Signal Processing: *Image Communication* 29 (2014) 274–292

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Methods for improving the tone mapping for backward compatible high dynamic range image and video coding

Alper Koz*, Frederic Dufaux

Multimedia Group, Signal and Image Processing Department (TSI), Télécom ParisTech, 75014 Paris, France

ARTICLE INFO

Available online 2 September 2013

Keywords:
High dynamic range video compression
Tone mapping
Perceptually uniform luminance values
Backward compatibility
H.264/AVC
Flickering

ABSTRACT

Backward compatibility for high dynamic range image and video compression forms one of the essential requirements in the transition phase from low dynamic range (LDR) displays to high dynamic range (HDR) displays. In a recent work [1], the problems of tone mapping and HDR video coding are originally fused together in the same mathematical framework, and an optimized solution for tone mapping is achieved in terms of the mean square error (MSE) of the logarithm of luminance values. In this paper, we improve this pioneer study in three aspects by considering its three shortcomings. First, the proposed method [1] works over the logarithms of luminance values which are not uniform with respect to Human Visual System (HVS) sensitivity. We propose to use the perceptually uniform luminance values as an alternative for the optimization of tone mapping curve. Second, the proposed method [1] does not take the quality of the resulting tone mapped images into account during the formulation in contrary to the main goal of tone mapping research. We include the LDR image quality as a constraint to the optimization problem and develop a generic methodology to compromise the trade-off between HDR and LDR image qualities for coding. Third, the proposed method [1] simply applies a low-pass filter to the generated tone curves for video frames to avoid flickering during the adaptation of the method to the video. We instead include an HVS based flickering constraint to the optimization and derive a methodology to compromise the trade-off between the rate-distortion performance and flickering distortion. The superiority of the proposed methodologies is verified with experiments on HDR images and video sequences.

1. Introduction

Video content represented by eight bits per pixel has been widely accepted in many applications such as internet/video streaming, DVD, HDTV, camcorders, personal computers, televisions and other consumer electronics products. The reasons of such a wide acceptance of eight-bit representation can be argued as the compact representation of one pixel value as a *byte* for the storage in memory chips and the sufficiency of 256 levels to cover the luminance range of commonly used low dynamic range (LDR) displays (0.1–80 cd/m^2) [2–5]. However, image capturing and display technologies have improved to span a wider dynamic range and true color representations in recent years. LCD and plasma displays with a peak luminance of 500–1000 cd/m^2 have dominated the market. HDR prototype displays are now available with a dynamic range of 1,000,000:1 and a peak luminance of 4000 cd/m^2 [6]. The cinema industries have also begun to capture movies as digital images, rather than film, by using modern digital cameras that have extended dynamic range feature [7]. Even more, the smart phones, which have reached 55% of usage penetration in EU5 countries (UK,

* Corresponding author. Tel.: + 33 698576207; fax: + 33 145817144.
E-mail addresses: alper.koz@telecom-paristech.fr, alperkoz@yahoo.com (A. Koz), frederic.dufaux@telecom-paristech.fr (F. Dufaux).

0923-5965/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.image.2013.08.017

DE, FR, ES, and IT) by October 2012 [8], now possess HDR photo utilities to generate more realistic views. Accordingly, standard eight-bit representation has become insufficient for these new technologies and related applications in digital cinema, medical imaging, and post-production.

HDR image and video formats are proposed to overcome the limitations of standard eight-bit representation. The main goal with these formats is to encode, transmit, and display the accurate physical luminance values (cd/m^2) in a real-world scene rather than the eight-bit intensity values and hence to generate a scene referred image independent of the display technology [1–5]. The proposed formats for such a goal obviously require higher bit depths than standard eight-bits when the entire luminance range of natural scenes ranging from extreme darkness (10^{-6} cd/m^2) to bright sunshine (10^8 cd/m^2) is considered. Current formats such as Radiance RGBE (*.hdr*), OpenEXR (*.exr*), and LogLuv TIFF (.tiff) use 16-bit floating point, 32-bit floating point or 32-bit integer pixels [1].

