of a step change in the input voltage of bipolar transistors is carried out.

4. The mechanisms of the input voltage jump are explained: in n-p-n – sorting of electrons, in p-n-p – mixing of electron flows with different average energy.

5. The ways of analyzing the operating mode of bipolar transistors are discussed, which make it possible to eliminate the shortcomings of the existing calculation models.

6. Taking into account the influence of contact phenomena in the transistor connection circuits can lead to an improvement in the operating parameters of the transistors.

7. Ideas about physical processes in transistors, which are included as an important section in textbooks and reference books on electronics and electrical engineering, will serve as a useful material for future engineers and researchers.

References

[1]. A. F. Ioffe, Semiconductor Thermoelements and Thermoelectric Cooling, *Infosearch*, London, 1957.

[2]. Jump up to: a b c Paul Horowitz and Winfield Hill, The Art of Electronics (2nd ed.), *Cambridge University Press*, 1989.

[3]. Juin Jei Liou and Jiann S. Yuan, Semiconductor Device Physics and Simulation, *Springer*, 1998.

[4]. Paolo Antognetti and Giuseppe Massobrio, Semiconductor Device Modeling with Spice, *McGraw–Hill Professional*, 1993.

[5]. Peter Ashburn, SiGe Heterojunction Bipolar Transistors, *Wiley*, New York, 2003, Chapter 10.

[6]. Paul Horowitz and Winfield Hill, The Art of Electronics (2nd ed.), *Cambridge University Press*, 1989, pp. 62–66.

[7]. 1947: Invention of the Point-Contact Transistor - The Silicon Engine – Computer History Museum, Retrieved August 10, 2016. http://www.computerhistory.org/semiconductor/timeline/1947-invention.html

[8]. Third case study – the solid state advent Archived September 27, 2007, at the Way back Machine (PDF). http://hm-treasury.gov.uk/media/B/C/queen_mary_ip_research_institute_p5_043_762kb.pdf

[9]. High Speed Switching Transistor Handbook (2nd ed.). *Motorola*, 1963.

[10]. Maupin, J. T., The tetrode power transistor, *IRE Transactions on Electron Devices*, 4, 1, 1957, pp. 1-5.

[11]. Kaden, H. E., Physical Principles of the Transistor, in Transistors Applied, *Palgrave*, London, 1965.

[12]. Gaydamachenko V. R., Beloglazkina E. K., Petrov R. A., Dagesyan S. A., Sapkov I. V., Soldatov E. S., Single-electron Transistor based on Single atom Charge Center, M. V. Lomonosov Moscow State University. Type: paper in conference proceedings, 2017.

[13]. John Bardeen, Walter H. Brattain, Physical Principles Involved in Transistor Action Institutions, *Physical Review*, Vol. 75, Issue 8, 1949, pp 1208 – 1225.

[14]. Irving Gottlieb, P. E., Fundamentals of transistor physics, *John F. Rider Publisher*, 1960, p. 41-52.

[15]. N. A. Parfentev, N. A. Parfentyeva, Physical Features of Branched Circuits with Contact Potential Differences, in *Proceedings of the 4th International Conference on Microelectronic Devices and Technologies (MicDAT '2022)*, Corfu, Greece, 21-23 September 2022, pp. 15-17.

Sensors & Transducers, Vol. 259, Issue 5, October 2022, pp. 99-108

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Towards Selective and Sensitive Detection of Carbon Monoxide with CuO/ZnO Heterojunction Nanocomposite Prepared by an Organometallic Approach

[1, *] Justyna Jońca, [2] Katia Fajerwerg, [2] Myrtil L. Kahn, [3] Philippe Menini, [1] Izabela Sówka and [2, 4, *] Pierre Fau

[1] Wroclaw University of Science and Technology, Faculty of Environmental Engineering, Department of Environment Protection Engineering, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
[2] Laboratoire de Chimie de Coordination, LCC-CNRS, Université de Toulouse, 205 Route de Narbonne, BP 44099, 31077 Toulouse cedex 4, France
[3] Laboratoire d'Analyse et d'Architecture des Systèmes, Université de Toulouse, CNRS, UPS, 7 Av. du Colonel Roche, 31400 Toulouse, France
[4] Laboratoire de Physique et Chimie de Nano-Objets, LPCNO-INSA, Université de Toulouse, 135 Avenue de Rangueil – INSA, 31077 Toulouse Cedex 4, France
[*] E-mail: justyna.jonca@pwr.edu.pl, pfau@insa-toulouse.fr

Received: 5 September 2022 Accepted: 10 October 2022 Published: 31 October 2022

Abstract: Nanometer size p-n heterojunction has been created from CuO and ZnO anisotropic nanoparticles prepared by a one-pot organometallic approach. The method is based on the hydrolysis or oxidation of an adequate metal-organic precursors in pure octylamine. The CuO and ZnO nanostructures were dispersed in ethanol and then, mixed at different mass ratios, *i.e.* CuO(75%)/ZnO(25%), CuO(50%)/ZnO(50%) and CuO(25%)/ZnO(75%). Finally, the CuO and ZnO suspensions and their mixtures were deposited on miniaturized gas sensors substrates by an ink-jet printing method and heated up gradually to 550 °C in ambient air. Then, the as-prepared sensors have been exposed to CO (100 ppm), C_3H_8 (100 ppm) and NH_3 (5 ppm) at different working temperatures (from 75 °C to 400 °C) and under 50 % of relative humidity (RH). Among all prepared sensors the one based on the mixture of CuO (75 %) and ZnO (25 %) presents a very sensitive and selective response to CO. Indeed, at the operating temperature of 165 °C, a high sensitivity towards CO was obtained (S_{CO}=624 %). In these conditions, the sensor exhibited low sensitivity to other tested gases ($K_{CO/C3H8}$= 14.5 and $K_{CO/NH3}$= 26) but its response time was quite long (t_{90}=2.3 min) and the recovery was very sluggish (t_{10}> 20 min). Therefore, it was better to increase the working temperature up to 300 °C. Although the sensitivity and selectivity towards CO worsened (S_{CO}=177%, $K_{CO/C3H8}$= 4.7 and $K_{CO/NH3}$= 9.8) the response/recovery time decreased to 50 s/4.5 min. The gas sensing performances of the CuO(75%)/ZnO(25%) composite was attributed to both high surface to volume ratio of the prepared nanostructures and the p-n heterojunction established between the CuO and ZnO nanoparticles.

Keywords: Organometallic approach, CuO/ZnO nanocomposite, Heterojunction, MOS gas sensors.

