

ADAPTATION OF THE METHOD OF ESTIMATION AND MOTION COMPENSATION IN THE TRANSMISSION OF THE VIDEO SEQUENCE IN JPEG 2000

Abdou Khadre Diop, Khaly Tall, Idy Diop and Sidi Mohamed Farssi

Laboratoire d'Imagerie Médicale et BioInformatique, Ecole Supérieure Polytechnique,
Cheikh Anta Diop university, Dakar, Senegal
djeylani2001@yahoo.fr

ABSTRACT

The traditional video compression standards like hybrid encoders MPEG-4 part10 or H.264 use the principle of estimation and motion compensation for inter-frame coding of video sequences with a lot of movement. Motion JPEG 2000 (Part 3 of the JPEG 2000 standard) has a higher performance compared to traditional video compression standards [1], however the problem encountered in the transmission of JPEG 2000 video sequences is the loss of packets in the codestream MJ2 file.

In this paper, we propose the use of the method based on the principle of estimation and motion compensation to reliably reconstruct the lost packets in the transmission of MJ2 codestream file with a lot of movement. An evaluation of the proposed method was performed on a number of video format MJ2 in terms of measurement of PSNR (Peak Signal to Noise Peak). The results are used to evaluate the performance of the method.

KEYWORDS

Motion JPEG 2000, conditional replenishment, estimation and motion compensation, packetization JPEG 2000 standard, order of progression, PSNR.

1. INTRODUCTION

A video signal is 3D (2D + t) with temporal but spatial correlations. Temporal correlations are mainly related to the motion of objects over time. Spatial correlations are related to the particular contours of objects and reasons for textures [2].

While traditional video coders like H.264 or MPEG4 predictive coders are based on a motion compensation loop and encoding 2D prediction residuals by Discrete Cosine Transform (DCT), Motion JPEG 2000 Part 3 of the JPEG 2000 standard that define the use of the JPEG 2000 standard for the treatment of video sequences is a concatenation of images in JPEG 2000 encoded using Discrete Wavelet Transform (DWT). And controlling the movement in video compression using the DWT is a major difficulty due to decomposition algorithms used by the DWT [7], [8].

Faced with the transmission of the codestream Motion JPEG 2000 file represented as a sequence of packets, there is packet loss in the codestream of these files. Therefore we can get a bad quality in the reconstruction of the video file transmitted especially if the sequence has a lot of movement.

To alleviate the consequences of these transmission errors, we use the method of estimation and motion compensation, a method well known in the inter-frame coding of video sequences using

traditional video compression, to reconstruct reliably lost packets in the codestream of Motion JPEG 2000.

To evaluate the performance of our proposed method, four video files specified by their movement were chosen to test the method.

To describe our work, the paper presents in the first part features of Motion JPEG 2000 and Motion JPEG 2000 packetization to locate the various packages in the codestream video file. The second part presents the method of estimation and motion compensation: method used to reconstruct lost packets. The final section presents the results obtained in terms of objective qualities. Finally we present the conclusion and future works.

2. MOTION JPEG 2000 FEATURES

Motion JPEG 2000 is defined as Part 3 of the JPEG 2000 standard. It is an extension of the JPEG 2000 for the treatment of video sequences. The principle is that each frame of video is encoded in JPEG 2000. A Motion JPEG 2000 format MJ2 or MJP2 is a concatenation of images in JPEG 2000: JP2 or J2K. (Figure 1)

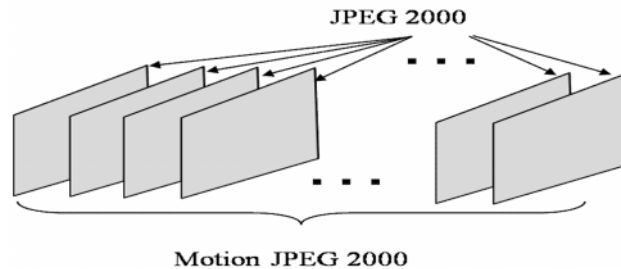


Figure 1. Composition of a Motion JPEG 2000

The allocation algorithm flow of the JPEG 2000 standard allows creating packages like all other allocation algorithms flow. Thus each packet corresponds to a quality level of a level of resolution of a component of the image [6].

The packetization is used to add to each packet codestream following a header containing all the coding parameters and in order of inclusion depending on the type of growth desired (e.g. by resolution, by quality level). Thus, it is possible to randomly have access to the codestream of each frame. So, we can determine the information in each packet (Figure 2).

