Signal Processing: *Image Communication* 42 (2016) 1–18

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Reduction of temporal distortion in video coding based on detection of just-noticeable temporal pumping artifact

Yanchao Gong [a,*], Shuai Wan [a], Kaifang Yang [a], Hong Ren Wu [b], Fuzheng Yang [c], Bo Li [a]

[a] School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China
[b] School of Electrical and Computer Engineering, Royal Melbourne Institute of Technology, Melbourne, Victoria 3001, Australia
[c] State Key Laboratory of ISN Xidian University, Xi'an, China

ARTICLE INFO

Article history:
Received 7 February 2015
Received in revised form
13 January 2016
Accepted 13 January 2016
Available online 22 January 2016

Keywords:
Just noticeable temporal pumping artifact
Human visual system
Hierarchical prediction structure
Quantization parameter cascading
Video coding

ABSTRACT

At low bit-rates, video coding under the hierarchical prediction structure induces a temporal distortion known as temporal pumping artifact (TPA) when the quantization parameter cascading strategy is used. TPA manifests itself as a pumping effect visually, due to severe quality fluctuations among adjacent pictures, which seriously affects the perceptual quality of a video. In this work, the causes and perception of the TPA is first analyzed. Then a metric of just-noticeable temporal pumping artifact (JNTPA) is formulated based on characteristics of the human visual system. Based on JNTPA, an efficient algorithm has been devised for TPA reduction (denoted as TRA-JNTPA for short). The experimental results have demonstrated that the proposed JNTPA measure is in line with human perception, and compared with JVT-P014, the proposed TRA-JNTPA achieves significant reduction of the TPA with superior subjective video quality at a given bit-rate.

1. Introduction

The temporal noise, distortion or artifact, which seriously affects visual quality of reconstructed videos, is generally defined as the temporal evolution or fluctuations of the spatial distortions in a video [1]. Commonly experienced temporal noises include temporal fluctuation or flickering, jerkiness, mosquito noise [2–4], and the temporal pumping artifact (TPA) which is a special type of temporal noise manifesting itself as a pumping effect [5,6]. Although the TPA may not attract observers' attention at middle to high bit-rates, as mentioned in [5], it should be noted that this annoying artifacts are very likely to be perceived at low bit-rates [6]. The TPA as an identified temporal artifact is caused by using inappropriate quantization parameter cascading (QPC) strategies in video coding using the hierarchical prediction structure (HPS), which leads to severe quality fluctuations from frame to frame among adjacent pictures.

The HPS is recently employed in video coding to make inter-frame prediction more efficient since it can greatly improve the objective rate-distortion performance. The HPS is therefore included in H.264/AVC, scalable video coding (SVC) [5], and in the latest high efficiency video coding (H.265/HEVC) as the random-access configuration or low-delay configuration [7]. When using the HPS, B frames or pictures are organized into temporal layers, and pictures in different layers are of different importance in terms of prediction. The coding efficiency for the HPS is related to how the quantization parameter (QP) is chosen

* Corresponding author.
E-mail addresses: ycgong@mail.nwpu.edu.cn (Y. Gong), swan@nwpu.edu.cn (S. Wan), kfyangnpu@mail.nwpu.edu.cn (K. Yang), henry.wu@rmit.edu.au (H.R. Wu), fuzhyang@mail.xidian.edu.cn (F. Yang), libo.npu@nwpu.edu.cn (B. Li).

