

Feedback-free Wavelet Based Distributed Coding for Video

¹ Chen Rui, ² Wu Minghu, ¹ Tong Ying

¹ School of Communications Engineering, Nanjing Institute of Technology,
Nanjing, China, 211167 China

² School of Electrical and Electronic Engineering, Hubei University of Technology,
Wu'han, 430068, China

¹ Tel.: 88413677, fax: 88413677

¹ E-mail: chenrui@njit.edu.cn

Received: 19 April 2013 / Accepted: 14 June 2013 / Published: 25 June 2013

Abstract: Distributed video coding can effectively reduce the complexity in encoder, so it adapt to the capacity-constrained scenarios, such as the wireless video sensor network. But in wireless networks, the transmission rate's instability caused by the fluctuant bandwidth will influence the video reconstructed quality. This paper proposes a wavelet domain distributed coder for video which allows for scalability and does not need feedback channel. At the encoder, a simplified integer wavelet transform algorithm is adopted to improve the computational efficiency. Then according to the channel condition, either the low-frequency coefficients or the low frequency plus several high frequency coefficients are selected. At the decoder, the video sequence is reconstructed by the received coding rate and the generated side information. The simulation results show that the proposed scheme can reduce the transmission bit rate with a good video transmission quality.

Copyright © 2013 IFSA.

Keywords: Distributed video coding, Wavelet transforms; Video coding.

1. Introduction

With the development of broadband wireless network, more and more mobile video terminals have been used in multimedia communications. These application scenarios require a light encoder in which the computation and memory capacity are scarce at the wireless terminals. A new video compression paradigm Distributed Source Coding (DSC) based on Slepian-Wolf lossless coding theory [1] and Wyner-Ziv lossy coding theory [2] has been proposed. For two correlated sources, X and Y are encoded independently and then jointly decoded to achieve compression. Source Y is compressed with traditional coding techniques and used at the decoder to estimate

source X . According to Slepian-Wolf theorem [1], if this estimate is used as a Side Information SI to aid decoding, the original source X can be lossless reconstructed at rates $R \geq H(X/SI)$ even if SI is not known at the encoder. The counterpart of the Slepian-Wolf theorem for lossy source coding is Wyner-Ziv's work [2]. Wyner proposed that such compression can be achieved using channel coding techniques, where X is compressed into its syndrome representation S , and SI is used to choose the most likely value of X from the other coset elements represented by S . Recently, several practical schemes based on the DSC principles have been proposed for video coding, i.e. Distributed Video Coding, such as the European DISCOVER codec architecture [3], Bernd Girod's

DVC scheme with feedback channel [4], Ramchandran's PRISM (Power-efficient Robust high-compression Syndrome-base Multimedia) scheme [5], and etc.

In the Distributed Video Coding (DVC) scheme based on Wyner-Ziv theorem, the video frames are divided into key frames and Wyner-Ziv frames. The key frames (i.e. K frames) adopt conventional intraframe coding decoding algorithm. The Wyner-Ziv frames (i.e. WZ frame) use inside coding and inter decoding. There are two approaches: the pixel domain Wyner-Ziv codec system and the transform domain Wyner-Ziv codec system [2]. Due to the low compression efficiency of the pixel domain WZ coding, many researches focused on the transform domain WZ coding, such as Stanford's Wyner-Ziv decoding scheme based on discrete cosine transform (DCT) [4], in which the video frame is divided into 8×8 or 16×16 sample blocks, and it leads to "block effects". So, the discrete wavelet transform (DWT) has been widely used in DVC paradigm due to its advantages of multi-resolution analysis and spatial de-correlation. In [6], the author proposed a linear prediction model to exploit the inter-band correlation of DWT coefficients of the hyper spectral imagery. In [7], the authors use the RWDT reference frames for finding matching blocks which satisfy CRCs to overcome the shift-variant of DWT domain. In [8], the author decomposed the original video into the reduced-resolution layer and the high-pass subbands by DWT, and the reduced-resolution layer is coded by an adaptive DPCM WZ coding and intra block coding strategy.