As a transmission technology is not 100% reliable, the transmission of MJ2 file can result in loss of packets in the codestream file.

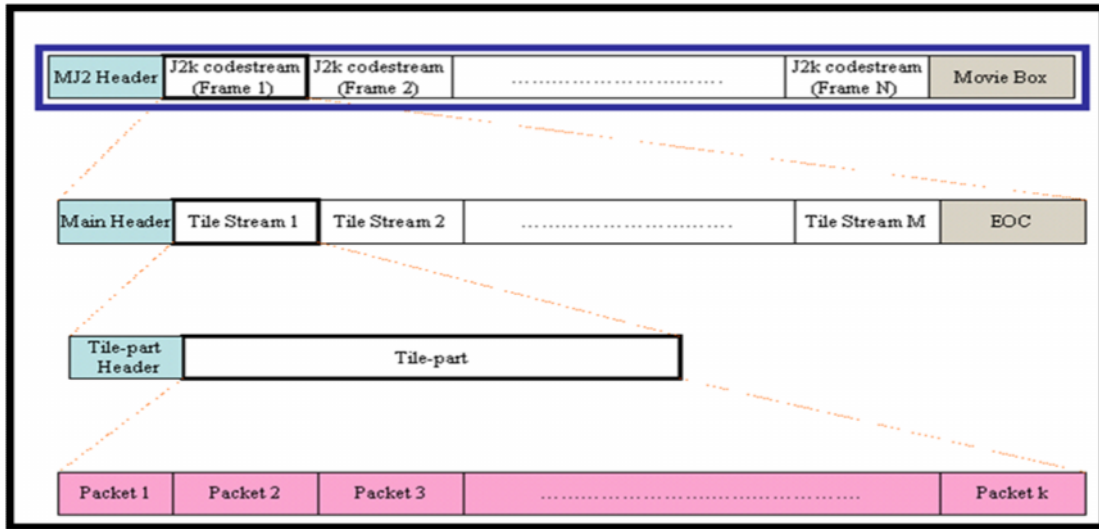


Figure 2. Structure of the Motion JPEG 2000 file

3. THE ESTIMATION AND MOTION COMPENSATION IN MOTION JPEG2000

We propose the method of estimation and motion compensation used in the inter-frame coding of video sequences in MPEG4 or H.264 for the recovery of lost packets during transmission. The purpose of the estimation and motion compensation is to estimate the motion field which corresponds to the conservation of the light intensity at each point along the path of movement.

To accomplish this task, we use the equation of Displaced Frames Difference (DFD). The term of the equation is given by: $DFD = I(x+dx, y+dy, t+ t) - I(x, y, t) = 0$, between times t and $t+ t$. The algorithm of the proposed method applied to two successive images can be shown in Figure 3 and 4.

