RECONSTRUCTION OF HIGH-RESOLUTION IMAGE FRAMES FROM A SEQUENCE OF LOW-RESOLUTION AND COMPRESSED OBSERVATIONS

C. Andrew Segall[†], Rafael Molina[*], Aggelos K. Katsaggelos[†] and Javier Mateos[*]

[†] Department of Electrical and Computer Engineering
Northwestern University, Evanston, IL 60208, USA
{asegall,aggk}@ece.nwu.edu

[*] Departamento de Ciencias de la Computación e I. A.
University of Granada, 18071 Granada, Spain
{rms, jmd}@decsai.ugr.es

ABSTRACT

A framework for recovering high-resolution information from a sequence of sub-sampled and compressed observations is presented. Compression schemes that describe a video sequence through a combination of motion vectors and transform coefficients are the focus (e.g. the MPEG and ITU family of standards), and we consider the influence of both the motion vectors and transform coefficients within the reconstruction algorithm. A Bayesian approach is utilized to incorporate the information, and results show a discernable improvement in resolution, as compared to standard interpolation methods.

1. INTRODUCTION

High-frequency information is often discarded during the acquisition and processing of an image. This data reduction begins at the image sensor, where the original scene is spatially sampled during acquisition, and continues through subsequent sampling, filtering or quantization procedures. Recovering the high-frequency information is possible though, as multiple low-resolution observations may provide additional information about the high-frequency data. This information is introduced through sub-pixel displacements in the sampling grid, which allows for the recovery of resolution.

Object motion within an image sequence provides one source of the necessary sub-pixel displacements [3,5,8]. (Global camera motion is another.) Here, we consider a sequence of observations that have been compressed prior to any resolution recovery. This mimics many practical applications, but it introduces several changes to the formulation. As a first change, the corrupting noise process becomes somewhat complex, as errors introduced during compression may be spatially varying, correlated or correspond to the well-known blocking or ringing artifacts. As a second difference, pixel intensities no longer comprise the observations. Instead, motion vectors and quantized transform coefficients are provided to the recovery algorithm. Since this data describes both the intensity and motion within the original sequence, we believe that it motivates a joint estimate of the sub-pixel displacement and high-resolution image data. This differs from other work where the estimation of the high-resolution data and the displacements are treated separately [1,2] and extends our previous efforts [6,7].

The rest of this paper is organized as follows: In section 2, we define the acquisition system to be considered. In section 3, we formulate the problem within the Bayesian framework and introduce an iterative algorithm for the joint estimation of the displacements and high-resolution information. In section 4, we present results from the proposed procedure.

2. SYSTEM MODEL

When images from a single camera are captured at closely spaced time instances, then it is reasonable to assume that the content of the frames are similar. That is, we can say that

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

where $f_l(x,y)$ and $f_k(x,y)$ are spatial locations in the high-resolution images at times l and k, respectively, $d_{l,k}^x(x,y)$ and $d_{l,k}^y(x,y)$ comprise the displacement that relates the pixel at time k to the pixel at time l, and $n_{l,k}(x,y)$ is an additive noise process that accounts for any image locations that are poorly described by the displacement model.

The expression in (1) can be rewritten with a matrix-vector notion. In this form, the relationship between two images becomes

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

where $\mathbf{f}_l$ and $\mathbf{f}_k$ are formed by lexicographically ordering each image into a one-dimensional vector, $\mathrm{C}(\mathbf{d}_{l,k})$ is the two-dimensional matrix that describes the displacement across the entire frame, $\mathbf{d}_{l,k}$ is the column vector defined by lexicographically ordering the complex values $d_{l,k}^x(x,y)+i*d_{l,k}^y(x,y)$, and $\mathbf{n}_{l,k}$ is the noise process. When the images are PMxPN arrays, then $\mathbf{f}_l$, $\mathbf{f}_k$, $\mathbf{d}_{l,k}$ and $\mathbf{n}_{l,k}$ are column vectors with length $PMPN$ and $\mathrm{C}(\mathbf{d}_{l,k})$ has dimension $PMPN$x$PMPN$.

This work has been partially supported by the "Comision Nacional de Ciencia y Tecnologia" under contract TIC2000-1275.

