16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

# MAXIMUM A POSTERIORI VIDEO SUPER-RESOLUTION WITH A NEW MULTICHANNEL IMAGE PRIOR

*Stefanos P. Belekos[1,2], Nikolaos P. Galatsanos[3], and Aggelos K. Katsaggelos[1,2]*

[1] Faculty of Physics, Department of Electronics, Computers, Telecommunications and Control,
National and Kapodistrian University of Athens
Panepistimiopolis, Zografos, 15784, Athens, Greece
phone: + (30) 2107276873, email: stefbel@phys.uoa.gr

[2] Department of Electrical Engineering and Computer Science, Northwestern University
Evanston, IL, 60208-3118, USA
[3] Department of Electrical and Computer Engineering, University of Patras
26500, Rio, Greece
phone: + (30) 2610969861, email: ngalatsanos@upatras.gr

## ABSTRACT

*Super-resolution (SR) is the term used to define the process of estimating a high resolution (HR) image or a set of HR images from a set of low resolution (LR) observations. In this paper we propose a class of SR algorithms based on the maximum a posteriori (MAP) framework. These algorithms utilize a new multichannel image prior model, along with the state-of-the art image prior and observation models. Numerical experiments comparing the proposed algorithms, demonstrate the advantages of the adopted multichannel approach.*

## 1. INTRODUCTION

Resolution enhancement of an image / frame or of a video sequence based on multiple LR observed frames, which is also referred to in the literature as *super-resolution (SR)*, is of critical importance in signal processing applications, such as video surveillance, remote sensing, medical imaging, cell phones, digital video cameras, portable Digital Versatile Disc (DVD), portable Global Positioning Systems (GPS), High Definition Television (HDTV) e.t.c. [1]-[2]. The super-resolution problem is an inverse problem that requires a regularized solution. The Bayesian framework, used in this work, offers many advantages (see [1] for example).

In most of the Bayesian formulations which have been used for the SR problem so far, single channel image priors have been adopted, based on a Gaussian stationary assumption for the residuals of the local image differences [3], whereas there have also been proposed non Bayesian total variation (TV) regularization techniques [4]-[5]. As far as the imaging models are concerned, many techniques are incorporating the motion field (MF) information [3]-[6], whereas others do not use this information at all.

The term multichannel [7] in the context of video recovery implies the use of the between frames correlations. Such approaches have been used successfully in the past for video restoration and compressed video reconstruction [6]-[8] and [9], respectively. However, these approaches were deterministic and the multichannel idea was basically imposed by using between frame regularization.

In this paper we address the video SR problem utilizing a MAP approach. One of the main contributions of this work is the use of a new multichannel prior that incorporates registration between frames information which is directly related with the accuracy of the motion field estimation.

This paper is organized as follows. Section 2 describes the appropriate mathematical background on all possible image priors and observation models used. Section 3 introduces a MAP problem formulation for the SR of uncompressed video for each one of the proposed models, along with the realizations of the corresponding algorithms. In section 4 we demonstrate the efficacy of each one of the models through simulation experiments which provides a comparison among them indicating the benefits of the new prior. Finally, section 5 presents the conclusions.

## 2. MATHEMATICAL BACKGROUND

### 2.1 Observation Models

In this paper we use two different observation models. In the first one the relationship between a LR observation $\mathbf{g}_i$ and its HR counterpart $\mathbf{f}_i$ is given in matrix-vector form (all images have been lexicographically ordered) by

$$\mathbf{g}_i = \mathbf{A}\mathbf{H}\mathbf{f}_i + \mathbf{n}_i \quad \text{for } i = 1,2\ldots P, \tag{1}$$

where $\mathbf{g}_i$ and $\mathbf{f}_i$ are of dimensions $MN \times 1$ and $LMLN \times 1$ respectively, $\mathbf{A}$ is the $MN \times LMLN$ down-sampling matrix which sub-samples the HR frame, $\mathbf{H}$ is the $LMLN \times LMLN$ blurring matrix, $\mathbf{n}_i$ of size $MN \times 1$, represents the additive white Gaussian noise (AWGN) term, which includes the acquisition errors, $P$ represents the total number of frames used and $L$ denotes the resolution enhancement factor.

