J. Vis. Commun. Image R. 25 (2014) 879–890

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

A novel total variation based frame layer rate control algorithm for H.264/AVC

Yadong Wu [a,b], Yu Sun [b,*], Zhidan Feng [c], Hongying Zhang [d]

[a] School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang 621010, China
[b] Department of Computer Science, University of Central Arkansas, Conway, AR 72035, USA
[c] IT Research and Development Group, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA
[d] School of Information and Engineering, Southwest University of Science and Technology, Mianyang 621010, China

ARTICLE INFO

Article history:
Received 31 May 2013
Accepted 5 February 2014
Available online 21 February 2014

Keywords:
H.264/AVC
Rate control
Total variation
Incomplete derivative PID controller
Quantization parameter
Video coding
Rate distortion model
Buffer control

ABSTRACT

Rate control (RC) plays a crucial role in controlling compression bitrates and encoding qualities for networked video applications. In this research, we propose a new total variation (TV) based frame layer rate control algorithm for H.264/AVC. One of its novelties is that a total variation measure, used in image processing field, is proposed to describe encoding distortion in video compression. For intraframes, we present a TV distortion–quantization (D_{TV}–Q_{step}) model to obtain accurate QP step size (Q_{step}). Using TV measure to represent frame complexity, we also present an analytic model to calculate Q_{step} for the initial frame, and develop an effective scene change detection method. In addition, an incomplete derivative proportional integral derivative (IDPID) buffer controller is proposed to reduce the deviation between the current buffer fullness and the target buffer fullness, and minimizes the buffer overflow or underflow. Extensive experimental results show that, compared with JVT-W042, the proposed algorithm successfully achieves more accurate target bit rates, reduces frame skipping, decreases quality fluctuation and improves the overall coding quality.

1. Introduction

As H.264/AVC achieves a significant improvement in coding performance in relation to prior coding standards, it has been widely used in telecom, broadcast and digital media storage industries for video coding in recent years. In video communications, compressed videos are delivered under certain real-world restrictions, such as bandwidth, delay, buffer sizes and so on. To fulfill these restrictions, rate control (RC), which regulates compression bitrates to meet fluctuating network bandwidths while obtaining optimal visual quality, must be employed in video compression. Due to limited and time-varying network bandwidths, RC is essential for achieving the best tradeoff between encoding quality and bandwidth utilization. Therefore, it plays a crucial role in video compression and communication.

There are two important issues that need to be carefully addressed when designing RC algorithms. One is how to determine the quantization parameter (QP) of an encoder to achieve the target bit rate and good visual quality. The other is buffer control to avoid buffer overflow and underflow. RC is well studied for various video coding standards and applications [1–4]. Recently, Li et al. [5] proposed an efficient RC algorithm, which has been adopted in H.264/AVC reference software. This algorithm adopts a linear prediction model and a quadratic rate–distortion model. Many other H.264/AVC RC algorithms, focusing on different aspects, have also been studied in the literature [5–7,11–13].

QP determination for intraframes is critical for RC performance. If an intraframe's QP is too small, a large number of bits will be over-spent in coding an intraframe, and correspondingly, the subsequent interframes cannot acquire enough bits, which could result in frame skipping and quality degradation. On the other hand, if this QP is too large, it will lead to large picture distortion of an intraframe. Since an intraframe is used as reference for motion prediction of subsequent interframes, the reconstructed visual qualities of these frames will be degraded accordingly. Thus, accurate QP estimation for intraframes is particularly important in H.264/AVC RC.

To determine the proper value of QP, rate–distortion (R–D) behavior of the video coding must be analyzed and estimated. R–D model represents relations among distortion, target coding bit rate and quantization parameter for a video sequence. Various R–D models have been proposed in the literature over the years, such as quadratic [4,5], linear [1], logarithm [2], exponential [7] and ρ-domain [10] models. However, most R–D models focus on building relationships among R, QP, and D for interframes, but not for intraframes.

* Corresponding author. Fax: +1 501 450 5615.
E-mail address: yusun@uca.edu (Y. Sun).

http://dx.doi.org/10.1016/j.jvcir.2014.02.006
1047-3203/

How to build R–D models for intraframes and how to determine an intraframe's QP through R–D models have not been well addressed.

Recently, several researches have been initiated to investigate R–D models for intraframes and use it to determine intraframe's QP. Wang and Kwong [13] investigated the relation between the distortion measured by mean square error (MSE) and QP step size (Q_{step}). Then, Yan and Wang [11] have proposed an improved model, $Q_{step} = \alpha \cdot MSE + \beta$, for intraframes, where MSE is estimated by the frame complexity, buffer states and its previous intraframe's MSE. In these models, MSE are used as distortion measures. However, MSE measure exhibits weak performance in assessing perceptual image quality based on error signal mathematically [16]. In addition to MSE, other distortion measures, like SAD and MAD, are frequently used in RC. These measures are very effective for expressing the fidelity between input and output signals, but are widely criticized for not correlating well with perceived quality [16]. We feel it is important to find a distortion measure which better fits human visual system.

