Neurocomputing 162 (2015) 180–190

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Model-based low bit-rate video coding for resource-deficient wireless visual communication

Xianming Liu [a,*], Xinwei Gao [a], Debin Zhao [a], Jiantao Zhou [b], Guangtao Zhai [c], Wen Gao [d]

[a] School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China
[b] Faculty of Science and Technology, University of Macau, E11 Avenida da Universidade, Taipa, Macau, China
[c] Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai 200240, China
[d] School of Electronic Engineering and Computer Science, Peking University, Beijing, China

ARTICLE INFO

Article history:
Received 9 October 2014
Received in revised form
16 January 2015
Accepted 24 March 2015
Communicated by Y. Yuan
Available online 14 April 2015

Keywords:
Wireless visual communication
Low bit-rate video coding
Low complexity

ABSTRACT

In this paper, an effective low bit-rate video coding scheme is developed to realize state-of-the-art video coding efficiency with lower encoder complexity, while supporting standard compliance and error resilience. Such an architecture is particularly attractive for application scenarios involving resource-deficient wireless video communications. At the encoder, in order to increase resilience to channel error, multiple descriptions of a video sequence are generated in the spatio-temporal domain by temporal multiplexing and spatial adaptive downsampling. The resulting side descriptions are interleaved with each other in temporal domain, while still with conventional square sample grids in spatial domain. As such, each side description can be compressed without any change to existing video coding standards. At the decoder, each side description is first decompressed, and then reconstructed to the original resolution with the help of the other side description. In this procedure, the decoder recovers the original video sequence in a constrained least squares regression process, in which 2D or 3D piecewise autoregressive model is adaptively chosen according to different predictive modes. In this way, the spatial and temporal correlation is sufficiently explored to achieve superior quality. Experimental results demonstrate that the proposed video coding scheme outperforms H.264/AVC and other state-of-the-art methods in rate–distortion performance at low bit-rates and achieves superior visual quality at medium bit rates as well, while with lower encoding computational complexity.

1. Introduction

In recent years, low-cost devices such as CMOS cameras that are able to ubiquitously capture video content from the environment have appeared in almost all small wireless mobile devices, such as smart-phones and tablet PCs. Furthermore, recent developments in sensor networking have encouraged the use of video sensors in these networks, which has fostered the development of Wireless Video Sensor Network (WVSN) [1–5]. Following the trends in low-power processing, wireless networking, and distributed sensing, WVSN has developed as a new technology with a number of potential applications, ranging from mobile multimedia, security monitoring, and advanced health care delivery to emergency response.

WVSN will be composed of inter-connected, battery-powered miniature video sensors. WVSN is usually mission driven and application specific, therefore it must operate under a set of unique constraints and requirements. The main concern for WVSN is that the energy provisioned for a video sensor is not expected to be renewed throughout its mission because sensor nodes may be deployed in a hostile or unpractical environment. And the bandwidth of wireless channels is usually limited. At the same time, it is necessary to provide some error-resilient mechanism against instability of wireless channels. More specially, there are several factors that mainly influence the design of a WVSN:

- *Compression efficiency, encoder complexity and bandwidth*: An effective video compression scheme can significantly reduce the amount of video data to be transmitted, which in turn saves a significant amount of energy in data transmission. However, more effective video compression methods often require higher computational complexity. These two conflicting effects imply that in practical system design there is always a tradeoff among compression efficiency, encoding complexity and bandwidth.
- *Resiliency to channel errors*: Wireless channels are unstable and noisy. Therefore, the source coder should provide some mechanism for robust and error-resilient coding of source data.

There exists a vast literature on video coding techniques. In traditional hybrid video compression standards, such as MPEG-2

* Correspondingauthor.
E-mail address: xmliu.hit@gmail.com (X. Liu).

http://dx.doi.org/10.1016/j.neucom.2015.03.054
0925-2312/

[6], H.264/AVC [7] and HEVC, consecutive frames are encoded jointly to achieve maximum coding efficiency, which are based on the idea of predictive coding to exploit spatio-temporal correlations. Although achieving the state-of-the-art rate-distortion performance, they may not be suitable for low-cost video sensors since predictive coding requires complex encoders and entails high energy consumption. Besides, the predictive coding system causes inter-frame dependency in decompression, which results in error propagation for an error-prone channel. Another approach in conventional wisdom is distributed video coding (DVC) [8–10], which shifts the complexity to the decoder end to allow the use of simple encoders. At the same time, DVC has an inbuilt robustness to channel losses because there is a duality between distributed source coding and channel coding. Clearly, DVC is very promising for WVSN. However, despite years of intensive research on DVC, the current systems still fail to meet the compression efficiency of their predictive coding counterparts [11–13]. Moreover, the feedback channel is required in existing mainstream DVC schemes, which will introduce delay and therefore is not suitable for practical WVSN applications.