In classical imaging scenarios, the noise introduced in (1) and (2) is the dominant noise component. However in many modern imaging systems, the sequence of images are further degraded by compression. This reduces the bandwidth required for transmission and storage, and it requires two transformations. First, each high-resolution frame is filtered and down-sampled. Then, the down-sampled result is compressed with a video compression algorithm. Several video compression algorithms are available, with the hybrid motion compensation-transform based coding techniques the most common (e.g. the MPEG and ITU family of standards). In these methods, each image is encoded by dividing it into a sequence of equally sized blocks. Then, the blocks are predicted from previously encoded frames, and the residual error is transformed and quantized. The prediction is then signaled to the decoder as a motion vector that defines the location of the predicted block, and the residual is transmitted as a sequence of variable length code words that define the quantized transform coefficients.

The conversion of a high-resolution frame to its low-resolution and compressed observation is therefore expressed as

$$\mathbf{g}_k = \mathbf{T}_{DCT}^{-1} Q\left[\mathbf{T}_{DCT}\left(\mathbf{AH}\mathbf{f}_k - \sum_{\forall i} C\left(\mathbf{v}_{k,i}\right)\mathbf{g}_i \right)\right] + \sum_{\forall i} C\left(\mathbf{v}_{k,i}\right)\mathbf{g}_i \,, \quad (3)$$

where $\mathbf{g}_k$ is a vector that contains the compressed low-resolution images with dimension MNx1, $\mathbf{f}_k$ is the high-resolution data, $\mathbf{v}_{l,k}$ is the motion vector transmitted by the encoder that signals the prediction of frame k from previously compressed frame i, $C(\mathbf{v}_{l,i})$ represents the prediction process with a matrix (for images said to be "intra-coded", the prediction from all frames is zero), $\mathbf{A}$ is an MNx$PMPN$ matrix that sub-samples the high-resolution image, $\mathbf{H}$ is an $PMPN$x$PMPN$ matrix that filters the high-resolution image, $\mathbf{T}_{DCT}$ and $\mathbf{T}_{DCT}^{-1}$ are the forward and inverse-DCT calculations, and $Q[\cdot]$ represents the quantization procedure. Combining (2) and (3), we state the relationship between any low and high-resolution image as

$$\mathbf{g}_l = \mathbf{T}_{DCT}^{-1} Q\left[\mathbf{T}_{DCT}\left(\mathbf{AH}C\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k - \sum_{\forall i} C\left(\mathbf{v}_{l,i}\right)\mathbf{g}_i \right)\right] + \sum_{\forall i} C\left(\mathbf{v}_{l,i}\right)\mathbf{g}_i \,, \quad (4)$$

where the compression noise is assumed dominant, and the noise term in (1) and (2) neglected.

The system model in (4) motivates an algorithm that recovers high-resolution frames from a sequence of low-resolution and compressed images. The approach should be apparent by noticing that information about a single high-resolution frame appears in multiple low-resolution observations. When this information is not redundant, as provided with sub-pixel displacements in the mapping of frame $\mathbf{f}_k$ to $\mathbf{f}_l$ and the introduction of aliasing by the sampling procedure $\mathbf{AH}$, then each observation provides additional information about the high-resolution image frame.

3. PROBLEM FORMULATION AND PROPOSED ALGORITHM

The Bayesian maximum *a posterior* (MAP) estimate provides the necessary framework for recovering high-resolution information from a sequence of compressed observations. Since information about the displacements is present in the compressed bit-stream, we consider the joint estimate

$$\begin{aligned} \hat{\mathbf{f}}_k, \hat{\mathbf{D}} &= \arg\max_{\mathbf{f}_k,\mathbf{D}} \left\{ p\left(\mathbf{f}_k, \mathbf{D} \mid \mathbf{G}, \mathbf{V}\right)\right\} \\ &\arg\max_{\mathbf{f}_k,\mathbf{D}} \left\{ \frac{p\left(\mathbf{G}, \mathbf{V} \mid \mathbf{f}_k, \mathbf{D}\right) p\left(\mathbf{f}_k, \mathbf{D}\right)}{p(\mathbf{G}, \mathbf{V})} \right\} \\ &\arg\max_{\mathbf{f}_k,\mathbf{D}} \left\{ \frac{p\left(\mathbf{V} \mid \mathbf{G}, \mathbf{f}_k, \mathbf{D}\right) p\left(\mathbf{G} \mid \mathbf{f}_k, \mathbf{D}\right) p\left(\mathbf{f}_k, \mathbf{D}\right)}{p(\mathbf{G}, \mathbf{V})} \right\}, \end{aligned} \quad (5)$$

