

IMAGE SEQUENCE CODING USING 3-D I.F.S.

M. Barakat[†], J.-L. Dugelay

Institut EURECOM, Multimedia Communications Dept.
2229, route des Crêtes, 06904 Sophia Antipolis
e-mail. dugelay@eurecom.fr, URL.http://www.cica.fr/~image

ABSTRACT

In this paper, we study the efficiency of the image sequence coding scheme based on 3-D I.F.S. This coding scheme is an extension of the still image coding scheme proposed by Jacquin [1]. After a brief presentation of the I.F.S. based coding and in particular the 3-D I.F.S. scheme for coding image sequences, we simulate and discuss the already published schemes. We propose some improvements based on the choice of isometries and the construction of domain cubes during the coding.

1. INTRODUCTION

Iterated Function Systems (I.F.S.) are sets of contractive functions in a complete metric space. Earlier studies on I.F.S. and their links to fractals are due to J. Hutchinson [2]. The main property of I.F.S. is the convergence of their successive iterations to a fixed point, or attractor, regardless of the starting point. This convergence is due to the contractivity of the functions composing an I.F.S. M. Barnsley *et al.* worked on the representation of functions, and more generally, of real world objects by means of I.F.S. [3]. A. Jacquin was the first to publish an I.F.S. based scheme for coding still images [1]. This scheme uses the notion of Local-I.F.S., and allows the construction of an I.F.S. representation of a still image by matching similarities at two scales, after a suitable partitioning of the image. See §2 for more

details. Several works have shown the efficiency of Jacquin's scheme and improved its features [4]. The first paper dealing with I.F.S. based moving picture coding was published by Beaumont in 1991 [5]. Several other papers have since treated this problem [6]. There are two major approaches to perform moving picture coding with I.F.S. (neglecting the frame-by-frame one). The first one is a combination between inter-frame coding using block-matching and intra-frame coding using I.F.S. [7]. This approach is similar to the MPEG video coding scheme in that it uses I.F.S. instead of D.C.T. The second one uses three-dimensional (3-D) I.F.S., and is an extension of the 2-D scheme [8], [9]. In this paper, we are concerned by the second approach. We present briefly the I.F.S. based still image coding and its extension to the three-dimensional case. We simulate the yet published schemes and propose some improvements.

2. STILL IMAGE CODING

In Local-I.F.S. based coding a still image is represented by a set of locally contractive transformations that map parts of the image to other parts. These transformations are generally affine of the form :

$$f \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ 0 & 0 & s \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ o \end{bmatrix}$$

where (x, y) denotes the spatial position of a pixel and z its grey scale value. A partition of the image support gives a pool of range blocks r_i . Also, a pool of domain blocks d_j is constructed (generally d_j have greater size than r_i). Then, the

[†]Current address : ENSEEIHT, GAPSE-LEN7, 2, rue Charles CAMICHEL, 31071 TOULOUSE CEDEX, e-mail. barakat@len7.enseeiht.fr

algorithm matches range and domain blocks by means of transformations $w = f \circ S$, where S is a sub-sampling operator. More clearly, for a given range block r_i , we search the domain block d and the transformation w which minimize the mean square error between r_i and $w(d)$, see Figure 1. α_{11} , α_{12} , α_{21} and α_{22} are fixed to the parameters of a set of authorized isometries (4 rotations and 4 reflections) applied to $S(d)$. A sufficient condition for the contractivity of w is $|s| < 1$. There exists a necessary and sufficient condition for the contractivity of the transformations, but it can not be tested *a priori*, and so its usefulness is limited. By Banach's fixed point theorem, iterating the I.F.S. code on an arbitrary initial image yields (an approximation of) the original image. This iteration constitutes the decoding process. The coding step computational cost is high. Several search strategies were proposed to reduce it [4].

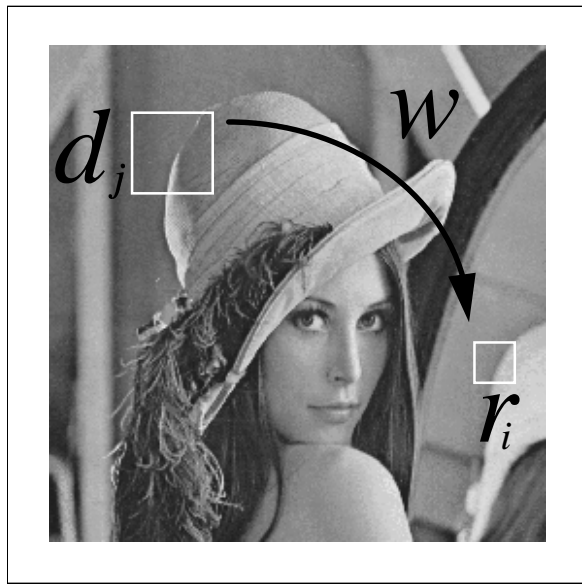


Figure 1: Matching between range and domain blocks.