In this paper, we investigate the sparse sampling based approach for wireless video communication. The fact that natural videos have an exponentially decaying power spectrum in spatial domain and strong correlations in temporal domain suggests the possibility of interpolation-based compact representation of video signals. We find that sparse sampling naturally fulfills the role of encoder because it greatly reduces the amount of data. To achieve maximum coding efficiency, the downsampled data should be further compressed. For this purpose we choose a uniform downsampling scheme to generate conventional square sample grids with smaller size. In this way, the information needed to be compressed and transmitted is significantly reduced, so that the proposed scheme can greatly reduce the encoder complexity, and naturally prolong the operational lifetime of video sensors. Similar to DVC, the proposed scheme provides an asymmetric video codec design, which shifts the associated computation burdens to the decoder. In this system design, the heavy-duty video decoding can be performed by powerful computers or perceivably in near future by cloud computing.

On the other hand, the needs for wireless-network-aware coding techniques to mitigate the problem of packet losses have generated much interest in multiple description coding. In the more simple architecture, MDC of a video source consists in generating two equivalent importance data streams that, all together, carry the input information. At the receiver side, when both the descriptions are available a high quality video is reconstructed. If only one bit-stream is available at the decoder end, a poorer but acceptable quality reconstruction is obtained. MDC has emerged as a promising approach to enhance the error resilience of a video delivery system. However, most of the existing video MDC techniques [14-17,19,27-29] are not compliant to existing video coding standards, either being completely different approaches or requiring a significant degree of modifications to an existing standard. Moreover, since introducing redundancy for error resilience, the compression efficiency of MDC based schemes is usually lower than the conventional video coding standard. A natural question is whether combing compact signal representation with MDC can be made a practical and competitive solution for wireless video communications. In this paper we will give an affirmative answer to the above question.

The rest of this paper is organized as follows. In Section 2, we overview the related work. Section 3 presents the framework of the proposed scheme. Sections 4 and 5 detail the main contribution of this paper: mode-dependent soft-decoding via constrained least squares regression. Section 6 presents the experimental results and comparative studies. Section 7 concludes the paper.

2. Related work

In the literature, many interpolation-based low bit-rate image/video compression algorithms have been proposed. In [29], an interpolation-dependent downsampling method was proposed to hinge the interpolation to the downsampling process. In [18], Shen et al. proposed a downsampling based video coding scheme, where an example based super-resolution technique is employed to restore the downsampled frames to their original resolutions. This work needs an offline training set with different video resolution. In some practical cases, the training set is not available. In [20], an adaptive downsampling mode decision in the encoder was proposed. The modes including different directions and sizes can be determined by the features of block contents. Ref. [30] proposes to find the optimal downsampling ratio that balances the distortions caused by downsampling and coding, thus achieving the overall optimal RD performance. However, this method will introduce heavy computation complexity at the encoder side to perform complex rate-distortion mode decision. All the above mentioned methods need modification on the current image/video coding standard, which limits their practicability. Compared with these methods, the proposed scheme is standard compliant. We envision that our scheme can be a useful enhancer of any existing video compression standard, by just adding pre- and post-processing modules, to improve low bit-rates compression performance.

3. The framework of model-based low bit-rate video coding scheme

3.1. Encoder

The major concern of our system design is to provide a light-duty encoder under energy consumption and bandwidth constraints, meanwhile, have the ability to mitigate the problem of packet losses during

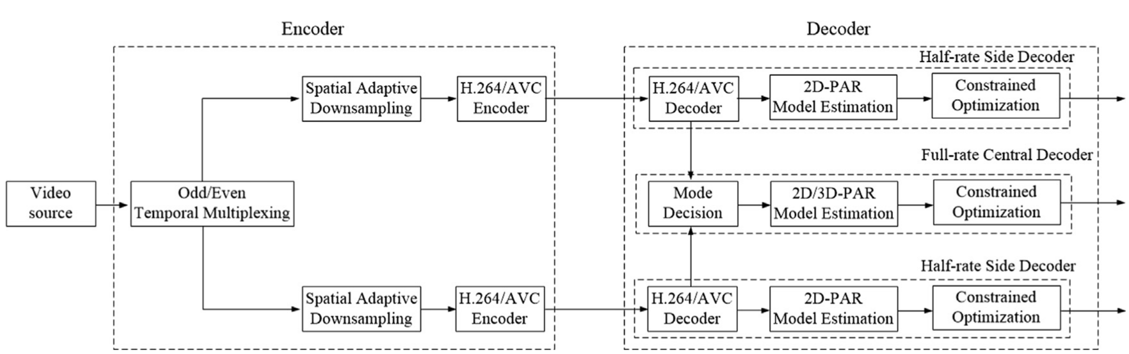

Fig. 1. The framework of the proposed video coding scheme.