ICIP 1999, Kobe, Japan, October 1999.

SUPER-RESOLUTION RECONSTRUCTION OF CONTINUOUS IMAGE SEQUENCES

Michael Elad

HP Laboratories - Israel (HPL-I)
The Technion City
Haifa 32000, Israel

Arie Feuer

Department of Electrical Engineering
The Technion—Israel Institute of Technology
Haifa 32000, Israel

Abstract

Super-resolution reconstruction algorithms perform a fusion of several low quality images of the same scene into a single improved quality image. As opposed to this STATIC recovery problem, in this paper we define a DYNAMIC super-resolution task: the restoration of a blurred, decimated, and noisy image sequence. We first model this problem through state-space equations, showing that this problem can be viewed as a sequence of STATIC super-resolution problems. Two efficient reconstruction algorithms are proposed, both being adaptive filtering approximations of the Kalman filter; the R-SD and the R-LMS. Computer simulations on synthetic sequences indicate the computational feasibility of these algorithms.

1. INTRODUCTION

In the Super-resolution reconstruction problem [1]–[4], several geometrically warped, blurred, decimated, and noisy images of an ideal image are given, and the objective is the recovery of this ideal image. Super-resolution reconstruction algorithms ([1]–[4]) effectively apply a fusion of the measurements into a single improved resolution image. As opposed to the this STATIC recovery problem, we define a DYNAMIC super-resolution task: the restoration of a blurred, decimated, and noisy image sequence. This paper focuses on this dynamic problem and its solution. Our approach can be described as the combination of ideas from two related problems: Super-resolution reconstruction of a single image (see [1]–[4]) and the restoration of an image sequence from blurred and noisy data (see e.g. [5]–[6]).

We use state-space equations for the description of both the relation between the measurements and the required output, and for the inter-relations that exists within the output sequence. The obtained state-space formalism reveals that the Kalman filter is needed for the reconstruction. However, due to the dimensions involved, approximations of it are required. From a different point of view, we show that this problem can be viewed as a sequence of STATIC super-resolution problems.

Two efficient reconstruction algorithms are proposed, both being adaptive filtering-based approximations of the Kalman filter; the R-SD and the R-LMS. The R-SD is a better approximation but requires both more computations and memory. The R-LMS is a very simple variation of the R-SD, based on the stochastic approximation algorithm. Both algorithms are computationally feasible, and computer simulations on synthetic sequences show promising results in terms of output quality.

Due to space limitations, this paper presents the basic results of our DYNAMIC super-resolution idea. More details can be found in [7]–[8].

2. MODELING THE PROBLEM

Consider a sequence of images $\{\underline{Y}(t)\}$, each image is of $M \times M$ pixels, as our measured data. We wish to generate a sequence $\{\underline{X}(t)\}$ of images of higher resolution, each image of $L \times L$ ($L > M$) pixels and of improved quality. For convenience of notation all images will be presented as vectors, ordered column-wise lexicographically. Namely, we have $\underline{Y}(t) \in \mathbb{R}^{M^2}$ and $\underline{X}(t) \in \mathbb{R}^{L^2}$. At each time instant t we assume that the two images are related via the following equation:

$$\underline{Y}(t) = DH(t)\underline{X}(t) + \underline{N}(t) \qquad (1)$$

which means that $\underline{X}(t)$ is blurred, decimated (i.e. down sampled) and contaminated by additive noise to give $\underline{Y}(t)$. $H(t)$ is the blur matrix which may be space and time variant, D the decimation matrix assumed constant, and $\underline{N}(t)$ is a zero mean Gaussian noise with

E-mail: elad@hpli.hpl.hp.com.
E-mail: feuer@ee.technion.ac.il.