Variational methods have been extremely successful in a wide variety fields in image processing and computer vision during last decades [20]. The total variation measure is first introduced by Rudin, Osher, and Fatemi in their pioneering work on edge preserving image denoising [21]. The most important advantage of TV measure is its good performance in describing image structure which is one of the main factors affecting human visual perception. Due to this advantage, TV measure has also been applied in video processing field recently. A fast video sequences restoration method based on TV is proposed in [22]. A TV regularization noise removable method for moving pictures compressed by MPEG-2 is given in [23]. To the best of our knowledge, there is no study in the literature about using TV measure in video rate control, particularly in building rate-distortion model. In this research, we conduct pioneer research on investigating how to apply TV measure in video rate control.

Buffer control is another important factor for RC. The aim of buffer control is to keep buffer fullness around the target level to reduce the chances of buffer overflow or underflow. To provide robust buffer control, Sun and Ahmad [18] first applied the famous proportional–integral–derivative (PID) technique in automatic control area to the field of video compression, and proposed a novel PID buffer controller. Their work achieved an excellent performance in terms of video quality, buffer control and quality fluctuation. Afterwards, Sun and Zhou [7,19] employed an incremental PID approach to improve H.264/AVC buffer control. However, the derivative part (D) in the PID controller has the effect to amplify the interfering signals in the PID control. It is too sensitive and easily leads to the control process oscillation when the control process is affected by high-frequency disturbances.

In this paper, we proposed a novel total variation based rate control algorithm for H.264/AVC, with the objective to achieve stable coding quality and robust rate control. The major contributions of this work are summarized as follows:

- According to our best knowledge, this work first introduces the TV measure in image processing to video rate control, and correspondingly proposes a new TV distortion measure to describe video encoding quality.
- To obtain accurate Q_{step} for intraframes, we propose a new D_{TV}–Q_{step} model based on TV distortion measure.
- Using TV measure to represent frame complexity, we also present an analytic model to calculate Q_{step} for the initial frame, and develop an effective scene change detection method.
- To reduce frame skipping and achieve robust buffer control, we propose an incomplete derivative proportional integral derivative (IDPID) buffer controller to adjust target bits allocated to interframes.

The rest of the paper is organized as follows. Section 2 presents the new TV distortion measure for video frames. The linear relationship between TV distortion measure and Q_{step} is also shown here. Section 3 presents the details how to decide intraframes' QP and initial QP, even under scene change situation. Section 4 discusses our IDPID buffer controller. Our algorithm is summarized in Section 5, and experimental results are shown in Section 6. Finally, Section 7 concludes this paper.

2. The proposed TV measure and TV distortion–quantization model

In order to achieve stable qualified coding results, how to evaluate a frame distortion is a key problem. Most papers [7,13] use the Peak Signal-to-Noise Ratio (PSNR) or MSE to measure distortion. However, MSE is not very good to measure the perceptual video quality, especially when dealing with perceptually important signals such as speech and images [24]. Because TV is well-known to be closer to human subjective evaluation, it shows the potential in evaluating video perceptual quality.

In our research, we observed that the distortion, the reconstructed error introduced by compression, can be described from two aspects. One aspect considers the distortion contained in a reconstructed frame. The quality of a reconstructed fame is decreased as Q_{step} increases. The decrease of reconstructed frame quality reflects loss of structure information, which can be described by TV measure. The other aspect considers the difference frame between an original frame and its reconstructed frame. With Q_{step} increases, structure information of the difference frame also increased. From the above analysis, we can tell that TV measure can be used to represent the distortion of an encoded frame. Based on this motivation, we plan to initiate the research of applying TV measure in building the relationship between distortion and Q_{step}.

For a frame u, the total variation model can be formulated by $TV(u) = \int_\Omega |\nabla u|$, where Ω denotes the frame domain, ∇ means the gradient operator. The discrete form of TV is

$$TV(u) = \sum_{\substack{i=1\dots m-1 \\ j=1\dots n-1}} \sqrt{(u_{i,j} - u_{i+1,j})^2 + (u_{i,j} - u_{i,j+1})^2}, \tag{1}$$

where $u_{i,j}$ is image intensity at pixel (i,j), and image size is $m \times n$.

In order to fully evaluate encoding distortion, we consider luminance information of both the reconstructed frame, and the difference between the original frame and reconstructed frame. Assuming u represents original frame's luminance component, and u' describes the luminance component of a reconstructed frame after decoding. The difference between these two luminance components is computed by $u_r = u - u'$. By applying TV model on u_r and u', we can obtain $TV(u_r)$ which represents the information loss of an encoded frame, and $TV(u')$ which measures the information remained after encoding.

To investigate the relationships among $TV(u_r)$, $TV(u')$ and Q_{step}, we have conducted extensive experiments by encoding the 1st I-frames for various sequences with different Q_{step} values ranged from 1.5 to 80.5 (corresponds to $8 \leqslant QP \leqslant 42$).

Fig. 1(a) and (c) show the relationship between $TV(u_r)$ and Q_{step} for the 1st I-frame (luminance component). Fig. 1(b) and (d) show the relationship between $1/TV(u')$ and Q_{step} for the 1st I-frame (luminance component). From Fig. 1, we can see that along the decrease of Q_{step}, $TV(u_r)$ and $1/TV(u')$ decreases.

According to extensive experimental results and analysis, we propose a new TV distortion measure, which is defined by:

$$D_{TV} = TV(u_r)/TV(u'). \tag{2}$$