1. Introduction

The use of miniaturized gas sensors is of increasing interest in areas such as environmental monitoring, safety, medical diagnostics and agriculture [1]. Metal oxide semiconductor gas sensors (MOS) are the most popular choice for these applications due to their numerous advantages, *i.e.* long lifetime, low cost,

small size and short response time. However, these devices have their limitations as well. They exhibit baseline drift, sensor poisoning, poor selectivity and low manufacturing reproducibility [2, 3]. Although much has been done in order to deal with these problems, the improvement of MOS sensors continues to attract researchers' attention.

The performance of MOS sensors is influenced by various parameters including: design of the sensing platform (*e.g.* geometry of electrodes and heating element), device operating temperature and physical and chemical properties of the sensing materials (*e.g.* grain size, defects density and presence of oxygen vacancies). The improvement of gas sensing properties of metal oxide semiconductors can be accomplished by designing them at the nanoscale [4]. Therefore, the synthesis of nanoparticles has attracted considerable interest. In comparison with high temperature vapor phase deposition method, solution-based chemistry protocols can be conducted at mild temperature and offer additional advantages such as straightforward processing, low cost, and eases of scale up. Among them, the organometallic approach, developed at the Laboratoire de Chimie de Coordination (LCC-CNRS), leads to well-controlled nanostructures in terms of size dispersion, chemical composition, surface chemistry properties, shape or organization [5]. Indeed, hydrolysis or oxidation of metal-organic precursors in the presence of alkylamine ligands produced several well-defined nanostructures of metal oxides by a one-step procedure [6-9]. Among them, the ZnO, CuO and SnO_2 nanoparticles were used for gas sensing applications [7-9]. For example, Ryzhikov et al., [7] investigated the ZnO morphology influence on the sensor response towards CO, C_3H_8 and NH_3 gases. More in details ZnO anisotropic nanoparticles, isotropic nanoparticles and cloudy-like aggregates were investigated as gas sensing layers. Sensors prepared with the anisotropic nanoparticles showed the highest response to both CO and C_3H_8, whereas sensors based on cloudy-like structures show the weakest response to C_3H_8. No effect of the ZnO morphology has been evidenced for NH_3 gas. It was concluded that the differences in sensors performance must be correlated to the exposed crystalline faces and their reactivity to the target molecules.

The dimensionality of the nanostructure can also have a significant impact on sensor response. For example, Jońca et al. [9] formed SnO_2 nanoparticles (0D structure) and hierarchical structures resembling a regular octahedron (3D structure) made of assembly of these nanoparticles. When exposed to low concentrations of CO (*i.e.* 0.25 to 20 ppm), hierarchical structures showed greater sensitivity than SnO_2 nanoparticles, but exhibited a lower upper limit of detection (*i.e.* 100 ppm *vs*. 500 ppm).

Gas sensing properties of metal oxide nanoparticles can be further improved by their decoration with noble metals or by doping them with 2D materials (*e.g.* graphene, g-C_3N_4, MoS_2, black phosphorus, etc.) [10]. Another approach is to combine two metal oxides in order to create a heterojunction [11]. Heterojunction refers to the area at the interface between two semiconductors with different band gap values. Due to the different chemical and physical parameters of both materials (*i.e.* band structure, dielectric constant, lattice constant, electron affinity), the phenomenon of mismatch at the interface gives the heterojunction new properties, which can find application, among others, in the construction of gas sensors.

MOS can be divided into two types, *i.e.* p-type (electron-hole concentration > free electron concentration, *e.g.* CuO, NiO, Fe_2O_3) and n-type (free electron concentration > electron-hole concentration, *e.g.* ZnO, WO_3, SnO_2). Therefore, three kinds of heterojunction can be distinguished: n-n, p-p and p-n. The gas sensing properties of the p-n heterojunction, including the CuO/ZnO one, have been explored by many research groups. Indeed, several CuO/ZnO-based gas sensors were prepared for the detection of HCHO, C_3H_6O, C_2H_5OH, NO_2, CO_2, H_2S, NH_3, etc. [12-18]. These sensors exhibited higher sensitivity, better selectivity and shorter response/recovery times than the devices based on single oxides. At the same time, it was found that the mass ratio of oxides used for the sensitive layer preparation plays a crucial role on the sensor performance. Moreover, mentioned above ZnO/CuO based sensors were operated at relatively low temperatures (some of them at room temperature) and thus, are suitable candidates to fulfill the market demand for near-zero power consumption devices.

This work explores gas sensing properties of the p-n heterojunction based on CuO and ZnO anisotropic nanoparticles prepared by the organometallic approach [6, 7]. The CuO and ZnO nanoparticle suspensions and their mixtures were evaluated as gas sensitive layers. These sensors have been exposed to different gaseous mixtures, *i.e.* 100 ppm CO, 100 ppm C_3H_8 and 5 ppm NH_3 at different working temperatures (from 75 ºC up to 400 ºC) under 50 % of relative humidity (RH). The optimal gas sensing performances were achieved for the CuO(75%)/ZnO(25%) configuration. Remarkably, in the case of CO detection, S_{CO}= 177%, $K_{CO/C3H8}$= 4.7 and $K_{CO/NH3}$= 9.8 and a response/recovery time of 4.5 min is obtained at 300 °C. The enhanced gas sensing properties of the CuO(75%)/ZnO(25%) sensors was attributed to the p-n heterojunction between the CuO and ZnO grains. The gas sensing mechanism of the mentioned heterojunction is also proposed herein.

2. Materials and Methods

2.1. Nanoparticles Synthesis

CuO nanoparticles (CuO Nps) were obtained from a mixture of (N,N'-diisopropylacetamidinato) Copper (I) (Cu(iPr-Me-amd)]$_2$) (0.125 mmol, 51 mg) and octylamine (0.625 mmol, 80.5 mg). Reagents were mixed in a small glass vial and then exposed to

ambient atmosphere. After 16 h, the obtained black product was washed 3 times with 5 mL of acetone using centrifuge (5000 rpm, 25 °C, 5 min).

ZnO nanoparticles (ZnO NPs) were obtained from a mixture of ($Zn(c\text{-}C_6H_{11})_2$ (0.25 mmol, 57.9 mg) and octylamine (0.5 mmol, 65 mg). The reagents were also prepared in a small glass vial, but then placed in a flat-bottomed reactor at room temperature and under argon atmosphere. The hydrolysis was performed by addition of degassed water to the reactor (18 µL). After 4 days, the obtained white product was washed one time with 5 mL THF and 3 times with 5 mL of acetone and centrifuge at 5000 rpm, 25 °C during 5 min.

2.2. Characterization

A JEOL JSM 1011 Transmission Electron Microscope (TEM) operating at 100 kV was used forimaging. TEM specimens were prepared by drop deposition of the washed and dispersed in ethanol nanomaterials on a carbon-supported copper or nickel grids. Analysis of the images were performed using Digital Micrograph program with at least 100 nanoparticles. The powder-diffraction patterns were obtained using SEIFERT XRD 3000TT X-Ray Diffractometer with Cu-Kα radiation, fitted with a diffracted-beam graphite monochromator. The data were collected in the 2θ configuration between 20 and 80°.

