J. Vis. Commun. Image R. 22 (2011) 313–324

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

No-reference analysis of decoded MPEG images for PSNR estimation and post-processing

Søren Forchhammer *, Huiying Li, Jakob Dahl Andersen

Department of Photonics Engineering, Technical University of Denmark, Bldg. 343, DK-2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:
Received 15 June 2010
Accepted 25 January 2011
Available online 2 February 2011

Keywords:
No reference PSNR estimation
Quantization parameter estimation
I-frame detection
Image post-processing
Video post-processing
DCT
MPEG

ABSTRACT

We propose no-reference analysis and processing of DCT (Discrete Cosine Transform) coded images based on estimation of selected MPEG parameters from the decoded video. The goal is to assess MPEG video quality and perform post-processing without access to neither the original stream nor the code stream. Solutions are presented for MPEG-2 video. A method to estimate the quantization parameters of DCT coded images and MPEG I-frames at the macro-block level is presented. The results of this analysis is used for deblocking and deringing artifact reduction and no-reference PSNR estimation without code stream access. An adaptive deringing method using texture classification is presented. On the test set, the quantization parameters in MPEG-2 I-frames are estimated with an overall accuracy of 99.9% and the PSNR is estimated with an overall average error of 0.3 dB. The deringing and deblocking algorithms yield improvements of 0.3 dB on the MPEG-2 decoded test sequences.

1. Introduction

Digital TV systems of today are operated using a range of image resolutions and qualities. Furthermore different video coding schemes including MPEG-2 and MPEG-4 part 10/H.264 are used. We address the problem of analyzing decoded DCT (Discrete Cosine Transform) images with focus on processing decoded MPEG-2 video, in a heterogenous environment. The goal is to achieve high visual quality and quality assessment. The DCT is also the basis of JPEG image coding and other video coding schemes as H.261-263, MPEG-4 part 2, and proprietary schemes. We consider the problems of frame type detection, PSNR estimation, post-processing and validation based solely on the decoded images. The aim is to increase performance by processing at TV receivers. Large high resolution flat panel displays and high quality projectors of today seem to magnify coding artifacts, so these also become visible in good quality images and video and thus put increased focus on image quality and post-processing. Besides broadcast video and video distributed on cable-net, also images and video material from the Internet, storage media and consumer cameras may be displayed on a large flat panel display. Decoded MPEG video may be post-processed to attenuate the coding artifacts and thereby increase the perceived quality, and it may also be re- or transcoded for storage or further transmission. For all these tasks, adaptive algorithms utilizing MPEG parameters extracted from the coded stream have been presented [1–4]. In some cases however, the coded stream is not accessible e.g. when encryption prevents access to the MPEG stream. In other cases, it may be desirable from an architectural point of view not to access the code stream parameters. One example is receiving the decoded digital video over a High-Definition Multimedia Interface (HDMI) connection, where the MPEG stream prior to decoding was only accessible in encrypted form, e.g. due to digital rights management (DRM) protection. Furthermore the decoded video may be delivered in a different resolution than it was coded. In all these cases, it may be necessary to base the processing and analysis solely on the decoded images and video signals.

MPEG-2 has been the work horse of digital TV. Now H.264/MPEG-4 is being deployed world wide, but MPEG-2 will play a major role in digital TV for years to come, e.g. on cable-net systems and Digital Terrestrial TV. We selected MPEG-2 as a prominent example out of a large class of DCT based image and video coders co-existing with H.264/MPEG-4. We focus on estimating the important MPEG-2 parameters based on the decoded video: DCT block size and position, I-frame detection and for the I-frames, estimation of the quantization step size, which determine the distortion. We shall utilize this information to estimate the PSNR of I-frames and to guide post-processing filters for deblocking and deringing. A measure is introduced to validate that the (decoded) video stream originates from an MPEG-2 coded stream as well as validating the parameter analysis. Thus decoded MPEG-2 I-frames may be separated from images originating from other formats e.g.

* Corresponding author.
E-mail address: sofo@fotonik.dtu.dk (S. Forchhammer).

1047-3203/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2011.01.006

decoded MPEG-4/H.264 frames. The identification of I-frames is also beneficial when transcoding MPEG streams [5], e.g. transcoding from MPEG-2 to MPEG-4/H.264.

MPEG-4/H.264 I-frames may differ from those of previous DCT based schemes by applying intra-prediction and an in-loop deblocking filter [6] as well as the smaller 4×4 DCT blocks. This means that the coding artifacts of MPEG-4/H.264 are different and less pronounced compared to those of MPEG-2 [7]. Therefore we suggest applying different post-processing for the two coding schemes and in this paper focus on MPEG-2 post-processing.

Methods for (decoder side) no reference PSNR estimation, i.e. without access to the original video, have been presented based on parameters extracted from the MPEG streams [4]. For MPEG-2 I-frames, an analysis just based on the decoded stream was presented in [8]. This was applied in an JPEG like setting with one fixed quantizer value for each I-frame.

We extend this to the general MPEG-2 case with variable quantizer values at the (16×16 luminance pixels) macroblock (MB) level. Using the estimated MPEG parameters, we shall perform postprocessing of MPEG-2 video coding artifacts originating from the DCT domain quantization and focus on the blocking and ringing artifacts, which are the major artifacts. Research on post-processing of DCT based coding as JPEG and MPEG has a long and active history, e.g. [1–3,9–15]. To pursue a goal of deringing algorithms, which may be useful for real-time video processing, we focus on the class of relatively simple (non-iterative) spatially adaptive post-filtering. To control this postfiltering, a popular approach [2,3,14], is to utilize the MPEG quantization scale parameter (Q_S). The parameter is read from the code stream and we shall refer to this as being embedded post-processing. Another goal is to perform the processing without access to code stream information. Therefore, the Q_S values are estimated based on the decoded video instead. We refer to this approach as pure post-processing to distinguish it from embedded processing. A new spatially adaptive deringing filtering based on texture analysis is developed for pure post-processing.