A major challenge for such a high-bit depth representation is an efficient compression, in particular for HDR video which requires significantly more storage size and transmission bandwidth than standard 8-bit LDR video. The size of the HDR video captured by the first HDR camera demonstrated at the annual ACM SIGGRAPH conference in 2009 was reported as 42 GB for a footage of one minute, compared with just 9 GB for its LDR counterpart [2]. The previous work has mainly handled this challenge in two ways (Fig. 1). The first approach [9–11] is to take the advantage of high profile of existing state-of-the-art H.264/AVC (Advanced Video Coding) codec, which can support the bit-depths more than 8 bit up to 14 bit. These methods first convert the real luminance values (cd/m^2) of the HDR video pixels into a bit-depth between [9–14] bits and apply the H.264/AVC high profile encoder. After the decoding and reconversion to luminance values (cd/m^2), the reconstructed HDR video is shown on a HDR display, or on a LDR display by performing a real-time tone mapping operation, as illustrated in Fig. 1-a.

The second approach [1,5,12,13] aims to provide the backward compatibility with existing standard 8-bit displays in order to enable a successful transition to HDR technology (Fig. 1-b). The methods in this approach first map the HDR video into 8-bit to produce an LDR video. The LDR video goes through the video encoding and decoding process by using a standard 8-bit (e.g. H.264/AVC) encoder. Then, the high dynamic range video is reconstructed from LDR video by applying a real-time inverse tone mapping operation. The parameters regarding the tone mapping and inverse tone mapping can be sent as a look-up table [1] or can be encoded using supplemental enhancement information [15]. The residue of the reconstructed HDR video can also be further encoded as an enhancement layer in the bit-stream. As this approach includes the LDR stream separately in the resulting bit-stream, it provides a direct compatibility to LDR displays. However, it requires real-time inverse-tone mapping operation to show the video content on HDR displays.

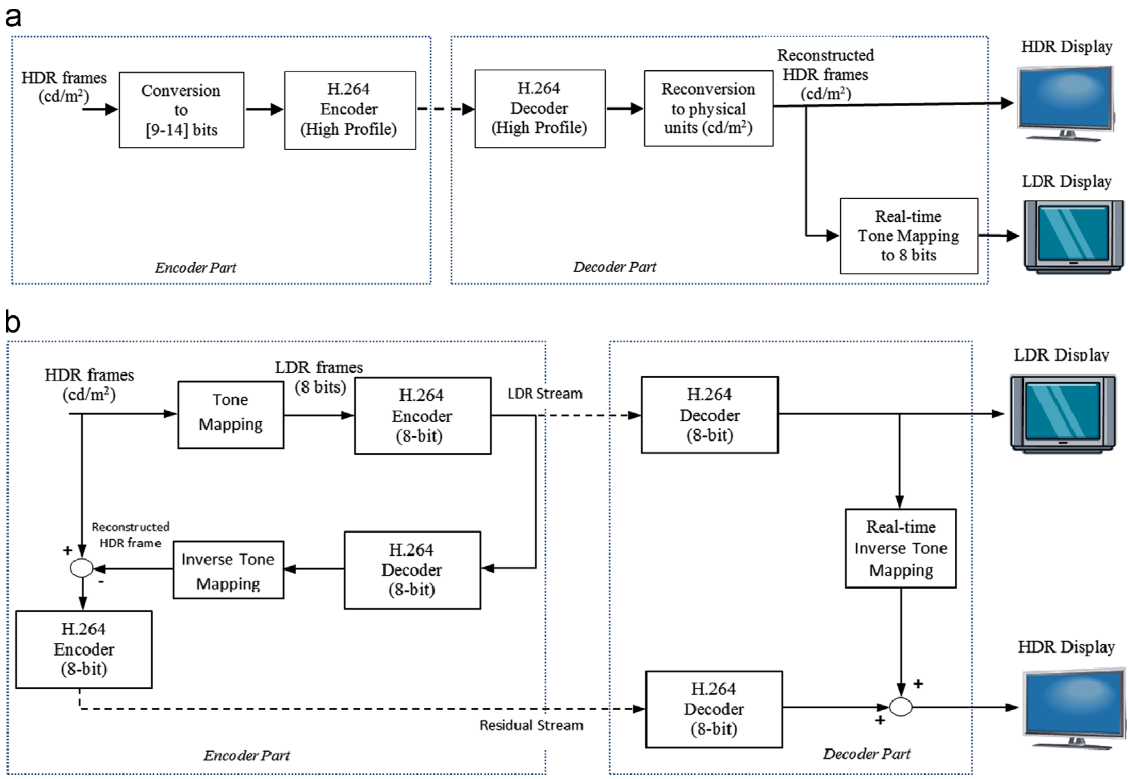

Fig. 1. A general scheme for the two approaches in HDR video compression, (a) using high profile H.264/AVC codec, and (b) using 8-bit H.264/AVC codec for backward compatibility to LDR displays.