In this paper, we propose a wavelet based DVC scheme with a feedback-free adaptive rate control algorithm according to the bandwidth. At the encoder, a simplified integer wavelet transform algorithm is adopted to improve the computational efficiency. Then according to the channel condition, either low-frequency coefficients or low frequency plus several high frequency coefficients are selected. At the decoder, the video sequence is reconstructed by the received coding rate and the generated side information. The rest of this paper is organized as follows. In Section 2, we describe our scheme in detail. In Section 3, we give the experimental results and compare the performance of the proposed coder with conventional intra-frame coding. Finally, conclusions and future works are discussed in Section 4.

2. The Proposed Feedback-free Wavelet Based DVC Architecture

In this section, we propose a feedback-free wavelet based DVC scheme, which is shown in Fig. 1. At the encoder, the video frames in one given GOP (group of picture) are divided into key frames (i.e. intra frames) and non-key frames (i.e. Wyner-Ziv frames, WZ frames for short). The key frames are intra coded with a traditional video coding approach like H.264/AVC, and the non-key frames are Wyner-Ziv coded.

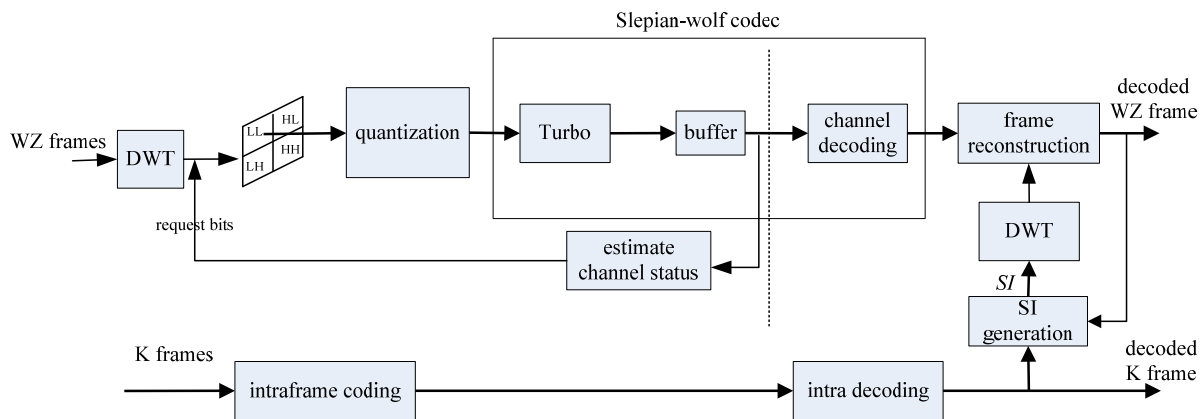


Fig. 1. Proposed scheme for feedback-free wavelet based DVC.

At the WZ encoder, a simplified integer discrete wavelet transform (DWT) is adopted to exploit the spatial correlation. After transformed by simplified integer DWT, there are two types of coefficients to be generated: the coefficients of the lowest frequency subband and the coefficients of high frequency subbands. Considering the coefficients of the lowest subband represent most information of one frame, so in our scheme, only the lowest frequency subband is

quantized, then the quantized coefficients are converted to their binary representations and a series of bitplanes are constructed by the same efficient location bit of them. These bitplanes are encoded using Turbo encoder, and only the parity bits are stored in buffer and ready for being transmitted on demands. All the coefficients of high frequency subbands are ignored due to their representation of the detailed information of the frame.

2.1. Simplified Integer Wavelet Transform for the Lowest Frequency Subband

Wavelet transform is used in DVC coding systems due to its flexibility in representing nonstationary image signals and its ability to adapt to human visual system. The wavelet transform can decompose a nonstationary signal into a set of multiscaled subbands from coarse to fine, as shown in Fig. 2. The coefficients are comparatively stationary and easier to be coded. After wavelet transform, the energy of the whole image is reallocated and most of energy is grouped together into a small location. Sweldens proposed wavelet transform theory based on division-forecast-update, resulted in the wavelet transform being calculated in the integer domain and avoiding the quantization error in pre-process [9]. Then Said. An adopted single ascension and double ascension method to improve the compression ratio by 10 %, compared with the predictive coding and Haar wavelet [10]. In [11], the authors proposed a new method to compute integer wavelet transform (IWT) by binary coefficient to avoid the multiplications, and the algorithm was easy to implement on FPGA or DSP hardware.