where $\hat{\mathbf{f}}_k$ and $\hat{\mathbf{D}}$ are estimates of the high-resolution image and motion field, respectively, $\mathbf{D}$ is the matrix defined as $(\mathbf{d}^T_{k\text{-}TB,k}, \ldots, \mathbf{d}^T_{k+TF,k})^T$, $\mathbf{G}$ is the matrix defined as $(\mathbf{g}^T_{k\text{-}TB}, \ldots, \mathbf{g}^T_{k+TF})^T$, $\mathbf{V}$ is the matrix defined as $(\mathbf{v}^T_{k\text{-}TB}, \ldots, \mathbf{v}^T_{k+TF})^T$, and $\mathbf{v}^T_{k+TF}$ is the column vector $(\mathbf{v}^T_{k\text{-}TB,0} \ \ldots \ \mathbf{v}^T_{k\text{-}TB,\infty})^T$ that contains all of the transmitted motion vectors utilized in the prediction of $\mathbf{g}_{k\text{-}TB}$. In the definitions, TF and TB describe the number of frames utilized along the forward and backward directions of the temporal axis.

Taking logarithms and recognizing that the optimization is independent of $p(\mathbf{G}, \mathbf{V})$, the MAP estimate becomes

$$\begin{aligned} \hat{\mathbf{f}}_k, \hat{\mathbf{D}} = \arg\max_{\mathbf{f}_k,\mathbf{D}} \{ &\log p\left(\mathbf{G} \mid \mathbf{f}_k, \mathbf{D}\right) + \log p\left(\mathbf{f}_k\right) \\ &+ \log p\left(\mathbf{V} \mid \mathbf{G}, \mathbf{f}_k, \mathbf{D}\right) + \log p\left(\mathbf{D}\right) \} \end{aligned}, \quad (6)$$

where $\mathbf{f}_k$ and $\mathbf{D}$ are assumed independent.

3.1 Distribution Models

The first density function in (6) defines the noise introduced by the quantization operator. This quantization noise is best understood in the DCT domain, where the noise distribution for each frequency index is independent but not necessarily identically distributed. Modeling the noise process as a jointly Gaussian random variable provides a reasonable approximation of the quantization error, as it can account for both the round-off error of a basic quantizer as well as the more sophisticated errors introduced by operational rate-distortion encoding methods.

The resulting conditional density is expressed as

$$\begin{aligned} p\left(\mathbf{G} \mid \mathbf{f}_k, \mathbf{D}\right) \propto \exp\Bigg\{ -\frac{1}{2} \sum_{l=k-TB}^{k+TF} &\left(\mathbf{AH}C\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k - \mathbf{g}_l\right)^T \\ &\mathbf{K}_{Q,l}^{-1} \left(\mathbf{AH}C\left(\mathbf{d}_{l,k}\right)\mathbf{f}_k - \mathbf{g}_l\right) \Bigg\} \end{aligned}, \quad (7)$$

where $\mathbf{K}_{Q,l}$ is the covariance of the quantization noise in the spatial domain at frame l. Since errors in the spatial domain are related to errors in the transform domain by the inverse-DCT operation, this covariance matrix is

$$\mathbf{K}_{Q,l} = \mathbf{T}_{DCT}^{-1} \, \mathbf{K}_{DCT,l} \, \mathbf{T}_{DCT}^{-1\ T} \,, \quad (8)$$

where $\mathbf{K}_{DCT,l}$ is the covariance matrix describing the noise in the DCT domain at frame l. One interesting property of this relationship is that when the noise is independent and identically distributed (i.i.d.) in the DCT domain, it is also i.i.d. in the spatial domain. Conversely, when the noise is not identically distributed in the DCT domain (but still

independent), then the distribution in the spatial domain becomes correlated. Both situations occur in practical applications, as the quantization intervals are equal for all frequency indices during inter-coding, which leads to an i.i.d. process in the spatial domain, while perceptually motivated quantization tables utilized during intra-coding result in correlated noise in the spatial domain.