Equation (1) can be rewritten for the multichannel case as

$$\tilde{\mathbf{g}} = \tilde{\mathbf{A}}\tilde{\mathbf{H}}\tilde{\mathbf{f}} + \tilde{\mathbf{n}}, \tag{2}$$

where

$$\tilde{\mathbf{g}} = \left[\mathbf{g}_{k-m}^T,\ldots,\mathbf{g}_k^T,\ldots,\mathbf{g}_{k+n}^T\right]^T,\ \tilde{\mathbf{f}} = \left[\mathbf{f}_{k-m}^T,\ldots,\mathbf{f}_k^T,\ldots,\mathbf{f}_{k+n}^T\right]^T,$$
$$\tilde{\mathbf{n}} = \left[\mathbf{n}_{k-m}^T,\ldots,\mathbf{n}_k^T,\ldots,\mathbf{n}_{k+n}^T\right]^T \tag{3}$$

and $n$, $m$ indicate respectively the number of frames used in the forward and backward temporal directions ($n + m + 1 = P$), with respect to the $k-th$ frame, $T$ denotes the transpose of a matrix-vector, $\mathbf{A}^T$ defines the up-sampling operation and

$$\tilde{\mathbf{H}} = diag\{\mathbf{H},\ldots,\mathbf{H},\ldots,\mathbf{H}\}, \tilde{\mathbf{A}} = diag\{\mathbf{A},\ldots,\mathbf{A},\ldots,\mathbf{A}\}, \tag{4}$$

are respectively of dimensions $PLMLN \times PLMLN$ and $PMN \times PLMLN$. Note that a generalization of models (1) and (2)

is to consider a different blur and down-sampling per frame, i.e. replace $\mathbf{H}$ by $\mathbf{H}_i$ and $\mathbf{A}$ by $\mathbf{A}_i$.

The second observation model utilizes the motion field, according to which

$$\mathbf{f}_i = \mathbf{M}(\mathbf{d}_{i,k})\mathbf{f}_k + \mathbf{n}_{i,k} = \hat{\mathbf{f}}_i^k + \mathbf{n}_{i,k},\ \mathbf{M}_{ik} = \mathbf{M}(\mathbf{d}_{i,k}), \quad (5)$$

where $\mathbf{f}_i$ and $\mathbf{f}_k$ are column vectors of size $LMLN \times 1$, $\mathbf{M}(\mathbf{d}_{i,k})$ is the 2-D motion compensation matrix of size $LMLN \times LMLN$, mapping frame $\mathbf{f}_k$ into frame $\mathbf{f}_i$ with the use of $\mathbf{d}_{i,k}$ (displacements at each pixel location), and $\mathbf{n}_{i,k}$ is assumed to be an AWGN process that accounts for the motion compensation (registration) errors and is also described by a column vector of size $LMLN \times 1$.

Combining (1) with (5), results in the definition of the second imaging model (warp – blur model [1]),

$$\mathbf{g}_i = \mathbf{AHM}(\mathbf{d}_{i,k})\mathbf{f}_k + \mathbf{w}_{i,k} \quad \text{for } i = k-m,...,k,...,k+n, \quad (6)$$

with $\mathbf{w}_{i,k} = \mathbf{n}_i + \mathbf{AHn}_{i,k}$ a column vector of size $MN \times 1$ representing *the total contribution* of the noise term (including both registration and acquisition errors) which is again modelled to be AWGN. Here, using the imaging model in (6) we are incorporating only the motion information that is relevant to the SR of the middle ($k-th$) frame.

### 2.2 Image Prior Models

In this work we consider two prior models. Although deterministic approaches have been developed in the past [10], the Bayesian formulation of a prior offers many advantages. A simple model based on a Gaussian stationary assumption (stochastic non-stationary assumptions [11] have also been used) is given by

$$\varepsilon_j = \mathbf{Q}\mathbf{f}_j \sim N(\mathbf{0}, \alpha_j^{-1}\mathbf{I}) \text{, or equivalently}$$

$$p(\mathbf{f}_j;\alpha_j) \propto \exp\left(-\frac{\alpha_j}{2}\left\|\mathbf{Q}\mathbf{f}_j\right\|^2\right), \quad (7)$$

where $\mathbf{Q}$ represents a convolutional operator (discrete Laplacian) of size $LMLN \times LMLN$, $\|\cdot\|$ denotes the $l_2$ norm and the parameter $\alpha_j$ accounts for the within channel (within frame *j*) inverse variance, with $\mathbf{0}$ being the $LMLN \times 1$ zero vector and $\mathbf{I}$ the $LMLN \times LMLN$ identity matrix.

The second image prior model introduced in this paper is a new *mulichannel prior* that takes into account both within frame smoothness captured by equation (7) and between-frame smoothness incorporated through the motion field information. More specifically the multichannel prior we propose, is given by

$$p(\tilde{\mathbf{f}};\tilde{\boldsymbol{\beta}},\tilde{\boldsymbol{\alpha}}) \propto Z \prod_{i=k-m}^{k+n} \prod_{\substack{j=k-m \\ j\neq i}}^{k+n} p(\mathbf{f}_i \mid \mathbf{f}_j;\beta_{ij}) p(\mathbf{f}_j;\alpha_j), \quad (8)$$