3. THREE DIMENSIONAL I.F.S.

Coding image sequences with 3-D I.F.S. can be considered as a direct extension of Jacquin's 2-D scheme. In fact, the image sequence is subdivided

into Groups Of Pictures (G.O.P.) and each one is coded as a still image, while replacing 2-D blocks (squares) by 3-D ones (cubes), see Figure 2. A 3-D block is made up of 2-D blocks, belonging to successive frames, and having the same spatial position. f is now an affine transformation on \mathbb{R}^4 and has the form:

$$f \begin{bmatrix} x \\ y \\ t \\ z \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & 0 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & 0 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & 0 \\ 0 & 0 & 0 & s \end{bmatrix} \begin{bmatrix} x \\ y \\ t \\ z \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ o \end{bmatrix}$$

where t is the time variable. S performs a sub-sampling in the spatial directions and the temporal domain. This approach, despite the relatively high computational cost (inherent to I.F.S. based coding methods), retains the most interesting properties of I.F.S. based coding, compared to the MPEG-like approach. In particular, the code we obtain is (theoretically) resolution independent [10], and so we can perform a sub/oversampling (at the decoding step) in the spatial directions and the temporal domain. Also, the iterative aspect of the decoding is maintained. Papers dealing with the 3-D approach differ mainly in their definitions of the transformation dictionary, and more precisely in the authorized isometries. Lazar proposed a combination of 2-D isometries (used in Jacquin's scheme) applied to each square of a domain cube, with temporal reverse to construct 3-D isometries [8]. Naemura, who worked on multi-view 3-D image (the third dimension being the views'one), used a rotation about a spatial axis, a reflection about a plane (parallel to the view axis and to one of the spatial axes), and their combination [9]. Both authors report encouraging results. We tested the two methods and we found that the regions with significant motion are poorly coded. However, backgrounds are well coded, which was to be expected considering that in still image coding the uniform areas are better coded than the textured ones. We simulated the sub/oversampling and found satisfying results with good codes. Figure 3 shows frames of a G.O.P. extracted from the reference sequence Miss America coded with respect to Lazar's scheme. The third column gives the localization of badly coded range cubes (with

high MSE). These cubes correspond to moving regions, which are eyes and mouse in our case. In the next section, we report the proposed improvements to the existing schemes in order to reduce the percentage of badly coded blocks.

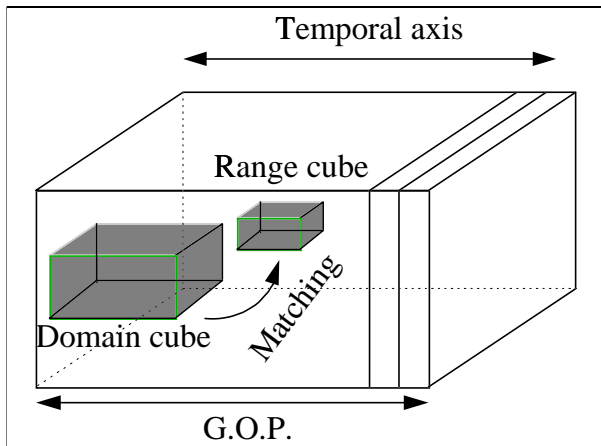


Figure 2: Matching between 3-D cubes in a G.O.P.

4. PROPOSED IMPROVEMENTS

Isometries play a prominent part in improving the coding quality. We combined some of Lazar’s isometries with other rotations and reflections about principal axes and planes (including Naemura’s isometries), while maintaining the total number of used isometries in order to obtain an objective comparison. Results are slightly better than the scheme with Lazar’s isometries. Also, we introduce a novel way to construct domain cubes. It consists of taking the successive frames constituting a given domain cube with some spatial shifting, having prefixed maximum amplitudes in the x and the y directions, see Figure 4. The aim of such a construction is to find self-similarities between the cubes containing motion. We thus obtain a sort of motion compensation without leaving the 3-D framework. In order to avoid time consumption, the domain cubes constructed with shifting were only used in the matching of badly coded range cubes. This scheme gives better results than the classical ones (without shifting), while reasonably

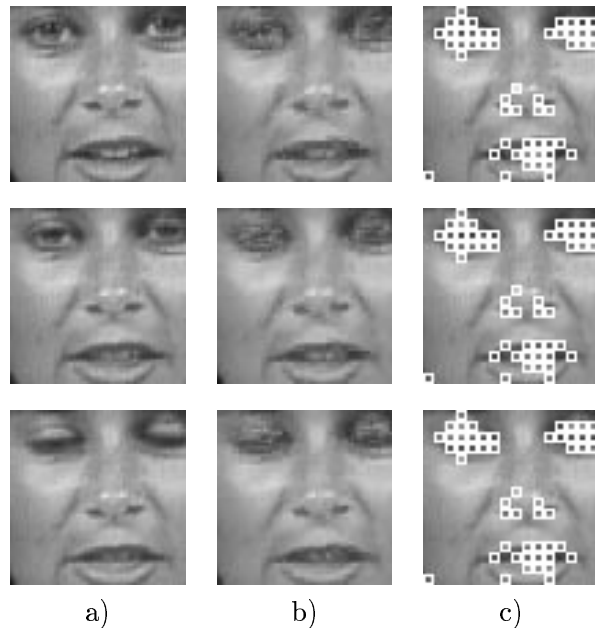


Figure 3: Three consecutive frames of a G.O.P.: a) original, b) coded-decoded (Lazar’s scheme), c) blocks with high MSE.

increasing the research domain and so the computing time. Figure 5 shows frames from a G.O.P. coded with shifting and isometries. As we can see in the third column, there are fewer blocks with high MSE.