http://dx.doi.org/10.1016/j.image.2016.01.006
0923-5965/

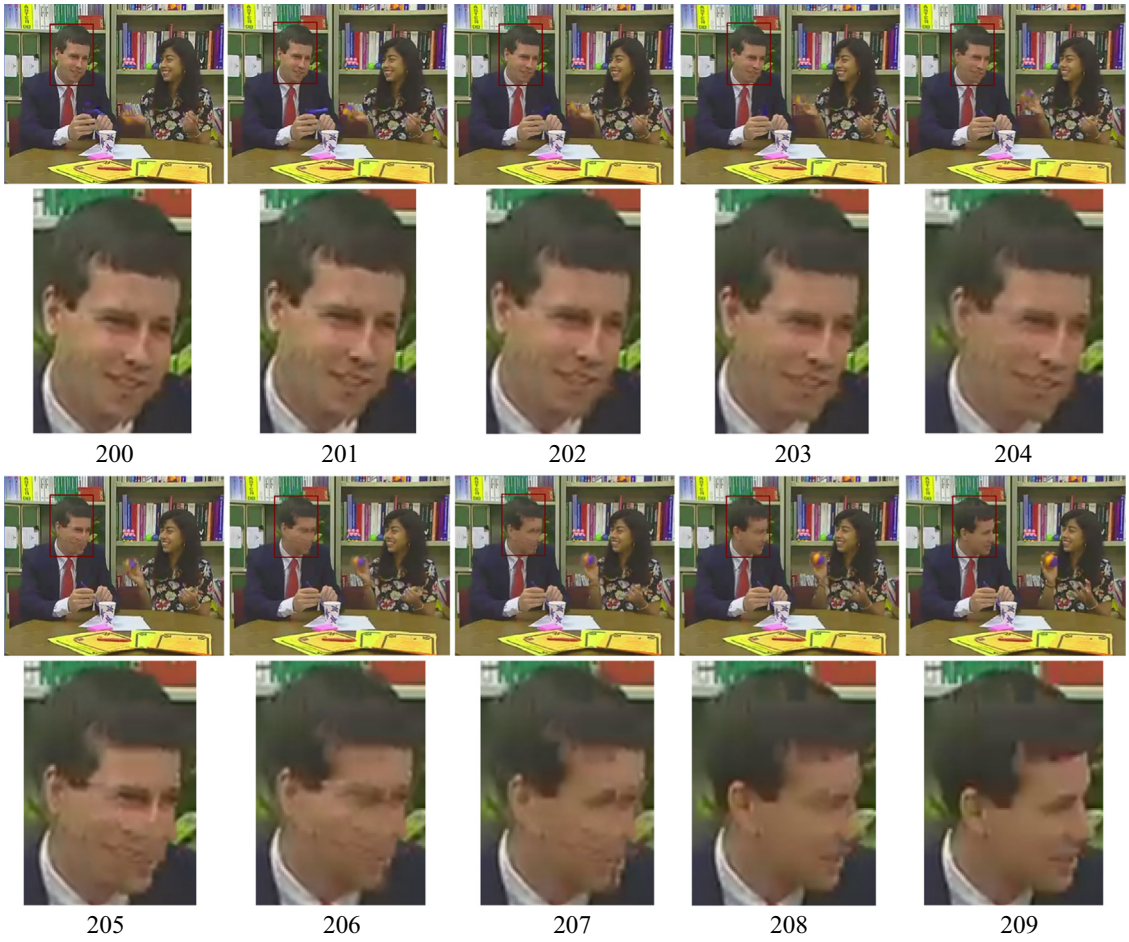

Fig. 1. Texture fidelity change of the 200th to 209th reconstructed frames of Paris encoded with HPS using the JVT-P014 QPC strategy [8] by JM17.2 codec [11] (bit-rate 101.30 kbit/s).

for pictures in different temporal layers. It has been proposed by JVT-P014 [8] for H.264/AVC/SVC and JCTVC-P1002 [9] for H.265/HEVC that the QP should be assigned with an increasing order of magnitude with the increasing of temporal layers, known as the QPC strategy. The QPC strategies are efficient in coding and have been widely acknowledged. However, the QPC strategy may result in severe quality fluctuation among adjacent pictures in a group of pictures (GOP) [6,8,10], and the related objective quality fluctuation may lead to a perceptually pumping effect, i.e., the TPA [6]. Fig.1 is an intuitive example to introduce the TPA, where a texture fidelity change can be found in the region indicated by the red box as enlarged, i.e., the man's head. That is, the region of man's head appears to be clear at first, then becomes blurry, and becomes clear again, which repeats periodically. When watch the video, most subjects will perceive a pumping artifact in the region of man's head.

Although temporal coding distortions such as flickering [12–15], jerkiness [16–18], and mosquito noise [19–21] have been studied over the years, there has been a lack of thorough interrogation into the TPA, and research efforts related to the HPS to date have mainly focused on improving the objective rate-distortion performance using a QPC strategy without considering its temporal visual effect [22,23]. The TPA was first mentioned in [5]. In our previous paper [6], the fundamental reason of perception of TPA is firstly analyzed in detail and the relevant factors which influence the perception of the TPA have also been discussed, including the quality of key frames, the content characteristics of different regions in the video, as well as the characteristics of the human visual system (HVS). Based on [6], in this paper, a metric for just noticeable temporal pumping artifact (JNTPA) is proposed based on the key factors which influence the TPA perception, which measures the threshold of the TPA that is just perceivable by human eyes. Finally, a TPA reduction algorithm based on the metric (denoted as TRA-JNTPA for short) is proposed.

The rest of this paper is organized as follows. Section 2 introduces the HPS and briefly summarizes the cause and perceptual features of the TPA. Subjective experiments are designed to identify the regions with perceptible TPA and to ascertain JNTPA in correspondence to human perception in Section 3. The metric for JNTPA is formulated in Section 4. The TRA-JNTPA algorithm is proposed in Section 5 with its performance evaluation provided in Section 6. Conclusions are drawn in Section 7.