2.3. Gas Sensor Preparation

Freshly prepared and washed ZnO and CuO nanostructures were dispersed in ethanol. The concentration of the nanostructures in the solution was 5 $mg.mL^{-1}$. From these two solutions, mixtures of ZnO/CuO were prepared at different mass ratios: CuO(75%)/ZnO(25%), CuO(50%)/ZnO(50%), CuO(25%)/ZnO (75%). Finally, the CuO and ZnO suspensions and their mixtures were deposited on miniaturized gas sensors substrates by an ink-jet printing method (Microdrop AG) [19, 20].

The silicon platform of this study has been developed by the Laboratoire d'Analyse et d'Architecture des Systèmes, LAAS-CNRS. The die size is 2×2 mm^2, and it integrates a 1.4 µm thick dielectric membrane (SiO_2/SiNx/) designed for an optimized mechanical behavior and thermal insulation of the heated area. A spiral shaped platinum heater is buried between the bottom dielectric membrane and the passivation top layer (silicon dioxide). This heater structure can stand temperatures up to 700 °C and the power consumption does not exceed 55 mW at the operating temperature of 500 °C. The interdigitated platinum electrodes for the measure of the sensitive layer are deposited as a final step on the top of the SiO_2 passivation layer and present a rounded shape. A distance of 10 µm between each electrode pole provides a reliable contact even for high resistive sensing layers [20].

2.4. Gas Test set-up

Gas tests have been performed using a setup composed of different gas bottles connected to mass flow controllers (QualiFlow) commanded by an Agilent Data Acquisition/Switch Unit 34970A. Sensors are placed in a measurement cell equipped with humidity and temperature sensors. The integrated heaters are driven by a HP6642A voltage controller. A National Instruments 6035E electronic card establishes the connection between a computing unit and the measurement cell. Freshly prepared sensitive layers were initially conditioned by a sequential *in situ* heating of the sensitive layer from ambient temperature up to 550 °C in air. Afterwards, the sensitive layer resistance is stabilized on the device by annealing at 500 °C in synthetic air (relative humidity, RH 50 %) at a total gas flow rate of $1L.min^{-1}$. Finally, the sensors were exposed to controlled levels of CO, C_3H_8 and NH_3. The tests reported herein have been performed at different operating temperatures from 75 °C up to 400 °C, at 50 % RH. Resistance is measured before (R_{air}) and after (R_{gas}) sensor exposure to reducing gas mixture, and the normalized responses to given gas, $S(\%)_i$, is calculated using equation (1):

$$S_i = \frac{\left|\left(R_{air} - R_{gas}\right)\right|}{R_{air}} \cdot 100\ \%, \quad (1)$$

The selectivity coefficient, $K_{i/j}$, is expressed as the ratio between the sensor normalized responses towards investigated gases (S_i, S_j):

$$K_{i/j} = \frac{S_i}{S_j} \quad (2)$$

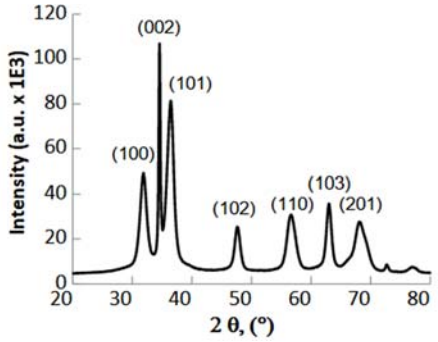
The response time is the time required for the sensor to reach a stable resistance value when it has been exposed to the target gas. The response time is given here as the time required to reach 90 % of the maximum signal value (t_{90}) obtained after exposure to a target gas. The recovery time, in turn, is the time necessary for the resistance to return to the value that the sensor had prior to an exposure to the target gas and is given here as the time required to achieve a resistance that is about 10% higher than the starting resistance (t_{10}). Results reported here have been performed by using at least 3 sensors prepared as described above.

3. Results and Discussion

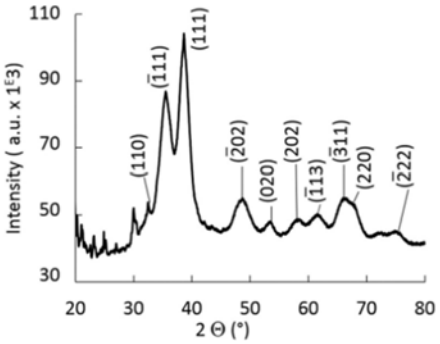
3.1. Morphology and Chemical Composition

Detailed characterization of the CuO and ZnO NPs has been reported earlier [7, 8]. Therefore, only a brief description of the morphology and chemical composition of the mentioned materials will be provided in this manuscript.

The X-ray diffraction analyses of as-prepared and washed ZnO nanostructures are presented in Fig. 1a [7]. The hydrolysis of the zinc precursor in the presence of octylamine and under argon atmosphere led to the formation of the hexagonal zincite structure (JCPDS 26-1451). The sharp and high intensity diffraction peak at 2θ = 34.6° corresponds to the (002) plane and is characteristic of the growth along the *c*-axis of the zincite crystals. The average crystallite size values, calculated by the Debye-Sherrer formula, are around 5 nm for the transversal axis and around 26 nm for the longitudinal one [7]. These results are in accordance with the TEM images analyses which depicted presence of anisotropic nanoparticles with the diameter of 5.3 ± 0.8 nm and length of 21.5 ± 5.6 nm [7] (Fig. 2a).



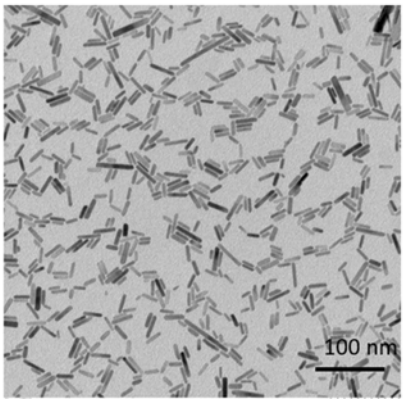
(a)



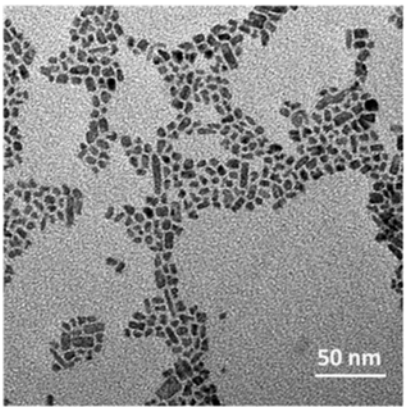
(b)

Fig. 1. X-rays diffractograms of ZnO (a) and CuO (b) nanoparticles [7, 8].