The structure of the compression system also motivates selection of the second density function in (6), $p(\mathbf{f}_k)$. The purpose of this density function is to incorporate *a priori* information about the original high-resolution images into the recovery method. Most critical here is that the original images rarely contain coding artifacts such as ringing or blocking. Both compression errors are penalized with the density

$$p(\mathbf{f}_k) \propto \exp\left\{-\left(\lambda_1 \left\|\mathbf{Q}_1\mathbf{f}_k\right\|^2 + \lambda_2 \left\|\mathbf{Q}_2\mathbf{A}\mathbf{H}\mathbf{f}_k\right\|^2\right)\right\}, \quad (9)$$

where $\mathbf{Q}_1$ responds to the high frequency content within each block, $\mathbf{Q}_2$ responds to significant differences across the horizontal and vertical block boundaries, and λ_1 and λ_2 control the influence of the different smoothing parameters.

The distribution for the motion vectors relies on both the structure of the compression system as well as the composition and encoding quality of previously compressed images. At the high-level, selection of the motion vectors are straightforward. For each block in the current down-sampled image, a block with similar content is found within a previously encoded frame. When the selected motion vector corresponds to the actual sub-pixel displacement within the image sequence, then the density function for the motion vectors follows from (7) and is stated as

$$p(\mathbf{V} \mid \mathbf{G}, \mathbf{f}_k, \mathbf{D}) \propto \exp\left\{-\frac{1}{2}\sum_{l=k-TB}^{k+TF}\left(\mathbf{A}\mathbf{H}C(\mathbf{d}_{l,k})\mathbf{f}_k - \sum_{\forall i} C(\mathbf{v}_{l,i})\mathbf{g}_i\right)^T \mathbf{K}_{MV,l}^{-1}\left(\mathbf{A}\mathbf{H}C(\mathbf{d}_{l,k})\mathbf{f}_k - \sum_{\forall i} C(\mathbf{v}_{l,i})\mathbf{g}_i\right)\right\}, \quad (10)$$

where $\mathbf{K}_{MV,l}$ is the covariance matrix of the quantization noise in the previously encoded frames "mapped" through the motion field. This is defined as

$$\mathbf{K}_{MV,l} = \sum_{\forall i} C(\mathbf{v}_{l,i})\mathbf{T}_{DCT}\mathbf{K}_{DCT,i}\mathbf{T}_{DCT}^T C(\mathbf{v}_{l,i})^T . \quad (11)$$

When match errors occur and can be modeled as a Gaussian process, then the covariance matrix becomes

$$\mathbf{K}_{MV,l} = \sum_{\forall i} C(\mathbf{v}_{l,i})\mathbf{T}_{DCT}^{-1}\mathbf{K}_{DCT,i}\mathbf{T}_{DCT}^{-1\ T} C(\mathbf{v}_{l,i})^T + \mathbf{K}_{\text{Match},l}, \quad (12)$$

where $\mathbf{K}_{Match,l}$ is the match error in the prediction of frame l.

The last distribution appearing in (6) provides an *a priori* model for the displacement between image frames. Like the distribution in (9), we utilize a prior that penalizes large undulations in the displacements across the image while also penalizing any synthetic blocking structure. The prior is given as

$$p(\mathbf{D}) \propto \exp\left\{\sum_{l=k-TB}^{k+TF} -\left(\lambda_3\left\|\mathbf{Q}_1\mathbf{d}_{l,k}\right\|^2 + \lambda_4\left\|\mathbf{Q}_2\mathbf{d}_{l,k}\right\|^2\right)\right\}, \quad (13)$$

where λ_3 and λ_4 control the influence of the different smoothing parameters. We note that dependencies between the displacements of different frames could also be incorporated into (13).