where

$$p(\mathbf{f}_i|\mathbf{f}_j;\beta_{ij}) \propto \exp\left(-(\beta_{ij}/2)\left\|\mathbf{f}_i - \mathbf{M}_{ij}\mathbf{f}_j\right\|^2\right) \quad (9)$$

and $p(\mathbf{f}_j;\alpha_j)$ is given by (7) with

$$\mathbf{M}_{ji} = (\mathbf{M}_{ij})^T = \mathbf{M}(\mathbf{d}_{j,i}), \quad (10)$$

where matrix $(\mathbf{M}_{ij})^T$ clearly represents the operation of backward motion compensation along the motion vectors (mapping frame $\mathbf{f}_i$ into frame $\mathbf{f}_j$ with the use of $\mathbf{d}_{j,i}$), whereas $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\alpha}}$ are also the column vectors that contain the parameters $\beta_{ij}$ and $\alpha_j$, respectively. The parameter $\beta_{ij}$ represents the inverse variance (precision) of the motion compensation error between frames *i* and *j*.

The joint pdf in the right hand side term of Eq. (8) can be also written as

$$p(\mathbf{f}_i,\mathbf{f}_j;\beta_{ij},\alpha_j) = p(\mathbf{f}_i \mid \mathbf{f}_j;\beta_{ij})\, p(\mathbf{f}_j;\alpha_j) \propto$$

$$\propto |\mathbf{R}_{ij}|^{-1/2} \exp\left(-\frac{1}{2}[\mathbf{f}_i^T,\mathbf{f}_j^T]\mathbf{R}_{ij}^{-1}\begin{bmatrix}\mathbf{f}_i\\ \mathbf{f}_j\end{bmatrix}\right), \quad (11)$$

where

$$\mathbf{R}_{ij}^{-1} = \begin{bmatrix} \beta_{ij}\mathbf{I} & -\beta_{ij}\mathbf{M}_{ij} \\ -\beta_{ij}(\mathbf{M}_{ij})^T & \alpha_j\mathbf{Q}^T\mathbf{Q} + \beta_{ij}(\mathbf{M}_{ij})^T\mathbf{M}_{ij} \end{bmatrix} \quad (12)$$

denotes the 'cross-channel' inverse covariance matrix. Then, we can rewrite (8) as

$$p(\tilde{\mathbf{f}};\tilde{\boldsymbol{\beta}},\tilde{\boldsymbol{\alpha}}) \propto Z \exp\left[-\frac{1}{2}\sum_{i=k-m}^{k+n}\sum_{\substack{j=k-m \\ j\neq i}}^{k+n}[\mathbf{f}_i^T,\mathbf{f}_j^T]\mathbf{R}_{ij}^{-1}\begin{bmatrix}\mathbf{f}_i\\ \mathbf{f}_j\end{bmatrix}\right] \quad (13)$$

$$= Z \exp\left[-\frac{1}{2}\tilde{\mathbf{f}}^T\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{f}}\right],$$

thus

$$Z = \left|\tilde{\mathbf{R}}^{-1}\right|^{1/2} = \left|\tilde{\mathbf{R}}\right|^{-1/2}, \quad (14)$$

with $\tilde{\mathbf{R}}^{-1}$ representing the inverse covariance matrix of the prior pdf ($\tilde{\mathbf{f}} \sim N(\mathbf{0},\tilde{\mathbf{R}})$) which is not given in closed form due to space constraints.

Evaluating $Z$ from (14) is cumbersome given the size of $\tilde{\mathbf{R}}^{-1}$ and mainly because of the need to find the derivative of this determinant with respect to all the involved parameters. Therefore, in the herein work we *approximate the functional relationship of the parameters* $\alpha_j$ and $\beta_{ij}$ and the normalizing constant $Z$ (the partition function) of the multichannel prior as (in this paper we choose $m = n$)

$$Z \propto \prod_{j=k-m}^{k+m}(\alpha_j^{P(LMLN\text{-}1)/2}) \prod_{i=k-m}^{k+m}\prod_{\substack{j=k-m \\ j\neq i}}^{k+m}(\beta_{ij}^{PLMLN/2}). \quad (15)$$

Clearly, the introduced prior is an improper one. Improper priors have been used in the past with success in image recovery problems, see for example [11].

## 3. MAP PROBLEM FORMULATION AND PROPOSED ALGORITHMS

Taking into account the main possible combinations of the observation models ((1) or equivalently (2) and (6)) with the prior models ((7) and (13)) we propose three formulations of the HR problem and derive the corresponding MAP algorithms.

