MPEG-4 COMPRESSION ARTIFACTS REMOVAL ON COLOR VIDEO SEQUENCES USING 3D NONLINEAR DIFFUSION

P. Bourdon, B. Augereau, C. Olivier and C. Chatellier

IRCOM-SIC, UMR-CNRS 6615, University of Poitiers
Bât SP2MI, Bvd M. et P. Curie, BP 30179
86962 Futuroscope Chasseuneuil Cedex, FRANCE
Tel : (+33) 5 49 49 65 85 ; Fax : (+33) 5 49 49 65 70
E-mail : {bourdon, augereau, olivier, chatellier}@sic.sp2mi.univ-poitiers.fr

ABSTRACT

Today's needs for low bitrates in data compression are responsible for MPEG and ITU family of video coding standards' success. However, it is important to realize that such standards involve lossy compression. This means a reconstructed image sequence will contain degradations relative to the original, because information is discarded during compression. Lower bitrate compression modes often introduce block edge, ringing, blurring, color bleeding, or mosquito noise effects. In this paper, we propose a Partial Differential Equations-based (PDE) method to reduce compression-related artifacts on MPEG-4 color image sequences. The main advantage of our approach is that both temporal and spatial artifacts removal are performed at the same time, by considering an image sequence as a 3D object. Although PDE-based methods for still images restoration are becoming quite popular, extensions to image sequences remain rare, especially in color. This is also the purpose of this paper to introduce one.

1. INTRODUCTION

MPEG and ITU family of standards emerged recently as an important development in the field of video coding. Because they addressed the growing demand for low bitrate compression, they quickly became popular among video and broadcast professionals. However, it should be recalled that such standards involve lossy compression, that is the reconstructed image contains degradations relative to the original, because information is discarded during compression, as opposed to lossless compression, in which the reconstructed image is identical to the original one. The lower the bitrate, the more pronounced the degradation. As an example, MPEG low bitrate compressions are known to introduce block edge effects (intensity discontinuities at the boundaries of adjacent blocks), ringing artifacts (rippling of high contrast edges), blurring (loss of spatial details in areas such as textures and edges), color bleeding (blurring effect in the color component of the image), and mosquito noise (temporal fluctuations of luminance/chrominance levels around high contrast edges or moving video objects).

In this paper, we propose a post-processing method, based on Partial Differential Equations (PDE), to reduce these artifacts. Such operation should result in a better looking image sequence, as well as in an improvement in terms of PSNR (Peak Signal to Noise Ratio). Compressed video enhancement techniques have already been proposed in literature [1, 2]. In [2], Choi introduces a method in which both temporal and spatial correlations in a video sequence are exploited. This method happens to be very interesting, since most techniques usually involve frame-per-frame enhancement. However, he only gives results for grayscale sequences. Our algorithm also exploits both spatial and temporal information, by considering the sequence as a 3D object, and is able to deal with color images. This paper will actually allow us to introduce a PDE-based method for image sequence restoration, which is an emerging domain compared to 2D image restoration, especially when it comes to color videos. This method is based on Perona-Malik's nonlinear filter [3], that we extend to 3D images, using our own diffusion function, as well as a 3D extension of Di Zenzo's gradient for color images [4] to detect discontinuities.

In section 2 , we present the principles of PDE-based still grayscale image denoising using a variational approach, to focus on the case of color image sequence and introduce our compression artifacts removal method in section 3. Finally, we will provide experimental results on MPEG-4 test sequences in section 4.

2. PDE's AND NONLINEAR FILTERING

We propose in this part a variational approach of noise reduction in image processing. This approach, proposed by Deriche and Faugeras [5], allows to unify most PDE-based methods in image enhancement and multi-scale analysis under the same formalism, especially Perona and Malik's [3].

2.1. Formulation

Let $I(x, y) = I$ be a still, grayscale, image, which can be represented by a function of $\Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ which associates with the pixel $(x, y) \in \mathbb{R}^2$ its gray level $I(x, y)$; Ω is the support of the image. We suppose I is describing a real scene from an observed (noisy) image $I_0(x, y) = I_0$. The reconstruction of I from I_0 can be written as the following minimization problem :

$$E(I) = \underbrace{\int_{\Omega} f(I, I_0) d\Omega}_{term\ 1} + \lambda \underbrace{\int_{\Omega} \Phi(\|\nabla I\|) d\Omega}_{term\ 2} \quad (1)$$

in which $E(I)$ is the energy to minimize, f an error function, λ a constant, and $\Phi(.)$: a regularization function to be defined. Equation (1) is composed of two terms : a data-fidelity term ($term\ 1$), which penalizes variations between the original image and the restored one, and a regularization term ($term\ 2$), which penalizes noise while preserving the image's edges. The idea is to perform heavy smoothing in low gradient areas (homogeneous areas), and light smoothing in high gradient areas (edges).