$W^{-1}(t) = E\{\underline{N}(t)\underline{N}^T(t)\}$. In addition, we assume that the sequence $\{\underline{X}(t)\}$ satisfies the following equation:

$$\underline{X}(t) = G(t)\underline{X}(t-1) + \underline{V}(t)\,. \qquad (2)$$

The matrix $G(t)$ stands for the geometric warp between the images $\underline{X}(t)$ and $\underline{X}(t-1)$, and $\underline{V}(t)$ is the system noise. Assuming the typical optical flow model with the nearest neighbor paradigm, most pixels in the image $\underline{X}(t)$ originate from pixels in the image $\underline{X}(t-1)$. Therefore, for each such pixel the corresponding row in the matrix $G(t)$ contains only one non-zero element at a position which reflects the address of the source pixel in the previous image and this entry equals 1. Other methods of interpolating can be used as well, resulting with other forms of row stochastic matrices $G(t)$. The vector $\underline{V}(t)$ contains the innovation sequence - all the new information which does not originate from the previous image. For the sake of our analysis here, we assume this vector to be a zero mean Gaussian process with $Q^{-1}(t) = E\{\underline{V}(t)\underline{V}^T(t)\}$. In the next section we show that this assumption is actually bypassed.

With equations (1) and (2), the problem we posed can be viewed as a state estimation problem and the most natural tool to consider is the Kalman filter. For the linear model we have assumed, the Kalman filter will provide the optimal solution in the Mean Square Error sense ([9]). However, because of the large dimensions involved - $L^2 \times L^2$ - the computation and storage required to use the Kalman filter makes its use in our case impractical. Hence, our goal is to develop algorithms which approximate the Kalman filter as far as performance but are significantly less demanding computation wise.

Throughout this paper we assume that the matrices D, $H(t)$, $W(t)$, $G(t)$, and $Q(t)$, which define the state-space system, are known. More details which justify such an assumption, and the way to actually obtain these matrices is described in [7]-[8].

When needed, equation (1) may include a regularization expression which represents prior knowledge on the spatial behavior of $\underline{X}(t)$. As typical in reconstruction problems, spatial smoothness may be used. Defining S as the Laplacian operator we propose:

$$\underline{0} = S\underline{X}(t) + \underline{U}(t) \qquad (3)$$

which simply means that applying the Laplacian on the image $\underline{X}(t)$ should give zeros up to some additive noise, $\underline{U}(t)$, which we assume again, to be zero mean Gaussian with $R^{-1}(t) = E\{\underline{U}(t)\underline{U}^T(t)\}$. Combining equation (1) and equation (3) we get:

$$\begin{bmatrix}\underline{Y}(t)\\ \underline{0}\end{bmatrix} = \begin{bmatrix}DH(t)\\ S\end{bmatrix}\underline{X}(t) + \begin{bmatrix}\underline{N}(t)\\ \underline{U}(t)\end{bmatrix} \qquad (4)$$

$$\Rightarrow \quad \underline{Y}_A(t) = H_A(t)\,\underline{X}(t) + \underline{N}_A(t)$$

with

$$W_A^{-1}(t) = \begin{bmatrix} W^{-1}(t) & 0 \\ 0 & R^{-1}(t)\end{bmatrix}.$$

The matrix $R(t)$ may be chosen such that $R(t) = \beta I$. This way, the parameter β controls the spatial smoothness of the resulting images. More complex choices of $R(t)$, introducing locally adaptive smoothness are also possible under this framework.

3. DYNAMIC ESTIMATION

3.1. The Pseudo-RLS Equations

Kalman filter may be presented either as the propagation of the mean-covariance pair or the information pair [9]. In our analysis we concentrate on the information pair. The prediction and the update equations are therefore:

$$\begin{aligned}\tilde{\underline{Z}}(t) &= \tilde{L(t)}G(t)\hat{L}^{-1}(t-1)\hat{\underline{Z}}(t-1) \\ \tilde{L(t)} &= \left[G(t)\hat{L}^{-1}(t-1)G^T(t) + Q^{-1}(t)\right]^{-1}\end{aligned} \qquad (5)$$