### 3.1 Model 1

The simplest approach to the super resolution problem is to use a single channel to obtain the observation model of (1). In this case no motion field information is used (in applications where more than one frames are to be restored, these frames are independently super resolved based on (1) without using any of the adjacent channels). Given (1), the fidelity pdf is defined as

$$p(\mathbf{g}_i \mid \mathbf{f}_i;\gamma_i) \propto \gamma_i^{\frac{MN}{2}} \exp\left(-(\gamma_i/2)\|\mathbf{g}_i - \mathbf{AHf}_i\|^2\right), \quad (16)$$

where the parameter $\gamma_i^{-1}$ is the acquisition noise variance (inverse precision), whereas the prior model is defined by (7).

In obtaining the MAP estimate the following objective function is minimized

$$\begin{aligned} J_{MAP}(\mathbf{f}_i \mid \mathbf{g}_i;\alpha_i,\gamma_i) &\propto -2\log p(\mathbf{g}_i,\mathbf{f}_i;\alpha_i,\gamma_i) = \\ &= -2\log p(\mathbf{g}_i \mid \mathbf{f}_i;\gamma_i) - 2\log p(\mathbf{f}_i;\alpha_i), \end{aligned} \quad (17)$$

resulting in

$$\alpha_i = \frac{(LMLN-1)}{\|\mathbf{Qf}_i\|^2}, \gamma_i = \frac{MN}{\|\mathbf{g}_i - \mathbf{AHf}_i\|^2}, \quad (18)$$

$$\left(\mathbf{H}^T\mathbf{A}^T\mathbf{AH} + (\alpha_i/\gamma_i)\mathbf{Q}^T\mathbf{Q}\right)\hat{\mathbf{f}}_i = \mathbf{H}^T\mathbf{A}^T\mathbf{g}_i\,. \quad (19)$$

### 3.2 Model 2

This model is based on [3], where the observation model is described by (6) and the prior model is also given by (7). In that case the fidelity pdf is given by

$$\begin{aligned} p(\mathbf{g}_i \mid \hat{\mathbf{f}}_i^k, \mathbf{d}_{i,k}, \mathbf{d}_{k,i};\gamma_{ik}) \equiv p(\mathbf{g}_i \mid \mathbf{f}_k,;\gamma_{ik}) \propto \\ \propto \gamma_{ik}^{\frac{MN}{2}} \exp(-(\gamma_{ik}/2)\|\mathbf{g}_i - \mathbf{AHM}_{ik}\mathbf{f}_k\|^2), \end{aligned} \quad (20)$$

where the parameters $\gamma_{ik}$ are the inverse variances (precisions related to the motion compensation errors / registration noise and also to the acquisition noise) and as is expected for $i = k$ it holds that $\gamma_{ik} = \gamma_i = \gamma_k$. Moreover, when $\mathbf{f}_k$ (and the motion field matrices) are given, the random variables $\mathbf{g_i}$ (observations), or else the respective error terms, are assumed to be *statistically* independent. Thus we have

$$p(\tilde{\mathbf{g}} \mid \mathbf{f}_k;\tilde{\boldsymbol{\gamma}}) = \prod_{i=k-m}^{k+m} p(\mathbf{g}_i \mid \mathbf{f}_k, \mathbf{d}_{i,k}, \mathbf{d}_{k,i};\gamma_{ik}), \quad (21)$$

where $\tilde{\boldsymbol{\gamma}}$ denotes the column vector that contains all the (scalar) parameters $\gamma_{ik}$.

With this model, the objective function that is minimized is given by

$$\begin{aligned} J_{MAP}(\mathbf{f}_k \mid \tilde{\mathbf{g}};\tilde{\boldsymbol{\gamma}},\alpha_k) &\propto -2\log[p(\tilde{\mathbf{g}} \mid \mathbf{f}_k;\tilde{\boldsymbol{\gamma}})p(\mathbf{f}_k;\alpha_k)] = \\ &= -2\log[\prod_{i=k-m}^{k+m} p(\mathbf{g}_i \mid \mathbf{f}_k, \mathbf{d}_{i,k}, \mathbf{d}_{k,i};\gamma_{ik})p(\mathbf{f}_k;\alpha_k)] = \\ &= -2\log[\prod_{i=k-m}^{k+m} p(\mathbf{g}_i \mid \mathbf{f}_k, \mathbf{d}_{i,k}, \mathbf{d}_{k,i};\gamma_{ik})] - 2\log[p(\mathbf{f}_k;\alpha_k)], \end{aligned} \quad (22)$$

resulting in

$$\gamma_{ik} = \frac{MN}{\|\mathbf{g}_i - \mathbf{AHM}_{ik}\mathbf{f}_k\|^2}, \quad (23)$$

$$(\tilde{\mathbf{J}} + a_k\mathbf{Q}^T\mathbf{Q})\hat{\mathbf{f}}_k = \tilde{\mathbf{Z}}, \quad (24)$$

where

$$\begin{aligned} \tilde{\mathbf{J}} &= \sum_{i=k-m}^{k+m} [\gamma_{ik}(\mathbf{M}_{ki}\mathbf{H}^T\mathbf{A}^T\mathbf{AHM}_{ik})], \\ \tilde{\mathbf{Z}} &= \sum_{i=k-m}^{k+m} [\gamma_{ik}(\mathbf{M}_{ki}\mathbf{H}^T\mathbf{A}^T\mathbf{g}_i)], \end{aligned} \quad (25)$$

whereas the estimation of the parameter $\alpha_k$ is given (by the left hand side term of (18)). Obviously with this model only the part of the motion field information which is relative to the HR frame $\hat{\mathbf{f}}_k$ is used and this contribution is attributed to the formulation of the observation model.