3.2 Optimization Procedure

By substituting the models presented in (7)-(13) into the estimate described in (6), a solution that simultaneously estimates the high-resolution motion fields as well as the high-resolution images evolves. Taking the negative of (6), the algorithm becomes a minimization that is expressed as

$$\begin{aligned}\hat{\mathbf{f}}_k, \hat{\mathbf{D}} = \arg\min_{\mathbf{f}_k,\mathbf{D}}\Bigg\{ & \sum_{l=k-TB}^{k+TF}\left(\mathbf{A}\mathbf{H}C(\mathbf{d}_{l,k})\mathbf{f}_k - \mathbf{g}_l\right)^T \mathbf{K}_{Q,l}^{-1}\left(\mathbf{A}\mathbf{H}C(\mathbf{d}_{l,k})\mathbf{f}_k - \mathbf{g}_l\right) \\ & + \sum_{l=k-TB}^{k+TF}\left(\mathbf{A}\mathbf{H}C(\mathbf{d}_{l,k})\mathbf{f}_k - \sum_{\forall i} C(\mathbf{v}_{l,i})\mathbf{g}_i\right)^T \\ & \mathbf{K}_{MV,l}^{-1}\left(\mathbf{A}\mathbf{H}C(\mathbf{d}_{l,k})\mathbf{f}_k - \sum_{\forall i} C(\mathbf{v}_{l,i})\mathbf{g}_i\right) \\ & + \lambda_1\left\|\mathbf{Q}_1\mathbf{f}_k\right\|^2 + \lambda_2\left\|\mathbf{Q}_2\mathbf{f}_k\right\|^2 + \lambda_3\left\|\mathbf{Q}_1\mathbf{d}_{l,k}\right\|^2 + \lambda_4\left\|\mathbf{Q}_2\mathbf{d}_{l,k}\right\|^2\Bigg\}. \end{aligned} \quad (14)$$

The minimization of (14) is accomplished with a cyclic coordinate-decent optimization procedure [4]. In this approach, an estimate for the motion field is found while the high-resolution image is assumed known. Then, the high-resolution image is estimated with the recently found motion field. The motion field is then re-estimated using the current solution for the high-resolution frame, and the process iterates by alternatively finding the motion field and high-resolution images. Treating the high-resolution image as a known parameter, the estimate for the motion field in (14) is found by the method of successive approximations

$$\begin{aligned}\hat{\mathbf{d}}_{l,k}^{m+1} = \hat{\mathbf{d}}_{l,k}^{m} - \alpha_d^{l,k}\Bigg\{ & \frac{\partial C(\hat{\mathbf{d}}_{l,k}^m)\mathbf{f}_k}{\partial \hat{\mathbf{d}}_{l,k}^m}\mathbf{H}^T\mathbf{A}^T\mathbf{K}_{Q,l}^{-1}\left(\mathbf{A}\mathbf{H}C(\hat{\mathbf{d}}_{l,k}^m)\mathbf{f}_k - \mathbf{g}_l\right) \\ & + \frac{\partial C(\hat{\mathbf{d}}_{l,k}^m)\mathbf{f}_k}{\partial \hat{\mathbf{d}}_{l,k}^m}\mathbf{H}^T\mathbf{A}^T\mathbf{K}_{MV,l}^{-1}\left(\mathbf{A}\mathbf{H}C(\hat{\mathbf{d}}_{l,k}^m)\mathbf{f}_k - \sum_{\forall i} C(\mathbf{v}_{l,i})\mathbf{g}_i\right) \\ & + \lambda_3\mathbf{Q}_1^T\mathbf{Q}_1\hat{\mathbf{d}}_{l,k}^m + \lambda_4\mathbf{Q}_2^T\mathbf{Q}_2\hat{\mathbf{d}}_{l,k}^m\Bigg\} \end{aligned} \quad (15)$$

where $\hat{\mathbf{d}}_{l,k}^{m+1}$ and $\hat{\mathbf{d}}_{l,k}^{m}$ are $(m+1)^{th}$ and m^{th} estimates of the displacement between frame k and l, respectively, $\mathbf{A}^T$ defines the up-sampling operation, and $\alpha_d^{l,k}$ controls the convergence and rate of convergence of the algorithm.

