Research on Distributed Video Compression Coding Algorithm for Wireless Sensor Networks

1, 2 HU Linna, 2 CAO Ning, 3 SUN Yu
1 Department of Dianguang, Nanjing University of Science & Technology Zijin College, No. 89 Wenlan Road, Xianlin University City, Nanjing, Jiangsu 210046, China
2 Computer & Information College, Hohai University, No. 8, Focheng West Road, Nanjing, Jiangsu 210098, China
3 China Mobile Jiangsu Corporation, No. 8, Dinghuamen Avenue, Nanjing, Jiangsu, 210036, China
1 Tel.: +8615951957966
1 E-mail: hulinna@139.com

Received: 13 May 2013 /Accepted: 19 July 2013 /Published: 31 July 2013

Abstract: The Distributed Video Coding (DVC) which has the feature of simple encoding and complex decoding can solve the problems that occur when Wireless Sensor Networks (WSN) nodes process and compress information in the resource-constrained conditions. Transitional Wyner-Ziv distributed video coding scheme adopting regional unified coding, may lead to the distortion problem of decoding estimation of the intense motion region. Aiming at the limitation of conventional Wyner-Ziv distributed video coding, an improved Wyner-Ziv distributed video coding algorithm which based on region of interest (ROI) extraction is proposed in this paper. The improved algorithm extracts the regions of the Wyner-Ziv frame which are intense motion area and cannot coding and decoding accurately, then compresses based on entropy coding. Simulation results show that the proposed algorithm can reduce the rate and improve the quality of decoded image, at the same time decrease the energy consumption of sensor nodes. Copyright © 2013 IFSA.

Keywords: Wireless sensor networks (WSN), Distributed video coding (DVC), Wyner-Ziv, Region of interest (ROI), Entropy coding.

1. Introduction

As the rapid development of wireless communication technology and the Interact, it is possible for real-time accessing to multimedia data, especially information-rich high-resolution video data. However, in wireless of networks, due to sensor network resources by the nodes, the transmission bandwidth constraints for the network video coding algorithm must be simple and easy to achieve, coding efficiency, but also strong anti-error capacity [1]. Nowadays, many typical Video Compression Coding Standards (like MPEG-X, H.26X) adopt motion compensation on the encoder side. So the complexity of Encoder is 5 to 10 times then the Decoder that can't meet the demand of Wireless Multimedia Sensor Networks which not only have a limited network resource but also the low complexity and low power consumption node [2]. In response to these requirements, distributed video coding (DVC) has been considered as an approach to shift the computational burden of the encoder to the decoder. Since the Distributed Video Coding adopt intra-frame encode and Inter-frame decode, and excavate the dependence of video sequence on the Decoder, so the complexity of Encoder is lower and make the
Encoder brief. The distributed video coding is very applicable to solve the video coding issues on the Wireless Multimedia Sensor Networks [3].

Distributed video coding (DVC) driven by emerging requirements in several applications, has sparked much research interest in recent years. Distributed video coding (DVC) shifts the computation complexity from the encoder to the decoder. Therefore, the correlation between the frames is not exploited anymore at the encoder but only at the decoder.

DVC is based on two results of information theory appeared in the 70’s [4]: the Slepian-Wolf theorem[5] for distributed lossless compression, and the Wyner-Ziv theorem [6] for the lossy compression of a signal X with receiver side information Y, a correlated signal of signal X. The latter theorem is the driving force for many practical DVC schemes, the so-called Wyner–Ziv (WZ) video encode.

In many Wyner–Ziv DVC implementations [7-10], a video sequence is divided into two interleaving sub-sequences: key frames and Wyner–Ziv (WZ) frames. The encoder uses intra-coding to encode the key frames while the decoder uses the reconstructed key frames to predict the WZ frames. The predictors of WZ frames are called side information (SI). SI prediction errors are then corrected by channel codes [11] to reconstruct the WZ frames. SI prediction, Slepian-Wolf encodes, and WZ reconstruction are the most critical components in a DVC codec.

In [8, 12, 13], the turbo-based or LDPC-based encoding has been presented. This coding scheme treats all regions of the Wyner-Ziv frames equally. However, it can’t accurately predict regions of drastic motions as well as edges of moving objects, with using the motion estimation technology. And, the decoding process would ask the encoder for a large amount of feedback information. In this way, not only the code rate would be increased but parts of the decoded images may still be not precisely enough.