Currently, fast WT algorithms include EZW, SPIHT SFQ, EPWIC, and EBCOT [12]. Two types of ascension wavelet are commonly used: one is 5/3 wavelet used for lossless image compression and lossy image compression; the other is 9/7 wavelet transform, used for high quality image lossy compression.

In this paper, we propose a simplified fast wavelet algorithm based on ascension algorithm. The experimental results shown in Table 1, the efficient of DWT is improved by 70 %. The proposed algorithm is based on the reversible integer 5/3 wavelet transform, which is one of the lifting wavelet. The basic arithmetic of lifting wavelet transformation is through a lazy wavelet to constructs a better nature new wavelet gradually. There are three steps of integer wavelet transform based on lifting scheme, which is described specifically as follows and shown in Fig. 2.

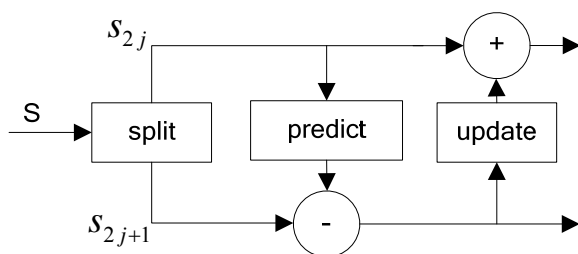


Fig. 2. Lifting wavelet transform diagram.

(1) Split

Subset the original signal S into even signal S_{2j} and odd signal S_{2j+1} , $F(S)=(S_{2j}, S_{2j+1})$.

(2) Predict

In view of data relevance, keep even signal S_{2j} constant, predict odd signal S_{2j+1} through interpolation method, d_j is the D -value between predictive value and the actual value, i.e. $d_j=S_{2j+1}-P(S_{2j+1})$, and P is the prediction operator.

(3) Update

Because decomposed into two subsets, some characteristics of the original set are lost, to make the subset data and the original set data have the same characteristics, update S_{2j} using d_j , in order to retain some characteristics of the original signal S , such as retain the same average value, this operation recorded as, $a_j=S_{2j}+U(d_{2j})$ and U is the update operator. Considering about 5/3 biorthogonal wavelet, d_j and a_j are calculated by formula (1) and (2) respectively.

$$d_j = s_{2j+1} - \left[\frac{1}{2}(s_{2j} + s_{2j+1}) \right] \tag{1}$$

$$a_j = s_{2j} + \left[\frac{1}{2}(d_{j-1} + d_j) \right] \tag{2}$$

The multiplication of lifting coefficient (1/2 and 1/4) can be replaced by shift operation to reduce computation. In the traditional fast algorithm, first step is to make a one-dimensional IWT transform in row, resulting in high and low frequency parts in horizontal. The second step is to do one-dimensional IWT in column. Due to the low frequency part containing the most energy image, the importance of low frequency part is far higher than the high frequency part.

Considering the limited processing ability of DVC encoder, the second step can be simplified by ignoring the high frequency part. A picture of original image is treated as shown in Fig. 3. The proposed method focuses on high-efficiency calculation of low-frequency region, and ignores the low-efficiency calculation of high-frequency region.

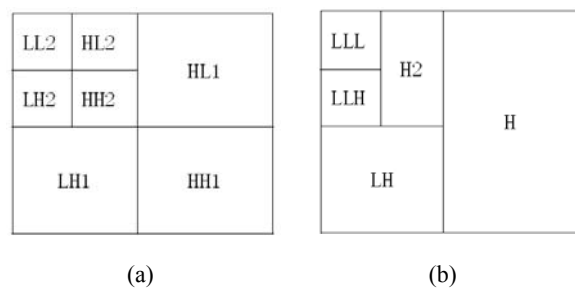


Fig. 3. Two-level wavelet decomposition (left: traditional algorithm, right: proposed simplified algorithm).