### 3.3 Model 3

With this model, the multichannel observation model described by (2) is combined with the new multichannel prior described by (8), (13) and (15).

Based on the above analysis the fidelity pdf term is given by

$$p(\tilde{\mathbf{g}} \mid \tilde{\mathbf{f}};\tilde{\boldsymbol{\gamma}}) \propto (Det\{\tilde{\boldsymbol{\Gamma}}\})^{-\frac{1}{2}} \cdot \exp\{-\frac{1}{2}(\tilde{\mathbf{g}} - \tilde{\mathbf{A}}\tilde{\mathbf{H}}\tilde{\mathbf{f}})^T\tilde{\boldsymbol{\Gamma}}^{-1}(\tilde{\mathbf{g}} - \tilde{\mathbf{A}}\tilde{\mathbf{H}}\tilde{\mathbf{f}})\}, \quad (26)$$

where $\tilde{\boldsymbol{\Gamma}} = diag\left\{\gamma_{k-m}^{-1}\mathbf{I},...,\gamma_k^{-1}\mathbf{I},...,\gamma_{k+m}^{-1}\mathbf{I}\right\}$ is the covariance matrix of size $PMN \times PMN$, $\mathbf{I}$ is the identity matrix of size $MN \times MN$, and $\tilde{\boldsymbol{\gamma}} = [\gamma_{k-m},...,\gamma_k,...,\gamma_{k+m}]^T$ is the column vector that contains the inverse noise variances for each one of the channels that are used.

Consequently, the objective function is given by

$$\begin{aligned} J_{MAP}(\tilde{\mathbf{f}} \mid \tilde{\mathbf{g}};\tilde{\boldsymbol{\beta}},\tilde{\boldsymbol{\alpha}},\tilde{\boldsymbol{\gamma}}) &\propto -2\log p(\tilde{\mathbf{g}},\tilde{\mathbf{f}};\tilde{\boldsymbol{\beta}},\tilde{\boldsymbol{\alpha}},\tilde{\boldsymbol{\gamma}}) = \\ &= -2\log p(\tilde{\mathbf{g}} \mid \tilde{\mathbf{f}};\tilde{\boldsymbol{\gamma}}) - 2\log p(\tilde{\mathbf{f}};\tilde{\boldsymbol{\beta}},\tilde{\boldsymbol{\alpha}}) \end{aligned} \quad (27)$$

and its minimization results in

$$\alpha_j = \frac{P(LMLN-1)}{(P-1)\|\mathbf{Qf}_j\|^2}, \beta_{ij} = \frac{PLMLN}{\|\mathbf{f}_i - \mathbf{M}_{ij}\mathbf{f}_j\|^2}, \quad (28)$$

$$(\tilde{\mathbf{G}} + \tilde{\mathbf{R}}^{-1})\hat{\tilde{\mathbf{f}}} = \tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{g}}, \quad (29)$$

with

$$\tilde{\mathbf{G}} = \tilde{\mathbf{H}}^T\tilde{\mathbf{A}}^T\tilde{\boldsymbol{\Gamma}}^{-1}\tilde{\mathbf{A}}\tilde{\mathbf{H}} =$$

$$= diag\{\gamma_{k-m}\mathbf{H}^T\mathbf{A}^T\mathbf{AH},...,\gamma_k\mathbf{H}^T\mathbf{A}^T\mathbf{AH},...,\gamma_{k+m}\mathbf{H}^T\mathbf{A}^T\mathbf{AH}\}$$

and

$$\tilde{\boldsymbol{\Lambda}} = \tilde{\mathbf{H}}^T\tilde{\mathbf{A}}^T\tilde{\boldsymbol{\Gamma}}^{-1} = diag\{\gamma_{k-m}\mathbf{H}^T\mathbf{A}^T,...,\gamma_k\mathbf{H}^T\mathbf{A}^T,...,\gamma_{k+m}\mathbf{H}^T\mathbf{A}^T\}.$$