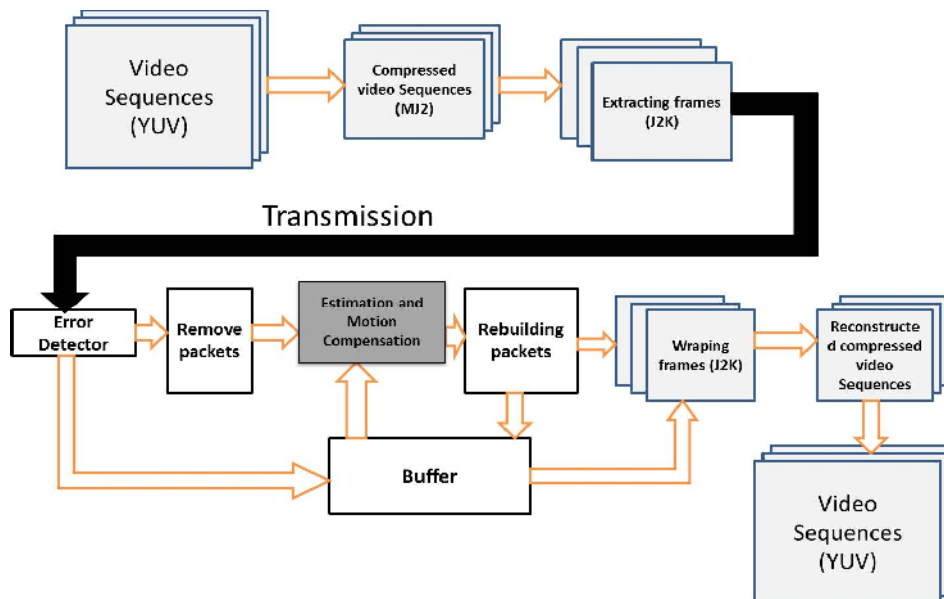


Figure 3: algorithm using estimation and compensation method

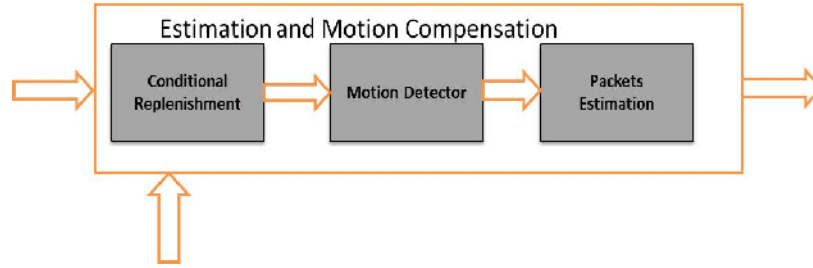


Figure 4. Different steps of the proposed method

Between these two images, one image is transmitted and the other image previously transmitted. If we lose packets in the codestream of the transmitted image, replenishment conditional method is used to recover packets of previously transmitted image for placement in the transmitted image [3], [9]. After packets are placed in the transmitted image, motion detector implemented from the equation Displaced Frames Difference (DFD) can locate the part of the image [4], [5]. Therein, we meet the movement to set a minimum range of motion at which we will apply our method to reconstruct objectively lost packets.

This minimum range of motion detected may be expanded to detect the maximum motion. For this we defined the variable X in order to widen the minimum range of motion by dilation.

Compression format video sequence MJ2, extracting images from a MJ2 to J2K file, wrapping J2K images in a MJ2 file are obtained using the OPENJPEG¹ platform.

4. RESULTS

4.1 Characterization of video sequences

To evaluate the proposed method, an evaluation is performed on four standard test video sequences. News sequence in CIF format (figure 5.1 and 5.2), 288 * 352 * 3 pixels per frame, 300 frames, is characterized by a quasi-stationary movement of the camera. In this sequence we see the movement with the dancer in the background.



Figure 5.1. Transmitted image



Figure 5.2. Previously transmitted image

The sequence Highway, in CIF format (figure 6.1 and 6.2), 288 * 352 * 3 pixels per image, 700 images, is characterized by a greater movement of the camera. This car traveling on a highway is followed by the camera. And there is movement in the part of the image where the car is.

¹ Platform of compressed images developed at the catholic Louvain University



Figure 6.1. Transmitted image



Figure 6.2. Previously transmitted image

Hall sequence in CIF format (figure 7.1 and 7.2), $288 * 352 * 3$ pixels per frame, 300 frames, is characterized by a smaller movement of the camera. In this sequence, a man moves in a corridor. So the movement is in the part of the image where this man is.



Figure 7.1. Transmitted image



Figure 7.2. Previously transmitted image

Foreman sequence in CIF format (figure 8.1 and 8.2), $288 * 352 * 3$ pixels per frame, 300 frames is characterized by a sudden and rapid movement of the camera. Since the movement of the camera is sudden and rapid, the movement is considered part of the image representing the man.



Figure 8.1. Transmitted image



Figure 8.2. Previously transmitted image

4.2. Experimental results and analysis

These compressed video sequences obtained in format MJ2, have been used to detect movement in a given image. Our method allows us to determine the minimum range of motion detection which is the portion of the image where we see the movement. Assuming that losses occur in the part where there is movement, the method can reconstruct the packets by using the previously transmitted image. In practice we find the figures (figure 9: $X=0$ and figure 10: $X=40$). We give the example of the sequence News.



Figure 9.1 Motion Detection, Figure 9.2 erroneous image, Figure 9.3 reconstructed image (X=0)



Figure 10.1 Motion Detection, Figure 10.2 erroneous image, Figure 10.3 reconstructed image (X=40)

To evaluate the objective quality of the proposed method, we used the PSNR between the original image and the reconstructed image or the original image and the erroneous image, given by equation (1):

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (1), \quad MSE = \frac{1}{HL} \sum_{i=1}^H \sum_{j=1}^L (I(i,j) - \hat{I}(i,j))^2 \quad (2)$$

Where MSE is the mean square error calculated by Equation (2):

Where H and L are respectively the height and width of each image in the video sequence, I is the original image and the image \hat{I} is the erroneous or reconstructed image.

