Variational approaches to super-resolution with contrast enhancement and anisotropic diffusion

Hyunwoo Kim
Ki-Sang Hong
POSTECH
Division of Electrical and Computer Engineering
Pohang 790-784
Korea

Abstract. *We present super-resolution methods that enhance image contrast and perform anisotropic diffusion simultaneously. Since we solve the super-resolution problem by encouraging Mach-band profiles while incorporating anisotropic diffusion, our technique not only reconstructs a high-resolution image from several overlapping noisy low-resolution images, but also enhances edges and image contrast while suppressing image noise during the restoration process. We apply this technique to a video stream, which can be aligned by 3×3 projective transformations.* © 2003 SPIE and IS&T.
[DOI: 10.1117/1.1556312]

1 Introduction

Image resolution and quality is limited to the characteristics of imaging sensors and/or image degradation due to lossy compression. Recently, to overcome resolution limitations due to imaging degradation, researchers have tried to reconstruct a high-resolution image from a collection of noisy low-resolution images, called *super-resolution*. Applications of super-resolution include the resolution improvement in remote sensing, enhancement of frame freeze in videos, overcoming resolution limitations in medical imaging, etc.

In Ref. 1, super-resolution algorithms are reviewed and classified into several approaches. After a comprehensive comparison, the authors conclude that two approaches are the most promising: the Bayesian [maximum a-posterior (MAP) and maximum likelihood (ML)] approach[2–6] and the set theoretic projection onto convex sets (POCS) methods,[7–9] in the sense that *a priori* information can be included in their formulas. More recently, hybrid methods are being developed.[5,10]

We focus our attention on the super-resolution of 2-D scenes, which can be approximated by a planar image (i.e., when videos are captured from planar objects or far away scenes, or obtained by rotating and zooming cameras). All frames from a 2-D scene can be aligned to a single frame.

This compact video representation is called an *image mosaic*. Since motion estimation in videos from 2-D scenes is easy and accurate compared to that of general videos and can be extended to general videos at the cost of difficult motion estimation, many algorithms have been developed and experimented with videos from 2-D scenes. Borman and Stevenson[1] mentioned that the super-resolution problem can be broken down into three major ones: 1. degradation modeling, 2. motion estimation, and 3. restoration algorithm.

Degradation modeling. The imaging process can be modeled by geometric warping, blurring, sampling, and uncorrelated additive Gaussian noises, which are added to the observed images.[10]

Motion estimation. The performance achievable using super-resolution restoration algorithms is dependent on accurate motion estimation and image registration. If frames of a video stream are well aligned by accurate motion estimation, only simple averaging of the aligned frames can construct a super-resolved image mosaic in the concept of image fusion.[11,12]

Restoration algorithm. To restore super-resolved images from videos from 2-D scenes, Irani and Peleg[13,14] introduced an iterative back-projection algorithm, which ensures convergence while suppressing spurious noise components in the solution, owing to the proper selection of a back-projection function. Mann and Picard[15] extended this to the projective case, and Zomet and Peleg[16] rendered the original implementation more efficient, and applied it to image mosaics. To preserve edge information while removing image noise, many algorithms have been developed. Schultz and Stevenson[5] proposed a Bayesian MAP estimator using a Huber–Markov random field model as an edge-preserving prior. Elad and Feuer[10,17] perform adaptive smoothing by giving a penalty to the regions with large gradients (i.e., edges). Capel and Zisserman[18] propose two estimators giving piecewise smooth results: a MAP estimator based on a Huber prior, and an estimator regularized by

using the total variation norm. Patti and Altunbasak[19] proposed a POCS algorithm to reduce ringing artifacts along the edges by preventing the inversion of a large blur in the direction orthogonal to the edges. Among these three major components, the development and improvement of restoration algorithms is focused on in this work.

We are interested in enhancing image contrast of the super-resolution solution during the restoration process. The goal of our contrast-enhancing super-resolution is to provide visually pleasing pictures with sharp and clear edge boundaries. This can be accomplished by generating over- and under-shoots, called *Mach-band profiles*, to noisy step edges. In this work, two contrast-enhancing algorithms are proposed. Method 1 incorporates a second-order smoothing term into a super-resolution functional. However, the approach fundamentally utilizes the super-resolution solution alone, not the whole low-resolution frames. To obtain feedback from all low-resolution frames, Method 2 uses the residuals between the Gaussian-convolved image of the super-resolution solution and each low-resolution image in each iterative optimization step. Compared with the second-order smoothing regularization, the method not only enhances image contrast by the Mach-band effect as before, but also provides better recovery of degraded details.

Moreover, we combine the contrast-enhancing functionals with anisotropic diffusion terms. Generally, the super-resolution algorithm can be performed by cooperating proper regularization terms. As we know, most previous works including the previous restoration algorithms can be categorized into edge-preserving regularization. However, when the target images include edges/lines, they can be easily aliased in low-resolution images through the imaging process. Therefore, to reduce the aliasing artifact, we need to *enhance* the structures while *preserving* them and reducing image noise. In this work, we introduce edge-enhancing regularization by incorporating anisotropic diffusion.

Degradation modeling is explained in Sec. 2, and motion estimation is explained in Sec. 3. In Sec. 4, a basic super-resolution restoration formulation and its solution, called the Irani-Peleg estimator, are reviewed. In Sec. 5, our proposed variational algorithms are presented. Experimental results are given in Sec. 6, and concluding remarks are given in Sec. 7.