In this model the motion field information is taken into account only through the prior and not through the observation term, whereas (18) also holds as far as the estimation of $\gamma_i$ is concerned. Moreover, in model 3 *simultaneous SR (and restoration) of all the HR frames is taking place*, which is not the case in models 1 and 2. Finally, it is clear that (19), (24) and (29) can not be solved in closed form, given that analytical inversion of matrices is not possible due to the non-circulant nature of matrices $\mathbf{A}, \mathbf{A}^T$ and $\mathbf{M}_{ij}$. Thus, we resorted to a numerical solution using *a conjugate-gradient (C.G.) algorithm.*

## 4. NUMERICAL EXPERIMENTS

In this section we present numerical experiments to evaluate our algorithms and also justify the benefits provided by the use of the new multichannel prior based on (13). In all experiments, a central $316 \times 316$ region of the sequence *Mobile* was selected, similarly to the experiments in [3]-[4] (these experiments have also been conducted using more sequences yielding similar results). Moreover, $m = n = 2$ ($k = 3$) and frames '018' - '022' were chosen.

Two cases were considered, where in the first one uniform $9 \times 9$ blur was used, whereas in the second one no blur was used ($\mathbf{H} = \mathbf{I}$). After blurring, sub-sampling by the factor of two ($L = 2$) in both dimensions took place and white Gaussian noise was added such that the blurred signal-to-noise ratio (BSNR) defined (in dB) as

$$BSNR = 10\log_{10}\left(\left\|\overline{\mathbf{AHf}_i} - \mathbf{AHf}_i\right\|^2 / (MN\gamma_i^{-1})\right),$$

or equivalently the SNR when $\mathbf{H} = \mathbf{I}$, for each LR frame equals to 20dB, 30dB and 40dB ( $\overline{\mathbf{AHf}_i}$ is the spatial mean of vector $\mathbf{AHf}_i$ ).

The metric used to quantify performance was the improvement in signal-to-noise ratio (ISNR). This metric (in dB) is defined as

$$ISNR = 10\log_{10}\left(\left\|\mathbf{f}_i - \mathbf{g}_{i,I}\right\|^2 / \left\|\mathbf{f}_i - \hat{\mathbf{f}}_i\right\|^2\right),$$

where $\mathbf{g}_{i,I}$ denotes the bicubic interpolation of the $i^{th}$ LR observation.

In model 1, considering both experimental cases, an iterative scheme was used (in our attempt to get the best possible initial conditions for models 2 and 3) where the bicubically interpolated LR observations served as initial conditions for the C.G. algorithm and for the estimation of the parameters.

In all experiments (in both cases) with models 2 and 3 the algorithm (including the estimation of the parameters) is initialized by a single frame stationary super-resolution algorithm (model 1) from which the motion field computation was also performed using a 3 level hierarchical block matching algorithm with integer pixel accuracy at each level. For the C.G. algorithm implementation, matrices $\mathbf{M}_{ij}$ were initially estimated and remained fixed (no iterative scheme was adopted and the same holds for the estimated precision parameters in the proposed algorithms of models 2 and 3).

In all the aforementioned models, the convergence criterion which was adopted for the termination of the C.G. algorithm was $\left\|\mathbf{f}_k^{new} - \mathbf{f}_k^{old}\right\|^2 / \left\|\mathbf{f}_k^{old}\right\|^2 < 5 \cdot 10^{-5}$ (in model 1 it was used for each frame independently).

Moreover, a noteworthy observation of our experiments was the robustness of model 3 in terms of initial conditions. More specifically the ISNR in all cases, when the (C.G.) algorithm was initialized by the bicubically interpolated LR observations and the motion fields along with the parameters were also estimated by them, proved to be lower but close enough to its respective values which are given in tables 1 and 2 below (for this model).

In table 1 the ISNR results for the blurry case are given, whereas in table 2 the ISNR results for no blur super-resolution are shown (for different (B)SNRs).

TABLE 1
ISNR (IN dB) COMPARISON AMONG MODELS WITH RESPECT TO THE MIDDLE FRAME

| Noise Level | BSNR=20dB | BSNR=30dB | BSNR=40dB |
|---|---|---|---|
| Model1 | 1.5489 | 3.4379 | 4.7620 |
| Model2 | 2.0390 | 3.5480 | 4.8175 |
| Model3 | 2.5756 | 4.1684 | 5.4582 |

TABLE 2
ISNR (IN dB) COMPARISON AMONG MODELS WITH RESPECT TO THE MIDDLE FRAME (NO BLUR CASE)

| Noise Level | SNR=20dB | SNR=30dB | SNR=40dB |
|---|---|---|---|
| Model1 | 1.9269 | 3.6800 | 4.1258 |
| Model2 | 2.2788 | 3.7974 | 4.1522 |
| Model3 | 2.8934 | 4.3088 | 4.5426 |

Figures 1, 2, 3 and 4 show representative examples of the proposed algorithms. In these figures, due to space constraints, we show the central segment of the corresponding (middle) frame for each adopted approach. As can be seen, several areas benefit from the recovery (*mainly by model 3*) i.e., the numbers in the calendar are sharper.