$$\begin{aligned}\hat{\underline{Z}}(t) &= \tilde{\underline{Z}}(t) + H_A^T(t)W_A(t)\underline{Y}_A(t) \\ \hat{L}(t) &= \tilde{L}(t) + H_A^T(t)W_A H_A(t)\end{aligned} \qquad (6)$$

and the output of the Kalman filter is the vector $\hat{\underline{X}}(t)$, obtained as the solution of the linear set of equations: $\hat{L}(t)\hat{\underline{X}}(t) = \hat{\underline{Z}}(t)$. The bottleneck in the above propagation equations is the need to invert a huge matrix twice for every time instance, as can be seen in the prediction stage (similar bottleneck appears in the propagation of the mean-covariance pair but on the update part of the propagation).

In order to simplify the prediction equations, and omit these inverses, we propose to replace the term $Q^{-1}(t)$ (which is assumed to be non-singular) with an approximation term of the form $\alpha(t)G(t)\hat{L}^{-1}(t-1)G^T(t)$, where $\alpha(t)$ is some positive scalar. One possibility is to choose this scalar so that $Q^{-1}(t) \leq \alpha(t)G(t)\hat{L}_{-1}(t-1)G^T(t)$. The above approach means that we assume a stronger system's noise $\underline{V}(t)$. Such approach is known as adding pseudo-noise to the system's equation ([9]), and is typically proposed for the treatment of model inaccuracies. Different methods for the choice of $\alpha(t)$, such as searching for $\alpha(t)$ which minimizes the estimation error, can be suggested [7]–[8].

Using the above approximation, the entire Information Kalman filter propagation equations simplify to:

$$\underline{\hat{Z}}(t) = \lambda(t)F^T(t)\underline{\hat{Z}}(t-1) + H_A^T(t)W_A(t)\underline{Y}_A(t) \quad (7)$$
$$\hat{L}(t) = \lambda(t)F^T(t)\hat{L}(t-1)F(t) + H_A^T(t)W_A(t)H_A(t)$$

where we have denoted $F(t) = G^{-1}(t)$ (or its pseudo-inverse if $G(t)$ is singular). This is the backward motion matrix representing the motion operator from the current image $\underline{X}(t)$ to the previous one $\underline{X}(t-1)$. This matrix therefore has the same properties as the matrix $G(t)$. We also denote $\lambda(t) = [1+\alpha(t)]^{-1}$. Quite clearly, these recursive equations are much simpler to implement, compared to the previous ones.

Another more intuitive approach, which yields these very same two recursive equations, is to totally omit the system's noise by assuming $Q^{-1}(t) = 0$. In this case, by putting the system's equation into the measurement equation, the two state-space equations can be combined into an infinitely long sequence of equations of the form:

$$\underline{Y}_A(t-k) = H_A(t-k)\underline{X}(t-k) + \underline{N}_A(t-k) \quad (8)$$
$$= H_A(t-k)\prod_{j=1}^{k} F(t-k+j)\underline{X}(t) + \underline{N}_A(t-k)$$

for $k = 0, 1, 2, \cdots t$. We can now define a Weighted Least Squares (WLS) problem, where we search for the image $\underline{X}(t)$ which minimizes the function:

$$\varepsilon^2(t) = \sum_{k=0}^{\infty}\left[\prod_{j=0}^{k-1}\lambda(t-j)\right]\cdot\Big\|\underline{Y}_A(t-k) - H_A(t-k)\cdot$$
$$\cdot\prod_{j=1}^{k} F(t-k+j)\underline{X}(t)\Big\|^2_{W_A(t-k)} \quad (9)$$

It turns out that the minimum of the above penalty function is the exact same solution that was shown in equation (7). Moreover, Looking closely at equation (9), this penalty term is exactly the one proposed for the static super-resolution problem in [4]. The difference here is that we are to solve this minimization problem per each instance t, this way creating a sequence of output images.

So far, the Kalman filter approximation we got involves solving a set of linear equations at each temporal point. We name this algorithm, as presented in equation (7), the Pseudo-RLS algorithm.