2 Degradation Modeling

The imaging process in 2-D videos can be modeled by geometric warping, blurring, sampling, and uncorrelated additive Gaussian noise added to the observed images. Elad and Feuer[10] introduced a matrix-vector formulation. Given N low-resolution images $\vec{y}_1, \ldots, \vec{y}_N$, the imaging process of $\vec{y}_k$ from the super-resolved image $\vec{x}$ can be formulated by

$$\vec{y}_k = \mathbf{D}_k \mathbf{C}_k \mathbf{F}_k \vec{x} + \vec{e}_k,\tag{1}$$

where $\vec{x}$ denotes a high-resolution image of size $[L \times L]$, reordered in a vector of size $[L^2]$. $\vec{y}_k$ and $\vec{e}_k$ denote the k'th low-resolution input image and the corresponding normally

distributed additive noise of size $[M_k \times M_k]$, reordered in a vector of size $[M_k^2]$, respectively. $\mathbf{F}_k$, $\mathbf{C}_k$, and $\mathbf{D}_k$ are the geometric warp matrix, the blurring matrix, and the decimation matrix, respectively. By considering all the frames and by stacking the vector equations, we obtain a matrix-vector formula:

$$\begin{bmatrix} \vec{y}_1 \\ \vdots \\ \vec{y}_N \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1 \mathbf{C}_1 \mathbf{F}_1 \\ \vdots \\ \mathbf{D}_N \mathbf{C}_N \mathbf{F}_N \end{bmatrix} \vec{x} + \begin{bmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_N \end{bmatrix},\tag{2}$$

which can be also represented by

$$\vec{y} = \mathbf{A}\vec{x} + \vec{e}.\tag{3}$$

3 Motion Estimation

We[20] introduced a robust, accurate global registration algorithm in the presence of moving objects with complex actions. When pairwise local registrations are performed, severe global misalignment may result: errors may accumulate during the process of recovering projective transformations between two pairs of consecutive frames and concatenating the transformations to align frames into an image mosaic. To overcome the problem, a robust global registration technique is adopted, which uses a graph to represent temporal and spatial connectivity among frames in an image sequence. The framework presented allows the automatic construction of the graph and the construction of a consistent mosaic from a collection of image sequences, which works well even in the presence of moving objects.

As a result, the algorithm gives accurate geometric warping matrices $\{\mathbf{F}_1, \cdots, \mathbf{F}_N\}$. In addition, at a position in each frame, an alignment measure (i.e., the error measure of the motion estimation) can be measured as a by-product. When low-resolution frames are overlapping at a pixel position (x, y) of the super-resolution image, the alignment measure can be represented by $W_i(x,y)$ in the i'th frame and calculated as follows. One can observe that a well-alignment frame has similar gray-level intensities with a reference frame at the overlapping pixel position. The similarity of gray-level values can be measured by calculating (normalized) cross-correlation between a frame and a reference frame, which ranges from 0 to 1. In the restoration procedure, the alignment measure is incorporated into our proposed framework and it weights well-aligned frames. When $W_j(x,y) = 1$, the corresponding pixel at the j'th frame is fully considered for super-resolution, whereas $W_k(x,y) < a\ specific\ threshold$, the corresponding pixel at the k'th frame is ignored.

4 Restoration Algorithm

Assuming that the noise process is uncorrelated and has a uniform variance in all observed images in Eq. (3), the maximum likelihood solution of a high-resolution image, $\vec{x}$, can be estimated by minimizing the functional:

$$E(\vec{x}) = \tfrac{1}{2}\|\vec{y} - \mathbf{A}\vec{x}\|^2.\tag{4}$$

The minimum of the functional occurs where the gradient of the functional is zero ($\nabla E = 0$). It is represented by

$$\nabla E = 0 \Leftrightarrow \mathbf{A}^T(\mathbf{A}\vec{\mathbf{x}} - \vec{\mathbf{y}}) = 0$$

$$\Leftrightarrow \sum_{k=1}^{N} \mathbf{F}_k^T \mathbf{C}_k^T \mathbf{D}_k^T (\mathbf{D}_k \mathbf{C}_k \mathbf{F}_k \vec{\mathbf{x}} - \vec{\mathbf{y}}_k) = 0. \tag{5}$$

Zomet and Peleg[16] mentioned that multiplication with $\mathbf{A}$ and $\mathbf{A}^T\mathbf{A}$ can be implemented using only image operations, such as warp, blur, and sampling. The matrix $\mathbf{A}^T\mathbf{A}$ operating on the vector $\vec{\mathbf{x}}$, and the matrix $\mathbf{A}^T$ operating on the vector $\vec{\mathbf{y}}$, can be represented by the following image operations.

$$\mathbf{A}^T\mathbf{A}\vec{\mathbf{x}} = \sum_{k=1}^{N} \mathbf{F}_k^T \mathbf{C}_k^T \mathbf{D}_k^T \mathbf{D}_k \mathbf{C}_k \mathbf{F}_k \vec{\mathbf{x}},$$

$$\mathbf{A}^T\vec{\mathbf{y}} = \sum_{k=1}^{N} \mathbf{F}_k^T \mathbf{C}_k^T \mathbf{D}_k^T \vec{\mathbf{y}}_k, \tag{6}$$

where $\mathbf{F}_k$, $\mathbf{C}_k$, and $\mathbf{D}_k$ are implemented by image warping, blurring, and subsampling, respectively. The matrix $\mathbf{D}_k^T$ corresponds to upsampling the image. The matrix $\mathbf{C}_k^T$ is implemented by convolution with the flipped kernel for a convolution blur, and $\mathbf{F}_k^T$ is forward warping of the inverse motion, if $\mathbf{F}_k$ represents backward warping.

Finally, to solve Eq. (3), Richardson iterations are used with an iteration step:

$$\vec{\mathbf{x}}^{(n+1)} = \vec{\mathbf{x}}^{(n)} - \nabla E = \vec{\mathbf{x}}^{(n)} + \sum_{k=1}^{N} \mathbf{F}_k^T \mathbf{C}_k^T \mathbf{D}_k^T (\vec{\mathbf{y}}_k - \mathbf{D}_k \mathbf{C}_k \mathbf{F}_k \vec{\mathbf{x}}), \tag{7}$$

where $\vec{\mathbf{x}}^{(n)}$ denotes the estimate at the n-th iteration with $\vec{\mathbf{x}}^{(0)}$ corresponding to the average value of all aligned images in the super-resolution domain. This algorithm is called the *iterative backprojection algorithm* or the *Irani-Peleg estimator*.[13,14]

5 Variational Methods for Super-Resolution

To enhance image contrast while suppressing image noise during the restoration process, two different variational formulations are given in this section. In Method 1, we introduce a super-resolution formulation by adopting second-order smoothing (Sec. 5.1). In Method 2, we propose an improved formulation by considering the residuals between the Gaussian-convolved image of the super-resolution solution and each low-resolution image (Sec. 5.2).