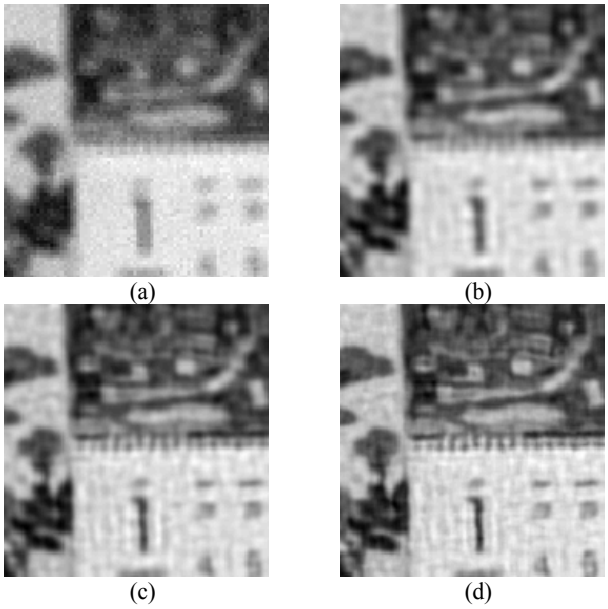



Figure 1 - SR results for BSNR=20dB (a) Segment of bicubically interpolated LR observation of middle frame, (b) SR model 1 (ISNR=1.5489dB), (c) SR model 2 (ISNR=2.0390dB), (d) SR model 3 (ISNR=2.5756 dB).

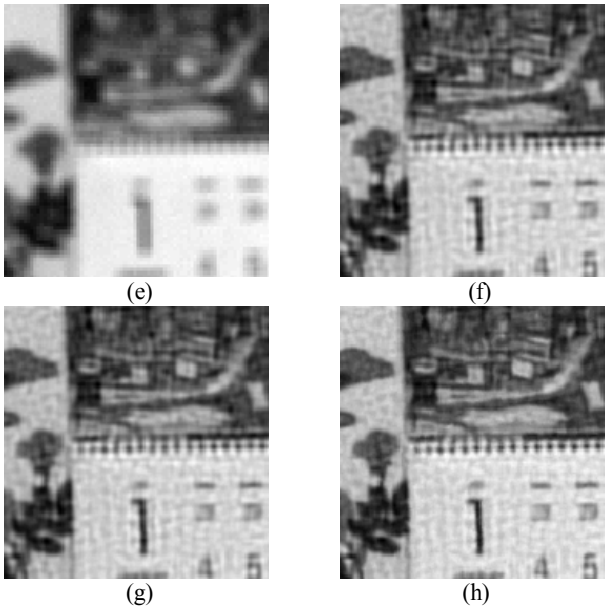
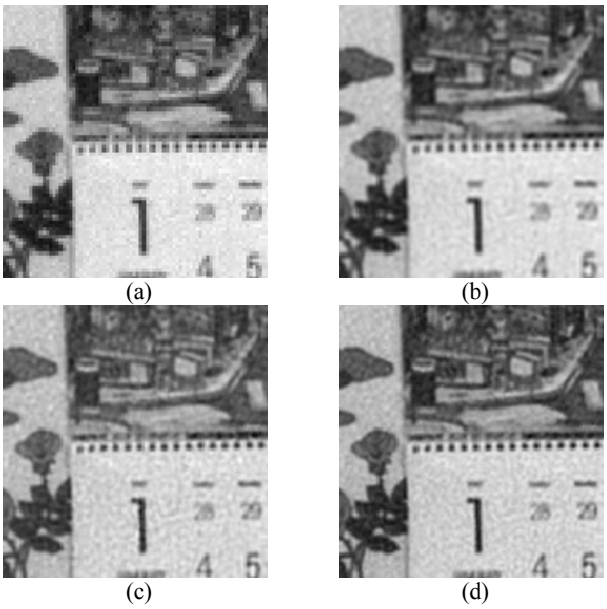
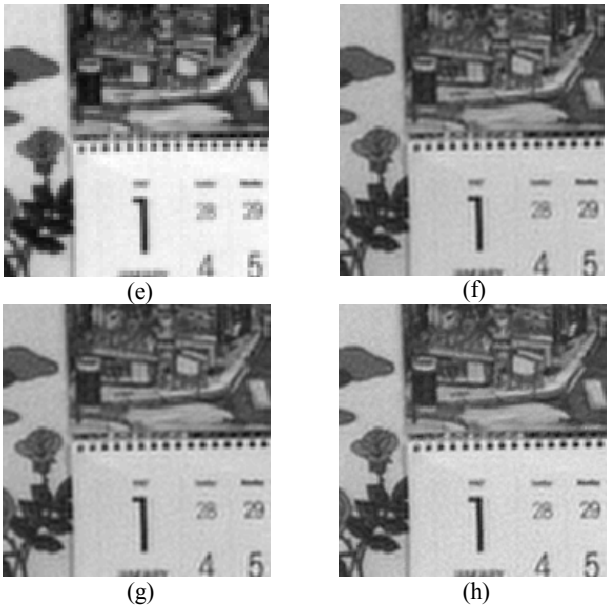