As revealed by the X-ray diffraction analyses (Fig. 1b), the oxidation of copper precursor under ambient atmosphere led to the formation of the CuO tenorite structure (ICDD: 96-410-5683) [8]. Determination of the crystallite size using Debye-Scherrer equation gives values of ca. 4–5 nm whereas the TEM images show the formation of rather anisotropic nanoparticles with uneven distribution of their size and shape (Fig. 2b). Analysis of these images through a 2D-plot representation revealed that three CuO populations of different size can be distinguished [21]. The first one, accounting for 50 % of the population, have an average width of 4.6±1.5 nm and length of 6.8 ± 2.7 nm. The second population possesses an average width of 5.8 ± 2.3 nm and length of 8.9 ± 3.5 nm and accounts for 43%. Finally, the third population, accounting for 6%, presents an average width of 6.5 ± 4.4 nm and length of 15.1 ± 10.0 nm [8]. The uneven distribution of the size and shape of the CuO NPs is probably associated with their growth conditions. Indeed, their synthesis is performed in ambient air and therefore the reaction parameters are not as well controlled as in the case of ZnO NPs.



(a)



(b)

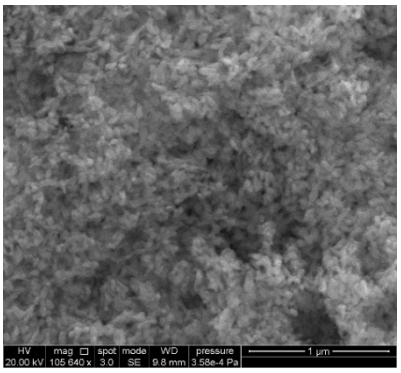
Fig. 2. TEM images of freshly prepared and washed ZnO (a), and CuO (b) nanoparticles [7, 8].

3.2. Gas Sensing Properties

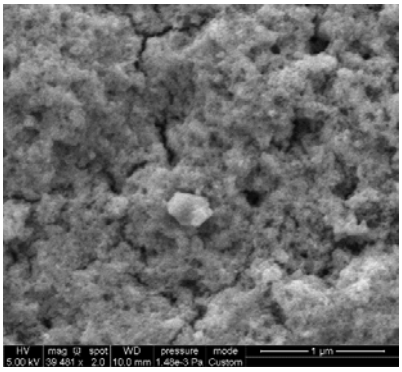
Our previous studies showed that from all ZnO nanostructures prepared through the organometallic approach, the anisotropic nanoparticles exhibited the highest sensitivity towards investigated gases at high temperatures [7], whereas CuO NPs gave high selectivity and sensitivity towards CO at low temperatures [8]. In order, to further improve the performance of our sensors, we investigate in the present work the behavior of sensitive layers composed of a mixture of both nanostructures.

The CuO, ZnO NPs and their mixtures have been deposited by an ink-jet printing method as gas

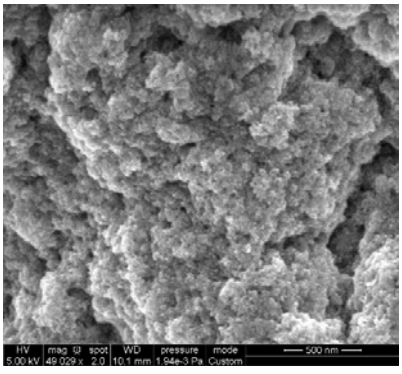
sensitive layers on silicon gas sensing devices. After deposition, the layers have been *in situ* annealed by operating the integrated Pt heater up to 550°C. This step allows the slow removal of solvent and ligand traces from the washed nanoparticles. The SEM images of these layers are presented Fig. 3. Good quality metal oxide semiconductor layers without large cracks or delamination were obtained. The morphology of the nanoparticles is preserved even after this treatment.



(a)



(b)



(c)

Fig. 3. SEM images of the ZnO (a), CuO (b) and CuO(50%)/ZnO(50%) (c) layers of the gas sensor devices.

In air, the CuO sensors exhibit a resistance of few kOhms. This value increases with the content of added ZnO, to reach MOhms for pure ZnO sensitive layers (Table 1). This is due to the differences in the band gap energies between these two materials, *i.e.* 1.2 eV and 3.7 eV for CuO and ZnO, respectively [22, 23]. In the presence of reducing gases, the resistance of CuO and CuO-ZnO based sensors increases which is characteristic for sensors based on p-type semiconductors. On the other hand, the resistance of ZnO based sensors decreases which is characteristic for n-type semi-conductors.

Normalized responses of all investigated sensors towards 100 ppm CO, 100 ppm C_3H_8 and 5 ppm NH_3 (RH=50 %) at different operating temperatures are presented in Fig. 4. The sensitivity of the ZnO sensor towards all investigated gases increases with the temperature, reaching the highest values at 400 °C. The highest sensitivity for the CuO sensor was achieved at 165 °C, but only towards 100 ppm CO and 100 ppm C_3H_8. The CuO sensor was not sensitive to 5 ppm NH_3 at all. This is in accordance with the results achieved earlier [7, 8]. Concerning the CuO-ZnO sensors, the highest sensitivity was also noticed at 165 °C for all tested gases and regardless the investigated metal oxide mixture. Overall, the sensitivity increases also with the CuO mass content in the sensing layer. However, for the CuO(75%)/ZnO(25%) sensor, the sensitivity is significantly higher, especially towards 100 ppm CO. This is due to the p-n heterojunction established between the CuO and ZnO NPs (Section 3.3).

From all investigated metal oxide mixtures, the CuO(25%)/ZnO(75%) sensor exhibited the lowest sensitivity. Interestingly, in the 340 °C – 400 °C temperature range, a switch from the p-type to n-type sensing behavior was observed (see Section 3.3).

Overall, the CuO(75%)/ZnO(25%) sensor is a suitable candidate for sensitive and selective CO detection. Indeed, the highest sensitivity towards 100 ppm of CO was achieved with sensitive layers composed of CuO (75 %) and ZnO (25 %) when operated at 165 °C, *i.e.* S_{CO} = 624 % (Table 2).

Moreover, the sensor showed relatively low response to 100 ppm C_3H_8 and 5 ppm NH_3, which indicates its high selectivity towards CO in these conditions ($K_{CO/C3H8}$=14.5 and $K_{CO/NH3}$=26). Therefore, the sensitivity and $K_{CO/C3H8}$ coefficient were improved as compared to ZnO (R_{CO}=39 % and K_{CO}/K_{C3H8}=1.2 at 400 °C) and CuO (S_{CO}=227 %, $K_{CO/C3H8}$=11.2 at 165 °C) sensors at their optimal operating temperatures (Table 2).