2.2. Solution

Conditions for minima of $E(I)$ are given by the Euler-Lagrange equation :

$$\nabla E(I) = \frac{\partial f(I, I_0)}{\partial I} - \lambda div(\Phi'(\|\nabla I\|) \frac{\nabla I}{\|\nabla I\|}) = 0 \quad (2)$$

that leads to the following PDE (without taking into account the data-fidelity term) :

$$\frac{\partial I(x, y, t)}{\partial t} = div(\Phi'(\|\nabla I\|) \frac{\nabla I}{\|\nabla I\|}) \quad (3)$$

$$= \Phi''(\|\nabla I\|) I_{\xi\xi} + \frac{\Phi'(\|\nabla I\|)}{\|\nabla I\|} I_{\eta\eta} \quad (4)$$

with $I_{\xi\xi}$ and $I_{\eta\eta}$ the second directional derivatives of I in the gradient's direction ξ and in its orthogonal direction η. t is the time of diffusion, and controls the smoothing strength. We can actually recognize in (3) Perona-Malik's PDE [3], $\frac{\Phi'(\|\nabla I\|)}{\|\nabla I\|}$ being the diffusion function.

2.3. Conditions of stability

Deriche-Faugeras' parabolic PDE (4) allows us to define the diffusion process' conditions of stability. These are :

$$\begin{cases} \Phi''(0) \geq 0 \quad and \quad \Phi'(0) \geq 0 \\ \lim_{\|\nabla I\| \to 0} \frac{\Phi'(\|\nabla I\|)}{\|\nabla I\|} = \lim_{\|\nabla I\| \to 0} \Phi''(\|\nabla I\|) = \Phi''(0) \\ \lim_{\|\nabla I\| \to \infty} \Phi''(\|\nabla I\|) = 0, \lim_{\|\nabla I\| \to \infty} \frac{\Phi'(\|\nabla I\|)}{\|\nabla I\|} = 0 \\ \lim_{\|\nabla I\| \to \infty} \frac{\Phi''(\|\nabla I\|)}{\frac{\Phi'(\|\nabla I\|)}{\|\nabla I\|}} = 0 \end{cases} \quad (5)$$

It means that smoothing is performed in all directions (isotropic diffusion) for low gradient areas, while it is only performed along the gradient's orthogonal direction (anisotropic diffusion) for high gradient areas. Several diffusion functions $\Phi(.)$ can be found in literature [5].

3. COMPRESSION ARTIFACTS REMOVAL

Previous section introduced the principles of PDE-based denoising in the case of a still grayscale image. In this section, we propose an extension to color image sequences, and apply it to compressed videos for artifacts removal. The use of Partial Differential Equations for image sequences restoration remains very rare, especially when it comes to color images [6]. Our method consists in considering an image sequence as a 3D object, and looking for local variations in this object to perform anisotropic diffusion in both spatial and temporal directions.

3.1. Extension to color image sequence

In section 2 , we considered the image as a still grayscale image, defining $I = I(x, y)$ as a $\mathbb{R}^2 \rightarrow \mathbb{R}$ application. In the case of color image sequences, we now define a 3D vectorial function $\vec{I}(x, y, z) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, which associates with the pixel (or voxel) $(x, y, z) \in \mathbb{R}^3$ its color levels $R(x, y, z)$, $G(x, y, z)$ and $B(x, y, z)$:

$$\vec{I}(x, y, z) = \begin{bmatrix} I_1(x, y, z) \\ I_2(x, y, z) \\ I_3(x, y, z) \end{bmatrix} = \begin{bmatrix} R(x, y, z) \\ G(x, y, z) \\ B(x, y, z) \end{bmatrix} \quad (6)$$

z being the temporal coordonate (not to be confused with t, the time of diffusion).

