# High-Resolution Color Image Reconstruction from Compressed Video Sequences

*Javier Mateos*[1], *Aggelos K. Katsaggelos*[2], *Rafael Molina*[1]

(1) Dep. de Ciencias de la Computación e I. A. Universidad de Granada. Granada, Spain
e-mail: {jmd, rms}@decsai.ugr.es
(2) Department of Electrical and Computing Engineering. Northwestern University. Evanston, IL
e-mail: aggk@ece.nwu.edu

**ABSTRACT**

In this work we propose an algorithm for the estimation of high resolution color frames from a low resolution compressed color video sequence. The algorithm exploits the existing correlation between the high and low resolution frames to obtain a high resolution frame reducing the artifacts introduced by the compression process. The performance of the proposed algorithm is demonstrated experimentally.

## 1 Introduction

High-resolution images are useful and often an important issue for many applications. Remote sensing applications, medical imaging, surveillance or frame freeze in video are some of the applications where high resolution images are crucial. An approach to obtain high resolution images is to increase the number of sensor in the camera that captures the images. Although this approach is feasible for some applications, obtaining a dense detector array may be very costly or simply unavailable. Another disadvantage of this approach is that the signal to noise ratio decreases as the sensor size decreases. An alternative approach is to estimate a high resolution image from a sequence of low resolution aliased images. This is possible if there exists subpixel motion between the acquired frames.

In many applications the available low resolution video has been compressed using a block based approach such as H.263 [1] or MPEG [2]. This is, for example, the case in most digital video cameras, in which the acquired data are compressed using one of the video compression standards, in order to reduce the storage requirements. This compression can introduce artifacts in the low resolution video sequence, such as blocking and mosquito artifacts [9], that should also be removed by the resolution enhancement algorithm.

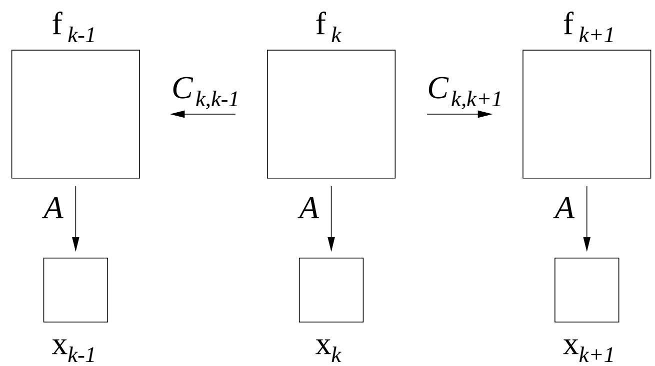


Figure 1: Relationship between high and low resolution frames.

Although a number of algorithms for the resolution enhancement of video have appeared in the literature (see, for example, [3, 6, 8] and the references therein), only very few results have been reported for the problem when the low resolution video sequence has been compressed [4, 5]. In this work we propose an algorithm for the estimation of high resolution color frames from a low resolution compressed color video sequence.

The paper is divided as follows. In section 2 the used notation and the mathematical problem formulation is introduced. Section 3 discusses the proposed estimation model and an algorithm for resolution enhancement of the compressed low resolution frames is proposed. In section 4 results with the proposed algorithm are presented and, finally, section 5 concludes the paper.

## 2 Problem formulation

Let $\mathbf{f}_k$ be the $k$th high resolution color frame, whose color components will be denoted by $\mathbf{f}_k^c$, $c \in \{Y, U, V\}$, $\mathbf{f}_k = \{(\mathbf{f}_k^Y)^t \ (\mathbf{f}_k^U)^t \ (\mathbf{f}_k^V)^t\}^t$, and let $\mathbf{x}_k$ be the corresponding non-compressed low resolution frame. If the original size of the low resolution frame $\mathbf{x}_k^Y$ is $M_1 \times N_1$ and we assume a 4:1:1 subsampling, we note that for the luminance band, $Y$, a $M^Y \times N^Y$ image is obtained with $M^Y = M_1$ and $N^Y = N_1$ and, for the chrominance bands, the images are $M^c \times N^c$, where $M^c = M_1/2$ and $N^c = N_1/2$ with $c \in \{U, V\}$.