Transitional scheme adopting regional unified coding, may lead to the distortion problem of decoding estimation of the intense motion region. In this paper, in order to overcome the limitation of conventional Wyner-Ziv distributed video coding, an improved Wyner-Ziv distributed video coding algorithm which based on region of interest (ROI) extraction is proposed. The improved algorithm extracts the regions of the Wyner-Ziv frame which are intense motion area and cannot coding and decoding accurately according to the criteria of region of interest, then compresses based on entropy coding. However, other regions which are encoding and decoding based on LDPC.

The paper is organized as follows. Section 2 describes the architecture of the Wyner-Ziv video coding system being used. In Section 3, an improved Wyner-Ziv distributed video coding algorithm is presented. Experimental results are presented and discussed in Section 4. Finally, Section 5 concludes the paper.

2. The Wyner–Ziv Video Coding Architecture

The Wyner–Ziv video codec is divide into (pixel domain Wyner–Ziv) and TDWZ (transform domain Wyner–Ziv).

The WZ codec discussed in this paper based on the transform domain coding with feedback channel. Fig. 1 illustrates the WZ video coding architecture. In a short, the adopted WZ codec works as follows: first, a video sequence is divided into key frames and WZ frames. The key frames are coded using the H.264/AVC Intra coding scheme while WZ frames are coded by applying a 4×4 H.264/AVC integer discrete cosine transform (DCT), followed by the construction of the DCT coefficients bands, which represent the DCT coefficients of the entire frame at the same position within the 4×4 blocks. Then, each DCT band is uniformly quantized, and bit planes are formed and sent to the turbo encoder which generates the corresponding parity sequences. Only the parity bits are stored in the buffer, punctured and transmitted in small amounts upon decoder request via the feedback channel.

![Fig. 1. The Wyner–Ziv video coding architecture.](image-url)
Distributed video coding (DVC) makes the decoding videos improved greatly, and meanwhile, it can obtain high compression efficiency. However, neither pixel-domain-based nor transform-domain-based DVC plans can avoid the weakness in decoding motion videos, which would often cause serious double images or ghost images after they have been decoded. Later, some researchers proposed in their literatures the spatial smoothing processing, which helped to enhance the quality of motion video decoding. Yet, some drastic motions still can’t be well treated, and these drastic motions are exactly the root causes of video problems. The estimation algorithm, used in for decoding by a DVC system, will not generate correct motion information because the motion amplitude may go beyond its prediction range, and thus decoding errors would appear.

3. The Improved DVC System Encoding and Decoding Scheme

In Wyner-Ziv video encoding, the video frames are generally divided into two types: key frame (K-frame) and Wyner-Ziv frame (WZ-frame), where, the K-frame is usually encoded by the conventional H.264, while the WZ-frame is encoded by Wyner-Ziv. At decoder side, WZ-frame uses check codes and the matched side information that sent from the decoder to make reconstruction. Here, the side information is a prediction to Wyner-Ziv frames. With strongly related to the Wyner-Ziv frames, these side information are obtained from neighboring frames, and can help decoder to decode the Wyner-Ziv frames. For encoding a Wyner-Ziv video, the motion estimation technology at the decoder side is to predict the motion vector in accordance with the relevance between frames. A weak relevance between side information of the decoder and Wyner-Ziv frames will lead to an inaccurate estimation, which hence affect the efficiency of decoding. In this paper, regions of drastic motions and edges of moving objects are extracted as ROI (regions of interest). Besides, entropy coding is employed to improve the encoding efficiency and the quality of decoded images.

3.1. The Improved DVC System Encoding and Decoding Framework

In the encoding and decoding framework, as shown in Fig. 2, the video stream that needs to be encoded is divided into K-frames, which are encoded by conventional method, and WZ-frames, which are encoded and decoded by a combination of LDPC and entropy. At encoder side, H.264/AVC encoding plan is employed for K-frames, while WZ-frames are encoded by ROI partition algorithm based on motion features; using ROI judging criteria to divide the micro-block into ROI micro-blocks, which take the entropy for encoding and decoding, and Non-ROI micro-blocks, which take LDPC for encoding and decoding. At decoder side, use the entropy to decode ROI micro-blocks, and use LDPC to decode Non-ROI micro-blocks. After finishing the decoding, the decoded data should be treated with IDCT (Inverse Discrete Cosine Transform) and Inverse Quantization to reconstruct the image and complete the video compression.