We perform our method on Foreman and Coastguard QCIF test sequences, comparing with ascension 3/5 fast wavelet algorithm under the same hardware and software environment. The results are shown in Table 1, the proposed simplified integer DWT is fast than the traditional lifting 3/5 wavelet algorithm by 57.4 %.

Table 1. The efficiency contrast of the traditional fast algorithm and the simplified algorithm.

Video sequences	Algorithms efficiency (millisecond/frame)		Efficiency compared
	Traditional algorithm	Proposed algorithm	
Foreman	26.69	16.95	+63.5%
Coastguard	33.35	18.31	+54.9%
Soccer	33.14	18.56	+56.0%
Hallmonitor	31.57	17.43	+55.2%

Notes: Test environments are ARM (model EBD9260, CPU 180 MHz) and Linux (kernel version 2.6.19), video quality is QCIF (176×144), processing frame number is 300 frames.

2.2. Filled Method for High Frequency Subbands

In the PRISM framework [5], which is characterized by a feedback-free architecture with a hash-based motion search at the decoder, at the encoder, hash codes are generated by reduced-resolution and coarse quantizations for each coding block (such as 8×8), then the mean-square deviation of hash code is calculated between the previous frame F_{i-1} and the current frame F_i at the same position. At the decoder, if the mean-square deviation is less than the threshold value G , the decoder directly copy the block from the previous frame F_{i-1} at the same position. According to this algorithm, few of the low frequency coefficients can meet the

condition, but most of the high frequency coefficients can meet the requirements [12].

Considering that if two adjacent frames in one video sequence are very close, the accurate of side information by interpolation or extrapolation is close to 30 dB in PSNR. Generally, human eyes are unable to distinguish two images if the PSNR value is more than 30 dB [13]. For *foreman* sequence, the experimental results of the wavelet coefficients statistics for the original image and SI are shown in Table 2 and Table 3 respectively. As we can see, the low frequency subband of the side information is very close to the original image in mean value and variance. Meanwhile, the low frequency subband contains most of the energy of original image while the high frequency subbands contains a small fraction.

Table 2. Wavelet coefficients statistics of *foreman* video sequence.

	Max	Min	Mean value	Variance	Energy
Original image	255	0	99.05	52.88	100 %
LLL	250.75	-4.2	99.02	52.56	99.7 %
LLH	34.25	-34.41	-0.45	2.92	0.04 %
H2	35.48	-31.32	-0.74	1.09	0.07 %
LH	20.11	-24.56	0.01	2.06	0.06 %
H	17.24	-15.17	0.03	2.14	0.13 %

Table 3. Wavelet statistics of *foreman* video sequence.

	Max	Min	Mean value	Variance	Energy
SI	255	0	92.05	47.88	100 %
LLL	250.75	5.0	101.02	69.56	99.7 %
LLH	30.25	-36.41	-0.72	2.11	0.04 %
H2	37.34	-28.17	-0.53	1.23	0.08 %
LH	22.04	-23.51	-0.01	2.17	0.06 %
H	16.64	-14.20	0.05	2.10	0.12 %

According to the above analysis, the LLL subband represents most information of one frame which is the most important part for transmission, and the high frequency subbands contains the detailed image information. Furthermore, according to the high frequency subbands of the original image and SI are

very close in mean variance, so we use the high frequency subbands of the SI to fill the decoding frame directly. The corresponding quality results of directly filled method of *foreman* and *coastguard* sequences are shown in Fig. 4, 5 and 6.

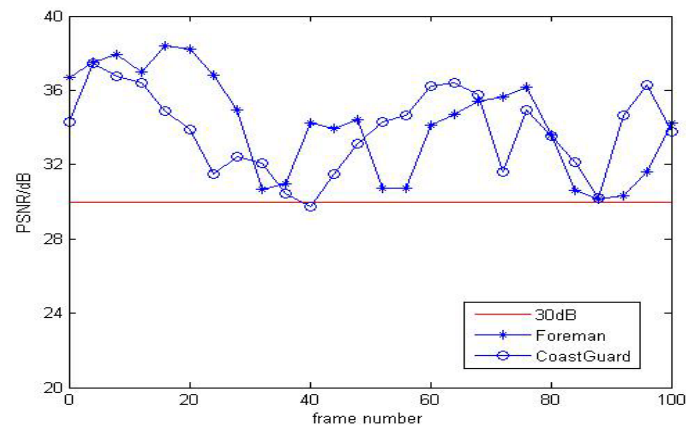


Fig. 4. RD performance of *foreman* and *coastguard* using filled method.