Then, each high resolution frame is $PM^c \times PN^c$, $c \in \{Y, U, V\}$, with $P$ an integer determining the degree of subsampling applied to obtain the low resolution image. The relationship between the bands of the high and low resolution frames is described by

$$\mathbf{x}_k^c = A^c \, \mathbf{f}_k^c,$$

where $A^c$, $c \in \{Y, U, V\}$, is a $(M^c \times N^c) \times (PM^c \times PN^c)$ matrix denoting the integration and subsampling operations. If we denote by $A$ the matrix $A = \{(A^Y)^t \ (A^U)^t \ (A^V)^t\}^t$, the relation between the high and low resolution frames is

$$\mathbf{x}_k = A\mathbf{f}_k. \qquad (1)$$

This relation is represented by the vertical arrows in figure 1. Notice that, while $A$ is a downsampling by averaging operator, $A^t$ is an upsampling and zero-order hold operator. Clearly there exists a relation between every two high resolution frames $\mathbf{f}_k$ and $\mathbf{f}_{k-i}$, represented by the horizontal arrows in figure 1, that can be expressed as

$$\mathbf{f}_{k-i} = C_{k,k-i}\mathbf{f}_k, \qquad \text{or} \qquad \mathbf{f}_{k-i}^c = C_{k,k-i}^c \mathbf{f}_k^c,$$

where the matrix $C_{k,k-i}^c$, $c \in \{Y, U, V\}$, represents the motion compensation operator for band $c$ of frame $k$ to frame $k - i$ and $C_{k,k-i}$ represents the motion compensation operator from the color frame $k$ to color frame $k - i$. Since the high resolution frames are related, the low resolution frames are also related according to

$$\mathbf{x}_{k-i} = A\mathbf{f}_{k-i} = AC_{k,k-i}\mathbf{f}_k. \qquad (2)$$

The low resolution frames $\mathbf{x}_k$, $k = 1, \ldots, L$, obtained according to Eq. (2), are now compressed using any of the video compression standards, such as H.263 [1], obtaining the observed compressed sequence of low resolution frames $\mathbf{g}_k$. So, the relation between a high resolution frame and the low resolution compressed sequence can be written as

$$\mathbf{g}_{k-i} = A\mathbf{f}_{k-i} + \mathbf{n}_{k-i} = AC_{k,k-i}\mathbf{f}_k + \mathbf{n}_{k-i}, \qquad (3)$$

which expresses that the frames in the low resolution compressed video sequence can be seen as different noisy subsampled realizations of a single high resolution frame via different motion compensation operations.

## 3 Recovery algorithm

Our objective is to find an estimate of frame $\mathbf{f}_k$, $\hat{\mathbf{f}}_k$, given $M1 + M2 + 1$ observed compressed frames $\mathbf{g}_i$, $i = k - M1, \ldots, k + M2$. The spatial resolution enhancement algorithm we propose uses a regularized reconstruction approach which takes into account that the observed frames were compressed and, hence, they may exhibit blocking and mosquito artifacts.

Based on the degradation model in Eq. (3), and in order to reduce the artifacts introduced by the compression process while enhancing the spatial resolution, the high resolution frame is selected as

$$\begin{aligned} \hat{\mathbf{f}}_k \;=\; & \arg\min_{\mathbf{f}} \Big\{ \sum_{i=-M2}^{M1} \| AC_{k,k-i}\mathbf{f}_k - \mathbf{g}_{k-i} \|^2 \\ & + \lambda_1 \| Q_1\mathbf{f}_k \|^2 + \lambda_2 \| Q_2\mathbf{f}_k \|^2 \\ & + \lambda_3 \| \mathbf{f}_k - \mathbf{f}_k^{\mathrm{mc}} \|^2 \Big\}, \end{aligned} \qquad (4)$$

where the first term, $\sum_{i=-M2}^{M1} \| AC_{k,k-i}\mathbf{f}_k - \mathbf{g}_{k-i} \|^2$, takes into account the fidelity to the received data while the other terms impose smoothness constraints on the high resolution image. More concretly, $Q_1$ and $Q_2$ are high pass operators that capture the within-block and between-block smoothness of the estimated frame $\mathbf{f}_k$, respectively, the term $\| \mathbf{f}_k - \mathbf{f}_k^{mc} \|^2$ enforces temporal smoothness between consecutive frames, with $\mathbf{f}_k^{mc}$ being the motion compensated neighboring frames onto $\mathbf{f}_k$, and $\lambda_1$, $\lambda_2$ and $\lambda_3$ are the regularization parameters that control the within-block, between-block and temporal smoothness, respectively. These parameters can be adjusted according to the macroblocks characteristics to obtain optimal results [7].

In addition, it is possible to use the information from the coder to impose other constraints such as:

- the fidelity to the quantization intervals, by projecting $A\mathbf{f}_k$ onto them;

- use of the received motion vectors for the low resolution sequence as an initial estimate of the motion vectors for the high resolution image in a hierarchical (multiscale) motion vector estimation algorithm;

- use of the macroblock type in adjusting the quantization intervals, as well as, the regularization parameter.