5.1 Contrast-Enhancing Super-Resolution: Method 1

We modify the energy functional in Eq. (4) by using an image representation instead of a matrix-vector representation. It can be written, in a continuous form, by

$$E(X) = \int \int \frac{1}{N} \sum_{k=1}^{N} \{Y_k(x,y) - a_k(X)(x,y)\}^2 dx dy, \tag{8}$$

where $a_k(X)$ is the transformed image of the super-resolution image X, with the image operation corresponding

to $\mathbf{A}_k (= \mathbf{D}_k \mathbf{C}_k \mathbf{F}_k)$, and Y_k is the k-th input image; and $N \equiv N(x,y)$ is the number of overlapping frames at the position (x, y). For simplification, we adopt a notation $R_k(X,Y_k)$ representing the residual; that is,

$$R_k(X,Y_k) \equiv Y_k(x,y) - a_k(X)(x,y), \tag{9}$$

and Eq. (8) is rewritten by

$$E(X) = \int \int \frac{1}{N} \sum_{k=1}^{N} R_k^2(X,Y_k) dx dy. \tag{10}$$

In Ref. 21, second-order smoothing terms are used to enhance image contrast of a single frame by generating over- and under-shoots, and we incorporate the regularization term in the super-resolution problem to enhance image contrast of the reconstructed image. After incorporating second-order smoothing terms, the super-resolution functional is extended to

$$E(X) = \int \int \frac{1}{N} \sum_{k=1}^{N} R_k^2(X,Y_k) dx dy + \lambda_1 \int \int \left(\frac{\partial^2 X}{\partial x^2}\right)^2$$

$$+ 2\left(\frac{\partial^2 X}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 X}{\partial y^2}\right)^2 dx dy. \tag{11}$$

Because the Euler–Lagrange equation corresponding to this functional is a fourth order differential equation, the evolution is very noise sensitive. Therefore, we introduce two new functions, β_x and β_y:

$$\beta_x \approx \frac{\partial X}{\partial x} \quad \text{and} \quad \beta_y \approx \frac{\partial X}{\partial y}. \tag{12}$$

Combining the new functions, we obtain a new functional:

$$E(X, \beta_x, \beta_y) = \int \int \frac{1}{N} \sum_{k=1}^{N} R_k^2(X,Y_k) dx dy$$

$$+ \lambda_1 \int \int \left(\frac{\partial \beta_x}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial \beta_x}{\partial y} + \frac{\partial \beta_y}{\partial x}\right)^2$$

$$+ \left(\frac{\partial \beta_y}{\partial y}\right)^2 dx dy + \lambda_2 \int \int \left(\beta_x - \frac{\partial X}{\partial x}\right)^2$$

$$+ \left(\beta_y - \frac{\partial X}{\partial y}\right)^2 dx dy. \tag{13}$$

It can be also rewritten by weighting the residual, $R_k(X,Y_k) \equiv Y_k(x,y) - a_k(X)(x,y)$, with the corresponding alignment measure, $W_k(x,y)$, as follows.

$$E(X) = \int \int \frac{1}{M} \sum_{k=1}^{M} W_k^2 R_k^2(X, Y_k) dx dy + \lambda_1 \int \int \left(\frac{\partial \beta_x}{\partial x}\right)^2$$

$$+ \frac{1}{2}\left(\frac{\partial \beta_x}{\partial y} + \frac{\partial \beta_y}{\partial x}\right)^2 + \left(\frac{\partial \beta_y}{\partial y}\right)^2 dx dy$$

$$+ \lambda_2 \int \int \left(\beta_x - \frac{\partial X}{\partial x}\right)^2 + \left(\beta_y - \frac{\partial X}{\partial y}\right)^2 dx dy, \quad (14)$$

where M is the number of overlapping frames after purging misaligned frames with poor alignment measure, and W_k denotes $W_k(x,y)$. [Note that $M \approx \Sigma_k W_k(x,y)^2$, because $W_k(x,y) \approx 1$, for well-aligned frames.]

Using the calculus of variations and the steepest descent method, we have corresponding coupled partial differential equations (PDEs), represented by

$$\frac{\partial X}{\partial t} = \frac{1}{M} \sum_{k=1}^{M} W_k^2 a_k^T[R_k(X, Y_k)] + \lambda \left(\frac{\partial^2 X}{\partial x^2} + \frac{\partial^2 X}{\partial y^2}\right)$$

$$- \lambda \left(\frac{\partial \beta_x}{\partial x} + \frac{\partial \beta_y}{\partial y}\right), \quad (15)$$

$$\frac{\partial \beta_x}{\partial t} = 2 \frac{\partial^2 \beta_x}{\partial x^2} + \frac{\partial^2 \beta_x}{\partial y^2} + \frac{\partial^2 \beta_y}{\partial x \partial y} - 2\xi\left(\beta_x - \frac{\partial X}{\partial x}\right), \quad (16)$$

$$\frac{\partial \beta_y}{\partial t} = 2 \frac{\partial^2 \beta_y}{\partial y^2} + \frac{\partial^2 \beta_y}{\partial x^2} + \frac{\partial^2 \beta_x}{\partial x \partial y} - 2\xi\left(\beta_y - \frac{\partial X}{\partial y}\right), \quad (17)$$

where

- $\lambda \equiv \lambda_2$,
- $\xi \equiv \lambda_2 / \lambda_1$,
- $a_k^T[R_k(X, Y_k)] \equiv a_k^T(Y_k) - a_k^T a_k(X)$,
- $a_k^T a_k(X)$ is the transformed image of the super-resolution image X, with the image operations corresponding to the matrix $\mathbf{A}_k^T \mathbf{A}_k$,
- $a_k^T(Y_k)$ is the transformed image of Y_k, with the image operations corresponding to the matrix $\mathbf{A}_k^T$.