3.2. Extraction of Regions of Interest (ROI)

In a traditional ROI model, the original gray or color images will be firstly extracted as feature distribution, for example, color brightness information along edge directions. Specifically, it is to extract the ROI eyes according to spatial distribution of each feature, and then integrate the ROIs of all extracted features to get the final ROI. In proposed algorithm, taking micro-blocks as the basic unit, calculates the motions of each micro-block to determine whether it belongs to the ROI motion region or not.
In this improved algorithm, the ROI extraction is implemented mainly by using the motion information of micro-blocks, and the regions of eyes visual system that commonly moves fast has also been paid close attention to. Motions of micro-blocks here are described by defining the motion factor $m(i, j)$:

$$m_x(i, j) = 1 + \frac{\max(V_x(m_{x, max}(i, j)}{m_{x, max}(i, j)}),$$

(1)

$$m_y(i, j) = 1 + \frac{\max(V_y(m_{y, max}(i, j)}{m_{y, max}(i, j)},$$

(2)

$$m(i, j) = m_x(i, j) \cdot m_y(i, j),$$

(3)

where, $(i, j)$ is the position of current micro-block, $m_x(i, j)$ and $m_y(i, j)$ expresses the horizontal direction motion factor and vertical direction motion factor of the micro-block; $V_x(i, j)$ and $V_y(i, j)$ represents the horizontal direction motion velocity and vertical direction motion velocity of the micro-block; and $V_x, \max(i, j)$ and $V_y, \max(i, j)$ is the maximum value along $V_x(i, j)$ and $V_y(i, j)$, respectively.

In this paper, we use an approach that based on absolute threshold to determine whether a micro-block belongs to the ROI motion region. After obtaining the motion factor $m(i, j)$ by calculating formula (3), we take ROI micro-blocks judging criteria to compare the factor $m(i, j)$ with the threshold $th(i, j)$ and make judgment by setting a threshold. If $m(i, j) \geq th(i, j)$, it means the micro-block belongs to ROI, otherwise it not belongs to ROI. Once the ROI micro-blocks have been judged, different micro-blocks shall be encoded by different ways, i.e. use the entropy to encode ROI micro-blocks, and use LDPC to encode Non-ROI micro-blocks. The value of threshold shall be manually set in accordance with different features of the video image. Generally, a best efficiency would be got when the ROI occupies 20 % of the whole video frames.

4. Experimental Results and Analysis

Assume that the sensor nodes have already collected video data that will be compressed and encoded in the experiment. The simulation is conducted on Matlab platform, which includes quantizer, DCT module, ROI extraction module, encoder, side information generator, decoder, and reconstruction module.

In this experiment, Wyner-Ziv frames are encoded and decoded by a combination of LDPC and entropy, in which, the check matrix of LDPC codes is generated by PEG method. Through the adjustment to different quantitative coefficients, different code rates can be outputted to obtain different compression rates. Meanwhile, key frames are encoded by using JPEG, and the encoding sequence is “K-W-K-W”, i.e. K-frames and WZ frames are encoded alternately.

This experiment takes two standard video sequences: Akiyo and Foreman, with image format is QCIF (176×144) and the encoding rate is 100 frames (30 fps). Further, comparisons have been made between the ROI partition-based algorithm, Wyner-Ziv encoding algorithm, H.264/AVC algorithm and the conventional JPEG algorithm.

![Fig. 3. The simulation results of improved DVC system.](image)

Fig. 4 (a) and (b) shows, respectively, the decoded images corresponding to the 25th frames of Akiyo that encoded by using the conventional LDPC and by using the proposed algorithm of this paper. It can be seen that, the improved ROI algorithm proposed in this paper has a better visual effect when compared to the LDPC encoding algorithm. Besides, the proposed ROI algorithm not only won’t increase extra consumption at encoder side, but also can compensate the motions at decoder side and make full use of redundant information of the frames, and thus obtaining a better rate-distortion performance.
5. Conclusions

This paper presents, on the ground of the Wyner-Ziv video encoding, a distributed video coding (DVC) plan that based on ROI motion regions which are extracted at the encoder side to use the entropy to encode those regions that can not be accurately predict by motion estimation. Next, apply entropy encoding for the ROI motion regions to protect these areas, so as to improve the quality of video decoding. Such method enriches Wyner-Ziv DVC’s theoretical system. DVC has such features as high compression efficiency, strong robustness, simple encoding, complex decoding, easy for classification and multiple description. Therefore, it is ideal for WMSN (Wireless Multimedia Sensor Networks) which is of low power consumption, small storage and limited computation capability. Effectiveness of the proposed algorithm can be verified by experimental results. For later works, we should take into consideration how to self-adaptively adjust the threshold in line with image features to further improve the quality of the decoded image.

References