The minimization in Eq. (4) can be carried out by an iterative gradient descendent algorithm described by

$$\begin{aligned}\mathbf{f}_k^{l+1} &= \mathbf{f}_k^l - \beta \left[ \sum_{i=-M2}^{M1} C_{k,k-i}^t A^t \left( A C_{k,k-i} \mathbf{f}_k^l - \mathbf{g}_{k-i} \right) \right. \\ &\quad + \lambda_1 Q_1^t Q_1 \mathbf{f}_k^l + \lambda_2 Q_2^t Q_2 \mathbf{f}_k^l \\ &\quad \left. + \lambda_3 (\mathbf{f}_k - \mathbf{f}_k^{\mathrm{mc}}) \right], \qquad (5)\end{aligned}$$

where $\mathbf{f}_k^l$ and $\mathbf{f}_k^{l+1}$ are the enhanced frames in the $l$th and $(l+1)$st iterations, respectively, and $\beta$ is the relaxation parameter that controls the convergence and the rate of convergence of the algorithm.

## 4 Experimental Results

In order to test this algorithm, the color *Mobile* sequence was used. Each frame, of size $720 \times 576$ pixels, was subsampled to obtain a CIF format frame. These frames are considered to be the original high resolution frames $\mathbf{f}^k$. They were further subsampled by two in each direction, according to Eq. (1), to obtain the QCIF low resolution frames. The first 40 frames of the sequence were compressed at 64kbp using the H.263 video coding standard [1].

For this experiment $M1 = M2 = 1$ in Eq. (5) was used. The smoothness parameters were set equal to $\lambda_1 = 0.05$, $\lambda_2 = 0.10$ and $\lambda_3 = 0.025$, and $\beta = 0.5$. In order to choose an initial image, $\mathbf{f}_k^0$, bilinear interpolation was used on the corresponding low resolution frame. The motion vectors needed by the algorithm, for the construction of $C_{k,k-i}$ and $\mathbf{f}_k^{\mathrm{mc}}$, were estimated from the initial images using a block matching algorithm with a block size of $4 \times 4$.

For comparison purposes, bilinear interpolation was used on these low resolution frames, without using motion information. Figure 2 shows the bilinearly interpolated frame 36 and the high resolution image obtained with the proposed algorithm. The PSNR of the Y-band for each image is 20.60dB and 22.30dB, respectively. By comparing these figures a good improvement is observed, both in visual quality and $PSNR$.

## 5 Conclusions

In this paper, a new method for enhancing the resolution of a compressed video color sequence is presented. The algorithm takes into account the special characteristics of the compressed video sequence reducing the artifacts introduced by the compression process.

Although the proposed method treats each band independently, the possibility of combining between band information is currently being investigated. Work is also in progress on the use of the motion vectors and other information provided from the coder.

## Acknowledgements

This paper was partially supported by the "Comisión Nacional de Ciencia y Tecnología" under contract TIC97-0989.

The work of Javier Mateos was partially supported by the Spanish "Ministerio de Educación y Cultura".

## References

[1] International Telecommunications Union, *ITU-T Recom. H.263. Video coding for low rate communication*, 1996.

[2] ISO/IEC JTC1 / SC 29 International Standard 13818-2, *Information technology – Generic coding of moving pictures and associated audio information: Video*, 1996.

[3] R. C. Hardie, K. J. Barnard and E. E. Amstrong, "Joint MAP Registration and High-Resolution Image Estimation using a Sequence of Undersampled Images", IEEE Trans. on Image Processing, vol. 6, pp. 1621–1633, 1997.

[4] A. J. Patti and Y. Altunbasak, "Super-Resolution Image Estimation for Transform Coded Video with Application to MPEG", in *Proceedings of the ICIP99*, n. 27A03.7, 1999.

[5] R. R. Schultz, "Extraction of High-Resolution Video Stills from MPEG Image Sequences" in *Proceedings of the ICIP98*, vol. II, pp. 465–469, 1998.

[6] B. C. Tom, *Reconstruction of a High Resolution Image from Multiple Degraded Mis-Registered Low Resolution Images*, PhD. thesis, Dept. of ECE, Northwestern University, Dec. 1995.

[7] C.-J. Tsai, P. Karunaratne, N. P. Galatsanos, and A. K. Katsaggelos, "A Compressed Video Enhancement Algorithm", in *Proceedings of the ICIP99*, n. 27AP5.1, 1999.