Fig. 5. Decoded frame #141 of *foreman* sequence (left: the original image; middle: Stamford scheme; right: the proposed filled method).



Fig. 6. Decoded frame #121 of *coastguard* sequence (left: the original image, middle: Stamford scheme, right: the proposed filled method).

2.3. Embedded Coding without Feedback

Compressed video streaming is sensible to the transmission errors in time-variable bandwidth and error-prone wireless communication networks, so the scalable coding technology was proposed. Scalability encoder is divided into two kinds: hierarchical coding and Fine Granularity Scalability (FGS). The hierarchical coding streaming is divided into several layers, and can only be truncated at the stratified place while transmitting or decoding. FGS can be cut off in any position, so its code rate is continuous variable and be called embedded coding.

Commonly, in DVC architecture, a feedback channel is required for the decoder to decide the parity bit rate according to SI, and feedback it to the encoder. But the feedback brings the additional delay, and it is unsuitable for real-time applications. For the

feedback free DVC, it is difficult for the encoder to estimate the optimal bit-rate of a specific coding block adaptively without access the side information available at the decoder. In this paper, the optimal number of parity bits is estimated according to the image signal's intensity. The encoder controls the parity rate according to the current signal intensity. Embedded coding can be accurate code rate control, which is suitable for multimedia communication network. The SI filled high frequency method, mentioned in the second part of this paper is fit for embedded codec. Encoder decides rate of code stream, according to real-time bandwidth, then use side information to fill the rest of the image. Before transmission, the wavelet coding plane should be transformed into binary bit plane, and the priority order of each subband is LLL-LLH-H2-LH-H, as shown in Fig. 7.

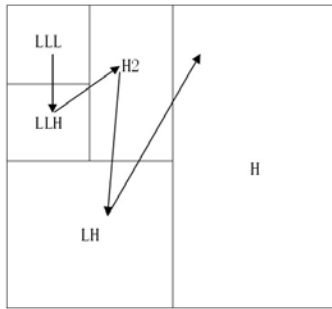


Fig. 7. Priority order of each frequency subband.

The data rate of coded LLL subband streaming is about 1/16 in one frame, but contains 99.7 % of the energy, so it has the highest priority. LLH and H2 subbands are similar to LH, HL and HH subbands. LH and HL represent the horizontal and vertical direction of the high frequency subbands respectively. HH represents the diagonal direction of high frequency subbands. According to human eyes are more sensitive to HL and LH than to HH subband, the priority of LLH subband is higher than H2, the priority LH priority is higher than H subband. In order to facilitate rate control, two extra bytes have been added into the bit streaming to identify the transmitted bit number. For example, the image resolution is QCIF (176×144), the coding rate is 4/5 Turbo code, and the video frame rate is 25 frame/s, then if all LLL, LLH, H2 subbands need to transmit, the data rate is close to 300 kbps, while the LLL subband need 79 kbps. In Fig. 8, the reconstructed quality of only transmitting LLL subband,

transmitting LLL+LLH subbands and transmitting LLL+LLH+H2 are shown respectively.

3. Experimental Results

To testify the coding efficiency of proposed wavelet domain DVC, we implemented it and assessed the performance on two QCIF video sequences, i.e. *foreman* and *coastguard* sequences. In each sequence, 300 frames are selected. For the sake of simplicity, we take WZ frames as the even frames and key frames as the odd frames of the sequence. So, we have the GOP structure as I-W-I-W. Because this paper mainly focuses on how to code the Wyner-Ziv frames efficiently, we assume that the key frames are available at the decoder perfectly reconstructed. The Wyner-Ziv frames are intra-frame encoded using the proposed scheme. The Turbo encoder is adopted for channel coding due to its inner robustness.