3.2. The R-SD algorithm

Since we need to solve a very large set of linear equations per each instance t, one immediate approach is to apply some iterative algorithm. Of-course, in order to get the exact solution one would need to apply many number of such iterations. However, since the propagation of $\hat{L}(t)$ and $\underline{\hat{Z}}(t)$ is independent of the estimation $\underline{\hat{X}}(t)$, any error caused by limiting ourselves to a finite number of iterations at $(t-1)$ would not propagate to time t.

The algorithm we propose consists of applying R iterations of the Steepest Descent (SD) algorithm at each time t. Hence we name this algorithm R-SD. A natural choice for the initialization of such algorithm is the vector $\underline{\hat{X}}_0(t) = G(t)\underline{\hat{X}}_R(t-1)$, where $\underline{\hat{X}}_R(t-1)$ is the result after the previous R iterations. This choice comes from the prediction step in the Kalman filter-equation. The R-SD algorithm is therefore:

Initialization:

$$\underline{\hat{X}}_R(0) = \text{Arbitrary}; \quad \hat{L}(0) = 0; \quad \underline{\hat{Z}}(0) = 0$$

For $\quad t \geq 1:$ (10)

$$\underline{\hat{Z}}(t) = \lambda(t)F^T(t)\underline{\hat{Z}}(t-1) + H_A^T(t)W_A(t)\underline{Y}_A(t)$$
$$\hat{L}(t) = \lambda(t)F^T(t)\hat{L}(t-1)F(t) + $$
$$+ H_A^T(t)W_A(t)H_A(t)$$
$$\underline{\hat{X}}_0(t) = G(t)\underline{\hat{X}}_R(t-1)$$
$$\underline{\hat{X}}_k(t) = \left[I - \mu\hat{L}(t)\right]\underline{\hat{X}}_{k-1}(t) + \mu\underline{\hat{Z}}(t); \quad 1 \leq k \leq R$$

This algorithm requires the propagation of the approximated information pair $\langle\underline{\hat{Z}}(t); \hat{L}(t)\rangle$ in time, and then use these terms in the recursive update equation of the estimated output vector $\underline{\hat{X}}_R(t)$.

The parameter μ in the above algorithm should be chosen so as to guarantee the Steepest Decent convergence ([7]-[8]).

An interesting question with respect to the R-SD algorithm is the density of the matrix $\hat{L}(t)$. In order to get that the R-SD is feasible, this matrix must be sparse and remain sparse for all t. In [8] an upper bound on the density of this matrix is obtained. The bound is shown experimentally to be tight, and very low, thus ensuring small number of memory cells and computations when dealing with this matrix.

3.3. The R-LMS algorithm

Taking the R-SD algorithm of the previous sub-section and further approximating the information pair $\langle\underline{\hat{Z}}(t); \hat{L}(t)\rangle$ by the instantaneous values in equation (7),

$$\langle H_A^T(t)W_A(t)\underline{Y}(t); H_A^T(t)W_A(t)H_A(t)\rangle \quad (11)$$

respectively, we get an algorithm which resembles the LMS algorithm, and thus the name R-LMS algorithm. After some algebra we get:

$$
\begin{aligned}
&\text{Initialization:}\\
&\underline{\hat{X}}_R(0) = \text{Arbitrary}\\
&\text{For } t \geq 1: \\
&\quad \underline{\hat{X}}_0(t) = G(t)\underline{\hat{X}}_R(t-1)\\
&\quad \underline{\hat{X}}_k(t) = \underline{\hat{X}}_{k-1}(t) + \mu H_A^T(t) W_A(t) \cdot\\
&\qquad \cdot \left[\underline{Y}(t) - H_A(t)\underline{\hat{X}}_{k-1}(t)\right]; \quad 1 \leq k \leq R
\end{aligned} \tag{12}
$$

Clearly, the R-LMS algorithm is simpler than the R-SD algorithm, both in the computational and the memory requirements. We note that the R-LMS algorithm can also be obtained from the R-SD algorithm by assigning $\lambda(t) = 0$. Presumably, such value for $\lambda(t)$ means no temporal memory, and thus no temporal smoothness. Indeed, this is the case if infinitely many iterations are performed per each time point ($R \to \infty$). However, since R is finite and relatively small, temporal smoothness is not discarded, although it comes from different origin.

