Signal Processing: Image Communication 47 (2016) 86–95

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Fast watermarking scheme for real-time spatial scalable video coding

Adamu Muhammad Buhari [a,**], Huo-Chong Ling [b], Vishnu Monn Baskaran [a], KokSheik Wong [c,*]

[a] *Faculty of Engineering, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia*
[b] *Faculty of Engineering & Science, Curtin University, 98000 Miri, Sarawak, Malaysia*
[c] *Faculty of Computer Science & Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia*

ARTICLE INFO

Article history:
Received 26 July 2015
Received in revised form
2 June 2016
Accepted 3 June 2016
Available online 4 June 2016

Keywords:
Watermarking
H.264/SVC
Spatially scalable
Fast watermarking
Video

ABSTRACT

Recent advancements in high resolution scalable video coding have significantly increased the computational complexities of watermarking solutions for real-time video encoding systems. This paper proposes a human visual system based watermarking algorithm for spatial scalable coding based on the H.264/SVC standard. The proposed algorithm extracts textural feature from a set of 7 high energy quantized coefficients in 4×4 luma INTRA-predicted blocks of all slices and embeds watermark into the highly textured block which has at least one non-zero coefficient in 6 selected locations. The same embedding process is performed for all layers of the video to improve robustness against common video processing attacks. Experiments were conducted by embedding up to 8192 watermark bits into a four-layer spatial scalable coded video. Results suggest that the proposed scheme produces watermarked video with an average visual quality degradation of ~0.36 dB at the expense of 2.18% bitrate overhead. In addition, the proposed watermarking scheme achieves an average detection rate of 0.98, 0.97 and 0.71 against re-encoding, recompression and Gaussian filtering attacks, respectively, using the base layer, and above 0.97, 0.94 and 0.66, respectively when using any of the enhancement layer.

1. Introduction

Digital watermarking [1] is a technique of inserting a signal (i.e., a watermark) into a digital media to form a variant of the original media containing the watermark information. The inserted watermark is imperceptible but it should be detectable for proof of ownership, authentication, tampering detection, and other applications. The presence of watermark in a digital media does not prevent the observer from either viewing, examining or changing the media content. However, the watermark is non-removable without significantly affecting the quality of watermarked media. Currently, watermarking for copyright protection has significantly reduced unauthorized or illegal distribution of digital media over the Internet [2,3].

In general, techniques for watermarking can be classified as either public or private. Public watermarking is typically used for applications requiring a robust watermark [4], such as identifying the buyers to prevent illegal duplication and distribution. On the other hand, private or non-blind watermarking algorithms require the original media to verify the watermark. Private watermarking is necessary for fragile watermarking applications, such as authentication and tamper detection. In terms of the domain in which the watermark is inserted, watermarking techniques can be further classified into the spatial, transformed and compressed domains. For real-time watermark embedding and extraction, the compressed domain watermarking technique is the best compromise [5,6] as it adopts fast watermarking process. In addition, most media also appear in the compressed form for efficient storage and transmission purposes.

Generally, there are three trade-off parameters when designing a watermarking algorithm, namely data payload, visual quality degradation, and security [7]. Specifically, data payload is the number of bits embeddable in the digital media, visual quality degradation refers to the perceptible difference between the original and the watermarked media (due to the insertion of the watermark), and security infers the level of resistance of the embedded watermark against both intentional and unintentional processings (i.e., robustness). Here, higher security implies greater robustness of the marked content against attacks. However, these parameters are mutually conflicting, where each parameter is selected to fulfill specific requirement(s) of an intended application [7,8]. For example, high watermark payload is preferred for screen annotation [9,10], low visual quality degradation watermark is preferred for medical and scientific media contents [11], while robust watermark is preferred for copyright protection [12].

* Corresponding author.
** Corresponding author.
E-mail addresses: adam_m.buhari@yahoo.com (A.M. Buhari), koksheik@um.edu.my (K. Wong).

http://dx.doi.org/10.1016/j.image.2016.06.003
0923-5965/

In view of the current digital video watermarking solutions for real-time coding systems with Advanced Video Coding (H.264/AVC) [13], several research concerning on this issue were performed. Then, the standardization of H.264/SVC [14] in 2007 further increased the computational complexities of watermarking solutions for real-time video encoding systems. To the best of our knowledge, the conventional schemes in H.264/AVC focuses on robustness and bitrate overhead, and there is still much room for improvement from the perspective of computational complexity. Although High Efficiency Video Coding (HEVC) [15] and the scalability extension of HEVC (SHVC) [16] are the latest video coding standards, most videos are still encoded and available in the H.264 format. In addition, many of the capturing devices, including surveillance camera, dash cam, smart phone, are capturing in the H.264 format. Therefore, fast watermarking algorithm specifically designed for the H.264 standard is much desired.

Therefore, this paper puts forward fast watermarking scheme for H.264/SVC compressed video. The rest of this paper is organized as follows: Section 2 reviews the conventional watermarking schemes, while Section 3 proposes the watermarking algorithm with fast HVS model. Section 4 proposes a fast watermark extraction algorithm for real-time detection. Section 5 presents the experimental results of the proposed watermarking algorithm. Finally, Section 6 concludes this paper.

