J. Vis. Commun. Image R. 24 (2013) 458–464

Contents lists available at SciVerse ScienceDirect

J. Vis. Commun. Image R.

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

Distributed video coding with progressive significance map

Yang Hu [1], William A. Pearlman *

Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

ARTICLE INFO

Article history:
Received 29 June 2012
Accepted 10 February 2013
Available online 17 February 2013

Keywords:
Distributed video coding
Set partition coding
SPIHT coding
Significance map
Wavelet transform
Slepian–Wolf coding
LDPCA codes
Error resilience

ABSTRACT

A distributed video coding (DVC) system based on wavelet transform and set partition coding (SPC) is presented in this paper. Conventionally the significance map (sig-map) of SPC is not conducive to Slepian–Wolf (SW) coding, because of the difficulty of generating a side information sig-map and the sensitivity to decoding errors. The proposed DVC system utilizes a higher structured significance map, named progressive significance map (prog-sig-map), which structures the significance information into two parts: a high-level summation significance map (sum-sig-map) and a low-level complementary significance map (comp-sig-map). This prog-sig-map alleviates the above difficulties and thus makes part of the prog-sig-map (specifically, the fixed-length-coded comp-sig-map) suitable for SW coding. Simulation results are provided showing the improved rate-distortion performance of the DVC system even with a simple system configuration.

1. Introduction

Distributed video coding (DVC) is a class of distributed source coding (DSC) applied to video source data. It is currently a subject of intense research. In a number of articles in the DSC literature, set partition coding (SPC), usually combined with wavelet transform, is used as the source coder preparing the data for Slepian–Wolf (SW) coding. SPC is a widely used technique for image and video compression, used in the JPEG 2000 [1,2] image compression standard and the new 2D lossy-to-lossless compression algorithm recently standardized by the Consultative Committee for Space Data Systems (CCSDS) [3,4]. It has also been successfully applied to several other well-known image coding algorithms including SPIHT [5], SPECK [6], SWEET [7], SBHP [8], and EZBC [9]. A comprehensive tutorial on SPC and its usage in wavelet coding systems can be found in the monographs [10,11] and the textbook [12].

A salient feature of SPC is that its compressed bitstream divides naturally into two components: a significance map (sig-map) that conveys location information; and a value bitstream that conveys intensity information of signs and lower order bits of wavelet coefficients. According to the DVC tradition, the module that reflects the DSC principle is referred to as a Slepian–Wolf codec, which includes an SW encoder and an SW decoder. The frame to be SW coded is referred to as an SW frame. In most cases of DSC, only the value data of an SPC bitstream are SW coded. For example, Guo et al. [13] codes the SW frame with single-pass ZTE [14] coding, where the significance map is coded in normal intra-frame or intra mode (sent out directly without SW coding), and only significant coefficients are further SW coded. A similar strategy is used when the SW frames are SPIHT coded [15,16]. In another work, a modified quadtree scanning is applied, wherein the decision bits from the quadtree scanning, similar to the significance map, is still intra coded [17].

The significance map occupies about 60% of the SPC bitstream for bit rates giving recovery quality measured by PSNR's in the 30–40 dB range, which covers almost all video viewing experiences. An efficient DSC system should try to SW code the sig-map in addition to the value information. However, the conventional sig-map is not conducive to SW coding, mainly because of two difficulties.

- It is hard or impossible to generate a qualified side information (SI) sig-map at the decoder, which should be of the same length as the original sig-map and be simultaneously bit-wise correlated with the original one.
- The conventional sig-map is highly error sensitive. Decoding errors are unavoidable for the SW decoder even though the error rate could be very low. Due to error propagation and the consequent loss of coding synchronization, even a single decoding error can be catastrophic.

To alleviate these difficulties, there are two existing strategies, both of which try to express (part of) the significance locations with fixed-

* Corresponding author. Fax: +1 518 276 8715.
E-mail addresses: yanghuyh@gmail.com (Y. Hu), pearlw@ecse.rpi.edu (W.A. Pearlman).

[1] Y. Hu is now with Cisco Systems, Santa Clara, CA, USA.

1047-3203/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jvcir.2013.02.003

length codes. In the first strategy the SPC is abandoned, and the significance locations, which are the locations of significant sets and pixels, are coded via a position-by-position scan of the DWT coefficients [18]. The second strategy extracts a fixed-length-coded part of the location information and SW codes just that part. For example, one method [19] separates the sig-map of SPIHT into two parts called SD and SP. The SD information, indicating whether or not a tree is a zerotree, is intra coded (without SW coding). For non-zerotrees, SP bits are sent out indicating the significance of every tested subset (coefficients and/or descendant subtrees) in the current tested tree. In this way, the SP is of fixed length. The decoder can produce side information SP′ with the help of SD. With either strategy, the sig-map alleviates the difficulty of SI generation, thanks to the fixed length property. Such a sig-map however still faces the hurdle of loss of synchronization. Even a single bit error in the decoded sig-map (or the SP) may destroy the whole decoding that follows.

The progressive significance map (prog-sig-map) [20] introduces a layered structure to represent the significance information. It consists of a summation significance map (sum-sig-map) and a *fixed-length-coded* complementary significance map (comp-sig-map). The sum-sig-map occupies about 60% and the comp-sig-map about 40% of the prog-sig-map for the common video bit rates. Thanks to the constraint of the sum-sig-map and the fixed-length property of the comp-sig-map, the decoding errors associated with the comp-sig-map would not propagate within the comp-sig-map or to the decoding of the sum-sig-map. In this paper, we propose a wavelet-based DVC system using SPC with the prog-sig-map to make both the comp-sig-map (the fixed-length part of the prog-sig-map) and the value information suitable for SW coding.

