High Resolution Images from Compressed Low Resolution Video: Motion Estimation and Observable Pixels

L. D. Alvarez[*], J. Mateos[*], R. Molina[*], and A. K. Katsaggelos[°]

Corresponding Author:

Professor A. K. Katsaggelos
Department of Electrical and Computer Engineering
Northwestern University
Evanston, Illinois 60208-3118, USA
E-mail: aggk@ ece.northwestern.edu

[*] Departamento de Ciencias de la Computación e I. A. Universidad de Granada, 18071 Granada, España. E-mails: ldac@decsai.ugr.es, jmd@decsai.ugr.es, rms@decsai.ugr.es

[°] Department of Electrical and Computer Engineering, Northwestern University, Evanston, Illinois 60208-3118. Email: aggk@ece.northwestern.edu

This work has been partially supported by the "Comisión Nacional de Ciencia y Tecnología" under contract TIC2003-00880.

High Resolution Images from Compressed Low Resolution Video: Motion Estimation and Observable Pixels

L. D. Alvarez, J. Mateos, R. Molina, and A. K. Katsaggelos

Abstract

In this paper, we address the problem of obtaining a high resolution (HR) image from a compressed low resolution (LR) video sequence. Motion information plays a critical role in solving this problem and we determine which pixels in the sequence provide useful information for calculating the high resolution image. The bit-stream of hybrid motion compensated video compression methods includes low resolution motion compensated images; we therefore also study which pixels in these images should be used to increase the quality of the reconstructed image. Once the useful (observable) pixels in the low resolution and motion compensated sequences have been detected, we modify the acquisition model to only account for these observations. The proposed approach is tested on real compressed video sequences and the improved performance is reported.

1. Introduction

Super resolution or resolution enhancement algorithms increase the resolution of an image utilizing multiple misregistered still images or a segment of a video sequence without changing the resolution of the image sensor (see [1], [2], [3], [4] for reviews). This is accomplished by exploiting the underlying sub-pixel shifts or motion among images of frames to provide multiple observations for each frame, and it mitigates the requirements for transporting and storing a high resolution (HR) sequence.

The high resolution image reconstruction problem is further complicated when the available low resolution (LR) video is compressed [5], [6]. Algorithms that attenuate the error due to compression of still images or video sequences belong to the field of post-processing

methods. As an example, filtering a decoded image with a spatially invariant kernel for removing blocking artifacts is proposed in [7]. Unfortunately such an approach also attenuates semantically meaningful edge features. Addressing this flaw is the goal of many subsequent enhancement approaches, such as [8-10], that reduce the amount of smoothing in the vicinity of edges.

Recovery methods can also be utilized for post-processing in super resolution from compressed video. These methods lead to a more rigorous approach to the problem, as prior knowledge about both the original image and the compression system are considered. In [11] this information is derived numerically, and the recovery procedure takes the form of a lookup table operation. In the majority of approaches though, explicit models are utilized. For example, distributions for the transform coefficients and original intensities are defined in [12-14]. These distributions lead to a maximum *a posteriori* (MAP) estimate for the post-processed image. In [15, 16], the available information is utilized to define convex sets the solution should belong to. The theory of projecting onto convex sets provides a feasible solution. As a third approach, constrained least squares solutions incorporate deterministic models and are considered in [17-19].

Post-processing and super resolution methods developed for uncompressed sequences have also been combined to tackle the super resolution from compressed video problem. In [20, 21], the quantization operator is incorporated into a super resolution procedure. The resulting algorithms consider the spatially varying noise process and treat the interpolation, restoration and post-processing problems. All necessary displacement values are assumed known. As a second example, motion vectors within the bit-stream influence the registration problem in [22, 23]. The post-processing problem is ignored though, and estimates for the motion and HR data are computed sequentially.

Building on our previous post-processing and super resolution work [24-29] a novel algorithm that solves the registration, interpolation, restoration, and post-processing problem simultaneously has been proposed in [6].

Accurate motion estimation and determination of the pixels in the LR sequence observable from the HR image under reconstruction are essential tasks in super resolution problems. Pixels in the LR data which cannot be predicted from the HR image under estimation or pixels associated with poor and conflicting motion estimates are of little or no use and therefore should not be taken into account. Not much work has been reported in the literature on this problem with the exception of Schultz and Stevenson [23].

In hybrid motion-compensated video compression a sparse vector field along with the transformed and quantized displaced frame difference (DFD) form the bit-stream. Both pieces of information can prove to be useful when estimating the HR image; in this case again we are faced with the problem of determining which pixels in the LR and/or DFD images are observable from the HR image under estimation.

Once these LR *valid* pixels have been detected the modeling of the observation process needs to be modified appropriately, so as to take into account only such pixels.

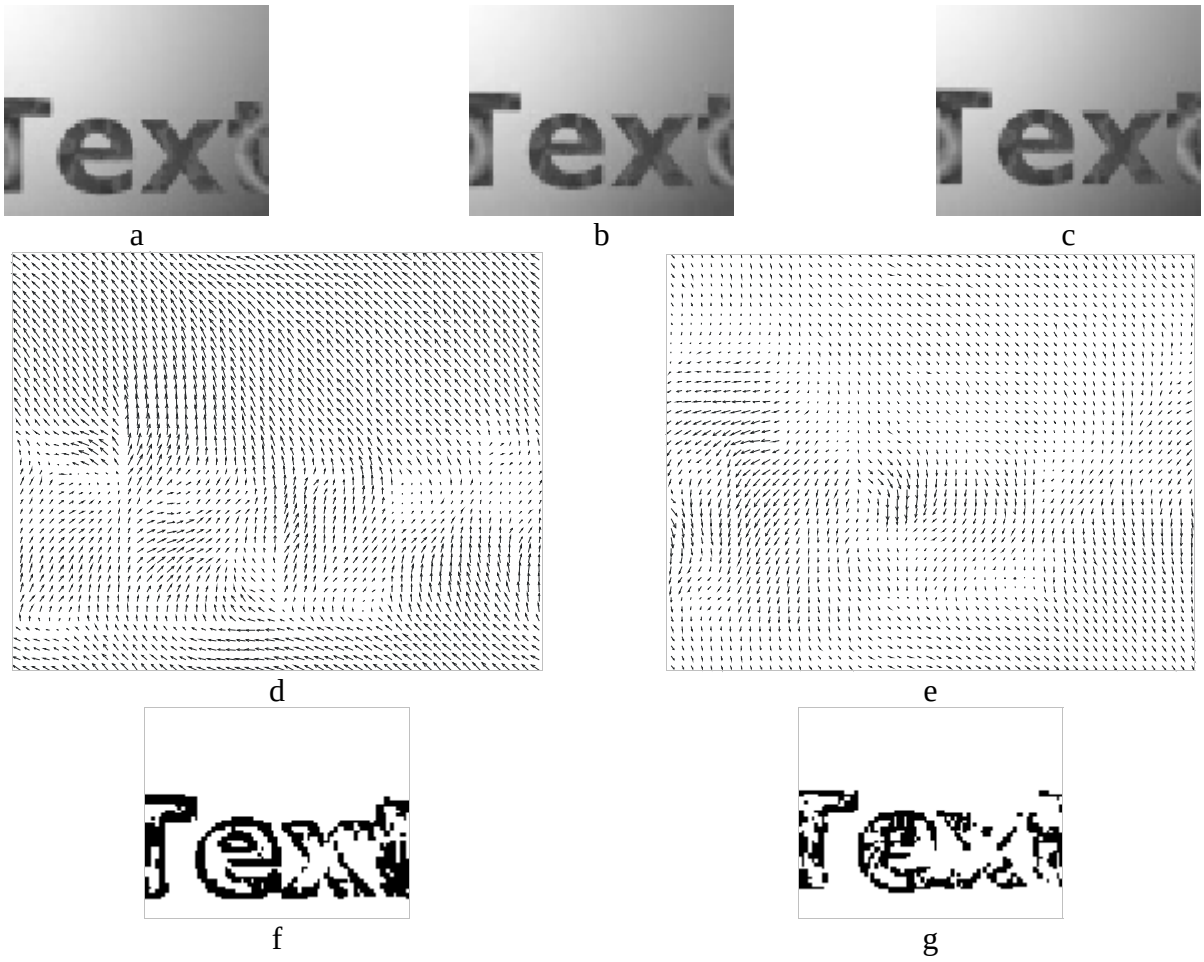


Fig. 1. (a)-(c): Three consecutive compressed LR frames.;(d),(e): Motion fields between the HR frame corresponding to (b) and the HR frames corresponding to (a) and (c), respectively; (f),(g): Black pixels represent pixels in (a) and (c) that are not observable or poorly predicted via motion compensation from the HR image under estimation.