To enhance image contrast, we make two modifications to these equations. First, to counteract the smearing out due to the second isotropic diffusion term ($\partial^2 X / \partial x^2 + \partial^2 X / \partial y^2$) and to enhance image contrast by sharpening and smoothing edges, anisotropic diffusion terms are introduced (i.e., $\nabla \cdot \mathbf{D} \nabla X$, where $\mathbf{D}$ denotes a 2×2 matrix derived from the gradient of X, called the *diffusion tensor*). Second, to control Mach-band effects for contrast enhancement, we employ a shooting parameter s_h. Finally, we have a contrast-enhancing version of the equation.

$$\frac{\partial X}{\partial t} = \frac{1}{M} \sum_{k=1}^{M} W_k^2 a_k^T[R_k(X, y_k)] + \lambda \nabla \cdot \mathbf{D} \nabla X$$

$$- s_h\left(\frac{\partial \beta_x}{\partial x} + \frac{\partial \beta_y}{\partial y}\right), \quad (18)$$

with Eqs. (16) and (17). The selection of the diffusion tensor determines the regularization property, which includes *a priori* knowledge in the formulation. The first term restores a super-resolution solution from the low-resolution images, the second term $\nabla \cdot \mathbf{D} \nabla X$ corresponds to anisotropic diffusion, and the third term ($\partial \beta_x / \partial x + \partial \beta_y / \partial y$) generates under- and over-shoots resulting in contrast enhancement. The parameter λ determines how much anisotropic diffusion contributes during the restoration process. Actually, both the second and third term contribute contrast enhancement simultaneously, and they are related to the super-resolution solution X alone, not all multiframe observations $\{Y_1, \dots, Y_N\}$ during the restoration process.

5.2 Contrast-Enhancing Super-Resolution: Method 2

While Method 1 encourages contrast enhancement on the super-resolution restoration alone, Method 2 aims to encourage contrast enhancement dependent on each of the low-resolution frames. To obtain feedback from all of the low-resolution frames, we modify the residual $R_k(X, Y_k)$ directly while turning off the third term by setting $s_h = 0$. It is achieved by computing the residual between pairs of the Gaussian-convolved image ($K_{\sigma_{sh}} * X$) of X with scale σ_{sh} and each low-resolution image Y_k. That is, the residual $R_k(X, Y_k) = [Y_k(x,y) - a_k(X)(x,y)]$ is changed to $R_k(X, Y_k, \sigma_{sh}) = [Y_k(x,y) - a_k(K_{\sigma_{sh}} * X)(x,y)]$. When we change the residual and set $s_h = 0$, we finally obtain a new contrast-enhancing functional utilizing multiframe images. It is represented by

$$\frac{\partial X}{\partial t} = \frac{1}{M} \sum_{k=1}^{M} W_k^2 a_k^T[R_k(X, Y_k, \sigma_{sh})] + \lambda \nabla \cdot \mathbf{D} \nabla X. \quad (19)$$

Note that it exaggerates the residual around noisy step edges compared with those in smooth regions and generates over-and under-shoots (i.e., Mach-band effect).

5.3 Incorporating Edge-Preserving and Edge-Enhancing Diffusion

Among many edge-preserving regularization techniques (for a review, see Refs. 22 and 23), we incorporate Perona-Malik regularization[24] into the contrast-enhancing methods described by Eqs. (18) and (19). Substituting $\mathbf{D}$ to $\mathbf{D}_{PM}$, defined by

$$\mathbf{D}_{PM} \equiv D(|\nabla X_\sigma|^2) = 1/(1 + |\nabla X_\sigma|^2/K^2), \quad (20)$$

edge-preserving properties are achieved by the super-resolution formulations. In this equation, ∇X_σ denotes the gradient of the Gaussian-convolved image ($K_\sigma * X$) with scale σ, and it is insensitive to structures with scales smaller than σ. K is a constant controlling diffusivity (for details, see Ref. 24). In regions containing edges or discontinuities, since $D(|\nabla X_\sigma|^2)$ is close to 0, $\mathbf{D}_{PM} \nabla X_\sigma \approx 0$ and this means the regularization is off and smoothing is prevented in edges. In smooth regions, since $D(|\nabla X_\sigma|^2)$ is close to 1, $\mathbf{D}_{PM} \nabla X_\sigma \approx \nabla X_\sigma$ and the regions are smoothed. Although this edge-preserving regularization eliminates

spurious noise within a smooth region very well, in the vicinity of noisy edges (e.g., aliased edges/lines), we need to introduce edge-enhancing regularization to enhance noisy edges/lines.

To enhance edge structures, the diffusion term takes into account not only the contrast of an edge, but also its direction during the evolution process.[24,25] This can be achieved by using an *edge-enhancing* regularization $\mathbf{D} \equiv \mathbf{D}_{EE}$, defined by

$$
\mathbf{D}_{EE} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} D(|\nabla X_\sigma|^2) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix}
$$
$$
= D(|\nabla X_\sigma|^2) \mathbf{v}_1 \mathbf{v}_1^T + \mathbf{v}_2 \mathbf{v}_2^T \quad (21)
$$

such that $\mathbf{v}_1 \| \nabla X_\sigma$ and $\mathbf{v}_2 \perp \nabla X_\sigma$. In smooth regions, since $D(|\nabla X_\sigma|^2)$ is close to 1, $\mathbf{D}\nabla X \approx \{\mathbf{v}_1 \mathbf{v}_1^T + \mathbf{v}_2 \mathbf{v}_2^T\} \nabla X$ and it means the regions are smoothed in all directions. In edge regions, since $D(|\nabla X_\sigma|^2)$ is close to 0, $\mathbf{D}_{EE} \nabla X \approx \mathbf{v}_2 \mathbf{v}_2^T \nabla X$. Therefore, edges are smoothed only in the edge direction $\mathbf{v}_2$, which results in edge enhancement in the restoration process. (Notice that $\mathbf{v}_1$ and $\mathbf{v}_2$ are perpendicular and parallel to an edge, respectively.)