4. SIMULATIONS AND ANALYSIS

The two tests presented in this part are based on two synthetic sequences, each containing 100 images of size [50 × 50] pixels. These sequences serve as the ideal images. The two measured image sequences were generated from these ideal sequences, by blurring each image using [3 × 3] uniform kernel, decimation using 2:1 decimation ratio on each axis, and adding zero mean Gaussian white noise with $\sigma = 5$ (the dynamic range of the gray level in the images is 0-255). Thus the measured sequences contain 100 images of size [25 × 25] pixels each. These dimensions were chosen in order to shorten the simulations run-time and to overcome the memory limitations posed by MATLAB.

The R-SD and the R-LMS algorithms were applied using 5 iterations per each time point. In all cases, the initialization image at $t = 0$ was chosen to be a bilinear interpolated version of the first measured image. The applied regularization in all the tests was the Laplacian operator, using relative weight $\beta = 0.02$. In the reconstruction process, the true motion, blur and decimation operators were assumed to be known.

In Figure 1, the results of the first test are given. In order to illustrate the temporal axis, the 1st, the 25th, the 50th, the 75th and the 100th images of each sequence are given. The motion in this sequence consist of global zoom in and out and global translation motion. The given sequences are A: The ideal sequence; B: The measured sequence; C: Bilinear interpolation of the measurement sequence; D: The R-LMS results without regularization; E: The R-LMS results with regularization; and F: The R-SD results with regularization. Similar to this figure, Figure 2 presents the results for the second sequence. The motion in this sequence consist of global constant rotation while zooming in and out.

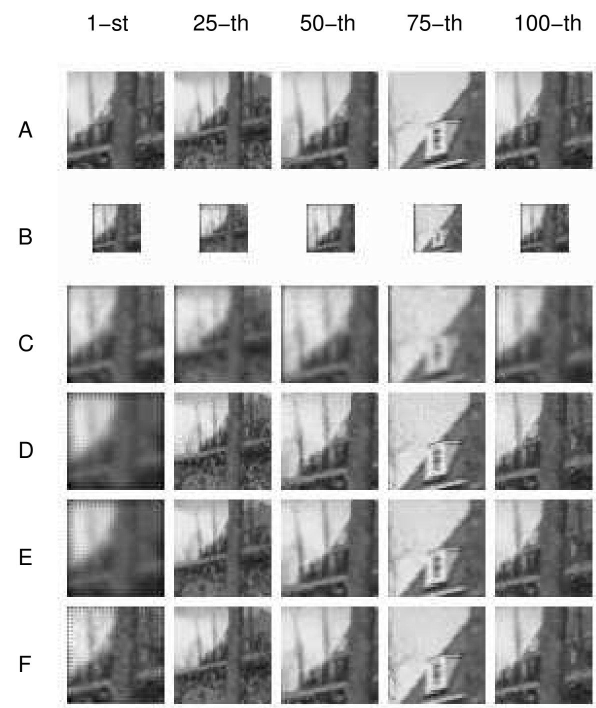

Figure 1: The first sequence results

From these results we can conclude the following:

1. In both examples, using any of the two proposed reconstruction methods, there is a clear improvement both in the resolution quality and in the suppression of noise and blur degradation effects.
2. As expected, the R-SD algorithm gives slightly better results. This is true both for the convergence rate at the initialization part and the steady state.
3. Regularization improves the performance in both examples.