To our knowledge, no work has been reported on how to determine the observable (predictable) pixels in an LR compressed sequence. This is a very important problem since super resolution algorithms assume that every frame in the sequence can be obtained by motion compensating a single HR frame. However, this is not the case in most dynamic sequences where objects move in front of a (moving) background thus uncovering areas that are not observable from the HR frame under estimation. Additionally, poor motion estimates may render data in the LR observed images unusable.

Figure 1 illustrates these problems. Figure 1(a)-(c) displays three consecutive compressed low resolution frames of a sequence where the background is moving three pixels up and to the left in each frame and the text is moving three pixels up and to the right in each frame. The objective is to obtain an HR estimate of the frame in Fig. 1(b). Estimated with a block matching technique motion vectors that relate the HR frame corresponding to Fig. 1(b) to the HR frames corresponding to Figs. 1(a) and 1(c) are plotted in Figs. 1(d) and 1(e),

respectively. Clearly the accuracy of these vectors is determined by the specific motion estimation algorithm used, as well as the accuracy of the HR estimates. Figures 1(f) and 1(g) show, in black, the pixels of Figs. 1(a) and 1(c) that are not observable or well predicted from the HR image via motion compensation. These pixels form regions around moving boundaries. There are also such not observable pixels in smooth areas of the objects (inside the letters), due probably to the aperture effect. The pixels shown in Figs. 1(f) and 1(g) should not be utilized in the resolution enhancement process because most probably they will lead to erroneous results. In this paper we examine how to determine which pixels provide useful information to the estimation of the HR image and modify the acquisition model to take into account only those observations.

The paper is organized as follows. The process to obtain a compressed low resolution video sequence from high resolution images is described in section 2. Prior information on the high resolution image and motion vectors under estimation is described in section 3. In section 4 we describe how to determine the low resolution and DFD pixels which are observable or predictable from the high resolution image under estimation. This process will lead to the modification of the observation model discussed in section 2. In section 5 the Bayesian paradigm to calculate the maximum *a posteriori* (MAP) high resolution image and motion vectors is described. Experimental results are presented in section 6 and section 7 concludes the paper.

2. Obtaining Low Resolution Compressed Observations from High Resolution Images

Let us denote the underlying high resolution (HR) video sequence in vector form by $\mathbf{f} = \{\mathbf{f}_1, \ldots, \mathbf{f}_k, \ldots, \mathbf{f}_L\}$, where the size of each high resolution image $\mathbf{f}_l$, l=1,…,L is (PM×PN)×1, with P>1 being the magnification factor. Each image vector resulting from the lexicographical ordering of a PM×PN image $f_l(a, b)$, $1 \le a \le M$, $1 \le b \le N$, that is, frames within the HR sequence are related through time,

$$f_l(a,b) = f_k\left(a + d_{l,k}^x(a,b), b + d_{l,k}^y(a,b)\right) + n_{l,k}(a,b) \tag{1}$$

where $d_{l,k}^x(a,b)$ and $d_{l,k}^y(a,b)$ denote, respectively, the horizontal and vertical components of the displacement between the l-th and the k-th frames, $d_{l,k}(a,b) = \left(d_{l,k}^x(a,b), d_{l,k}^y(a,b)\right)$, and $n_{l,k}(a,b)$ represents the noise introduced due to motion compensation or registration. The above equation maps a gray level pixel value at location (a,b) at time l to a gray level pixel value at location $\left(a + d_{l,k}^x(a,b), b + d_{l,k}^y(a,b)\right)$ at time k.

We rewrite (1) in matrix-vector notation as

$$\mathbf{f}_l = \mathbf{C}(\mathbf{d}_{l,k})\mathbf{f}_k + \mathbf{n}_{l,k} \tag{2}$$

where $\mathbf{C}(\mathbf{d}_{l,k})$ is the ($PM\times PN$)×($PM\times PN$) matrix that maps frame $\mathbf{f}_l$ to frame $\mathbf{f}_k$, and $\mathbf{n}_{l,k}$ the registration noise.

The HR sequence, through filtering and downsampling, produces an (unobserved) uncompressed low resolution (LR) sequence, denoted by $\mathbf{g} = \{\mathbf{g}_1, \ldots, \mathbf{g}_L\}$, of size $M\times N\times 1$. Each LR image $\mathbf{g}_l$, l=1, …, L is related to the corresponding HR image $\mathbf{f}_l$ by

$$\mathbf{g}_l = \mathbf{AHf}_l + v_l, \quad l = 1,2,3,\ldots \tag{3}$$

where $\mathbf{H}$ of size ($PM\times PN$)×($PM\times PN$) describes the filtering of the HR image, $\mathbf{A}$ of size MN×($PM\times PN$) denotes the downsampling matrix, and $\boldsymbol{v}_l$ the noise process. Matrices $\mathbf{A}$ and $\mathbf{H}$ are both assumed to be known. We assume here for simplicity that all blurring matrices $\mathbf{H}$ are the same, although they can be time dependent.

Combining (3) and (2) we obtain the equation describing the acquisition system, that is,

$$\mathbf{g}_l = \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k + \mathbf{e}_{l,k} \tag{4}$$

where $\mathbf{e}_{l,k}$ is the acquisition noise representing the combination of the registration and filtering and downsampling noise processes. The problem at hand is to estimate the HR image frame $\mathbf{f}_k$, given $\mathbf{A}$, $\mathbf{H}$ and a set of LR images $\mathbf{g}_l$. In providing a solution to this estimation problem the motion vectors $\mathbf{d}_{l,k}$ are required and therefore need to also be estimated.

Let us now briefly describe the compression process. The LR frames are compressed with a hybrid motion compensated compression system resulting in $\mathbf{y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_L\}$, also of size $M\times N$. During compression, frames are divided into blocks that are encoded with one or two available methods, intracoding or intercoding. For the first one, a linear transform such as the Discrete Cosine Transform (DCT) is applied to the block. The operator decorrelates the intensity data, and the resulting transform coefficients are independently quantized and transmitted to the decoder. For the second method, predictions for the blocks are first generated by motion compensating previously transmitted image frames. The compensation is controlled by motion vectors $v_{l,m}(i,j)$ that predict pixel $y_l(i,j)$ from the previously coded frame $\mathbf{y}_m$. These motion vectors that predicts $\mathbf{y}_l$ from $\mathbf{y}_m$ are represented by the (2×M×N)×1 vector $\mathbf{v}_{l,m}$ that is formed by stacking the transmitted horizontal and vertical offsets. The transformed by a linear transformation (such as DCT) and quantized prediction error along with the motion vectors form the bit-stream generated by the encoder, which in turn is used by the decoder to provide an estimate of the original video.

Using all this information, the relationship between the acquired low resolution frame $\mathbf{g}_l$ and its compressed observation $\mathbf{y}_l$ becomes

$$\mathbf{y}_l = T^{-1}\{Q[T(\mathbf{g}_l - \mathbf{m}_l)]\} + \mathbf{m}_l \quad l = 1, \ldots, L \tag{5}$$

where $Q[\ ]$ represents the quantization procedure, T and T^{-1} are the forward and inverse-transform operations, respectively, and $\mathbf{m}_l$ is the motion compensated prediction of $\mathbf{g}_l$

formed by motion compensating the appropriate previously decoded frame/frames depending on whether the current frame at l is an *I*, *P* or *B* frame (see [30]). Note that, to be precise, we should make clear that $\mathbf{m}_l$ depends on $\mathbf{v}_l$ and only a subset of $\mathbf{y}_1,\ldots,\mathbf{y}_L$; however we will keep the above notation for simplicity and generality.

To model the compression error, we assume that the quantization noise is dominant and approach the quantity $T^{-1}Q\left[T\left(\mathbf{g}_l-\mathbf{m}_l\right)\right]$ in (5) by $\mathbf{g}_l-\mathbf{m}_l+\boldsymbol{\varepsilon}_{Q,l}$, where $\boldsymbol{\varepsilon}_{Q,l}\sim N\left(0,\mathbf{K}_{Q,l}\right)$ and $\mathbf{K}_{Q,l}$ is the covariance matrix of the quantization noise in the spatial domain at frame l. Using (4) , we then have in (5)

$$\mathbf{y}_l=\mathbf{AHC}\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k+\boldsymbol{\varepsilon}_{Q,l}\ . \tag{6}$$

Estimates for $\mathbf{K}_{Q,l}$ can be derived from the compressed bit-stream. For example, the noise variance σ_l^2 for transform index l is defined as $\sigma_l^2=q_l^2/12$ when the quantization interval is uniform with step-size q_l. Since errors in the spatial domain are related to errors in the transform domain by the inverse-transform operation, the needed covariance matrix is therefore expressed as

$$\mathbf{K}_{Q,l}=T^{-1}\mathbf{K}_{Transform,l}\left(T^{-1}\right)^T \tag{7}$$

where $\mathbf{K}_{Transform,l}$ is the covariance matrix describing the noise in the transform domain. Of course, other definitions could also be constructed. When additional noise is present in the observation though, the compressed bit-stream does not completely specify $\mathbf{K}_{Q,l}$. Instead, the information about these other corrupting processes must also be included.