5.4 How to Solve the PDE

The differential equation for nonlinear diffusion cannot be solved analytically, so numerical approximations are used. Mainly finite difference (FD) methods are used, because they are easy to implement and a digital image provides information on a regular rectangular pixel grid.[26] First, in an image domain, FD methods replace all derivatives on an image by finite differences between neighboring values. Next, for the iteration step, we set the time step size τ and the solution is updated by

$$
X^{(n+1)} = X^{(n)} + \tau \frac{\partial X}{\partial t}, \quad (22)
$$

$$
\beta_x^{(n+1)} = \beta_x^{(n)} + \tau \frac{\partial \beta_x}{\partial t}, \quad (23)
$$

$$
\beta_y^{(n+1)} = \beta_y^{(n)} + \tau \frac{\partial \beta_y}{\partial t}, \quad (24)
$$

for Eq. (18), and

$$
X^{(n+1)} = X^{(n)} + \tau \frac{\partial X}{\partial t}, \quad (25)
$$

for Eq. (19), where $X^{(n)}$, $\beta_x^{(n)}$, $\beta_y^{(n)}$ denote the estimate at the nth iteration, and $X^{(0)}$ is set to be the average value of the aligned low-resolution images in the domain of the super-resolution solution X.

6 Experimental Results

The first three experiments are applied to simulated images and the others are applied to real video streams.
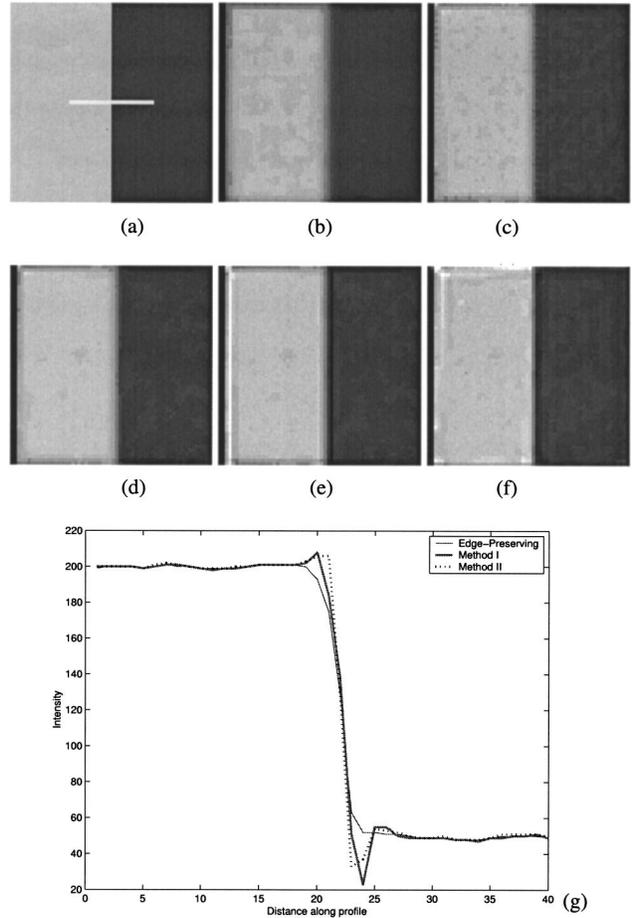
Fig. 1 Simulation result 1 (2×2 enlargement). (a) step function and the profile line, (b) average blending, (c) Irani-Peleg estimator, (d) edge-preserving method ($\sigma=0.5$, $K=3.0$, $\lambda=0.2$, $s_h=0.0$), (e) contrast-enhancing Method 1 with edge-preserving diffusion ($\sigma=0.5$, $K=3.0$, $\lambda=0.2$, $s_h=0.15$), (f) contrast-enhancing Method 2 with edge-preserving diffusion ($\sigma=0.5$, $K=3.0$, $\lambda=0.2$, $\sigma_{s_h}=1.2$), and (g) profile views of the super-resolution results.

6.1 Simulations

In the first simulation, the response to a step function is presented and an example image of the step function is shown in Fig. 1(a). To obtain low-resolution images from the step image, the original image is first warped with transformation parameters ($\mathbf{F}_k$) with ten different rotations and translations, smoothed with a Gaussian filter with scale 0.8 ($\mathbf{C}_k$), 2×2 downsampled ($\mathbf{D}_k$), and added Gaussian noise with standard deviation 5.0 gray level ($\vec{\mathbf{e}}_k$). As a result, we obtain ten low-resolution images. In addition, we simulate a misalignment error by generating uniform random deviation in the horizontal direction ranging from −0.5 to 0.5 pixels in low-resolution image alignment. That is, F_k deviates from the known groundtruth by the translational motion in the horizontal direction. In Figs. 1(b) through 1(f), the results of average blending, the Irani-Peleg estimator, the edge-preserving method, Method 1, and Method 2 are shown. In Methods 1 and 2, the edge-preserving diffusion is used as an anisotropic regularization. The Irani-Peleg estimator in Fig. 1(c) enhances the degraded slope of the noisy step compared with average
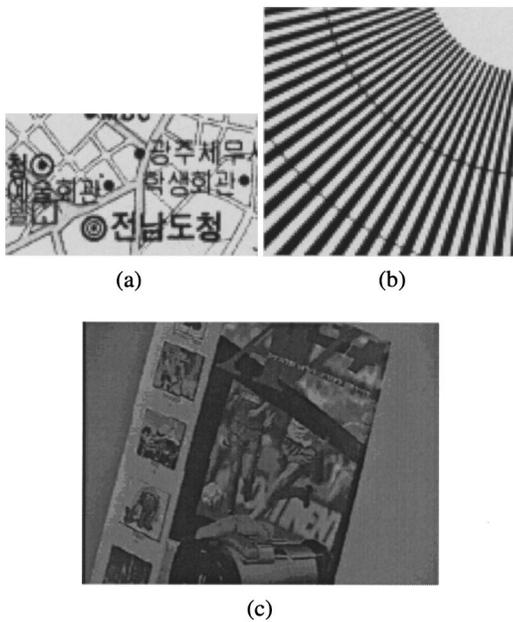
Fig. 2 Example images: (a) a part of a map image (110×62), (b) a part of a chart image (100×100), and (c) a reference frame from 8 frames (240×180).