The rest of the paper is organized as follows: Section 2 introduces notation and the MPEG-2 parameter estimation based on the decoded video for block size and position estimation, quantization step size estimation, and I-frame detection. Section 3 presents no-reference PSNR estimation for I-frames based on the estimated MPEG parameters. Section 4 describes the use of estimated quantization values for deblocking and deringing of MPEG-2 without code stream access. A new deringing filter developed for this set-up is also introduced. Experimental results are given in Sections 5 and 6 concludes the paper.

2. MPEG-2 parameter estimation

The discrete cosine transform is widely used in image and video coding. The quality is determined by the quantization of the DCT coefficients. We shall focus our analysis on intra decoded MPEG-2 I-frames as one example among the DCT based standards and consider estimation of MPEG-2 parameters based on the decoded video. To facilitate the MPEG analysis, we shall reconstruct the MPEG (quantized) DCT values based on the decoded pixel values. The DCT is reversible, but rounding, clipping and lack of exact specification of DCT/IDCT in the MPEG-2 specification leave uncertainties.

2.1. Notation for MPEG-2 decoding

To provide the notation for the analysis, selected parts in the decoding process of MPEG-2 I-frames are briefly described. The basic processing unit is the 16×16 pixels (luminance) macroblock (MB), which is further divided into four 8×8 DCT blocks. The DCT transformed coefficients are locally quantized specified by one quantizer scale value, Q_S, per MB.

The variable length decoder outputs the integer values, $I_Q(u,v)$, which represent the indices of the quantization interval for the DCT coefficient at frequency (u,v). Based on $I_Q(u,v)$, a DCT coefficient $F''(u,v)$ is reconstructed in conformance with [16]. For an intra MB, i.e. no motion-compensation is used, the AC coefficients, i.e. $(u,v) \neq (0,0)$, are reconstructed with an absolute value given by

$$|F''(u,v)| = \left\lfloor \frac{|I_Q(u,v)| \times Q_M(u,v) \times Q_S}{16} \right\rfloor, \quad (1)$$

where $\lfloor\ \rfloor$ denotes the floor function and $Q_M(u,v)$ denotes the frequency dependent quantization matrix values. The four luma DCT blocks in one macroblock are quantized using the same Q_S value, but Q_S may change from one MB to the next. After $F''(u,v)$ is reconstructed at the decoder, the inverse DCT will transform $F''(u,v)$ to an inverse transformed value, which is rounded, and, if necessary clipped, to obtain reconstructed integer values $r'(i,j)$ in the range [0,255], for intra blocks, and thereafter $r'(i,j)$ is output as the decoded video, $d(i,j)$.

Based on the details of dequantization of coefficients before as well as rounding and clipping after the IDCT, analysis of the decoded MPEG-2 video may be established. The focus is on estimating or detecting three important (sets of) MPEG parameters: Position of I-frames, DCT blocksize and position, and quantization step sizes (in I-frames). First the DCT block boundary positions are estimated both horizontally and vertically. Based on that, the DCT is applied to each 8×8 DCT block to obtain the recalculated DCT coefficients $F'(u,v)$ as an approximation of $F''(u,v)$. (If the detected block size is not 8×8, the image shall be scaled such that DCT blocks are rescaled to 8×8, prior to the DCT transformation.) Both frame and field DCT may be applied on MBs. The MB type (frame or field MB) can be estimated by selecting the type having the minimum number of zero DCT coefficients within the MB. Thereafter, estimation of Q_M at frame level and Q_S at MB level is performed based on the recalculated DCT coefficients $F'(u,v)$ (frame DCT if the MB is evaluated to be a frame MB, otherwise field DCT). Furthermore, measures of mismatch at MB level (M_{MB}) and frame level are calculated for detection of I-frames and validation of the analysis on detected I-frames. Details of the estimation tasks are given below.

2.2. Blocksize estimation

In [17], the size of DCT blocks was estimated by calculating absolute differences between adjoining pixels, as part of blocking artifact analysis. In order to increase independence of image content, we instead calculate a difference of absolute differences, *DAD*, horizontally and vertically [18]. Let a, b, c, d, e, f denote the value of 6 consecutive values horizontally. Define the absolute difference $D'_{cd} = |d - c|$ and D_{cd} as the sum of D'_{cd} plus the two corresponding differences D' of the two pixels above and for the two pixels below c and d, respectively. Based on D values of neighboring pixels, an initial difference of difference is calculated as $DAD'_{cd} = 2D_{cd} - 2D_{bc} - 2D_{de} + D_{ab} + D_{ef}$, where subscripts specifies the pixels involved. This value is thereafter thresholded to form $DAD_{cd} = DAD'_{cd}$ if $3 < DAD'_{cd} < 120$ and 0 otherwise. The *DAD* values are thereafter projected by summation onto the horizontal and vertical axis, respectively.

2.3. Quantization parameter estimation

Given the DCT blocksize and position, the DCT coefficients may be recalculated, based on the decoded frames. The next objective is