From the above discussion we finally rewrite (6) as (see [5], [6], [31]),

$$P\left(\mathbf{y}_l\mid\mathbf{f}_k,\mathbf{d}_{l,k}\right)\propto\exp\left[-\frac{1}{2}\left(\mathbf{y}_l-\mathbf{AHC}\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k\right)^T\mathbf{K}_{Q,l}^{-1}\left(\mathbf{y}_l-\mathbf{AHC}\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k\right)\right] \tag{8}$$

where $\mathbf{K}_{Q,l}$ is the covariance matrix that describes the noise as defined in (7).

As already mentioned $\mathbf{m}_l$ is the motion compensated prediction of $\mathbf{g}_l$ formed by motion compensating the appropriate previously decoded frame/frames. We then expect that given $\mathbf{d}_{l,k}$ and $\mathbf{f}_k$

$$\mathbf{m}_l=\mathbf{AHC}\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k+\boldsymbol{\eta}_{MV,l}\ , \tag{9}$$

where

$$\boldsymbol{\eta}_{MV,l}\sim N\left(0,\mathbf{K}_{MV,l}\right) \tag{10}$$

and $\mathbf{K}_{MV,l}$ is the covariance matrix describing the error between the uncompressed low resolution image and its motion compensated estimate.

Like the previous covariance matrix, an estimate for $\mathbf{K}_{MV,l}$ can be extracted from the compressed bit-stream. This covariance matrix describes the transmitted displaced frame difference, and the variance at each transform index l could be defined as $\sigma_l^2=c_l^2+q_l^2/12$,

where c_l is the decoded transform coefficient and q_l is the width of the quantization interval. The relationship in (7) then maps the variance information to the spatial domain. So, from the above discussion we can write (see [6], [31])

$$P(\mathbf{m}_l \mid \mathbf{f}_k, \mathbf{d}_{l,k}) \propto \exp\left[-\frac{1}{2}(\mathbf{m}_l - \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k)^T \mathbf{K}_{MV,l}^{-1}(\mathbf{m}_l - \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k)\right] \tag{11}$$

where $\mathbf{K}_{MV,l}$ is the covariance matrix for the prediction error between the original LR frame $(\mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k)$ and its motion compensation estimate $\mathbf{m}_l$.

We can now write the joint acquisition model of the low resolution compressed and motion compensated images given the high resolution image and motion vectors as

$$P(\mathbf{y}, \mathbf{m} \mid \mathbf{f}_k, \mathbf{d}) = \prod_l P(\mathbf{y}_l \mid \mathbf{f}_k, \mathbf{d}_{l,k}) P(\mathbf{m}_l \mid \mathbf{f}_k, \mathbf{d}_{l,k}), \qquad l = 1, \ldots, L \tag{12}$$

where $\mathbf{m} = \{\mathbf{m}_1, \ldots, \mathbf{m}_L\}$ and $\mathbf{d} = \{\mathbf{d}_{1,k}, \mathbf{d}_{2,k} \ldots, \mathbf{d}_{L,k}\}$.

In examining equations (8) and (11) it should be clear that there are pixels in the LR frame $\mathbf{y}_l$ that are unobservable in frame $\mathbf{f}_k$. These pixels represent, for instance, uncovered background or pixels entering the scene in $\mathbf{f}_k$. Analogously, since $\mathbf{m}_l$ is a prediction of $\mathbf{y}_l$ (which may not be the same as the one obtained from $\mathbf{f}_k$) there are pixels in the LR frame $\mathbf{m}_l$ that are unobservable in frame $\mathbf{f}_k$. As a consequence, either (8) or (11) or both are not applicable acquisition models for some low resolution pixels. Furthermore, even being observable, due to the inaccuracies of the motion estimator, there are additional pixels in $\mathbf{y}_l$ and/or $\mathbf{m}_l$ which are not predictable from $\mathbf{f}_k$ and $\mathbf{d}_{l,k}$, to which models (8) and/or (11) are also not applicable. We examine in section 4 how such low resolution pixels can be detected and removed from the observation model.

3. Regularization in High Resolution Estimation from Compressed Low Resolution Video

The distribution $P(\mathbf{f}_k)$ we use for $\mathbf{f}_k$, reflects the facts that we expect the images to be smooth within homogeneous regions and free of blocking artifacts; that is

$$P(\mathbf{f}_k) \propto \exp\left[-\left(\frac{\lambda_1}{2}\|\mathbf{Q}_1\mathbf{f}_k\|^2 + \frac{\lambda_2}{2}\|\mathbf{Q}_2\mathbf{AHf}_k\|^2\right)\right] \tag{13}$$

where $\mathbf{Q}_1$ represents a linear high-pass operator that penalizes non-smooth, $\mathbf{Q}_2$ represents a linear high-pass operator that penalizes estimates with block boundaries, and λ_1 and λ_2 control the relative contribution of the two terms. For a complete study of all the prior models used for this problem see Segall *et al.* [5].

Constraints are also imposed on the high resolution motion vectors. Assuming that the displacements are independent between frames, an assumption that should be reviewed in the future, we can write

$$P(\mathbf{d}) = \prod_{l} P(\mathbf{d}_{l,k}) \,. \tag{14}$$

In addition, we can enforce $\mathbf{d}_{l,k}$ to be smooth within each frame, that is,

$$P(\mathbf{d}_{l,k}) \propto \exp\left[-\frac{\lambda_3}{2} \left\| \mathbf{Q}_3 \mathbf{d}_{l,k} \right\|^2 \right] \tag{15}$$

where $\mathbf{Q}_3$ represents a linear high-pass operator that, once again, penalizes the displacement estimates that are not smooth and λ_3 controls the variance of the distribution (see again Segall *et al.* [5] for details).

4. Observable and Predictable Pixels from the High Resolution Image

In order to determine which low resolution pixels will be used to reconstruct the high resolution image we make use of a fixed estimate, $\bar{\mathbf{f}}_l$, of the high resolution image $\mathbf{f}_l$, l=1,…,L. Following [32] and [33] we calculate $\bar{\mathbf{f}}_l$ as

$$\bar{\mathbf{f}}_l = \arg\min_{\mathbf{x}} \left(\alpha \left\| \mathbf{Q}\mathbf{x} \right\|^2 + \beta \left\| \mathbf{y}_l - \mathbf{A}\mathbf{H}\mathbf{x} \right\|^2 \right) \tag{16}$$

where $\mathbf{Q}$ denotes the Laplacian operator and α and β are calculated following the approach in [32], [33]. We note here that since $\mathbf{y}_l$ is a blurred image, its bilinearly interpolated version could not serve as $\bar{\mathbf{f}}_l$.

Let us now assume that $\tilde{\mathbf{d}}_{l,k}$ and $\tilde{\mathbf{f}}_k$ are the current estimates of the high resolution motion vectors and image. The absolute value of the DFD is then given by

$$DFD(\bar{\mathbf{f}}_l, \tilde{\mathbf{d}}_{l,k}, \tilde{\mathbf{f}}_k)(m,n) = \left| \bar{\mathbf{f}}_l(m,n) - \left(C(\tilde{\mathbf{d}}_{l,k}) \tilde{\mathbf{f}}_k \right)(m,n) \right| \tag{17}$$

will serve as a criterion for determining whether $\bar{\mathbf{f}}_l(m,n)$ is also observable or predictable by the current estimates $\tilde{\mathbf{f}}_k$ and $\tilde{\mathbf{d}}_{l,k}$ (and eventually by the original $\mathbf{f}_k$ and $\mathbf{d}_{l,k}$ at convergence). A large value of the DFD detects a pixel in the HR image $\bar{\mathbf{f}}_l$ that is either not present in the HR image $\tilde{\mathbf{f}}_k$ or is represented by a poor motion estimate. If $\bar{\mathbf{f}}_l(m,n)$ is detected as an unobservable or unpredictable pixel in $\tilde{\mathbf{f}}_k$ then all the observations in $\mathbf{y}_l$ that depend on it using the model defined in (3) should not be taken into account, in estimating the HR image $\mathbf{f}_k$. From now on we will use the term observable to denote both, observable and predictable pixels, but it is important to take into account that they correspond to two different scenarios.