We obtained different values of PSNR depending on the variable X which broadens the minimum range of motion scaling. In the phase where we lose packets in the minimum range of motion detection, the following results (Table 1) are used to assess the objective quality between the original image and the erroneous image transmitted.

Table 1. PSNR in terms of X with erroneous image for each sequence

X	0	5	10	15	20	25	30	35	40
News	23,9	21,7	20,3	19,0	17,9	16,6	15,5	14,6	13,7
Highway	33,4	28,2	26,2	23,9	21,5	19,6	18,0	16,6	16,4
Hall	27,4	23,4	20,9	18,5	16,6	15,1	13,9	12,9	11,9
Foreman	24,1	21,4	19,2	17,4	15,9	14,6	13,5	12,4	11,6

Figure 11 shows the movement in different sequences we used. We note that the objective quality of the sequence Highway is better because the field of motion detection is smaller with this sequence. Therefore the loss of packets in the range of motion detection in this sequence is smaller than the other sequences.

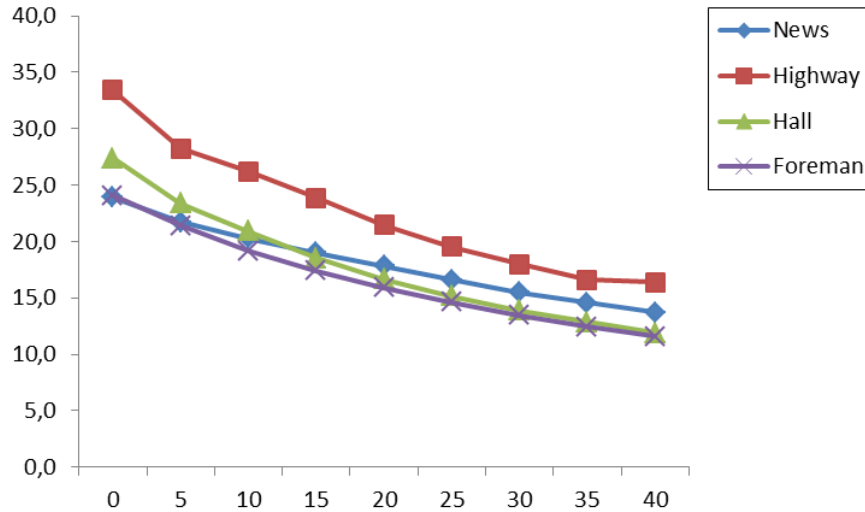


Figure 11. PSNR in terms of X with erroneous image for each sequence

The second phase of our practice allows the reconstruction of lost packets by using the previously transmitted image as the method permits us to do so. The following results (Table 2) allow us to assess the objective quality between the original image and the reconstructed image transmitted.

Table 2. PSNR in terms of X with reconstructed image for each sequence

X	0	5	10	15	20	25	30	35	40
News	36,2	33,9	33,1	32,8	32,7	32,5	32,4	32,4	32,3
Highway	40,5	35,8	34,4	33,8	33,2	32,8	32,5	32,2	32,1
Hall	38,1	36,0	34,2	33,2	32,3	31,7	31,1	30,6	29,9
Foreman	36,0	34,2	33,4	32,6	31,5	31,1	30,5	29,9	29,5

The minimum range of motion detection is even smaller in the sequence Highway in the other sequences. The reconstitution of lost packets in the minimum range has much more reliability for this sequence from Figure 12.