To our knowledge, there was only one previous attempt to SW encode the significance map. In this work [21], the significance map was divided into bits from direct and indirect descendants, which were separately SW encoded in each bit plane. Sychronization of these relatively short bitstream segments can be maintained between the key and SW frames only when they are very highly correlated. For the multi-spectral images in their simulations, that was the case, but for video sequences, the target source in this paper, such synchrony could not be maintained. Thus, this method is inapplicable to video.

The rest of the paper is organized as follows. A brief overview of the progressive significance map is provided in Section 2. We introduce the proposed DVC system in Section 3. Modules in the system are detailed in Section 4. Section 5 provides the simulation results. Finally, Section 6 concludes the paper.

2. Overview the progressive significance map

A significance map conveys the location information of significant sets and pixels, generated from the significance tests (sig-tests), and indicates the execution path of an SPC process. There are various set structures and partitioning rules in different SPC coders (such as the spatial orientation tree in SPIHT [5], the quadtree and octave banding partitioning in SPECK [6] and SWEET [7], etc.). They all use the core sig-test that can be uniformly referred to as the (c,w) test. The c denotes the number of *candidates*, defined as the sets (including singleton sets) to be sig-tested; the w denotes the number of *winners*, defined as the newly found significant sets in the sig-test. Conventionally each candidate is assigned a binary indicator I signalling whether or not it is a winner ($I = 1$, if winner; $I = 0$, otherwise). In this way, a (c,w) test generates c indicators $\{I_1, I_2, \ldots, I_c\}$, which go to the conventional sig-map.

The progressive sig-map also consists of the results of a series of correlated (c,w) tests, but in a layered structure, using a sum-sig-map and a comp-sig-map. The sum-sig-map conveys how many of the c candidates are winners. The number of winners w is actually the summation of the c indicators as $w = \sum_{i=1}^{c} I_i$. If $w = 0$ or $w = c$, the c indicators must be all 0's or all 1's, respectively. If $0 < w < c$, there are $M = \binom{c}{w}$ different choices for the w winners out of the c candidates. In this case, the comp-sig-map is introduced. For most 2D SPC methods, the candidate number c satisfies $c \in \{1, 2, 3, 4\}$, so six types of (c,w) tests ((2,1),(3,1),(3,2),(4,1), (4,2), and (4,3)) need a comp-sig-map. In the variable-length comp-sig-map, Huffman codes are used to code a set of M equally probable symbols for each (c,w) type. To build the fixed-length-coded comp-sig-map, additional bits are appended to the shorter Huffman codewords so as to make them of the same length as the longer ones for each (c,w) type. In this way, a (c,w) test has a fixed number $L(c, w)$ of complementary bits, according to $L(c, w) = \lceil \log_2 M \rceil$.

The fixed-length feature of the comp-sig-map enables the generation of a qualified SI comp-sig-map. In addition, if the sum-sig-map is correct, it will produce correct winner numbers. Under the condition of correct winner numbers, the fixed-length comp-sig-map is robust to bit errors, in the sense that bit errors in the comp-sig-map will impact only locally instead of propagating within the comp-sig-map or to the sum-sig-map and value data. In this way, the prog-sig-map alleviates the drawbacks (SI generation difficulty and high sensitivity to decoding errors) of the conventional sig-map in DVC application. Thus, the fixed-length comp-sig-map is a good source for SW coding. For details concerning the compression performance of the prog-sig-map and the error resilience of the fixed-length comp-sig-map, please refer to its original paper [20].

3. Proposed DVC system

The proposed DVC system is illustrated in Fig. 1. This system is an implementation of Wyner's DSC scheme[22] using punctured or rate adaptive low-density parity check (LDPC) codes. Explanations of Wyner's scheme and principles of distributed source coding may be found in the textbook by Pearlman and Said [12]. Different from existing DSC schemes using coders that generate significance maps (sig-maps), part of the sig-map, the fixed-length part called the comp-sig-map, serves as the SW coding object and is robust to decoding errors. In order to highlight the key point (the SW coding of the comp-sig-map), we adopt very simple techniques for other modules. A video sequence is split into *key frames* and *SW frames*. The key frames are coded with conventional intra coding methods, such as H.264 intra coding, JPEG 2000, normal 2D SPIHT, and so forth. The SW frames are coded in a distributed way. A GOF (group of frames) of length 2 is used — the even frames are key frames and the odd frames are SW frames.

An SW frame is first set partition coded to generate its prog-sig-map, comprising the sum-sig-map S and the fixed-length comp-sig-map M, and significant value information V, comprising signs and refinement bits. The sum-sig-map S is sent to the output bitstream directly; the fixed-length comp-sig-map M and the value information V are SW coded, *separately*. An SW encoder includes a channel encoder and a parity generator. The channel encoder computes a set of syndrome bits. These syndrome bits are buffered and incrementally sent to the decoder which iteratively asks for augmented syndrome bits via the feedback channel until successful decoding is achieved. The parity generators calculate error-checking bits, that are sent to the decoder to help detect false positive decoding.

The decoding of a SW frame is more complex. The key frames are assumed to be losslessly available at the decoder. An estimation of the original SW frame is calculated from the available key