Equation (3) can be easily extended from I (scalar, grayscale 2D image) to $\vec{I}$ (vectorial, color 3D image) :

$$\begin{cases} R(x, y, z, t) = div(\frac{\Phi'(\|\nabla \vec{I}\|)}{\|\nabla \vec{I}\|} \nabla R) \\ G(x, y, z, t) = div(\frac{\Phi'(\|\nabla \vec{I}\|)}{\|\nabla \vec{I}\|} \nabla G) \\ B(x, y, z, t) = div(\frac{\Phi'(\|\nabla \vec{I}\|)}{\|\nabla \vec{I}\|} \nabla B) \end{cases} \quad (7)$$

Marginal approaches seem to be the easiest solution for color image diffusion. They consist in processing a k-bands multispectral image as k scalar images. Diffusion would then be performed on each spectral band I_i using $\|\nabla I_i\|$ as an edge detector. Alas, such method introduces false colors, since diffusion directions can change from one spectral band to another. We decided to use a common representation of the image's discontinuities, to have common diffusion directions for R, G, and B. $\left\|\nabla \vec{I}\right\|$ can be defined as a multispectral 3D gradient, which is calculated using an extension of Di Zenzo's gradient for still color images [4]. Di Zenzo's gradient norm is based on differential geometry of surfaces, and has already been successfully applied for anisotropic diffusion of still color images. It consists in defining a tensor gradient, associated with a vector field, to look for local variations in the image. The extension of Di Zenzo's tensor to our 3D space can be written as follows :

$$\left\|d\vec{I}\right\|^2 = \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}^T . \underbrace{\begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{12} & g_{22} & g_{23} \\ g_{13} & g_{23} & g_{33} \end{bmatrix}}_{G} . \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} \quad (8)$$

with : $g_{ij} = \sum_{k=1}^{3} \frac{\partial I_k}{\partial x_i} . \frac{\partial I_k}{\partial x_j}$ $(x_1 = x, x_2 = y, x_3 = z)$

G is defined as the multispectral tensor. It is a symmetric, positive semidefinite matrix, which highest eigenvalue corresponds to the squared gradient norm.

3.2. Diffusion function

In [3], Perona-Malik use the following anisotropic diffusion function in their PDE (3) :

$$D(\left\|\nabla \vec{I}\right\|) = \frac{\Phi'(\left\|\nabla \vec{I}\right\|)}{\left\|\nabla \vec{I}\right\|} = \exp(-(\left\|\nabla \vec{I}\right\|/k)^2) \quad (9)$$

in which k is a threshold that defines the frontier between low gradient ($\|\nabla I\| < k$) and high gradient ($\|\nabla I\| > k$) areas. Alas, this threshold happens to be hard to define and varies with the image to restore. Unlike Perona-Malik's, our diffusion function $D(.)$ isn't directly linked to $\|\nabla I\|$, but to a normalized version of it, thus allowing us to define a threshold α, which values remain the same no matter what image is being processed. We also decided to use a function that only performs forward diffusion, since its only purpose is denoising, not edge enhancement (2.3). $D(.)$ is defined as :

$$D(s) = \frac{\Phi'(s)}{s} = (1 + s^2).e^{-s} \quad (10)$$

with : $s = \alpha . \frac{Max(\left\|\nabla \vec{I}\right\| - Min(\left\|\nabla \vec{I}\right\|)}{Var(\left\|\nabla \vec{I}\right\|)} . \left\|\nabla \vec{I}\right\|$

Threshold α allows us to introduce an anisotropy level for $D(.)$. We can easily notice that for $\alpha = 0$, $D(s)$ equals 1, discarding s : in this case, the diffusion process turns into pure isotropic diffusion.

3.3. Parameters estimation

Another difference between Perona-Malik's method and ours is the fact that instead of using a constant threshold, we decide to make α evolve with time, from purely isotropic diffusion (heavy denoising) to highly anisotropic diffusion (softer denoising and better edge preservation). The algorithmic implementation then requires the determination of two parameters :

– N, the number of iterations (discrete equivalent to T, the total time of diffusion)

– α, which value evolves according to iteration n. We rename it α_n, with $n \in [0, N-1]$

We are looking for an image quality criteria, in order to determine optimal parameters α_n and N by judging the restored image quality. Many quality measures have been proposed for that purpose, including psychovisual studies. We decided to use the well-known Mean Square Error (MSE), which main advantage is its simplicity. The parameters' determination can be written as a minimization problem :

$$\widehat{\alpha_n} = \arg\min_{\alpha_n} \left(\frac{\sum_{x=0}^{P-1} \sum_{y=0}^{Q-1} \sum_{z=0}^{R-1} \sum_{i=1}^{3} (I_i(x,y,z,n) - I_{0i}(x,y,z))^2}{3 \times P \times Q \times R} \right) \quad (11)$$

for $n \in [0, N-1]$, with $P \times Q \times R$ the size of the sequence.

As the diffusion process goes on, the MSE gets smaller, indicating that the restored sequence is getting closer to the original. α_n values start from 0 (isotropic diffusion), and increase as the number of iterations grows, so to minimize the MSE at each iteration. N is the number of iterations above which the MSE tends to be steady (experiments have shown N barely exceeds 20).