Table 1. The resistance (kOhm) of the ZnO, CuO and ZnO/CuO gas sensors at different operating temperatures.

T °C	CuO	CuO/ZnO (75%/25%)	CuO/ZnO (75%/25%)	CuO/ZnO (75%/25%)	ZnO
400	1.20	3.55	845	1094	3603
340	1.29	3.96	982	1921	4438
300	1.42	4.68	1326	2945	4626
165	2.44	17.8	4712	10366	9742
75	13.3	230	19547	22389	87341

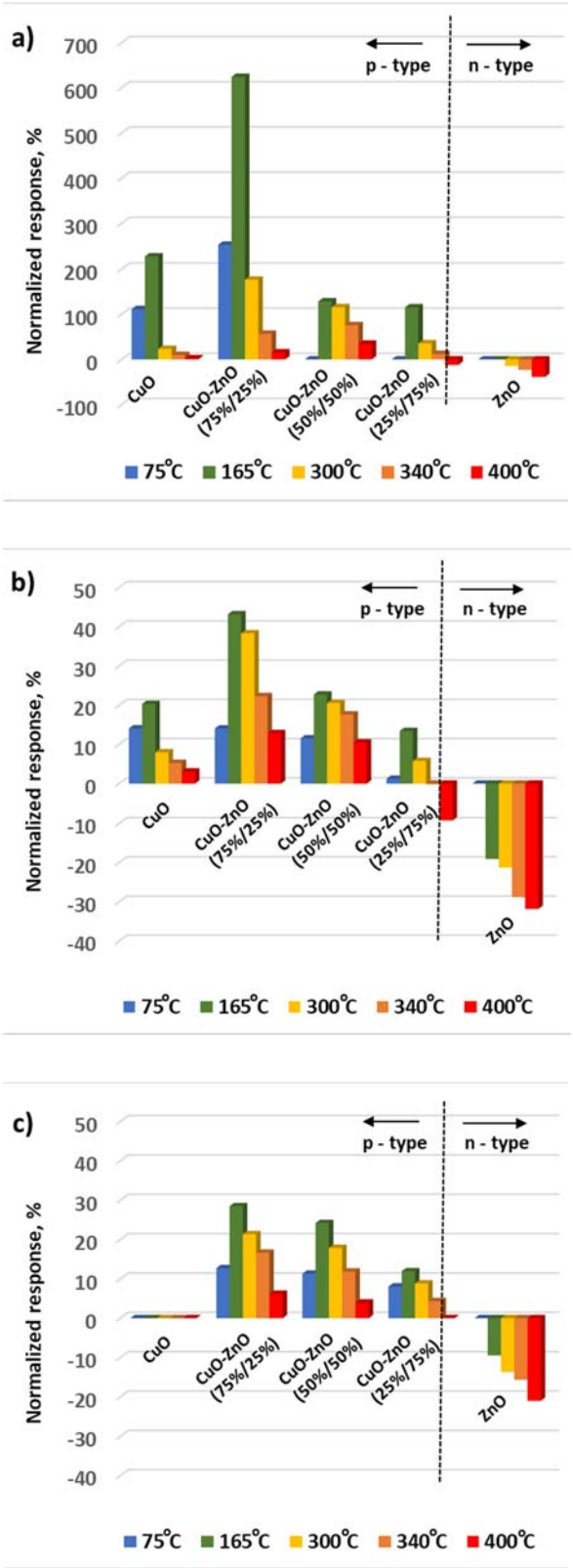


Fig. 4. Normalized response of the CuO, ZnO, and CuO/ZnO based gas sensors towards 100 ppm CO (a), 100 ppm C_3H_8 (b) and 5 ppm NH_3 (c) at different operating temperatures (50 % RH). The gas responses at 75°C present a very slow response time and the normalized response is measured before the complete resistance variation is reached, this value is therefore a transitory one and is under-evaluated. Please note, that the scale at (a) is different from the scale at (b) and (c).

Table 2. The normalized response (S_{CO}), selectivity coefficients ($K_{CO/C3H8}$ and $K_{CO/NH3}$), response (t_{90}) and recovery time (t_{10}) for the ZnO, CuO and CuO(75%)/ZnO(25%) gas sensor towards 100 ppm CO at different operating temperatures.

T, °C	S_{CO}, %	$K_{CO/C3H8}$	$K_{CO/NH3}$	t_{90}	t_{10}
CuO MOS sensor					
75*	111.4	7.9	selective	> 20 min	> 20 min
165	226.7	11.2	selective	6.2 min	20 min
300	23.4	2.9	selective	20 s	2 min
340	9.9	1.9	selective	30 s	30 s
400	2.5	0.8	selective	59 s	30 s
CuO(75%)/ZnO(25%) MOS sensor					
75*	252.7	18.0	19.9	> 20 min	> 20 min
165	624.3	14.5	21.9	2,3 min	> 20 min
300	176.7	4.6	8.3	50 s	4.5 min
340	56.9	2.6	3.4	30 s	1.8 min
400	15.8	1.2	2.5	18 s	20 s
ZnO MOS sensor					
75	-	-	-	-	-
165	-	-	-	-	-
300	14.3	0.7	1.1	1.2 min	1.8 min
340	22.8	0.8	1.5	50 s	1.8 min
400	38.6	1.2	1.8	20 s	1.2 min

(-) No response towards investigated gaseous mixture.

* The gas responses at 75°C present a very slow response time and the normalized response is measured before the complete resistance variation is reached, this value is therefore a transitory one and is under-evaluated.

The $K_{CO/NH3}$ coefficient was also improved as compared to the ZnO sensor ($K_{CO/NH3}$=1.8 at 400°C). Since, the CuO sensor is not sensitive to NH_3 at all, the $K_{CO/NH3}$ for this sensor is higher than for the CuO(75%)/ZnO(25%) one.

The response time of the CuO(75%)/ZnO(25%) sensor to 100 ppm CO at 165 °C is quite long (t_{90}=2.3 min) but shorter than for the CuO one (t_{90}=6.2 min at 165 °C) (Table 2). Unfortunately, at this temperature the recovery is very slow, *i.e.* t_{10} > 20 minutes (Fig. 5a). Therefore, to shorten the analysis time it is better to operate at 300°C. Even if, for these conditions, the sensor exhibits lower sensitivity (S_{CO} = 177%) and selectivity ($K_{CO/C3H8}$ = 4.7 and $K_{CO/NH3}$ = 9,8), the response/recovery time is significantly shortened (t_{90} = 50 s, t_{10} = 4.5 min) (Fig. 5b). It is noteworthy, however, that at 300 °C, the CuO(75%)/ZnO(25%) sensor still presents better sensitivity and selectivity than the devices based on pure CuO (S_{CO}=2 3%, $K_{CO/C3H8}$ =2.9 at 300 °C) or ZnO (S_{CO}=14 %, $K_{CO/C3H8}$ =0.67 and $K_{CO/NH3}$=1.1 at 300 °C) sensitive layers.