Once the estimate for the motion field is found, then the high-resolution image is computed. For a fixed $\mathbf{D}$, the minimization of (14) is expressed as

$$\begin{aligned}\hat{\mathbf{f}}_k^{n+1} = \hat{\mathbf{f}}_k^{n} - \alpha_f\Bigg\{ & \sum_{l=k-TB}^{k+TF}\mathbf{C}^T(\mathbf{d}_{l,k})\mathbf{H}^T\mathbf{A}^T\mathbf{K}_{Q,l}^{-1}\left(\mathbf{A}\mathbf{H}\mathbf{C}(\mathbf{d}_{l,k})\hat{\mathbf{f}}_k^n - \mathbf{g}_l\right) \\ & + \sum_{l=k-TB}^{k+TF}\mathbf{C}^T(\mathbf{d}_{l,k})\mathbf{H}^T\mathbf{A}^T\mathbf{K}_{MV,l}^{-1}\left(\mathbf{A}\mathbf{H}\mathbf{C}(\mathbf{d}_{l,k})\hat{\mathbf{f}}_k^n - \sum_{\forall i}\mathbf{C}(\mathbf{d}_{l,k})\mathbf{g}_i\right) \\ & + \lambda_1\mathbf{Q}_1^T\mathbf{Q}_1\hat{\mathbf{f}}_k^n + \lambda_2\mathbf{Q}_2^T\mathbf{Q}_2\hat{\mathbf{f}}_k^n\Bigg\} \end{aligned} \quad (16)$$

where $\hat{\mathbf{f}}_k^{n+1}$ and $\hat{\mathbf{f}}_k^{n}$ are the enhanced frames at the n^{th} and $(n+1)^{th}$ iteration, α_f is a relaxation parameter that determines the convergence and rate of convergence of the algorithm, and $C^T(\mathbf{d}_{k,l})$ compensates an image backwards along the motion vectors.

4. EXPERIMENTAL RESULTS

The performance of the algorithm is illustrated by processing frames from the *Mobile* sequence. Each original image is 704x576 pixels, and it is decimated by a factor of two in each dimension, cropped to a size of 176x144 pixels and compressed with an MPEG-4 encoder operating at 1Mbps. Three frames from the compressed bit-stream are then sequentially provided to the proposed algorithm, where *TB*,*TF*=1, $\mathbf{Q}_1$ is a 3x3 discrete Laplacian, $\mathbf{Q}_2$ is a difference operation across the horizontal and vertical block boundaries, $\lambda_1=5\text{x}10^{-2}$, $\lambda_2=2\text{x}10^{-2}$, $\lambda_3=1\text{x}10^{5}$, $\lambda_4=0$, $\alpha_f=1.25\text{x}10^{-1}$, and $\alpha_d^{l,k}=1\text{x}10^{-6}$. The algorithm is stopped when $\| \hat{\mathbf{f}}_{k+1} - \hat{\mathbf{f}}_k \|^2 / \| \hat{\mathbf{f}}_k \|^2 < 2\text{x}10^{-7}$, and a new estimate for the motion field is computed whenever $\| \hat{\mathbf{f}}_{k+1} - \hat{\mathbf{f}}_k \|^2 / \| \hat{\mathbf{f}}_k \|^2 < 1\text{x}10^{-6}$.

A representative result from the algorithm appears in Figure 1, where (a) is the original image before decimation or compression, (b) is the compressed observation after bi-linear interpolation, (c) is the compressed observation after bi-cubic interpolation, and (d) is the image provided by the proposed algorithm. As can be seen from the figure, the proposed method is able to recover a significant amount of the high-frequency information. This is most evident in the text at the top and bottom of the frame as well as the horizontal stripes in the upper left of the images. The improvement also manifests itself in a PSNR of 33.20dB for the entire frame, an increase over the PSNRs of 30.41dB and 30.58dB for the bi-linear and bi-cubic methods, respectively.

5. REFERENCES

[1] D. Chen and R.R. Schultz, "Extraction of High-Resolution Still from MPEG Sequences," *Proc. of the IEEE ICIP*, pp.465-69, Chicago, IL, Oct. 4-7, 1998.