Fig. 4 Simulation result (2×2 enlargement): (a) bicubic interpolation of a low-resolution frame, (b) average blending, (c) Irani-Peleg estimator, (d) contrast-enhancing Method 2 with edge-preserving diffusion ($\sigma = 0.5$, $K = 3.0$, $\lambda = 0.2$, $\sigma_{s_h} = 1.2$), (e) contrast-enhancing Method 2 with edge-enhancing diffusion ($\sigma = 0.5$, $K = 3.0$, $\lambda = 0.2$, $\sigma_{s_h} = 1.2$), and (f) Irani-Peleg estimator followed by unsharp masking.

blending in Fig. 1(b). An edge-preserving method in Fig. 1(d) removes image noise in the smooth region, and contrast-enhancing methods in Figs. 1(e) and 1(f) result in a steeper response than others. In Fig. 1(g) the difference between the two contrast-enhancing methods are compared by showing their profile views. We can see that, in the similar over- and under-shoots, Method 2 gives a steeper response than Method 1.

In the second simulation, the map image in Fig. 2(a) is used as an original high-resolution image. Using the same imaging process, we obtain 20 simulated low-resolution images. However, there is no warping error in this simulation because the misalignment is not added, and the known warping parameters are used for our restoration process. Figures 3(a)–3(f) show the super-resolution images mentioned in the caption. We use edge-preserving regulariza-
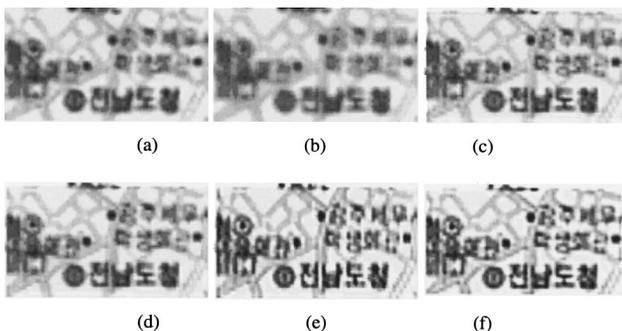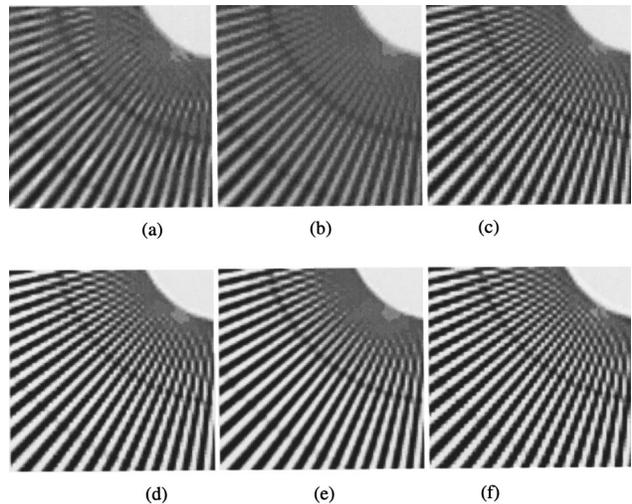

Fig. 3 Simulation result 2 (2×2 enlargement): (a) bicubic interpolation, (b) average blending, (c) Irani-Peleg estimator, (d) edge-preserving method ($\sigma = 0.5$, $K = 3.0$, $\lambda = 0.2$, $s_h = 0.0$), (e) contrast-enhancing Method 1 with edge-preserving diffusion ($\sigma = 0.5$, $K = 3.0$, $\lambda = 0.2$, $s_h = 0.25$), and (f) contrast-enhancing Method 2 with edge-preserving diffusion ($\sigma = 0.5$, $K = 3.0$, $\lambda = 0.2$, $\sigma_{s_h} = 1.2$).

tion as the anisotropic term. We can see that the results by Methods 1 and 2 are superior, owing to their contrast-enhancing property. Compared with the original image, Method 2 gives better restoration results of the details of the location circles in the map than Method 1. Note that there is a more rounded circle.

In the last simulation, the chart image in Fig. 2(b) is used as an original high-resolution image to present the edge-enhancing property of our methods. We obtain 20 simulated low-resolution images using the simulated imaging. Figures 4(a)–4(f) show the super-resolved images mentioned in the caption. In the diagonal lines of the fan-shaped pattern, one observes that the recovery of the proposed edge-enhancing method [Fig. 4(e)] is better than others as compared with the original image. The edge-preserving super-resolution results in aliased diagonal lines, because the lines are aliased during the simulated imaging process and the algorithm is inherently based on an edge-preserving regularization. Also, in the results of the Irani-Peleg estimator [Figs. 4(c) and 4(f)], the diagonal lines are more aliased than the edge-preserving regularization. However, our edge-enhancing method overcomes edge aliasing by smoothing line boundaries (i.e., edge enhancing).

6.2 Real Videos

Eight images were captured using a hand-held camera. Figure 2(c) shows a reference frame from the images. We can see that the boundaries are aliased and the straight lines are distorted during the imaging process. Using the motion estimation method mentioned in Sec. 3, the other frames are aligned to the reference frame and several super-resolution algorithms are applied to those images. In this experiment, the magnification ratio is 2×2. Figures 5(a)–5(h) show super-resolution images, obtained by several methods as mentioned in the caption. The bicubic interpolated image [Fig. 5(a)] shows that single frame interpolation cannot en-
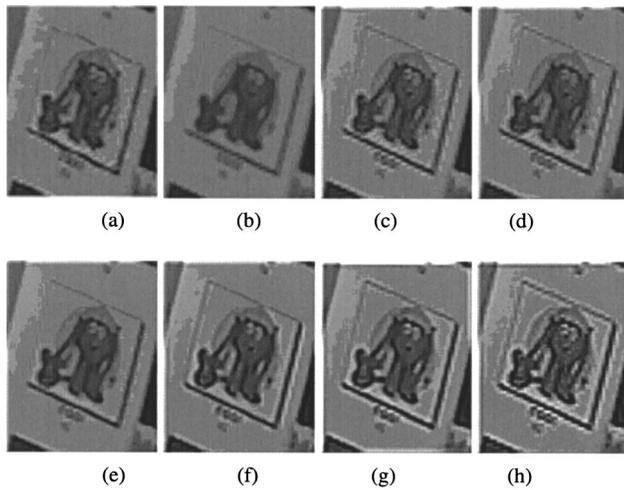
Fig. 5 Experimental result (2×2 enlargement): (a) bicubic interpolation, (b) average blending, (c) Irani-Peleg estimator, (d) edge-preserving method ($\sigma=0.5$, $K=3.0$, $\lambda=0.2$, $K_{\sigma_{sh}}*X=X$), (e) edge-enhancing method ($\sigma=0.5$, $K=3.0$, $\lambda=0.2$, $s_h=0.0$), (f) contrast-enhancing Method 1 with edge-enhancing diffusion ($\sigma=0.5$, $K=3.0$, $\lambda=0.2$, $s_h=0.25$), (g) contrast-enhancing Method 2 with edge-enhancing diffusion ($\sigma=0.5$, $K=3.0$, $\lambda=0.2$, $\sigma_{s_h}=1.2$), and (h) Irani-Peleg estimator followed by unsharp masking.