Sensitivity and response/recovery time are closely related to the sensor operating temperature. At low temperature the adsorption is favored, with the

formation of strong chemical bonds between the analyzed gas and the surface of the sensitive metal oxide semiconudctor layer. This enhanced interaction increases the change in sensor resistance and thus its sensitivity. However, at such low temperature, the process of gas diffusion in the material and its desorption process are less efficient, negatively affecting other sensor parameters (*e.g.* recovery/regeneration time). Of course, as the temperature increases, the increased thermal movement of the gaseous molecules can be observed, leading to an increase in the desorption rate, which, in turn, may reduce the sensitivity of the sensor [24].

Data described above are in agreement with results published earlier. Indeed, several CuO-ZnO based sensors have been described for selective and sensitive detection of HCHO, C_3H_6O, C_2H_5OH, NO_2, CO_2, H_2S or NH_3 [12-18]. These results showed that the ZnO/CuO heterojunction sensors exhibits higher sensitivity, better selectivity and shorter response/recovery time than the devices based on single oxides (Table 3).

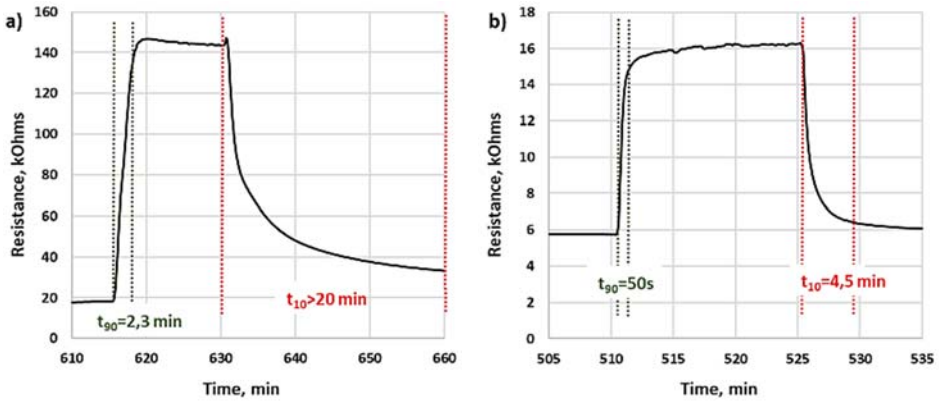


Fig. 5. Response of the CuO(75%)/ZnO(25%) MOS sensor at 165ºC (a) and 300ºC (b) towards 100 ppm of CO

Table 3. Examples of ZnO / CuO based gas sensors

Sensing material	T_{op}	Target gas	Gas conc.	S (%)*	$t_{(90)}/t_{(10)}$	Comments	Ref
CuO nanoparticles/ ZnO nanoparticles	225°C	H_2S	1 ppm	800%	30/100s	Stronger and faster response as compared to CuO and ZnO nanoparticle sensors	[12]
ZnO nanorods/ CuO nanoparticles	RT	NH_3	1 ppm	159%	2,3/2,1s	Better sensitivity and selectivity as compared to ZnO and CuO sensors	[13]
ZnO / CuO nanorods	RT	CO_2	1000 ppm	9,7%	4.2/3,5 min	Sensor is not sensitive to environmental gases such as CO, NO_2 and H_2S.	[14]
ZnO nanowires/ CuO nanoparticles	150°C	NO_2	100 ppm	175%	14 /197s	Low sensitivity (below 20%) towards CH_3OH, SO_2, CO, H_2S, Cl_2	[15]
CuO@ZnO microcubes	240°C	C_2H_5OH	50 ppm	360%	5/18s	CuO@ZnO cubes offer 2.6 times higher response as compare to CuO sensor.	[16]
ZnO hexagonal prisms/ CuO nanoparticles	RT	HCHO	1 ppm	188%	1,78/2,9s	The ratio of CuO and ZnO in the sensitive layer influences greatly its performance	[17]
ZnO branched p-CuxO@n-ZnO nanowires	250°C	C_3H_6O	50 ppm	85%	-	The nanocomposite exhibits 6 times higher response as compare to CuO nanowires.	[18]
ZnO / CuO anisotropic nanoparticles	165°C	CO	100 ppm	624%	3s/20min	The ratio of CuO and ZnO in the sensitive layer influences greatly its performance	This work

* In this work the sensitivity is expressed as $S=(R_{air}-R_{gas})/R_{air}*100$. However, the sensitivity can be calculated in different ways, *i.e.* $S=R_{air}/R_{gas}$, $S=R_{gas}/R_{air}$, $S=(R_{air}-R_{gas})/R_{air}$, etc. Therefore, in some cases presented in the table we estimated the normalized response using response curves or data provided within the cited papers. This allows a better comparison of sensitivity parameter between different studies.

3.3. Gas Sensing Behavior

The response of the CuO and CuO/ZnO MOS sensors to reducing gas mixtures shows p-type sensing behavior. For p-type MOS sensors, the widely accepted gas sensing mechanism is based on the change of conductivity due to the interactions between the surface adsorbed oxygen species and the analyzed gas [25]. When CuO sensing layer is exposed to the air atmosphere, oxygen molecules will adsorb on its

surface to form oxygen species, and then captured free electrons of CuO to form a hole accumulation layer (HAL). When exposed to the CO, the gas molecules react with pre-adsorbed oxygen species on the CuO surface and return the captured electrons back to the conduction band (CB) of CuO. Meanwhile, the thickness of the HAL becomes thinner, and the resistance of the CuO increases.

According to previous reports [26, 27], the improvement of the gas sensing properties of p-type semiconductors combined with n-type semiconductors can be described as follows. When the CuO contacts in air with the ZnO, an internal self-built electrical field at the interface will be formed *via* charge carries diffusion due to their different work functions, electron affinities and band gaps. In the electrical field, the electrons transfer from ZnO to CuO, and holes move in the opposite direction until the systems achieve equalization at the Fermi level (E_F), as displayed on the Fig. 6a. The process will lead to development of the potential barrier at the heterojunctions as the band bending and a wide depletion layer appears at the interface.