5. CONCLUSION AND FURTHER WORK

The 3-D I.F.S. based method to compress image sequences is presented in this paper. Simulation results on the already published schemes and on an application of this method (sub/oversampling) are given and discussed. Some variations of these schemes are proposed, in particular, a novel way to construct domain cubes, allowing better coding quality than the previous schemes. Given the high information quantity in the range cubes, non-affine models can be envisaged in further works, but they raise the problem of contractivity and so of the convergence of the code.

Acknowledgments. This study is in part supported by FRANCE TELECOM, CNET. The authors thank J. Signès for helpful discussions.

6. REFERENCES

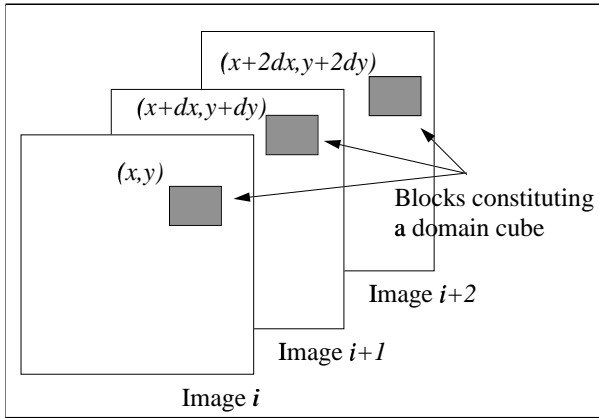


Figure 4: Construction of domain cubes with shifting.

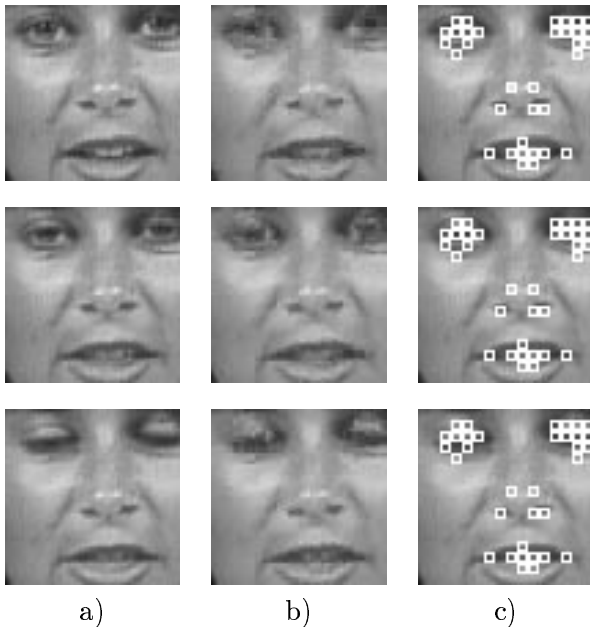


Figure 5: Three consecutive frames from a G.O.P.: a) original, b) coded-decoded with our scheme, c) blocks with high MSE.

- [1] A. E. Jacquin, "Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations," *IEEE Trans. on Image Processing*, Vol. 1, No. 1, 1992.
- [2] J. E. Hutchinson, "Fractals and Self Similarity," *Indiana University Mathematics Journal*, Vol. 30, No. 5, 1981.
- [3] M. F. Barnsley, L. P. Hurd, *Fractal Image Compression*. AK Peters Books in Computer Science and Related Areas, 1993.
- [4] Y. Fisher (Ed.), *Fractal Image Compression: Theory and Application to Digital Images*, Springer Verlag, New York, 1994.
- [5] J. M. Beaumont, "Image Data Compression using Fractal Techniques," *British Telecom technology Journal*, Vol. 9, No. 4, pp. 93-109, Oct. 1991.
- [6] J.-L. Dugelay, J.-M. Sadoul, M. Barakat, "A Comparative Study of Existing Approaches to Moving Picture Coding using I.F.S.," *Int. Symp. on Multimedia Communications & Video Coding*, New York, Oct. 95.
- [7] B. Hürtgen, P. Büttgen, "Fractal Approach to low rate Video Coding," *VICIP'93 Boston*, SPIE, Vol 2094, pp. 120-131, Nov. 1993.
- [8] S. Lazar, L.T. Bruton, "Fractal Block Coding of Digital Video," *IEEE Trans on Circuits and Systems for Video Technology* 4,3 (1994), pp. 297-308.
- [9] T. Naemura, H. Harashima, "Fractal coding of a multi-view 3d image," *IEEE ICIP'94 Austin, Texas*, Nov. 1994.
- [10] E. Polidori, J.-L. Dugelay, "Zooming using I.F.S.," *Fractal Image Encoding and Analysis*, a NATO Advanced Study Institute, July 1995, Trondheim.