To remove these LR pixels, we first note that $\mathbf{K}_{Q,l}$ in (8) is a block matrix, consisting of 64×64 sub-matrices, which they either represent the covariance of the 8×8 image blocks used by the Discrete Cosine Transform or they are zero representing the zero correlation of motion compensating errors in different blocks.

Let $\mathbf{K}_{Q,l}^{u,v}$ be the (u,v) sub-matrix of the covariance matrix $\mathbf{K}_{Q,l}$ corresponding to the image block where the unobservable low resolution pixel (p,q) in $\mathbf{y}_l$ is located. In order not to take into account this observation when estimating the high resolution image, we remove the row r and column r of $\mathbf{K}_{Q,l}^{u,v}$, where $r=(p-1)\cdot 8+q$. We apply the same procedure to all covariance sub-matrices corresponding to blocks with unobservable low resolution pixels. We then obtain new covariance sub-matrices $\left(\mathbf{K}_{Q,l}^{u,v}\right)^o$ and consequently a new covariance matrix, $\mathbf{K}_{Q,l}^o$, for the compensating errors of all the low resolution pixels in $\mathbf{y}_l$ observable or predictable from $\mathbf{f}_k$.

We can now write the conditional distribution of the LR observation given the HR image we want to estimate as (updated form of (8))

$$P\left(\mathbf{y}_l^o \mid \mathbf{f}_k, \mathbf{d}_{l,k}\right) \propto \exp\left[-\frac{1}{2}\left(\mathbf{y}_l^o - \mathbf{A}_l^o\mathbf{H}\mathbf{C}\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k\right)^T\left(\mathbf{K}_{Q,l}^o\right)^{-1}\left(\mathbf{y}_l^o - \mathbf{A}_l^o\mathbf{H}\mathbf{C}\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k\right)\right] \tag{18}$$

where $\mathbf{y}_l^o$ is the LR image consisting of all the pixels in $\mathbf{y}_l$ observable from $\mathbf{f}_k$, and $\mathbf{K}_{Q,l}^o$ and $\mathbf{A}_l^o$ are, respectively, the covariance matrix and the downsampling matrix that are formed taking into account only observable pixels.

The image $\mathbf{y}_l^o$ is determined utilizing a threshold T for the DFD in (17), and represent the *observable pixel map for T*.

The same ideas are applied to the LR motion compensated observations in (11). We first upsample $\mathbf{m}_l$ by replacing $\mathbf{y}_l$ by $\mathbf{m}_l$ in (16) to obtain $\overline{\mathbf{m}}_l$. We then use

$$DFD\left(\overline{\mathbf{m}}_l, \tilde{\mathbf{d}}_{l,k}, \tilde{\mathbf{f}}_k\right)(m,n) = \left|\overline{\mathbf{m}}_l(m,n) - \left(\mathbf{C}\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k\right)(m,n)\right| \tag{19}$$

to detect unobservable pixels and following the same steps as with the LR compressed observation define

$$P\left(\mathbf{m}_l^o \mid \mathbf{f}_k, \mathbf{d}_{l,k}\right) \propto \exp\left[-\frac{1}{2}\left(\mathbf{m}_l^o - \mathbf{A}_l^o\mathbf{H}\mathbf{C}\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k\right)^T\left(\mathbf{K}_{MV,l}^o\right)^{-1}\left(\mathbf{m}_l^o - \mathbf{A}_l^o\mathbf{H}\mathbf{C}\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k\right)\right] \tag{20}$$

where $\mathbf{m}_l^o$ is the LR image consisting of all the pixels in $\mathbf{m}_l$ observable from $\mathbf{f}_k$, $\mathbf{K}_{MV,l}^o$ is the covariance matrix from $\mathbf{K}_{MV,l}$ in a similar fashion of obtaining $\mathbf{K}_{Q,l}^o$ from $\mathbf{K}_{Q,l}$, and $\mathbf{A}_l^o$ is the downsampling matrix formed as before (the set of pixels resulting from (17) and (19) need not be the same). As before the image $\mathbf{m}_l^o$ represents the *observable pixel map for T*.

Note that in this section we have defined two observable pixel maps. One corresponding to the compressed LR observations and the other to the LR motion compensated observations.

5. Estimating High Resolution Images

Having described in the previous sections the HR image prior and the acquisition model, we turn now our attention to computing the HR frame and motion vectors. Our goal is then to find $\hat{\mathbf{f}}_k$ and $\hat{\mathbf{d}}$ that satisfy

$$\hat{\mathbf{f}}_k, \hat{\mathbf{d}} = \arg\max_{\mathbf{f}_k, \mathbf{d}} \left[P(\mathbf{f}_k, \mathbf{d}) P\left(\mathbf{y}^o, \mathbf{m}^o \mid \mathbf{f}_k, \mathbf{d}\right) \right] \tag{21}$$

where the distribution of HR intensities is given in section 3 and the acquisition model is described in section 2 and modified in section 4 to consider only the LR pixels in $\mathbf{y}$ and $\mathbf{m}$ observable from $\mathbf{f}_k$ and $\mathbf{d}$.

Following [5], given $\mathbf{y}^o$ and $\mathbf{m}^o$ equation (21) is solved with the use of the cyclic coordinate descent procedure [34]. An estimate for the displacements is first found by assuming that the HR image is known, so that

$$\hat{\mathbf{d}}^{q+1} = \arg\max_{\mathbf{d}} \left[P(\mathbf{d}) P\left(\mathbf{y}^o, \mathbf{m}^o \mid \hat{\mathbf{f}}_k^q, \mathbf{d}\right) \right] \tag{22}$$

where q is the iteration index for the joint estimate.

Treating the high resolution image as a known parameter, the estimate for the motion field in (22) can be found by the method of successive approximations

$$\begin{aligned} \overline{\mathbf{d}}_{l,k}^{(i+1)} = \overline{\mathbf{d}}_{l,k}^{(i)} - \alpha_d^{l,k} \Bigg\{ & \left. \frac{\partial \mathbf{C}\left(\mathbf{d}_{l,k}\right) \hat{\mathbf{f}}_k^q}{\partial \mathbf{d}_{l,k}} \right|_{\mathbf{d}_{l,k} = \overline{\mathbf{d}}_{l,k}^{(i)}} \mathbf{H}^T \left(\mathbf{A}_l^o\right)^T \Big[\left(\mathbf{K}_{Q,l}^o\right)^{-1} \left(\mathbf{y}_l^o - \mathbf{A}_l^o \mathbf{H} \mathbf{C}\left(\overline{\mathbf{d}}_{l,k}^{(i)}\right) \hat{\mathbf{f}}_k^q \right) \\ & + \left(\mathbf{K}_{MV,l}^o\right)^{-1} \left(\mathbf{m}_l^o - \mathbf{A}_l^o \mathbf{H} \mathbf{C}\left(\overline{\mathbf{d}}_{l,k}^{(i)}\right) \hat{\mathbf{f}}_k^q \right) \Big] + \lambda_3 \mathbf{Q}_3^T \mathbf{Q}_3 \overline{\mathbf{d}}_{l,k}^{(i)} \Bigg\} \end{aligned} \tag{23}$$

where, given $\hat{\mathbf{f}}_k^q$, $\overline{\mathbf{d}}_{l,k}^{(i+1)}$ and $\overline{\mathbf{d}}_{l,k}^{(i)}$ are respectively the (*i+1*)-*th* and *i-th* estimates of the displacement between frame *k* and *l*, $\left(\mathbf{A}_l^o\right)^T$ defines the up-sampling operation, and $\alpha_d^{l,k}$ controls the convergence and rate of convergence of the algorithm. At convergence, *i=end*, $\hat{\mathbf{d}}_{l,k}^{q+1}$ is set to $\overline{\mathbf{d}}_{l,k}^{(end)}$.