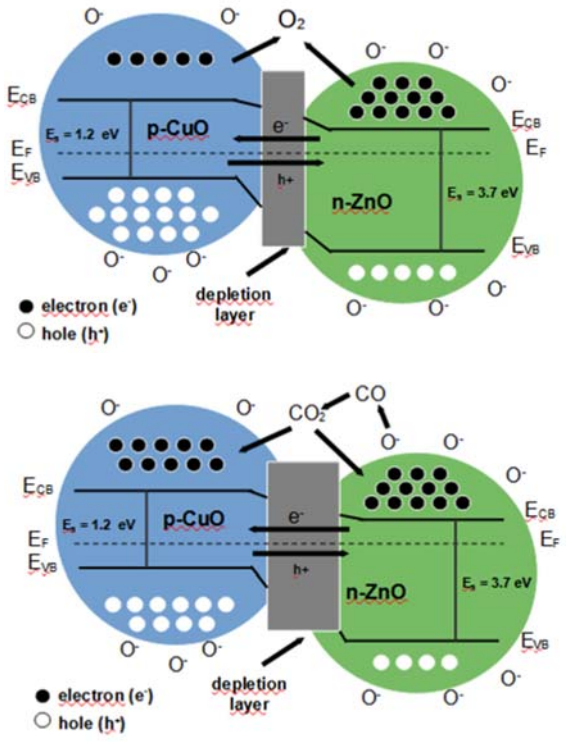


Fig. 6. The band diagram of the ZnO/CuO heterojunction (a) in the air (RH 50 %); (b) in the CO.

When the reducing gas is injected in, its molecules react with pre-adsorbed oxygen species and the electrons return back to the interfaces of the CuO and ZnO. The electrons return to the CB of the CuO and combine with the holes in the valence band (VB), which reduces the hole concentration in the CuO HAL and increases the resistance of CuO. On the other side, the electrons that have returned to the ZnO increase the electron concentration in the ZnO layer. The electrons transfer from ZnO to CuO until a new equilibrium is obtained (Fig. 6b). As a result, the hole quantity in the VB of the CuO is further reduced, which greatly increases the resistance of the CuO. This results in a significant improvement in the response of the CuO/ZnO gas sensor.

Taking into account the results obtained in this work, it is obvious that the CuO(75%)/ZnO(25%) sensor exhibits the highest number of heterojunctions. With increasing amount of the ZnO nanoparticles within the sensitive layer the amount of heterojunctions decreases and thus, the sensitivity decreases as well reaching the lowest values for the CuO(25%)/ZnO(75%) sensor. Moreover, at 340 °C the sensor does not respond (or the response is very week). The signal generated by the n-type ZnO is balanced by the response produced by the p-type CuO and CuO/ZnO heterojunction. A change from a p-type to n-type sensing behavior is thus observed (Fig. 4). Such transition has been observed before [28-31]. For example, temperature driven p- to n- type switching was described by Sinha et al. for sensitive layers made of CNT/ZnO composite [28]. Indeed, the CNT/ZnO composite is able to operate as a dual mode sensor, in which CNT dominates in mild temperature region (≤ 150 °C) and ZnO at high temperature region (>150 °C) with sensitivity of 72.6 % and 23.2 %, respectively.

4. Conclusions

The ZnO and CuO anisotropic nanoparticles and their mixtures were deposited on miniaturized gas sensors substrates as gas sensitive layers and exposed to different reducing gases (CO, C_3H_8, NH_3). Several conclusions can be drawn from these investigations:

- The ratio of ZnO and CuO nanoparticles within the sensitive layer play a crucial role on MOS sensor performances;
- The CuO(75%)/ZnO(25%) sensor exhibits the highest sensitivity and selectivity towards CO from all investigated sensors;
- Although the CuO(75%)/ZnO(25%) sensor can operate at relatively low temperature (165 °C), its recovery is very slow;
- The analysis time can be shorten by increasing the temperature up to 300 °C, but this will decrease the sensitivity and selectivity of the CuO(75%)/ZnO(25%) sensor;
- The enhanced gas sensing properties of CuO(75%)/ZnO(25%) sensor is attributed to the formation of the p-n heterojunction and its gas sensing mechanism is described;
- The CuO(25%)/ZnO(75%) sensor exhibits the lowest gas sensing performance towards tested gases,
- A shift from p-type to n-type sensing behavior is observed for the CuO(25%)/ZnO(75%) at 340 °C / 400 °C.

These results highlight the influence of the ZnO/CuO heterojunction on gas sensor sensitivity and selectivity, as well as on its response and recovery times. The CuO(75%)/ZnO(25%) sensor is a suitable candidate for an efficient CO sensing in the presence of C_3H_8 and NH_3.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101033564. This work was partly supported by LAAS-CNRS micro and nano technologies platform member of the French RENATECH network.

References

[1]. M. V. Nikolic, V. Milovanovic, Z. Z. Vasiljevic, Z. Stamenkovic, Semiconductor Gas Sensors: Materials, Technology, Design, and Application, *Sensors,* 20, 2020, 6694.

[2]. B. Szulczyński, J. Gębicki, Currently Commercially Available Chemical Sensors Employed for Detection of Volatile Organic Compounds in Outdoor and Indoor Air, *Environments,* 4, 2017, 21.

[3]. Yaqoob, U., Younis, M. I., Chemical Gas Sensors: Recent Developments, Challenges, and the Potential of Machine Learning—A Review, *Sensors,* 21, 2021, 2877.

[4]. S. Sharma and M. Madou, A new approach to gas sensing with nanotechnology, P*hilosophical Transactions: Mathematical, Physical and Engineering Sciences,* 370, 2012, pp. 2448–2473.

[5]. M. L. Kahn et al., Size- and Shape-Control of Crystalline Zinc Oxide Nanoparticles: A New Organometallic Synthetic Method, *Adv. Funct. Mater*, Vol. 15, Issue 3, 2005, pp. 458-468.

[6]. G. Casterou, V. Collière, P. Lecante, Y. Coppel, P. A. Eliat, F. Gauffre, M. L. Kahn, Improved Transversal Relaxivity for Highly Crystalline Nanoparticles of Pure γ-Fe_2O_3 Phase, *Chemistry*, 14, 2015, pp. 18855-18861.

[7]. A. Ryzhikov et al., Organometallic synthesis of ZnO nanoparticles for gas sensing: towards selectivity through nanoparticles morphology, *Journal of Nanoparticle Research*, 17, 2015, 280.

[8]. J. Jońca et al., Organometallic Synthesis of CuO Nanoparticles: Application in Low-Temperature CO Detection. *ChemPhysChem,* Vol. 18, Issue 19, 2017, pp. 2658-2665.

[9]. J. Jońca et al., SnO_2 "Russian Doll" Octahedra Prepared by Metalorganic Synthesis: A New Structure for Sub-ppm CO Detection, *Eur. J. Chem.,* Vol. 22, Issue 29, 2016, pp. 10127-10135.

[10]. V. S. Bhati, M. Kumar, R. Banerjee, Gas sensing performance of 2D nanomaterials/metal oxide nanocomposites: a review, *Journal of Materials Chemitry C*, 9, 2021, pp. 8776-8808.