In our experiments, we compare the RD performance of our scheme with Stanford scheme. The reconstruction quality is shown in Fig. 9 and the RD performance of each frame at a constant bit rate is shown in Fig. 10. As we can see, the RD performance of our scheme is better than the SPIHT scheme. The reconstructed image is acceptable according to the PSNR is about 30 dB [6]. Fig. 10 shows the comparison between and Stanford method [4] in different channel bandwidth. As we can be seen, the proposed method can solve the DCT “block effect”, and the video output quality is improved by 1~2dB in PSNR.



Fig. 8. Decoded frame #60 of Foreman (from left to right: the original image, transmitting LLL, transmitting LLL+LLH, transmitting LLL+LLH+H2).



Fig. 9. Reconstruction quality at 80 Kbps (PSNR = 28.6 dB) (left: the original image; middle: Stamford scheme; right: the proposed scheme).

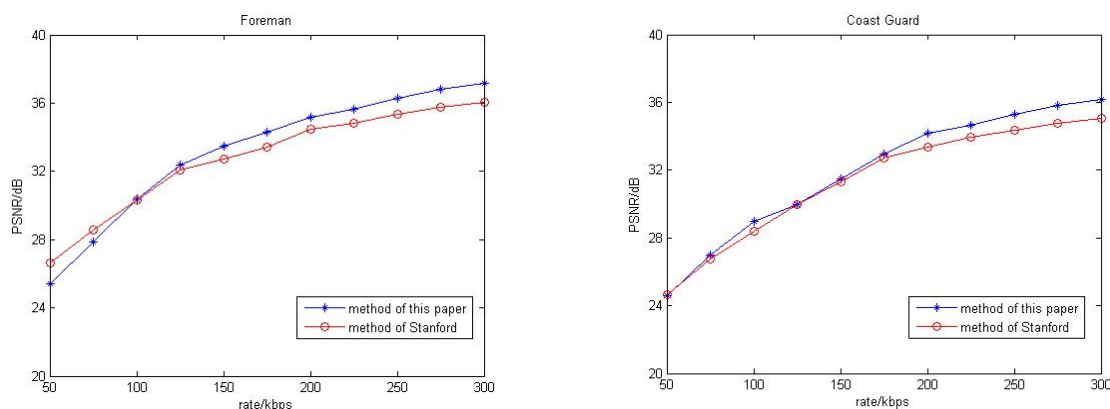


Fig. 10. RD performance (left: *foreman* sequence; right: *coastguard* sequence).

4. Conclusions

In this paper, we propose a wavelet domain distributed coder for video which allows for scalability and does not need feedback channel. At the encoder, a simplified integer wavelet transform algorithm is adopted to improve the computational efficiency. Then according to the channel condition, either the low-frequency coefficients or the low frequency plus several high frequency coefficients are selected. At the decoder, the video sequence is reconstructed by the received coding rate and the generated side information. The simulation results show that the proposed scheme can reduce the transmission bit rate with a good video transmission quality.

Acknowledgments

This work is financially supported by National Science Foundation of China (60872018), the Science and Technology Research Program of Hubei Provincial Department of Education (D20121408), and Program for Research Innovation of Nanjing Institute of Technology (CKJ20110006, QKJA201203).

References

- [1]. J. Slepian, J. Wolf., Noiseless Coding of Correlated Information Sources, *IEEE Transactions on Information Theory*, Vol. 19, No. 4, 1974, pp. 471- 480.
- [2]. A. D. Wyner, J. Ziv, The Rate-distortion Function for Source Coding with Side Information at the Decoder, *IEEE Transactions on Information Theory*, Vol. 22, No. 1, 1976, pp. 1-10.
- [3]. X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, M. Ouaret, The DISCOVER codec: Architecture, Techniques and Evaluation, in

Proceedings of the Picture Coding Symposium, 2007, pp. 144-150.