The intensity information is then estimated by assuming that the displacement estimates are exact, that is

$$\hat{\mathbf{f}}_k^{q+1} = \arg\max_{\mathbf{f}_k} \left[P(\mathbf{f}_k) P\left(\mathbf{y}^o, \mathbf{m}^o \mid \mathbf{f}_k, \hat{\mathbf{d}}^{q+1}\right) \right]. \tag{24}$$

Once the estimate for the motion field is found, then the high resolution image is computed, the solution of (24) is found by using

$$\bar{\mathbf{f}}_k^{(i+1)} = \bar{\mathbf{f}}_k^{(i)} - \alpha_f \left\{ \sum_{l=1}^{L} \left(\mathbf{C}\left(\hat{\mathbf{d}}_{l,k}^{q+1}\right)\right)^T \mathbf{H}^T \left(\mathbf{A}_l^o\right)^T \left[\left(\mathbf{K}_{Q,l}^o\right)^{-1} \left(\mathbf{y}_l^o - \mathbf{A}_l^o \mathbf{H} \mathbf{C}\left(\hat{\mathbf{d}}_{l,k}^{q+1}\right) \bar{\mathbf{f}}_k^{(i)} \right) \right.\right.$$
$$\left.\left. + \left(\mathbf{K}_{MV,l}^o\right)^{-1} \left(\mathbf{m}_l^o - \mathbf{A}_l^o \mathbf{H} \mathbf{C}\left(\hat{\mathbf{d}}_{l,k}^{q+1}\right) \bar{\mathbf{f}}_k^{(i)} \right) \right] + \lambda_1 \mathbf{Q}_1^T \mathbf{Q}_1 \bar{\mathbf{f}}_k^{(i)} + \lambda_2 \mathbf{H}^T \left(\mathbf{A}_l^o\right)^T \mathbf{Q}_2^T \mathbf{Q}_2 \mathbf{A}_l^o \mathbf{H} \bar{\mathbf{f}}_k^{(i)} \right\} \quad (25)$$

where, given $\hat{\mathbf{d}}_{l,k}^{q+1}$, $\bar{\mathbf{f}}_k^{(i+1)}$ and $\bar{\mathbf{f}}_k^{(i)}$ are respectively the (*i+1*)-*th* and *i-th* estimates of the HR image $\mathbf{f}_k$, α_f is a relaxation parameter that determines the convergence and rate of convergence of the algorithm, and $\left(\mathbf{C}\left(\hat{\mathbf{d}}_{l,k}^{q+1}\right)\right)^T$ compensates an image backwards along the motion vectors. At convergence, *i=end,* $\hat{\mathbf{f}}_k^{q+1}$ is set to $\bar{\mathbf{f}}_k^{(end)}$.

The displacement information is re-estimated with the result from equation (23) and the process iterates until convergence.

In our experiments we have found that the initial estimate of the HR image is quite important in determining the quality of the final solution. In the experimental section we will compare the use of the bilinear interpolation of $\mathbf{y}_k$ and the solution of (16) for $\mathbf{y}_k$.

The process of finding the HR image and motion vectors is summarized as follows:

Let $\hat{\mathbf{f}}_k^0$ *and* $\hat{\mathbf{d}}^0$ *be the initial estimates of the HR image and motion vectors. Find* $\mathbf{y}^o$ *and* $\mathbf{m}^o$ *following the procedure described in section 4 for the estimates* $\tilde{\mathbf{f}}_k = \hat{\mathbf{f}}_k^0$ *and* $\tilde{\mathbf{d}} = \hat{\mathbf{d}}^0$,

For q=0,…, *and until a convergence criterion is met*

1.Find $\hat{\mathbf{d}}^{q+1}$ *by solving (23).*

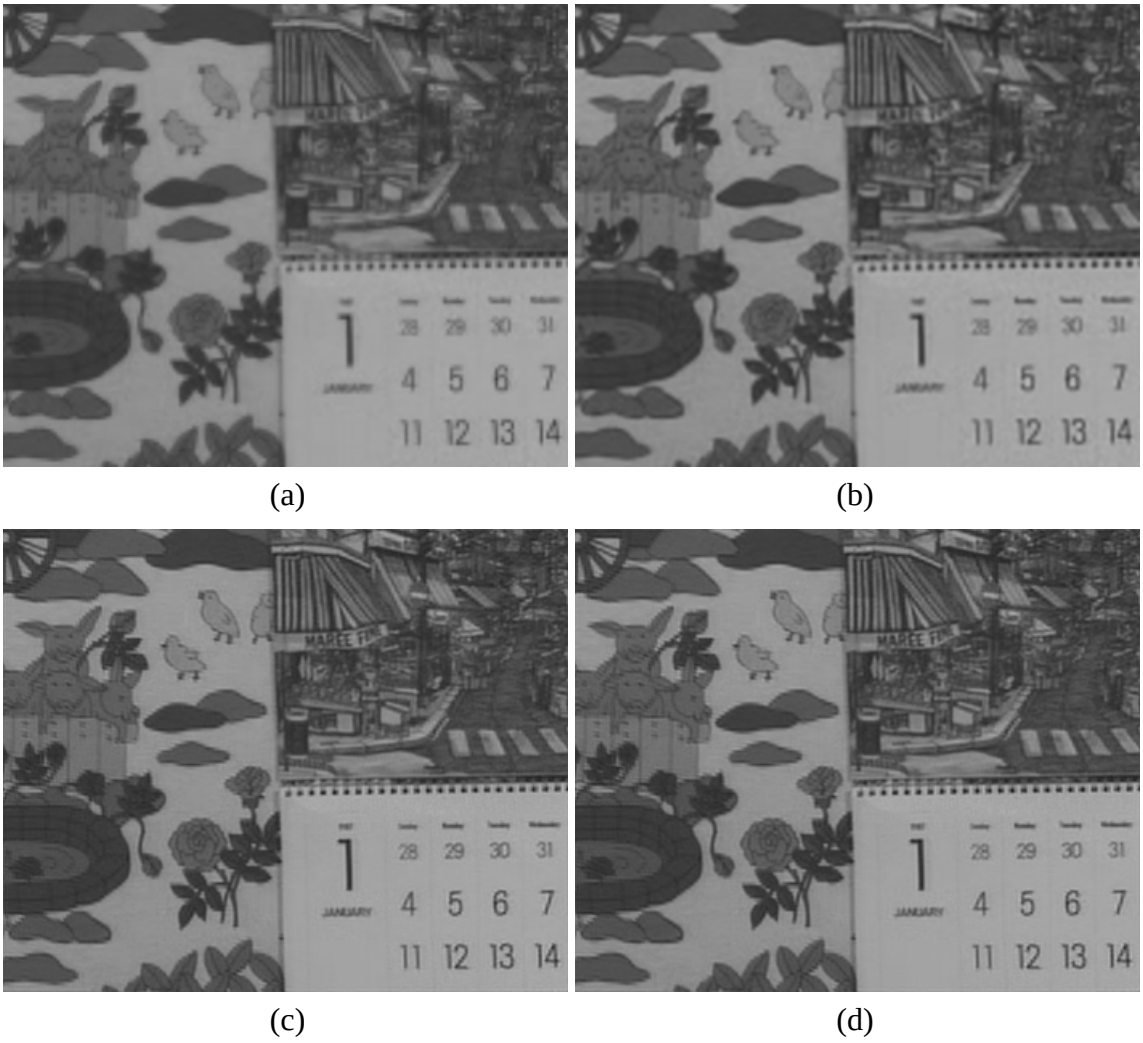
2.Find $\hat{\mathbf{f}}_k^{q+1}$ *by solving (25).*

3.Following the procedure described in section 4, use $\tilde{\mathbf{f}}_k = \hat{\mathbf{f}}_k^{q+1}$ *and* $\tilde{\mathbf{d}} = \hat{\mathbf{d}}^{q+1}$ *to find the new observable pixels* $\mathbf{y}^o$ *and* $\mathbf{m}^o$.

6. Experimental Results

In order to illustrate the performance of the proposed method to obtain an HR image from a compressed LR sequence, a sequence of experiments has been performed on the 352x288 central part of the "mobile" sequence. Figures 2(a)-(c) show three original HR frames of this sequence. Each frame in the sequence is blurred with a 2x2 mean filter and downsampled by a factor of two, thus obtaining a sequence of images of size 176x144. This LR sequence was then compressed using the MPEG-4 coder. The sequence was encoded at 30 frames per second using an IPPP... strategy so that all frames, but the first one which is intra-coded, are inter-coded as P-frames. The VM5+ rate control was used to achieve a bit rate of 1024 kbps. Once decoded, the reconstructed frames and the information from the decoder were used as inputs to the resolution enhancement method. The resulting decoded frames corresponding to the frames shown in figures 2(a)-(c) are depicted in figures 2(d)-(f).

Fig. 2. (a)-(c) Three consecutive frames (numbers 8, 9 and 10) of the original “mobile” sequence. (d)-(f) Their corresponding compressed low resolution frames. The frames have been scaled by zero-order hold to the high resolution image size.