[11]. L. Liu, Y. Wang, Y. Liu, S. Wang, T. Li, S. Feng, S. Qin and T. Zhang, Heteronanostructural metal oxide-based gas microsensors, *Microsystems & Nanoengineering,* 8, 2022, 85.

[12]. S. Park, S. Kim, H. Kheel, S. K. Hyun, C. Jin, C. Lee, Enhanced H_2S gas sensing performance of networked CuO-ZnO composite nanoparticle sensor, *Materials Research Bulletin*, Vol. 82, October 2016, pp. 130-135.

[13]. C. Cheng, C. Chena, H. Zhang, Y. Zhang, Preparation and study of ammonia gas sensor based on ZnO/CuO heterojunction with high performance at room temperature, *Materials Science in Semiconductor Processing*, Vol. 146, 2022, 106700.

[14]. S. Keerthana, K. Rathnakannan, Hierarchical ZnO/CuO nanostructures for room temperature detection of carbon dioxide, *Journal of Alloys and Compounds*, Vol. 897, 2022, 162988.

[15]. Y. H. Navale, S. T. Navale, F. J. Stadler, N. S. Ramgir, V. B. Patil, Enhanced NO_2 sensing aptness of ZnO nanowire/CuO nanoparticle heterostructure-based gas sensors, *Ceramics International,* 45, Issue 2, Part A, 2019, pp. 1513-1522.

[16]. M. Yin, F. Wang, H. Fan, L. Xu, S. Liu, Heterojunction CuO@ZnO microcubes for superior p type gas sensor application, *Journal of Alloys and Compounds,* Vol. 672, 2016, pp. 374-379.

[17]. J. Liu, Y. Chen and H. Zhang, Study of Highly Sensitive Formaldehyde Sensors Based on ZnO/CuO Heterostructure via the Sol-Gel Method, *Sensors,* 21, 14, 2021, 4685.

[18]. X.-T. Xue, L.-Y. Zhu, K.-P. Yuan, C. Zeng, X-X. Li, H.-P.Ma, H.-L. Lu, D. W. Zhang, ZnO branched p-CuxO @n-ZnO heterojunction nanowires for improving acetone gas sensing performance, *Sensors and Actuators B: Chemical,* Vol. 324, 2020, 128729.

[19]. Ph. Menini et al., High performances of new microhotplate for gas sensors, in *Proceedings of the 22nd European Conference on Solid state Transducers (EUROSENSORS XXII),* September 2008, Dresden, Germany, hal-02099850.

[20]. Microdrop Technologies: http://www.microdrop.de/wDeutsch/technology/microdrop.shtml?navid=28

[21]. Z. Zhao, Z. Zheng, C. Roux, C. Delmas, J.-D. Marty, M. L. Kahn, C. Mingotaud, Importance of the Correlation between Width and Length in the Shape Analysis of Nanorods: Use of a 2D Size Plot To Probe Such a Correlation, *Chemistry - A European Journal*, Vol. 22, Issue 35 2016, pp. 12424-12429.

[22]. A. A. Radhakrishnan, B. B Beena, Structural and Optical Absorption Analysis of CuO Nanoparticles, *Indian Journal of Advances in Chemical Science,* 2, 2, 2014, pp. 158-161.

[23]. M. El-Kemary, H. El-Shamy, I. El-Mehasseb, Photocatalytic degradation of ciprofloxacin drug in water using ZnO nanoparticles, *Journal of Luminescence,* 130, 2, 2010, pp. 2327-2331.

[24]. N. A. Isaac, I. Pikaar, G. Biskos, Metal oxide semiconducting nanomaterials for air quality gas sensors: operating principles, performance, and synthesis techniques, *Microchimica Acta,* 189, 2022, 196.

[25]. H. J. Kim, J. H. Lee, Highly sensitive and selective gas sensors using p-type oxide semiconductors: Overview. *Sensors and Actuators B*, 192, 2014, pp. 607-627.

[26]. H. Zhang, H. Li, L. Cai, Q. Lei, J. Wang, W. Fan, K. Shi, G. Han, Performances of In-doped CuO-based heterojunction gas sensor, *Journal of Materials Science: Materials in Electronics*, 31, 2020, pp. 910-919.

[27]. L. Cai, H. Li, H. Zhang, W. Fan, J. Wang, Y. Wang, X. Wang, Y. Tang, Y. Song. Enhanced performance of

the tangerines-like CuO-based gas sensor using ZnO nanowire arrays, *Materials Science in Semiconductor Processing*, Vol. 118, 2020, 105196.

[28]. M. Sinha S. Neogi, R. Mahapatra, S. Krishnamurthy, R. Ghosh, Material dependent and temperature driven adsorption switching (p- to n- type) using CNT/ZnO composite-based chemiresistive methanol gas sensor, *Sensors and Actuators B: Chemical,* 336, 2021, 129729.

[29]. R.-C. Wang, Y.-R. Hou, J.-Y. Liu, and Y.-W. Chen, Differentiating Ammonia from Other Reducing Gases via Response Reversal Phenomena by Varied ZnO/CuxO Nanorod Arrays, *Journal of the Electrochemical Society*, 165, 2018, B484.

[30]. X. Zhou, T. Tao, Y. Bao, X. Xia, K. Homewood, Z. Wang, M.Lourenço, Z. Huang, G. Shao, and Y. Gao, Dynamic Reaction Mechanism of P–N-Switched H_2-Sensing Performance on a Pt-Decorated TiO_2 Surface, *ACS Applied Materials and Interfaces* 13, 21, 2021, pp. 25472–25482.

[31]. S. Rani, M. Kumar, P. Garg, R. Parmar, A. Kumar, Y. Singh, V. Baloria, U. Deshpande, and V. N. Singh, Temperature-Dependent n–p–n Switching and Highly Selective Room-Temperature n-$SnSe_2$/p-SnO/n-SnSe Heterojunction-Based NO_2 Gas Sensor, *ACS Applied Materials and Interfaces*, 14, 13, 2022, pp. 15381–15390.

Call for Chapters

IFSA Publishing

Your Chapter may be in the next Volume of the popular Book Series

ADVANCES IN SENSORS: REVIEWS 8

Open Access Book Series

Sinse 2012

- Physical, chemical and biosensors
- Sensor uncertainty, accuracy, calibration
- Smart sensors and systems
- Intelligent sensors and systems
- Artificial Intelligence enabled sensors
- Virtual and soft sensors
- Sensor instrumentation
- Sensor buses and networks
- Signal processing, sensor fusion
- Nanosensors
- Photonic sensors
- Sensor interfacing and signal conditioning
- Materials for sensors
- MEMS, BioMEMS and NEMS
- Remote Sensors and Telemetry
- Sensor packages and reliability

https://www.sensorsportal.com/HTML/IFSA_Publishing.htm

IFSA

International Frequency Sensor Association Publishing

www.sensorsportal.com

ISSN 1726- 5479