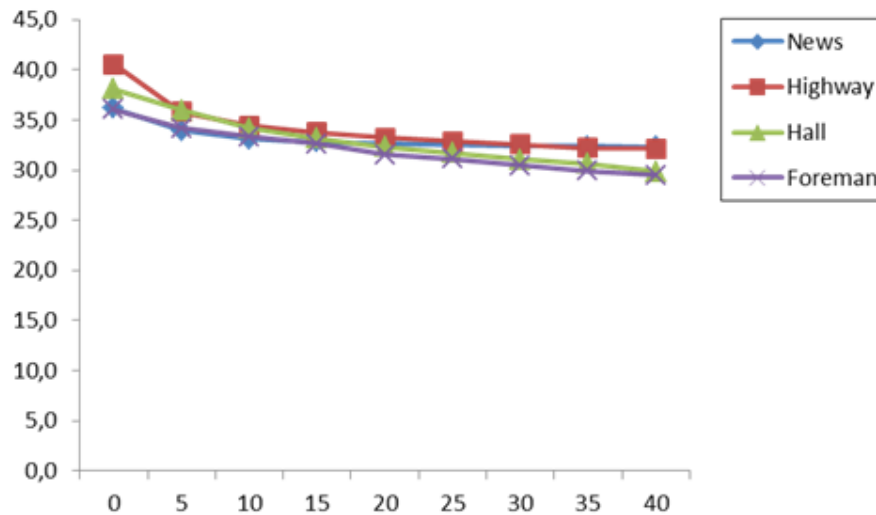


Figure 12. PSNR in terms of X with reconstructed image for each sequence

Our method for the recovery of lost packets during transmission of JPEG 2000 video has high reliability; however, the movement must be contained in the minimum field of motion detection.

5. CONCLUSION

We have presented an approach for robust transmission of JPEG 2000 movie based on the method of estimation and motion compensation. Although it is well known in the inter-frame coding sequences MPEG-4 and H.264, the method allows reconstructing lost packets in an image transmitted by using the previously transmitted Image.

An objective assessment of the method is performed on four standard test sequences. The results allowed highlighting the effects of movements in different movies. Thus, the method allows for greater transmission reliability if any movement is contained in the minimum field of motion detection in JPEG 2000 video.

As future works, we intend to resort to other methods of assessment of subjective visual quality of the image to reconstruct markedly the lost packets.

REFERENCES

- [1] Detlev Marpe, Valeri George, Hans L. Cycon, and Kai U. Barthel, (2004) " Performance evaluation of Motion-JPEG2000 in comparison with H.264/AVC operated in pure intra coding mode" , *Wavelet Applications in Industrial Processing, Providence, RI, USA*, vol. 5266 of Proceedings of SPIE, pp. 129-137.
- [2] Wissal Hassen, Mbainabeye Jérôme, Noureddine Ellouze & Christian Olivier, (2009) "Compression vidéo couleur basée sur le codage entropique du signe des coefficients en ondelettes en utilisant la compensation de mouvement", *TAIMA*.
- [3] François-Olivier Devaux, Jérôme Meessen, Christophe Parisot, Jean-François Delaigle, Benoît Macq & Christophe De Vleeschouwer, (2007) "A flexible video transmission system based on JPEG 2000 conditional Replenishment with multiples references", *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*.

- [4] Benjamin Le Guen, Stéphane Pateux & Jacques Weiss, (2009) “Compensation spatio-temporelle Globale pour le Codage Vidéo par Ondelettes 3D”, *Compression et Représentation des Signaux Audiovisuels (CORESA)*.
- [5] Marie Andrée Agostini, Thomas Andre, Marc Antonini & Michel Barlaud (2005) “Codage scalable des vecteurs mouvements pour la compression de vidéos basée ondelettes”, *Compression et Représentation des Signaux Audiovisuels (CORESA)*.
- [6] Majid Rabbani & Rajan Joshi (2002) “An overview of the JPEG 2000 still image compression standard”, *Signal Processing: Image Communication*, Vol. 17 pp3-48.
- [7] James E. Fowler & Béatrice Pesquet-Popescu (2007) “An Overview on Wavelets in Source Coding, Communications, and Networks”, *EURASIP Journal on Image and Video Processing* Article ID 60539.
- [8] François G. Meyer, Amir Averbuch & Ronald R. Coifman (1997) “Motion compensation of wavelet coefficients for very low bit rate video coding”, *International Conference on Image Processing (ICIP)*.
- [9] Abdou Khadre Diop, Idy Diop, Mamadou Niang, Khaly Tall & Sidi Mohamed Farssi (2009) “Rafraîchissement conditionnel dans la transmission de séquences vidéo JPEG 2000”, *Journal des Sciences Pour l’Ingénieur*.

Author

Abdou Khadre DIOP was born the 10th of July 1979 in Pikine, Dakar, Senegal. He received the degree in Telecommunication and Electronics Engineering from the Gaston Berger University of Saint-Louis (Senegal) in 2005. In 2012, he was employed as ATER at the Department of Technology in Information Communication of the Alioune Diop University of Bambey. In 2008, he joined the Laboratory of Biomedical Imaging & Bioinformatics for a joint PhD thesis. His current research areas cover transmission of the video sequence in JPEG 2000.