Figure 2 - SR results for BSNR=30dB (e) Segment of bicubically interpolated LR observation of middle frame, (f) SR model 1 (ISNR=3.4379dB), (g) SR model 2 (ISNR=3.5480dB), (h) SR model 3 (ISNR=4.1684dB).



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

Figure 3 - SR results for SNR=20dB (a) Segment of bicubically interpolated LR observation of middle frame, (b) SR model 1 (ISNR=1.9269dB), (c) SR model 2 (ISNR=2.2788dB), (d) SR model 3 (ISNR=2.8934dB).



(e) (f) (g) (h)

Figure 4 - SR results for SNR=30dB (e) Segment of bicubically interpolated LR observation of middle frame, (f) SR model 1 (ISNR=3.6800dB), (g) SR model 2 (ISNR=3.7974dB), (h) SR model 3 (ISNR=4.3088dB).

## 5. CONCLUSIONS

In this paper, we presented a MAP approach of *a new multichannel image prior* applied in the digital video SR problem, along with three proposed algorithms and their comparative study. These algorithms have been tested in different cases (in the presence or the absence of blur for different BSNR and SNR values respectively). The experimental results show in both cases that the algorithm which is based on the new proposed model (model 3) *performs better* than previous ones in terms of both *robustness with respect to the initial conditions and improvement in SNR (ISNR).* Moreover, the efficacy of this algorithm is further established by the fact that it provides the maximum gain with respect to the single frame SR algorithm (model1) at low BSNR / SNR values, which is also the case with respect to model 2 when no blur is used (resolution enhancement efficacy). Finally, the comparison between models 2 and 3 serves as a strong indication that the use of motion field in the prior term is much more efficient both in terms of restoration capability and resolution enhancement with respect to its use in the observation term (model).

## ACKNOWLEDGEMENT

This paper is part of the 03ED-535 research project, implemented within the framework of the "Reinforcement Programme of Human Research Manpower" (PENED) and co-financed by National and Community Funds (25% from the Greek Ministry of Development-General Secretariat of Research and Technology and 75% from E.U.-European Social Fund).

## REFERENCES

[1] A. K. Katsaggelos, R. Molina, and J. Mateos, *Super Resolution of Images and Video.* 1st ed. Ed. Morgan and Claypool, 2007, pp. 5-9, 29-30, 77-88.

[2] 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 Norte Dame, Norte Dame, IN 46556, USA, July 1998.

[3] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos, "Bayesian resolution enhancement of compressed video," *IEEE Transactions on Image Processing*, vol. 13, no. 7, pp. 898-910, 2004.

[4] C. Wang, P. Xue, and W. Lin, "Improved super-resolution reconstruction from video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 11, pp. 1411-1422, November 2006.

[5] M. K. Ng, H. Shen, E. Y. Lam, and L. Zhang, "A total variation regularization based super-resolution algorithm for digital video," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 74585, doi: 10.1155/2007/74585.

[6] S. Borman and R. Stevenson, "Simultaneous multi-frame MAP super-resolution video enhancement using spatio-temporal priors," in *Proc. ICIP 1999*, Kobe, Japan, October 24-28. 1999, pp. 469-473.

[7] N. P. Galatsanos and R. T. Chin, "Digital restoration of multichannel images," *IEEE transactions on Acoustics Speech and Signal Processing*, vol. 37, no. 3, pp. 415-421, March 1989.

[8] M. G. Choi, N. P. Galatsanos, and A. K. Katsaggelos, "Multichannel regularized iterative motion compensated restoration of image sequences," *Journal of Visual Communications and Image Representation*, vol. 7, no. 3, pp. 244-258, September 1996.

[9] M. G. Choi, Y. Yang, and N. P. Galatsanos, "Multichannel regularized recovery of compressed video sequences," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 48, no. 4, pp. 376-387, April 2001.

[10] A. K. Katsaggelos, J. Biemond, R. Mersereau, and R. Schaefer, "A regularized iterative restoration algorithm," *IEEE Trans. Signal Processing,* vol. 39, pp. 914-929, 1991.

[11] G. Chantas, N. Galatsanos, and A. Likas, "Bayesian restoration using a new non-stationary edge-preserving image prior," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 2987-2997, October 2006.