(a) (b)

(c) (d)

Fig. 3. Both, compressed low resolution and motion compensated images are considered in the observation model with $\mathbf{y}^o = \mathbf{y}$ and $\mathbf{m}^o = \mathbf{m}$. (a) Initial high resolution image obtained by bilinear interpolation of image $\mathbf{y}_k$. (b) Initial high resolution image obtained by the solution of (16) for $\mathbf{y}_k$ (Method II). (c) Reconstruction using bilinear interpolation as initial image estimation. (d) High resolution estimate using the image provided by method II as initial high resolution image estimate.

In order to reconstruct frame 9, two previous and two frames after frame 9 were considered. Following [6], for all the tests, the covariance matrices $\mathbf{K}_{Q,l}$ and $\mathbf{K}_{MV,l}$ were calculated using the parameters from the bit-stream. Matrices $\mathbf{Q}_1$ and $\mathbf{Q}_3$ correspond to the Laplacian operator while matrix $\mathbf{Q}_2$ performs the difference between pixels on the 8x8 block boundaries. Parameters for the prior models were empirically chosen to be $\lambda_1 = 0.01$, $\lambda_2 = 0.023$ and $\lambda_3 = 10^5$.

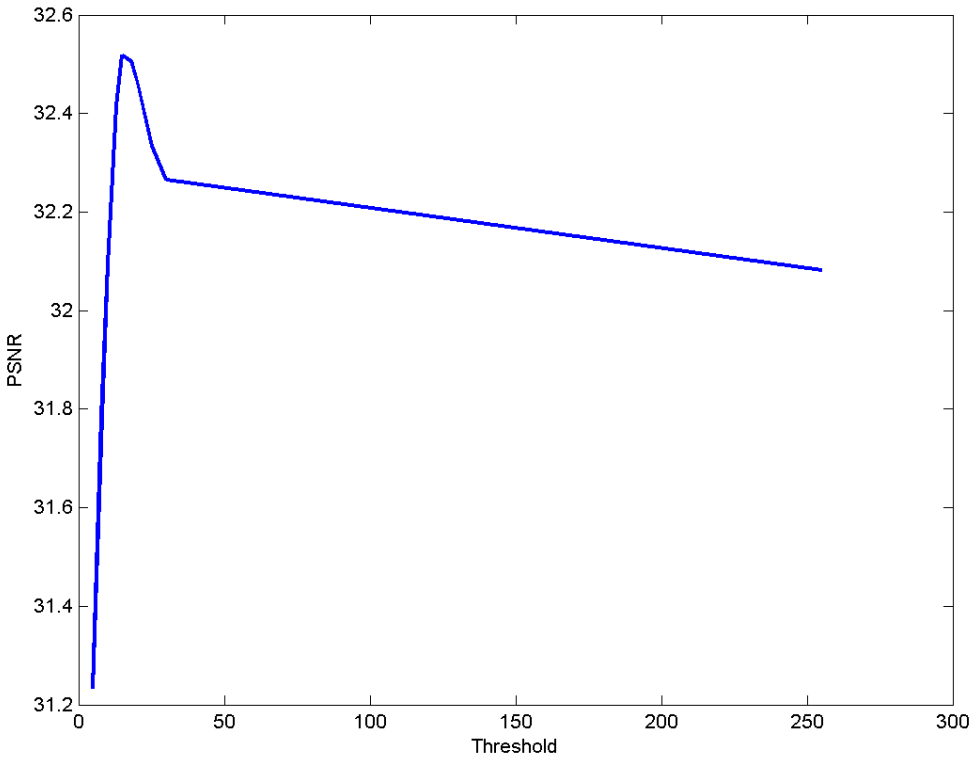


Fig. 4. PSNR evolution with threshold values for the observable pixel map for $\{\mathbf{y}_l\}$, when compressed low resolution motion compensated images $\{\mathbf{m}_l\}$ are not used in the observation model.

The initial motion estimates were obtained using the motion estimation method described in section 5 with four resolution levels. A gradient descent algorithm was used to solve (22) in an iterative fashion until two consecutive motion estimates, $\hat{\mathbf{d}}_{l,k}^{new}$ and $\hat{\mathbf{d}}_{l,k}^{old}$, satisfy $\left\|\hat{\mathbf{d}}_{l,k}^{new}-\hat{\mathbf{d}}_{l,k}^{old}\right\|^2 / \left\|\hat{\mathbf{d}}_{l,k}^{old}\right\|^2 < 10^{-9}$, obtaining then the new high resolution displacements $\hat{\mathbf{d}}_{l,k}^{q+1}=\hat{\mathbf{d}}_{l,k}^{new}$. Using the new motion estimate $\hat{\mathbf{d}}^{q+1}$, the gradient descent method used to solve (24) iterates until $\left\|\hat{\mathbf{f}}_{k}^{new}-\hat{\mathbf{f}}_{k}^{old}\right\|^2 / \left\|\hat{\mathbf{f}}_{k}^{old}\right\|^2 < 10^{-6}$. The new estimate of the high resolution image, $\hat{\mathbf{f}}_{k}^{q+1}$, is set to $\hat{\mathbf{f}}_{k}^{new}$. The whole coordinate descent iterative procedure stops when $\left\|\mathbf{f}_{k}^{q+1}-\mathbf{f}_{k}^{q}\right\|^2 / \left\|\mathbf{f}_{k}^{q}\right\|^2 < 10^{-7}$ or a maximum number of 25 iterations is reached.

In order to test the influence of the initial image estimate on the estimation process we chose two different initial images. First, the decoded low resolution image was bilinearly interpolated (Method I) to the size of the high resolution image (figure 3(a)), resulting in a PSNR of 29.19dB. As described in section 5, we also used as initial high resolution image estimate the solution of (16) for $\mathbf{y}_k$ (Method II) resulting in a PSNR of 30.63dB (figure 3(b)).

(a) (b)

Fig. 5. (a) High resolution estimate with a threshold of 255 for the DFD defined in (17). (b) High resolution estimate with a threshold of 15 in (17). The compressed low resolution motion compensated images $\{\mathbf{m}_l\}$ are not taken into account in the observation model.

The algorithm summarized at the end of section 5 was first executed using both compressed low resolution and motion compensated images and $\mathbf{y}^o = \mathbf{y}$ and $\mathbf{m}^o = \mathbf{m}$. Using as initial HR images the ones provided by Methods I and II above. The resulting HR reconstructed images are displayed in figures 3(c) and 3(d), respectively. Their corresponding PSNRs are 31.95dB (after 25 iterations) and 32.20dB (after 17 iterations), respectively. These figures support the intuitive idea that the initial point should not drastically condition the result of the estimation algorithm. In fact, both resulting images are visually indistinguishable. The solution obtained by starting with the image provided by Method II seems to provide slightly better results with a lower number of iterations since it was already closer to the real solution. The remaining results will be presented using as initial high resolution estimate the one provided by Method II.

We also tested the influence of the observable pixel maps on the solution, as well as, the importance of the use of the conditional distribution $P(\mathbf{m}_l \mid \mathbf{f}_k, \mathbf{d}_{l.k})$ in the observation model. First, we did not use this conditional distribution and executed the algorithm with different thresholds for the DFD defined in (17). The obtained PSNRs are plotted in figure 4. From this plot we note that the PSNR for a threshold of 255 in (17), that is, when all pixels are observable, is 32.08dB while the maximum obtained PSNR is 32.52dB which is achieved for a threshold of 15. Figure 5 shows the corresponding reconstructed images where this improvement is also visible, especially in the upper right part of the image.

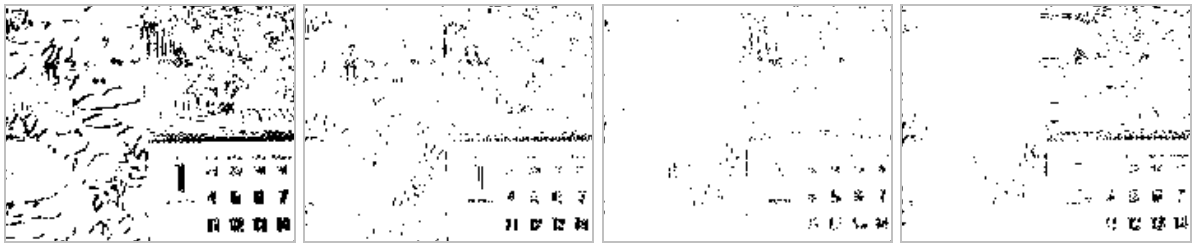


Fig. 6. In white, observable pixels $\mathbf{y}_l^o$ in $\mathbf{y}_l$ in frames 7, 8, 10 and 11 for a threshold of 15 in (17).