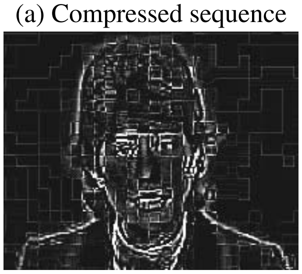
The parameters' determination is a pre-processing, that occurs before the encoded image sequence is transmitted. Once estimated, diffusion parameters are placed in the compressed video file's header, to be later extracted during decompression and used for post-processing.

4. EXPERIMENTAL RESULTS

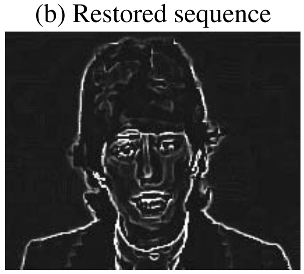
Figure 1 shows results obtained for the "Claire" sequence (352 × 288, 160 frames, 15 fps) encoded at 24 kb/s using Microsoft MPEG-4 Visual Reference Software version 2 FDAM1-2.3-001213. We can notice visible enhancement between the compressed sequence (Figure 1.a an 1.c) and the restored one (Figure 1.b an 1.d) : blockiness and ringing artifacts have been removed (mosquito noise as well, although temporal artifacts can't be observed on paper). The results are summarized in Figure 1.e, where the PSNR values of the recovered frames are shown (as we already mentioned, there are better image quality measures than this one, alas none seems to have been proposed as a freely available software application).

(a) Compressed sequence

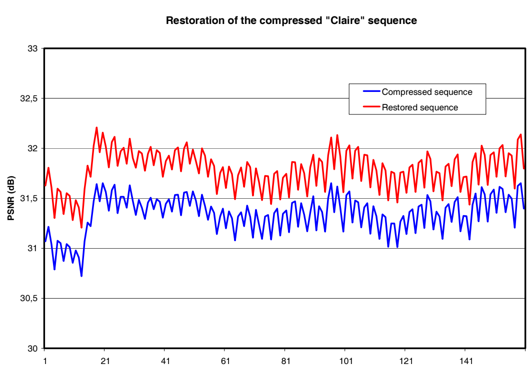
(b) Restored sequence



(c) Edges of (a)



(d) Edges of (b)



(e) PSNR of recovered frames

Fig. 1. Restoration of "Claire" sequence

5. CONCLUSION

In this paper, we've presented a PDE-based color image sequence restoration method, and used it in order to remove spatial and temporal compression artifacts on low bitrate videos. Extensions of Perona-Malik's nonlinear filter and Di Zenzo's gradient for color images to a 3D space have been proposed, in order to adapt them to color image sequences. We've also introduced a forward diffusion function, and a diffusion parameters estimation technique. The main advantages of our method are its simplicity, the fact that it includes a parameters' estimation that allows it to work a totally unsupervised way, and a low processing time (less than 5 seconds for 1 iteration on a 352 × 288 pixels, 10 frames, color video[1], knowing the complete restoration process hardly requires more than 20 iterations). Results have been shown on a MPEG-4 compressed test sequence. These results may open up new perspectives, since we've only focused on post-processing. Pre-processing diffusion may also improve compression, by removing spurious noise and other insignificant features from the original data.

6. REFERENCES

[1] C. Segall and A. Katsaggelos, "Pre- and post-processing algorithms for compressed video enhancement," in *Proceedings of the 34th Asilomar Conference on Signals, Systems and Computer*, Pacific Grove, California, USA, October 2000.

[2] M. C. Choi, Y. Yang, and N. P. Galatsanos, "Multichannel regularized recovery of compressed video sequences," *IEEE Transactions on Circuits and Systems II*, vol. 48, no. 4, pp. 376–387, April 2001.

[3] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[4] S. Di Zenzo, "Note : A note on the gradient of a multi-image," *Computer Vision, Graphics, and Image Processing*, vol. 33, no. 1, pp. 116–125, January 1986.

[5] R. Deriche and O. Faugeras, "Les EDP en Traitement des Images et Vision par Ordinateur," *Traitement du Signal*, vol. 13, no. 6, 1996.

[6] P. Kornprobst, R. Deriche, and G. Aubert, "A pde based coupled method for image restoration and motion segmentation," in *Proceedings of the 5th European Conference on Computer Vision*, Freiburg, Germany, June 1998, vol. II, pp. 548–562.

[1] Results obtained using a Pentium III 450Mhz 128Mb