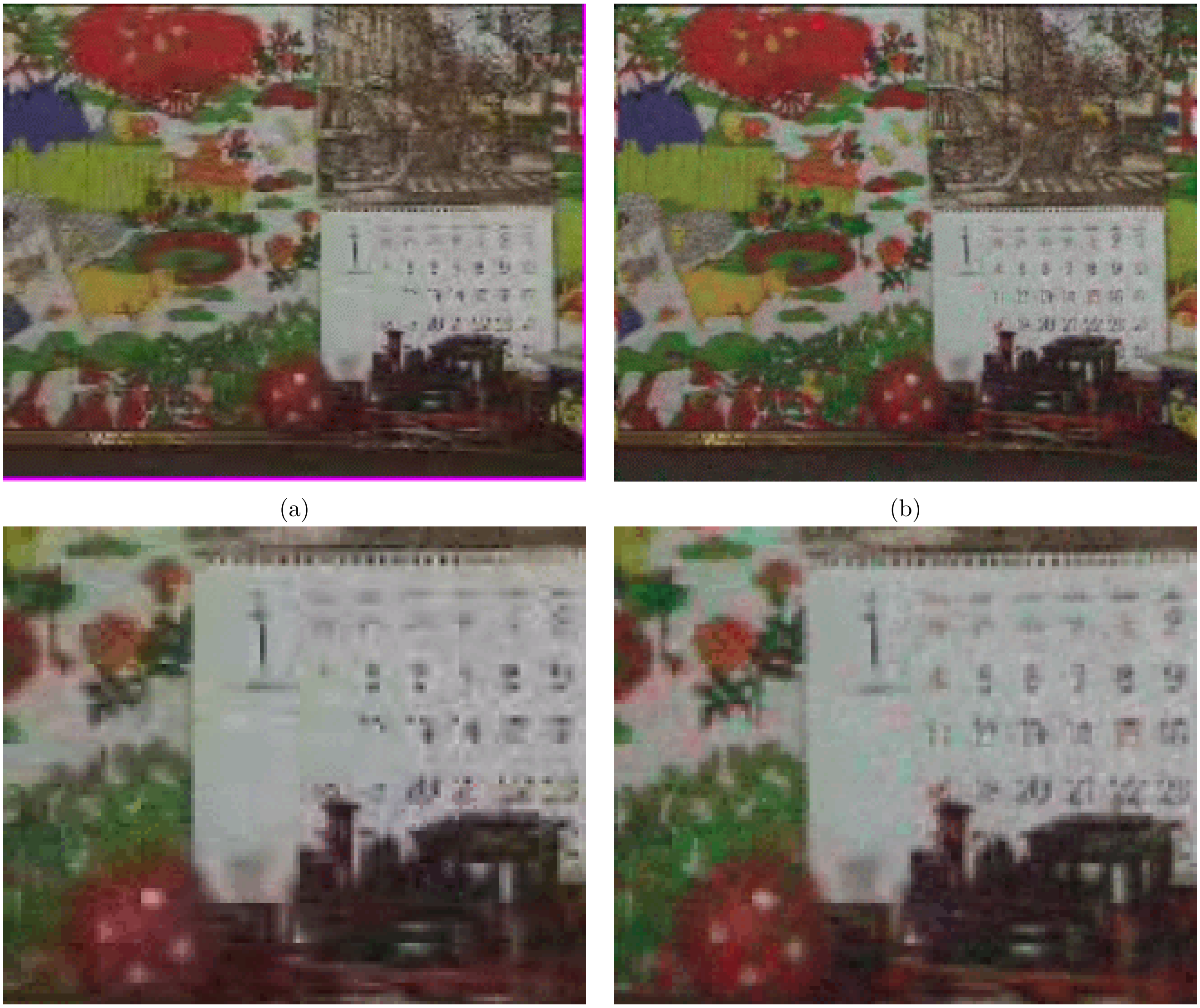


Figure 2: High resolution estimation of frame 36. (a) By bilinear interpolation, (b) With the proposed algorithm. (c) A part of the image (a) enlarged, and (d) a part of the image (b) enlarged.

[8] T. R. Tuinstra and R. C. Hardie, "High-Resolution Image Reconstruction from Digital Video by Explotation of Nonglobal Motion", *Optical Eng.* vol. 38, n. 5, pp. 806–814, 1999.

[9] M. Yuen and H. R. Wu, "Reconstruction Artifacts in Digital Video Compression", in *Proceedings of the Digital Video Compression: Algorithms and Technologies 1995 Conf., SPIE Proc.*, vol. 2419, pp. 455–465, 1995.