New Non Predictive Wavelet Based Video Coder: Performances Analysis

Tarek Ouni, Walid Ayedi, and Mohamed Abid

National Engineering School of Sfax, Road Sokkra, Km 3 Sfax, 3078 Tunisia
tarek.ouni@gmail.com, walid.ayedi@gmail.com,
Mohamed.abid@enis.rnu.tn

Abstract. A non-predictive video coding is a new branch of emerging research area in video coding, where the motion estimation/compensation or prediction step in the temporal domain is omitted. One direction was to look for the exploitation of temporal decomposition of video frames. The proposed method consists on 3D to 2D transformation of the temporal frames that allows exploring the temporal redundancy of the video using 2D wavelet transforms and avoiding the computationally demanding motion compensation step. Although the many advantages presented by the proposed coder, some annoying artifacts still exist. In this paper, we will explore the performances of the proposed method and try to better show what it actually offers to users. The paper presents also the extensions chosen in order to reduce the perceived artifacts and increase the perceptual as well as objective (PSNR) decoded video quality, which is actually competitive with state-of-the-art video coder algorithms, especially when low computational demands of the proposed approach are taken into account.

Keywords: video coding, temporal decomposition, wavelet, correlation.

1 Introduction

Video compression has generated a lot of discussion and increasing attention from the research in recent years.

Among many proposed methods, motion compensated coding has taken the most attention and taken its place in many standards. These include Mpeg, H.26L, etc. Such encoders exploit inter-frame correlation in order to further improve its compression. However, the main challenge of these methods lies on the motion estimation process which is known to be computationally intensive. Besides, its real time implementation is difficult and costly [1],[2]. Nevertheless, new applications such as sensor networks and portable video devices necessitate a low processing capability for the compression, which makes the encoding complexity a big burden. To deal with this problem, motion-based video coding standard MPEG was primarily developed for stored video applications, where the encoding process is typically carried out off-line on powerful computers. With the explosive growth of video devices ranging from hand-held digital cameras to low-power video sensors, a new class of multimedia devices is required which includes the following architectural requirements: Low

A. Campilho and M. Kamel (Eds.): ICIAR 2010, Part I, LNCS 6111, pp. 344–353, 2010.

power, less- complexity encoding and real time constraint. Therefore, there have been extensive research efforts in video coding in order to give response to the new requirements of video applications different than those targeted by conventional coding schemes in the past years [1]. A non-predictive video coding is a new branch of emerging research area in video coding, where the motion estimation/compensation or prediction step in the temporal domain is omitted.

In [3]-[4], authors exploit 3D transforms in order to exploit temporal redundancy. Coder based on 3D transform produces video compression ratio which is comparable to some motion estimation based coding one but with lower processing complexity [5]. However, 3D transform based video compression methods process temporal and spatial redundancies in the 3D video signal in the same way. This can reduce the efficiency of these methods as pixel's values variations in spatial or temporal dimensions are not uniform and hence, temporal and spatial redundancies have not the same pertinence. It is known that the temporal redundancies are more relevant than spatial one [2], hence its practical importance. It is more beneficial to utilize the proposed method rather than the 3D based methods, because it is able to achieve higher compression by more exploiting the redundancies in the temporal domain.

This method consists on 3D to 2D transformation of the video frames; it will then explore the temporal redundancy of the video using 2D transforms and avoids the computationally demanding motion compensation step. In particular, the used method projects temporal redundancy of each group of pictures into spatial domain and combines it with spatial redundancy in one representation with high spatial correlation [6]. Then, the new representation will be compressed as still image using wavelet transform based coder (JPEG2000). Actually, the proposed approach presents many advantages. It exploits objectively temporal and spatial redundancy. It omits the temporal prediction step and transforms a 3D processing into 2D one while reducing considerably the complexity processing. Furthermore, it inherits the JPEG 2000 proprieties such as scalability ROI and error resilience.

In this paper, we focus on the analysis of experimental results, solutions and extensions proposed to remove some annoying artifacts of the presented method. Experimental results will show the efficiency of the proposed method at an expense of some annoying artifacts.

The rest of paper is organized as follows. In section 2, we review the basics of the used approach. In section 3, we present some experimental results and explore the method performances and limitations. Section 4 presents the extensions chosen to further improve the compression ratio. Finally, conclusions are drawn in Section 5.

2 Description of the Used Approach

Actually, the used method relies on the following assumption: high frequency data is more difficult to compress compared to low frequency one.

The main idea of the proposed method is to make some geometric transformation of the 3D data in order to make one representation with very high correlation, and consequently without high frequencies data. We will play on the disposition of pixel’s data in the video cube.