5. SUMMARY AND CONCLUSIONS

This paper presents the dynamic super-resolution reconstruction problem, and algorithms to solve it. In this problem, an image sequence is to be recovered from down-sampling, blurring, and noise degradations. We show that this problem is both a generalization of

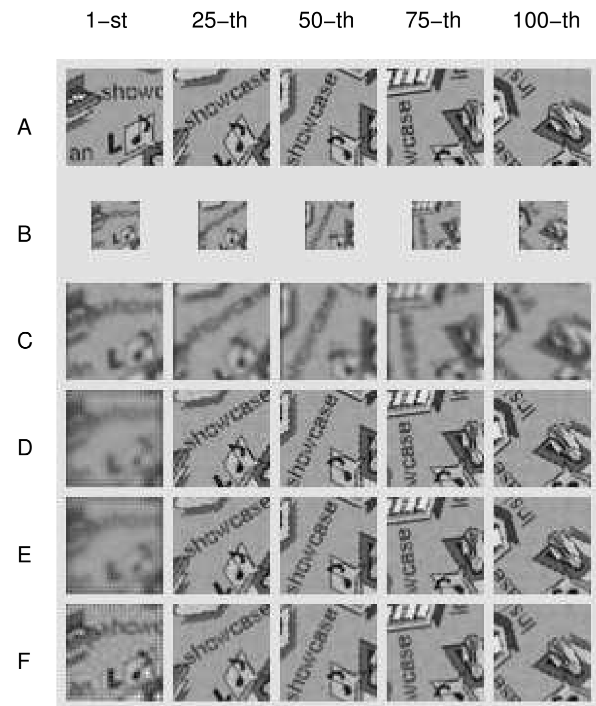

Figure 2: The second sequence results

the image sequence restoration and the static super-resolution recovery problems. Two novel reconstruction algorithms are proposed, the R-SD and the R-LMS. These two algorithms are shown to be approximations of the well-known Kalman filter. Their performance on synthetic image sequences is found to be promising.

Acknowledgments

This work was supported by the The Israel Science Foundation founded by The Israel Academy of Sciences and Humanities, and by the Technion V.P.R. fund, N. Haar and R. Zinn Research Fund.

6. REFERENCES

[1] Irani and S. Peleg, "Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency", *J. of VCIR*, Vol. 4, pp. 324–335, December 1993.

[2] A.J. Patti, M.I. Sezan and A.M. Tekalp, "High-Resolution Image Reconstruction from a Low-Resolution Image Sequence in the Presence of Time-Varying Motion Blur", *Proc. ICIP, Austin - Texas*, pp. 343–347, November 1994.

[3] R.S. Schultz and R.L. Stevenson, "Extraction of High-Resolution Frames from Video Sequences", *IEEE Trans. Image Processing*, Vol. 5, pp. 996–1011, June 1996.

[4] M. Elad and A. Feuer, "Restoration of Single Super-Resolution Image From Several Blurred, Noisy and Down-Sampled Measured Images ", *IEEE Trans. Image Processing*, Vol. 6, pp. 1646–58, December 1997.

[5] T. Patti, A.M. Tekalp and A.M Sezan, "Image Sequence Restoration and De-interlacing by Motion-Compensated Kalman Filter", *Proc. SPIE - The International Society for Optical Engineering*, Vol. 1903, pp. 59-70, 1991.

[6] A.K. Katsagellos, J.N. Driessen, S.N. Efstratiadis and L.J. Lagendijk, "Spatio-temporal Motion Compensated Noise Filtering of Image Sequences", *Proc. SPIE - International Society for Optical Engineering*, Vol. 1199, pp 61–70, 1989.

[7] M. Elad and A. Feuer, "Super-Resolution Restoration of Continuous Image Sequence - Adaptive Filtering Approach", *IEEE Trans. Image Processing*, March 1999.

[8] M. Elad and A. Feuer, "Super-Resolution Restoration of Image Sequences", accepted to the *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, March 1999.

[9] P.S. Maybeck, *Stochastic Models, Estimation and Control*, Vol. I & II, Academic Press, New-York, 1979.