A New Object-Based Fractal Compression of Monocular and Stereo Video Sequences

Kamel Belloulata[1,*] and Zhu Shiping[2]

[1] Département D'électronique, Faculté des Sciences de l'Ingénieur, BP89,
Université Djillali Liabès de Sidi Bel Abbès, 22000, Algérie
k_belloula@yahoo.fr
[2] Department of Measurement Control & Information Technology,
School of Instrumentation Science & Optoelectronics Engineering, BeiHang University,
100083, Beijing, P.R.China
Shiping.Zhu@buaa.edu.cn

Abstract. A novel object-based fractal monocular and stereo video compression scheme with quadtree-based motion and disparity compensation is proposed in this paper. Fractal coding is adopted and each object is encoded independently by a *prior* image segmentation *alpha* plane, which is defined exactly as in MPEG-4. The first *n* frames of right video sequence are encoded by using the Circular Prediction Mapping (CPM) and the remaining frames are encoded by using the Non Contractive Interframe Mapping (NCIM). The CPM and NCIM methods accomplish the motion estimation/compensation of right video sequence. According to the different coding or user requirements, the spatial correlations between the left and right frames can be explored by partial or full affine transformation quadtree-based disparity estimation/compensation, or simply by applying CPM/NCIM on left video sequence. The testing results with monocular and stereo video sequences provide promising performances at low bit rate coding. We believe it will be a powerful and efficient technique for the object-based monocular and stereo video sequences coding.

Keywords: Monocular and stereo video coding, fractal coding, object-based coding, low bit rate coding.

1 Introduction

The next generation visual communications must address the application of capture, transmission, and display of 3D visual information and then realize one of the most desired features of high quality telecommunication services, which is in terms of ''The sensation of 3D reality''. Although holographic and volumetric 3D displays may provide full 3D perception of the scene, but the vast amount of the optical information prevent their practical uses for the time being and also because their state-of-art presentation abilities of only still images.

---

[*] Kamel Belloulata was with the Département de génie électrique et de génie informatique, Faculté de génie, Université de Sherbrooke, Sherbrooke, Québec, Canada J1K 2R1. (e-mail: Kamel.Belloulata@USherbrooke.ca).

M. Kamel and A. Campilho (Eds.): ICIAR 2009, LNCS 5627, pp. 103–115, 2009.

Alternatively, 3D stereoscopic displays can supply the 3D representation through the human brain to fuse the left and right views of the same scene, which are captured from slightly different viewing angles. Undoubtedly, it will be a very attractive and effective direction and technique to realize the 3D visual communication in the near future [1]. Consequently, monocular video sequence [2], and especially stereo video sequences [3], or in general multiview video sequences [4], which can provide vivid and plentiful information about the 3D scene, have been researched recently to generate virtual synthesized intermediate views [5] and thus supply the 3D feeling of the scene through a flat 2D screen. They have wide applications in many fields, such as in the 3D visual communication, telemedicine [6], etc. However the initial challenge of delivery of huge amount of video sequence data should be overcome to allow them to be transmitted economically and efficiently. Meanwhile efficient coding techniques should be developed to reduce bit rate in the transmission of monocular [2], stereo [7] and multiview [8] video sequences.

Two major approaches are used for video sequence coding: block-based and object-based. The most currently used block-based approach has the advantages of simplicity, efficiency and robustness. It has achieved great success and is commonly adopted in many video compression standards. But the subjective quality of reconstructed images may be bad at low bit rate. Object-based video sequence coding has been intensely investigated in the last few years and is also supported by the new MPEG-4 standard [9], [10]. It has an important advantage over the block-based coding: it allows manipulation of image objects without complete decoding of the stream, and then improves the coding quality and reduces the bit rate. In such a scheme, a *prior* segmentation map (*alpha* plane) [11] of the image, which segments the image into objects, is known in advance [12].

The object-based approach has been considered as a very promising alternative to the block-based approach. It alleviates the problem of annoying coding effects, such as blocking artifacts and mosquito effects than the block-based approach at low bit rate, especially when the blocks coincide with boundaries of different objects. The object-based approach can also provide more natural representation of the scene and potential benefit of acquiring the depth information of the semantically meaningful objects [8].

In such a scheme, the task of automatic extraction and modeling of objects directly from the image intensities require sophisticated image analysis techniques to segment the image into homogeneous regions [13, 14] or semantically meaningful objects [12], or even user interaction will be needed to segment the image into regions corresponding to the real objects before the automatic segmentation algorithm [11]. Several methods for coding of the binary *alpha* plane have been considered during the development of MPEG-4, such as chain coding of the object contours, quadtree coding [2], modified modified reed (MMR) and context-based arithmetic encoding (CAE).

However, little work has been reported on the fractal video coding technique [15], [16]. A scheme, which is not truly object-based, has been proposed for object-based coding system [17], and it is based on quadtree partitioning [18].

Recently, we have proposed a fractal based image codec with region-based functionality [19]. It permits new functionalities at the decoder, such as independent transmission/decoding of each object in the video, object/background replacement, object- based video retrieval, and especially the gain of better video visual quality