Fig. 6 Super-resolution of an MPEG video (2×2 enlargement): (a) four frames from image sequence, (b) interpolation followed by intensity averaging, and (c) super-resolved image.

hance degraded details and aliased boundaries due to sensor resolution. The simple fusion by averaging aligned interpolated frames is blurred due to motion estimation error and smoothing by interpolation [Fig. 5(b)]. Figure 5(c) shows the result of the Irani-Peleg estimator. The method can fuse all multiple frames well, but the edges are still aliased with a ringing effect around the edges. To overcome the ringing artifact, the edge-enhancing method is used. The method can reduce the ringing artifact, but image details are smoothed. The method is compared to the edge-preserving method in Figs. 5(d) and 5(e). To compensate for the smoothing, contrast-enhancing algorithms with edge-enhancing regularization are applied. The two methods give improved results from the viewpoint of contrast enhancing and artifact reduction, as shown in Figs. 5(f) and 5(g). One can see that Method 2 reconstructs degraded details well and results in clearer boundaries than Method 1. In addition, Fig. 5(h) shows the result of the Irani-Peleg estimator followed by unsharp masking. The postprocessing can enhance image contrast, but it also enhances image noise.

We applied our super-resolution algorithm to image mosaics.[16] We particularly applied it to image mosaics of dynamic scenes, because the alignment measure $W_k(x,y)$ removes moving objects during the super-resolution process. Figure 6(a) shows four selected frames from 24 observed frames. Figure 6(b) shows the results of bilinear interpolation followed by intensity averaging, and Fig. 6(c) is the result of Method 2 with edge-preserving diffusion. In the circled region, the result of our super-resolution method gives clearer and sharper boundaries than average blending.

Regarding computational complexity, it does highly depend on the number of iterations, since super-resolution algorithms have iterative optimization routines. In terms of the number of iterations, our algorithm iterates approximately five times more than the Irani-Peleg algorithm, be-
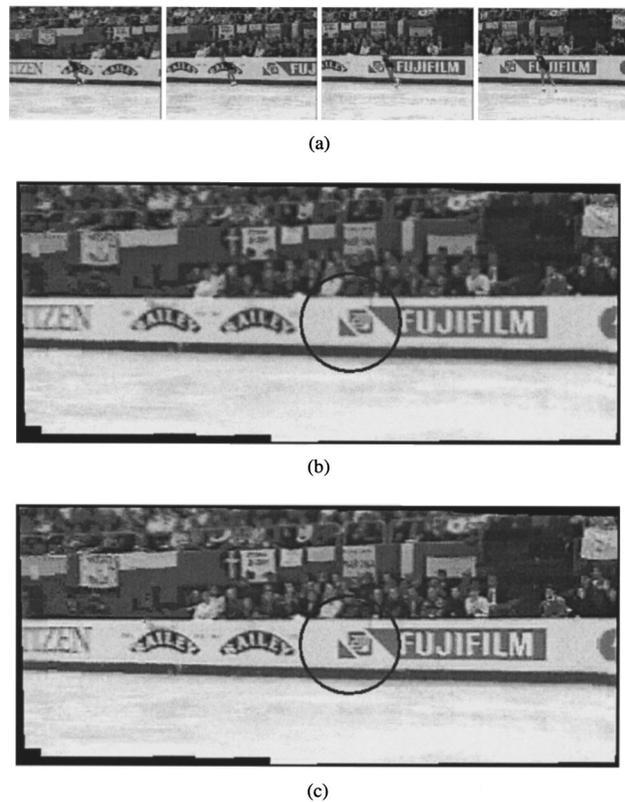
cause our algorithm needs smaller step size to guarantee the convergence of our PDE formulation.

7 Concluding Remarks

We deal with the problem of producing a super-resolution mosaic image from low-resolution video frames. Toward this goal, we have proposed two contrast-enhancing super-resolution algorithms, Method 1 and Method 2, combined with anisotropic diffusion using variational approaches. The proposed algorithms were found useful in the presence of noisy and/or aliased edges. Our simulations and real experiments show that our proposed algorithms can enhance image contrast while recovering degraded details due to sensor resolution limitation. Moreover, Method 2 is superior to the Method 1, because Method 2 appears to give a sharper step response and restore degraded details well, compared with the original image. In addition, we have incorporated edge-enhancing regularizations in our framework, so that noisy and/or aliased edges can be enhanced while reducing image noise. Experiments show that the proposed algorithms can produce super-resolution images with quality much better than the existing algorithms.

8 Appendix: Anisotropic Nonlinear Diffusion

Diffusion is a physical process that equilibrates concentration differences without creating or destroying mass. The diffusion process is expressed by the diffusion equation:

$$\frac{\partial X}{\partial t} = \nabla \cdot \mathbf{D} \nabla X, \tag{26}$$

where X denotes concentration and $\mathbf{D}$ denotes the diffusion tensor, a positive-definite symmetric matrix. In image processing, nonlinear diffusion filters regard the original input image as the initial state of a diffusion process that adapts itself to the evolving image. The fact that nonlinear adaptation may enhance interesting structures, such as edges, relates them to image enhancement and image restoration methods.[26] The design of the nonlinear diffusion filters may be reduced to design the diffusion tensor $\mathbf{D}$, and the tensor is closely related to the differential structure of the evolving image. We are particularly interested in anisotropic cases where $\mathbf{D} \nabla X$ and ∇X are not parallel.