[2] B.K. Gunturk, Y. Altunbasak and R.Mersereau, "Bayesian Resolution-Enhancement Framework for Transform-Coded Video," *Proc. of the IEEE ICIP*, Thessaloniki, Greece, Oct. 7-10, 2001.

[3] R.C. Hardie, K.J. Barnard and E.E. Armstrong, "Joint MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images," *IEEE Trans. IP*, vol.6, no.12, pp.1621-1633, 1997.

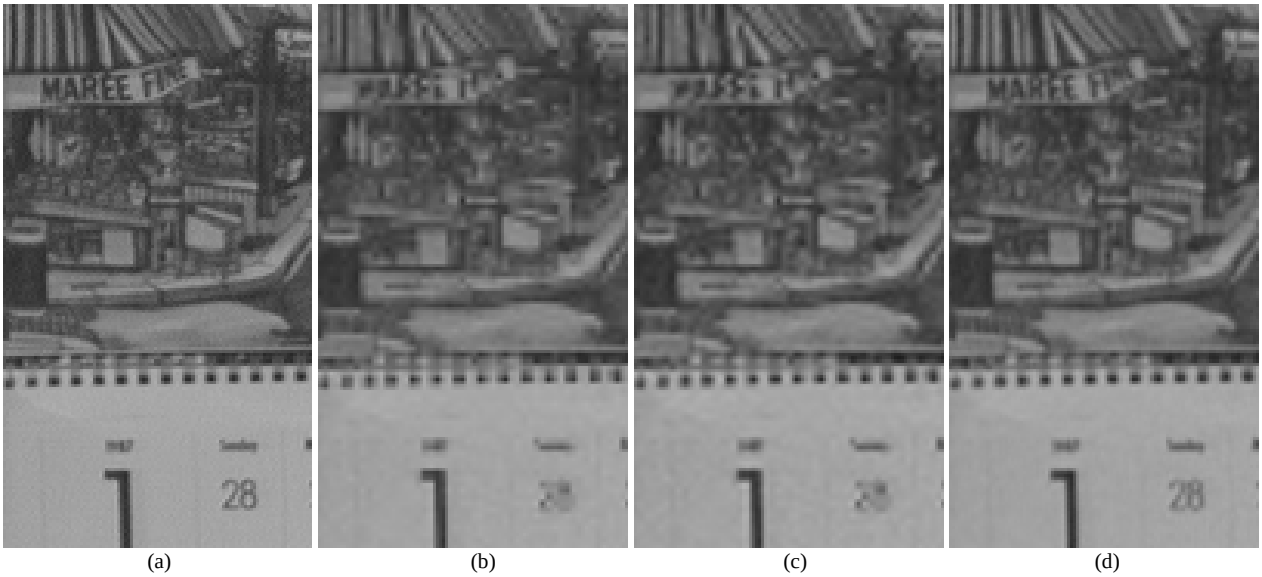
[4] D.G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, 1984.

[5] R.R. Schultz and R.L. Stevenson, "Extraction of High Resolution Frames from Video Sequences," *IEEE Trans. IP*, vol.5, no.6, pp.996-1011, 1996.

[6] C.A. Segall, R. Molina, A.K. Katsaggelos and J. Mateos, "Bayesian High-Resolution Reconstruction of Low-Resolution and Compressed Video," *Proc. of the IEEE ICIP*, Thessaloniki, Greece, Oct. 7-10, 2001.

[7] C.A. Segall, A.K. Katsaggelos, R. Molina and J. Mateos, "Super-Resolution from Compressed Video," in Super-Resolution Imaging, S. Chaudhuri, editor, Chapter 9, pp.211-242, Kluwer Academic Publishers, forthcoming.

[8] B. Tom and A.K. Katsaggelos, "Resolution Enhancement of Monochrome and Color Video Using Motion Compensation," *IEEE Trans. IP*, vol.10, no.2, pp.278-287, Feb. 2001.



(a) (b) (c) (d)

Figure 1. Recovery of high-frequency information from a sequence of low-resolution and compressed observations: (a) original image before decimation and compression; (b) decoded observation after bi-linear interpolation; (c) decoded observation after bi-cubic interpolation, and (d) result of the proposed method. PSNRs for the frames are (b) 30.41dB, (c) 30.58dB and (d) 33.20dB.