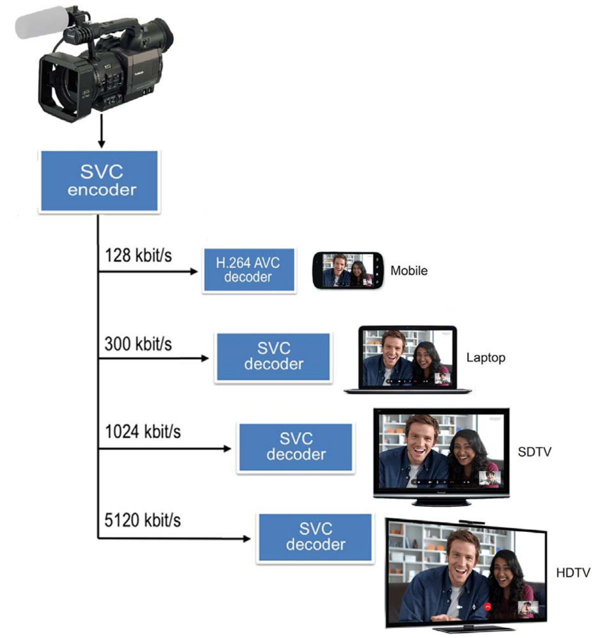

Fig. 1. The principle of SVC stream for spatial scalability.

2. Related work

This section briefly reviews the H.264/SVC coding structure, then surveys the conventional watermarking schemes designed for H.264/SVC compressed video.

2.1. Overview of H.264/SVC

The structure of H.264/SVC encoder is similar to that of H.264/AVC with additional inter-layer prediction and multiplexer to form a scalable bitstream. H.264/SVC supports three main types of scalability which are temporal, quality and spatial scalability [17]. Each of the supported scalability forms a bitstream in such a way that when part of the stream is not available, the resulting sub-stream forms another format-compliant (i.e., valid) bitstream. Specifically, temporal scalable coding forms a bitstream to provide layers with a different frame rate, quality scalable coding forms a bitstream that provides layers with a different visual quality, and spatial scalable coding forms a bitstream that provides layers with a different display resolution.

Fig. 1 shows the principle of scalable coding, which is the type of scalability considered in this work. Specifically, a source video is initially downsampled to produce videos with lower resolutions for low-end playback devices. The lowest resolution video represents the base layer, and the original video represents the $(X-1)-$ th enhancement layer, where X is the total number of available layers.

Similar to H.264/AVC, frames are segmented into regions called slices. A slice is defined as either INTRA coded slice (*I*-slice), predictive coded slice (*P*-slice) or bi-predictive coded slice (*B*-slice). Specifically, *I*-slice codes the image region independently, *P*-slice relies on the forward predicted image region (mainly estimated from *I*-slice), while *B*-slice relies on the bi-directionally predicted image region.

Each slice is sub-divided into a number of 16×16 pixel blocks called macroblock. In H.264/SVC encoder, the macroblocks are predicted using INTRA-prediction and INTER-prediction for the base layer. For the enhancement layer(s), INTRA-prediction, INTER-prediction and an additional INTER-layer prediction (i.e., previously encoded layer's predicted data) are used to form the predicted macroblocks. Each predicted macroblock is then subtracted from its reference macroblock to form the residual macroblock. Each residual macroblock is transformed using Discrete Cosine Transformation (DCT) and quantized, then entropy coded using either Context-adaptive variable-length coding (CAVLC) or Context-adaptive binary arithmetic coding (CABAC) to produce a network abstract layer (NAL) sub-streams. Upon completion of layered coding, a multiplexer then constructs a single scalable bitstream by using these sub-streams.

For playback, the decoder has the option to extract any number of layers (that are available) from a scalable bitstream to form a sub-stream based on the hardware/network capabilities of an endpoint video recipient [18,19].

2.2. Conventional watermarking schemes

One of the earliest work on digital watermarking scheme for H.264/SVC is presented by Park and Shin [20], where encryption and watermarking are combined to simultaneously provide access control and authentication of the video content. This scheme is appropriate for real-time applications as it is implemented within the encoding process. However, the reported results in [20] show bitrate overhead of 1.59% when merely 190 bits (i.e., watermark information) is embedded into a QCIF video. Later Shi et al. [21] investigated the detection strength of the base layer and each enhancement layer of a SVC bitstream, where a watermarking algorithm is proposed based on the combination of contrast masking, frequency masking, temporal masking and luminance adaption. Shi et al.'s scheme achieve insignificant false alarm rates, with an average of 0.5 dB in visual quality degradation for the case of 4CIF input video. As reported in [21], this watermarking algorithm is not suitable for real-time video encoding, because the frame-rate of the watermarked videos is reduced. In [22], Meerwald and Uhl suggested a blind watermark detection algorithm for both high and low resolution decoded video, where the reported results show that the bitrate was reduced when extending the base layer watermark to the enhancement layers. The authors then extended their work to cater for both the spatial and quality scalability models [12].

While the aforementioned watermarking schemes are