Acknowledgments

References

1. S. Borman and R. L. Stevenson, "Super-resolution from image sequences—A review," *Proc. 1998 Midwest Symp. Circuits and Systems*, Notre Dame, IN (1998).
2. P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-resolved surface reconstruction from multiple images," in *Maximum Entropy and Bayesian Methods*, pp. 293–308, Kluwer, Santa Barbara, CA (1996).
3. F. Dellaert, S. Thrun, and C. Thrope, "Jacobian images of super-resolved texture maps for model based motion estimation and tracking," *Proc. Workshop Appl. Computer Vision* (Session 1A) (1998).
4. 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. Image Process.* **6**(12), 1621–1633 (1997).
5. R. R. Schultz and R. R. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Process.* **5**(6), 996–1011 (June 1996).
6. B. C. Tom and A. K. Katsaggelos, "Reconstruction of a high resolution image from multiple degraded mis-registered low resolution images," *Proc. SPIE* **2308**, 971–981 (1994).
7. P. E. Eren, M. I. Sezan, and A. Tekalp, "Robust, object-based high-resolution image reconstruction from low-resolution video," *IEEE Trans. Image Process.* **6**(10), 1446–1451 (1997).
8. A. J. Patti, M. I. Sezan, and A. M. Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Trans. Image Process.* **6**(8), 1064–1076 (Aug. 1997).
9. B. C. Tom and A. K. Katsaggelos, "An iterative algorithm for improving the resolution of video sequences," *Proc. SPIE* **2727**, 1430–1438 (1996).
10. M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Trans. Image Process.* **6**(12), 1646–1658 (1997).
11. M. C. Chiang and T. E. Boult, "Efficient super-resolution via image warping," *Image Vis. Comput.* **18**(10), 761–771 (2000).
12. R. A. Smith, A. W. Fitzgibbon, and A. Zisserman, "Improving augmented reality using image and scene constraints," in *Proc. British Machine Vision Conf.*, pp. 295–304, Nottingham, (1999).
13. M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graph. Models Image Process.* **53**, 231–239 (1991).
14. M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion, and transparency," *J. Visual Commun. Image Represent* **4**, 324–335 (1993).
15. S. Mann and R. W. Picard, "Virtual bellows: Constructing high quality stills from video," *Proc. IEEE Intl. Conf. Image Process.* (1994).
16. A. Zomet and S. Peleg, "Efficient super-resolution and applications to mosaics," *Proc. Intl. Conf. Patt. Recog.*, pp. 579–583 (2000).
17. M. Elad and A. Feuer, "Superresolution restoration of an image sequence: Adaptive filtering approach," *IEEE Trans. Image Process.* **8**(3), 387–395 (1999).
18. D. Capel and A. Zisserman, "Super-resolution enhancement of text image sequneces," *Proc. Intl. Conf. Patt. Recog.*, pp. 600–605 (2000).
19. A. J. Patti and Y. Altunbasak, "Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants," *IEEE Trans. Image Process.* **10**(1), 179–186 (2001).
20. H. Kim, I. Cohen, and K. S. Hong, "Robust global registration in the presence of moving objects," POSTECH Technical Report 2000-01 (2000).
21. M. Proesmans, E. Pauwels, and L. van Gool, "Coupled geometry-driven diffusion equations for low-level vision," in *Geometry-Driven Diffusion in Computer Vision*, M. Bart and H. Romeny, eds., pp. 191–228, 1994.
22. S. Teboul, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Variational approach for edge-preserving regularization using coupled PDEs," *IEEE Trans. Image Process.* **7**(3), 387–397 (1995).
23. S. Z. Li, "On discontinuity-adaptive smoothness priors in computer vision," *PAMI* **17**(6), 576–586 (1995).
24. J. Weickert, "Theoretical foundations of anisotropic diffusion in image processing," in *Theoretical Foundations of Computer Vision, Computing Suppl. 11*, W. Kropatsch, R. Klette, and F. Solinar, Eds., pp. 221–236 (1996).
25. H. Kim, J. H. Jang, and K. S. Hong, "Edge-enhancing super resolution using anisotropic diffusion," *Proc. IEEE Intl. Conf. Image Process.* pp. 130–133, Thessaloniki, Greece, Oct., 2001.
26. J. Weickert, "Nonlinear Diffusion Filtering," Chap. 15, *Handbook of Computer Vision and Applications*, Academic Press, New York (1999).
27. J. Weickert, "Coherence-enhancing diffusion filtering," *Int. J. Comput. Vis.* **31**(2/3), 111–127 (1999).
28. D. Capel and A. Zisserman, "Automated mosaicing with super-resolution zoom," *Proc. Intl. Conf. Computer Vision Patt. Recog.*, pp. 885–891 (1998).

Hyunwoo Kim received his BS degree in electronic communication engineering from Hangyang University, Seoul, Korea, in 1994, and his MS and PhD degrees in electrical and computer engineering from POSTECH, Pohang, Korea, in 1996 and 2001, respectively. In 2000 he worked in the Institute for Robotics and Intelligent Systems at the University of Southern California, Los Angeles, as a visiting scientist. In 2001 he joined Samsung Advanced Institute of Technology, Korea, where he is currently a research scientist. His current research interests include computer vision, virtual reality, augmented reality, and computer graphics.

Ki-Sang Hong received his BS degree in electronic engineering from Seoul National University, Korea, in 1977, and the MS degree in electrical and electronic engineering from KAIST, Korea, in 1979. He also received the PhD degree from KAIST in 1984. From 1984 to 1986, he was a researcher in the Korea Atomic Energy Research Institute, and in 1986, he joined POSTECH, Korea, where he is currently a professor of electrical and electronic engineering. During 1988 to 1989, he worked in the Robotics Institute at Carnegie Mellon University, Pittsburgh, Pennsylvania as a visiting professor. His current research interests include computer vision, augmented reality, pattern recognition, and SAR image processing.