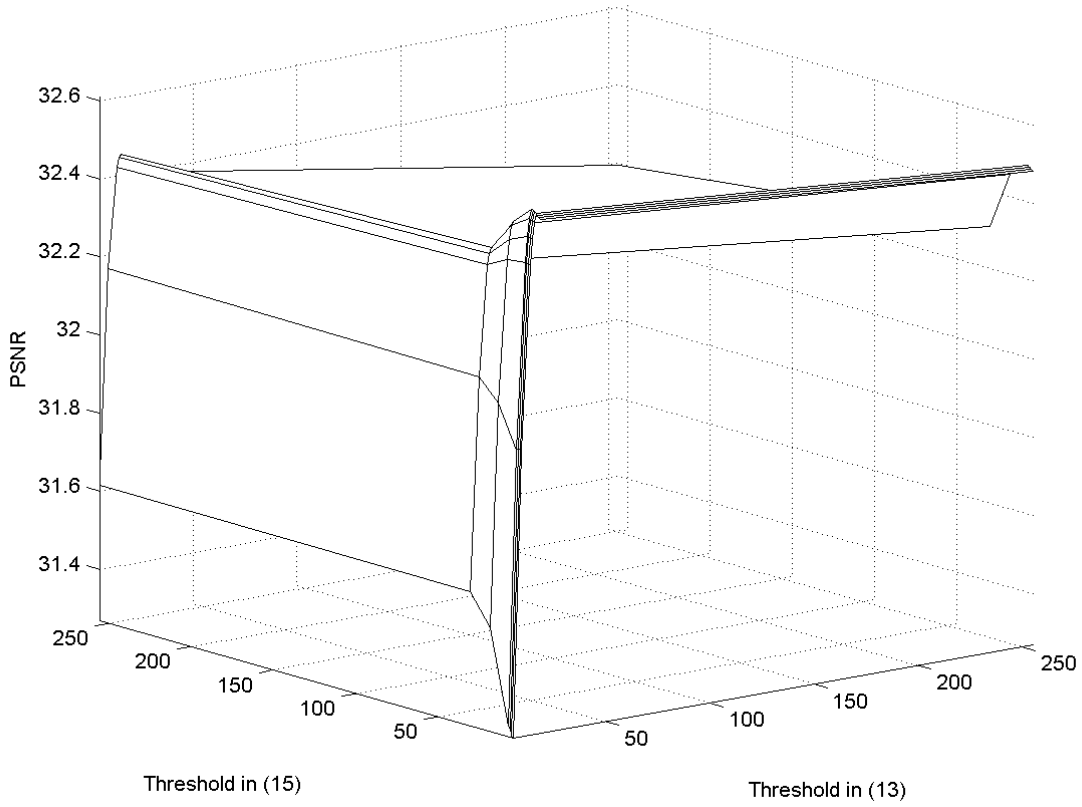


Fig. 7. PSNR evolution with the value of the thresholds in (17) and (19) for the observable pixel maps.

The observable pixel maps for frames 7, 8, 10 and 11 for a threshold T=15 are depicted in figure 6. These figures show that the best PSNR is obtained by considering as unobservable not only the pixels that have no prediction from the reference frame, that is, the borders of the image that appear or disappear but also the pixels where the predicted value via motion compensation is not very close to the current frame value.

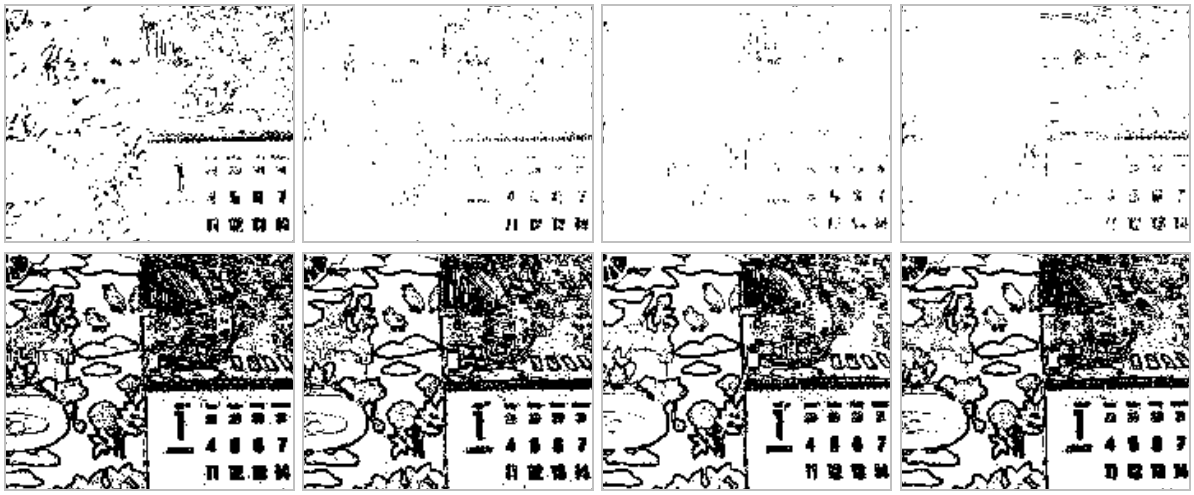
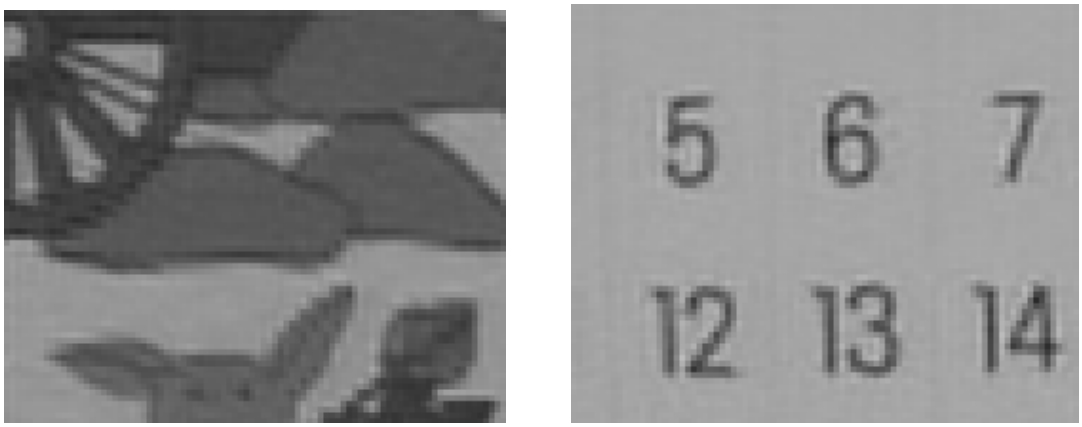


Fig. 8. Upper figures, observable pixel maps for the compressed low resolution frames 7, 8, 9 and 10 for a threshold in (17) of 16; Lower figures, observable pixel maps for the low resolution motion compensated frames 7, 8, 9 and 10 for a threshold in (19) of 7.

When we include both $P(\mathbf{m}_l \mid \mathbf{f}_k, \mathbf{d}_{l.k})$ and $P(\mathbf{y}_l \mid \mathbf{f}_k, \mathbf{d}_{l.k})$ in the observation model, the results are slightly better. As in the previous experiment, we consider different thresholds for the DFDs in (17) and (19). Since we have two different DFDs we combined the values of the thresholds for both of them obtaining the PSNR values plotted in figure 7. A threshold of T=255 for both compressed low resolution observations and motion compensated images produces $\mathbf{y}^o = \mathbf{y}$ and $\mathbf{m}^o = \mathbf{m}$ and a PSNR of 32.20dB. The highest PSNR is obtained when the thresholds in (17) and (19) are set to 16 and 7, respectively, obtaining a PSNR value of 32.61dB. Figure 8 shows the observable pixel maps for thresholds 16 and 7 in (17) and (19), respectively. Figure 9 depicts the reconstructed image for these thresholds values. The image corresponding to a threshold value of 255 for both compressed low resolution and motion compensated observations has already been shown in figure 3(d).