- [4]. B. Girod, A. M. Aaron, S. Rane, D. Rebollo-Monedero, Distributed video coding, in *Proceedings of the IEEE*, Vol. 93, No. 1, 2005, pp. 71-83.
- [5]. R. Puri, A. Majumdar, K. Ramchandran, PRISM: a video coding paradigm with motion estimation at the decoder, *IEEE Trans. on Image Processing*, Vol. 16, No. 10, 2007, pp. 2436-2448.
- [6]. C. Tang, Ngai-Man Cheung, Antonio Ortega, C. S. Raghavendra, Efficient Inter-band Prediction and Wavelet Based Compression for Hyperspectral Imagery: A Distributed Source Coding Approach, in *Proceedings of the IEEE Data Compression Conference (DCC'05)*, 2005, pp. 437-446.
- [7]. J. E. Fowler, M. Tagliasacchi, B. Pesquet-Popescu, Wavelet-Based Distributed Source Coding of Video, in *Proceedings of the European Signal Processing Conference, Antalya, Turkey, September 2005*, pp. 342-346.
- [8]. Bo Wu, Xiangyang Ji, Debin Zhao, Wen Gao, Wavelet Based Distributed Video Coding with Spatial Scalability, in *Proceedings of the IEEE International Symposium on Circuit and Systems*, 2008, pp. 3458-3461.
- [9]. Daubechies I., Sweldens W., Factoring Wavelet Transforms into Lifting Steps, *J. Fourier Anal*, Vol. 4, No. 3, 1994, pp. 247- 269.
- [10]. Said A, Pearlman W, A new fast and efficient image codec based on set partitioning in hierarchical trees, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 3, 1996, pp. 243-250.
- [11]. R. Bernardini, R. Rinaldo, P. Zontone, D. Alfonso, Wavelet domain distributed coding for video, in *Proceedings of the IEEE International Conference on Image Processing*, October 2006, pp. 245-248.
- [12]. Feng Ye, Aidong Men, He Xiao, and Jinhong Di., Feedback-free Distributed Video Coding

Using Parallelized Design, In *Proceedings of the Picture Coding Symposium (PCS'12)*, February 2012, pp. 217-220.

[13]. Minghu Wu, Xiuchang Zhu, Zongliang Gan, Xin Li, Adaptive Dictionary Learning for Distributed Compressive Video Sensing,

Journal of JDCTA, Vol. 6, No. 4, 2012, pp. 141-149.

[14]. Fu Xiao, Jinkai Liu, Lijuan Sun, Ruchuan Wang, A New Side Information Estimate Based Distributed Video Coding Algorithm for Wireless Multimedia Sensor Networks, *Journal of IJACT*, Vol. 4, No. 1, 2012, pp. 320-326.

2013 Copyright ©, International Frequency Sensor Association (IFSA). All rights reserved.
(<http://www.sensorsportal.com>)



International Frequency Sensor Association Publishing Call for Books Proposals

Sensors, MEMS, Measuring instrumentation, etc.



Benefits and rewards of being an IFSA author:

1

Royalties

Today IFSA offers most high royalty in the world: you will receive 50 % of each book sold in comparison with 8-11 % from other publishers, and get payment on monthly basis compared with other publishers' yearly basis.

2

Quick Publication

IFSA recognizes the value to our customers of timely information, so we produce your book quickly: 2 months publishing schedule compared with other publishers' 5-18-month schedule.

3

The Best Targeted Marketing and Promotion

As a leading online publisher in sensors related fields, IFSA and its Sensors Web Portal has a great expertise and experience to market and promote your book worldwide. An extensive marketing plan will be developed for each new book, including intensive promotions in IFSA's media: journal, magazine, newsletter and online bookstore at Sensors Web Portal.

4

Published Format: printable pdf (Acrobat).

When you publish with IFSA your book will never go out of print and can be delivered to customers in a few minutes.

You are invited kindly to share in the benefits of being an IFSA author and to submit your book proposal or/and a sample chapter for review by e-mail to editor@sensorsportal.com. These proposals may include technical references, application engineering handbooks, monographs, guides and textbooks. Also edited survey books, state-of-the art or state-of-the-technology, are of interest to us. For more detail please visit: http://www.sensorsportal.com/HTML/IFSA_Publishing.htm

Promoted by IFSA

Status of the CMOS Image Sensors Industry Report up to 2017

The report describes in detail each application in terms of market size, competitive analysis, technical requirements, technology trends and business drivers.

Order online:

http://www.sensorsportal.com/HTML/CMOS_Image_Sensors.htm