Note that the proposed algorithm clearly improves the quality of the reconstruction when using appropriate thresholds for the observable pixel maps since it removes the prediction errors that introduce inaccurate information in the HR image. Note also that most of the removed pixels correspond to frame 7, a frame at distance 2 from the objective frame under reconstruction, where the motion estimate is not very accurate. Note also that most of the pixels around edges are removed from $\mathbf{m}_l$ when using the maps in figure 8. This is expected since, in general, prediction errors in $\mathbf{m}_l$ are concentrated in zones with high spatial activity and contours of image objects. Figure 10 shows two zoomed up portions of Fig. 3(d) and Fig. 9 where it is clearly appreciated how the incorporation of the observable pixel maps helps to reduce artifacts around edges while obtaining a crisp and well resolved high resolution estimation, see however the number 4 in figure 10(a) and 10(b).

Fig. 9. High resolution image reconstruction with observable pixel maps corresponding to thresholds 16 and 7 for compressed low resolution and motion compensated observations respectively.



(a)

(b)

Fig. 10. (a) Two details of the image in Fig. 3(d). (b) Two details of the image in Fig. 9.

7. Conclusions

In this paper we have examined which pixels in a compressed low resolution sequence provide useful information in calculating the high resolution image we are trying to estimate. We have also studied which low resolution motion compensated pixels should be used to increase the quality of image we are reconstructing. The process to remove unobservable or poorly predicted pixels has led to modify the acquisition model in order not to take into account those observations. The proposed approach was tested on real compressed video sequences with satisfactory results.

References

[1] S. Borman and R. Stevenson, *Spatial Resolution Enhancement of Low-Resolution Image Sequences a Comprehensive Review with Directions for Future Research,* Technical Report, Laboratory for Image and Signal Analysis (LISA), University of Notre Dame, Notre Dame, IN 46556, USA, July 1998.

[2] A. K. Katsaggelos and N. P. Galatsanos (ed.), *Recovery Techniques for Image and Video Compression,* Kluwer Academic Press, 1998.

[3] S. Chaudhuri, Ed., *Super-Resolution Imaging*, Kluwer Academic Publishers, 2001.

[4] M. G. Kang and S. Chaudhuri, Eds., *Super-Resolution Image Reconstruction*, IEEE Signal Processing Magazine, vol. 20, no. 3, 2003.

[5] C. A. Segall, R. Molina, and A. K. Katsaggelos, *High-Resolution Images from Low-resolution Compressed Video*, IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 37–48, 2003.

[6] C. A. Segall, R. Molina, A. K. Katsaggelos, and J. Mateos, *Bayesian Resolution Enhancement of Compressed Video*, IEEE Transactions on Image Processing, to appear, 2004.

[7] H. C. Reeves and J. S. Lim, *Reduction of Blocking Effects in Image Coding*, Optical Engineering, vol. 23, no. 1, pp. 34-37, 1984.

[8] B. Ramamurthi and A. Gersho, *Nonlinear Space Variant Post-Processing of Block Coded Images*, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 34, no. 5, pp. 1258-1267, 1986.

[9] K. Sauer, *Enhancement of Low Bit-Rate Coded Images using Edge Detection and Estimation*, CVGIP: Graphical Models and Image Processing, vol. 53, no.1, pp. 52-62, 1991.

[10] C. J. Kuo and R. J. Hsich, *Adaptive Postprocessor for Block Encoded Images*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 5, no. 4, pp. 322-336, 1995.

[11] S. W. Wu and A. Gersho, *Improved Decoder for Transform Coding with Application to the JPEG Baseline System*, IEEE Transactions on Communications, vol. 40, pp. 251-254, 1992.

[12] J. Luo, C. W. Chen, K. J. Parker, and T. Huang, *Artifact Reduction in Low Bit Rate DCT-based Image Compression*, IEEE Transactions Circuits and Systems for Video Technology, vol. 5, no. 9, pp. 1363-1368, 1996.

[13] T. P. O'Rourk and R. L. Stevenson, *Improved Image Decompression for Reduced Transform Coding Artifacts*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 5, no. 6, pp. 490-499, 1995.

[14] T. Ozcelik, J. C. Brailean, and A. K. Katsaggelos, *Image and Video Compression Algorithms Based on Recovery Techniques using Mean Field Annealing*, Proceedings of the IEEE, vol. 83, no. 2, pp. 304-316, 1995.

[15] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, *Projection-Based Spatially Adaptive Reconstruction of Block-Transform Compressed Images*, IEEE Transactions on Image Processing, vol. 4, no. 7, pp. 896-908, 1995.

[16] Y. Yang and N. P. Galatsanos, *Removal of Compression Artifacts Using Projections onto Convex Sets and Line Process Modeling*, IEEE Transactions on Image Processing, vol. 6, no. 10, pp. 1345-1357, 1998.

[17] S. J. Reeves and S. I. Eddins, *Comments on Iterative Procedures for Reduction of Blocking Effects in Transform Coding*, IEEE Transactions Circuits and Systems for Video Technology, vol. 3, no. 6, pp. 439-440, 1993.

[18] R. Rosenholtz and A. Zakhor, *Iterative Procedures for Reduction of Blocking Effects in Transform Image Coding*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 2, no. 1, pp. 91-94, 1992.

[19] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, *Regularized Reconstruction to Reduce Blocking Artifacts of Block Discrete Cosine Transform Compressed Images*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 3, no. 6, pp. 421-432, 1993.

[20] B. K. Gunturk, Y. Altunbasak, and R. Mersereau, *Bayesian Resolution-Enhancement Framework for Transform-Coded Video*, in Proceedings of the IEEE International Conference on Image Processing, vol. 2, pp. 41-44, Thessaloniki, Greece, 2001.

[21] A. J. Patti and Y. Altunbasak, *Super-Resolution Image Estimation for Transform Coded Video with Application to MPEG*, in Proceedings of the IEEE International Conference on Image Processing, vol. 3, pp. 179-183, Kobe, Japan, 1999.

[22] R. R. Schultz and R. L. Stevenson, *A Bayesian Approach to Image Expansion for Improved Definition*, IEEE Transactions on Image Processing, vol. 3, no. 3, pp. 233-242, 1994.

[23] R. R. Schultz and R. L. Stevenson, *Extraction of High-Resolution Frames from Video Sequences*, IEEE Transactions on Image Processing, vol. 5, no. 6, pp. 996-1011, 1996.

[24] J. Mateos, A. K. Katsaggelos, and R. Molina, *Simultaneous Motion Estimation and Resolution Enhancement of Compressed Low Resolution Video*, in Proceedings of the IEEE International Conference on Image Processing, vol. 2, pp. 653-656,Vancouver, Canada, 2000.

[25] J. Mateos, A. K. Katsaggelos, and R. Molina, *Resolution Enhancement of Compressed Low Resolution Video*, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4, pp. 1919-1922, Istanbul, Turkey, 2000.

[26] C. A. Segall, R. Molina, A. K. Katsaggelos, and J. Mateos, *Bayesian High-Resolution Reconstruction of Low-Resolution Compressed Video*, in Proceedings of the IEEE International Conference on Image Processing, vol. 2, pp. 25-28, Thessaloniki, Greece, 2001.

[27] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos, *Super-Resolution from Compressed Video*, chapter 9 in Super-Resolution Imaging, S. Chaudhuri, Ed., pp. 211–242, Kluwer Academic Publishers: Boston, 2001.

[28] C. A. Segall and A. K. Katsaggelos, *Enhancement of Compressed Video using Visual Quality Metrics*, in Proceedings of the IEEE International Conference on Image Processing, vol. 2, pp. 645-648, Vancouver, Canada, 2000.

[29] J. Mateos, A. K. Katsaggelos, and R. Molina, *A Bayesian Approach for the Estimation and Transmission of Regularization Parameters for Reducing Blocking Artifacts*, IEEE Transactions on Image Processing, vol. 9, no. 7, pp. 1200-1215, 2000.

[30] A. M. Tekalp, *Digital Video Processing*, Prentice Hall, Signal Processing Series, 1995.

[31] C. A. Segall, R. Molina, A. K. Katsaggelos, and J. Mateos, *Reconstruction of High-resolution Image Frames from a Sequence of Low-Resolution and Compressed Observations*, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 2, pp. 1701 –1704, 2002.

[32] R. Molina, M. Vega, J. Abad, and A. K. Katsaggelos, *Parameter Estimation in Bayesian High-resolution Image Reconstruction with Multisensors*, IEEE Transactions on Image Processing, vol. 12, no. 12, pp. 1655–1667, 2003.

[33] J. Mateos, R. Molina, and A. K. Katsaggelos, *Bayesian High-resolution Image Reconstruction with Incomplete Multisensor Low-Resolution Systems*, in Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing, vol. III, pp. 705–708, 2003.

[34] D. G. Luenberger, *Linear and Nonlinear Programming*, Reading, MA: Addison-Wesley Publishing Company, Inc., 1984.