An EM Based Resolution Enhancement Algorithm Considering Partially Known Point Spread Function

Hu He, Cibu Jose, Lisimachos Kondi
Dept. of Electrical Engineering, SUNY at Buffalo, Buffalo, NY 14260
{huhe, cibujose, lkondi}@eng.buffalo.edu

***Abstract*— In this paper, we propose a resolution enhancement algorithm based on the Expectation Maximization (EM) framework. The objective of resolution enhancement (super-resolution) is to reconstruct a high-resolution image from a sequence of low-resolution (LR) images, under the assumption that there exists subpixel motion between the low-resolution frames.**

EM based image restoration has been studied in previous work, while its results can not be directly applied to the resolution enhancement scenario because of the subsampling (decimation) process in the acquisition model. This leads to a non-square degradation matrix which is not circulant as in the restoration case and hence cannot be diagonalized and operated on by the EM algorithm in the frequency domain.

To overcome this difficulty, we propose to reorganize and interlace the low-resolution frames to construct an interlaced image using the registration parameters. This interlaced image is equivalent to a uniform blur process of the PSF blurred image. Now the resolution enhancement problem reduces to a restoration problem with two low-pass filters to deblur: one is the blur due to the point spread function (PSF) of the optical lens, and the other the uniform blur due to the decimation matrix. EM based restoration algorithm is thus computed efficiently in the frequency domain, considering the inaccurate estimate of the PSF and unknown power spectrum of both the high-resolution image and noise. Simulations using synthetic images are implemented to verify the proposed algorithm and conclusions are drawn.

I. Introduction

The goal of resolution enhancement (super-resolution) is to estimate a high-resolution image from a sequence of low-resolution images while also compensating for blurring due to the point spread function of the camera lens and the effect of the finite size of the photo-detectors, as well as additive noise introduced by the capturing process. Resolution enhancement using multiple frames is possible when there exists subpixel motion between the captured frames. Thus, each of the frames provides a unique look into the scene.

In many practical applications, the blurring process is unknown or only partially known. Deblur, or deconvolution, is an important step to reconstruct a high-resolution image from the degraded low-resolution image sequence. There has been extensive work on deblur with partially known blur or totally unknown blur (blind deconvolution) [1]. The deconvolution process can be categorized into two classes: methods with separate blur identification as a disjoint procedure from restoration/super-resolution, and methods which combines blur identification and restoration/super-resolution in one process. The methods in the first class tend to be computationally simpler.

In this paper, we concentrate on a special super-resolution case. We assume that the blur is space invariant and partially known for the low-resolution image sequence; the subpixel motion of each low-resolution frame is pure global translation on the high-resolution grid. These assumptions are limiting but quite practical in some real applications. For this special case, we propose a super-resolution algorithm using a first class EM based deconvolution. The rest of the paper is presented as follows: in section II, we propose to construct a squared and semiblock circulant (SBC) matrix from the low-resolution frames; the obtained data are interlaced to the high-resolution grid; in section III, EM based restoration algorithm is implemented in the high resolution grid with two blurring processes to deconvolve. In section IV, experimental results are presented to verify the proposed algorithm and finally conclusion and future work are discussed in section V.

II. Interlaced High-resolution Image From Low-resolution Frames

The image degradation process is modeled by a motion, linear blur, subsampling by pixel averaging and an additive Gaussian noise process [2], [3], [4]. All vectors are ordered lexicographically. Assume that p low-resolution frames are observed, each of size $N_1 \times N_2$. The desired high-resolution image $\mathbf{z}$ is of size $N = L_1N_1 \times L_2N_2$ where L_1 and L_2 represent the down-sampling factors in the horizontal and vertical directions, respectively. Thus, the observed low-resolution images are related to the high-resolution (HR) image through motion shift, blurring and subsampling. Let the kth low-resolution frame be denoted as $\mathbf{y}_k = [y_{k,1}, y_{k,2}, \ldots, y_{k,M}]^T$ for $k = 1, 2, \ldots, p$ where $M = N_1N_2$. The full set of p observed low-resolution images can be denoted as

$$\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_p^T]^T = [y_1, y_2, \ldots, y_{pM}]^T. \quad (1)$$

The observed low-resolution frames are related to the high-resolution image through the following model:

$$y_{k,m} = \sum_{r=1}^{N} w_{k,m,r}(\mathbf{s}_k) z_r + \eta_{k,m}, \quad (2)$$

for $m = 1, 2, \ldots, M$ and $k = 1, 2, \ldots, p$. The weight $w_{k,m,r}(\mathbf{s}_k)$ represents the "contribution" of the rth high-resolution pixel to the mth low-resolution observed pixel of the kth frame. The vector $\mathbf{s}_k = [s_{k,1}, s_{k,2}, \ldots, s_{k,K}]^T$, is the K registration parameters for frame k, measured in reference to a fixed high resolution grid. The term $\eta_{k,m}$ represents additive noise samples that is assumed to be independent and identically distributed (i.i.d.) Gaussian noise samples with variance σ_η^2. The system can be modeled in matrix notation as

$$\mathbf{y} = \mathbf{W_z z} + \mathbf{n}. \tag{3}$$

In equation (3), the degradation matrix

$$\mathbf{W_z} = [\mathbf{W_{z,1}}, \mathbf{W_{z,2}}, \cdots, \mathbf{W_{z,k}}]^T \tag{4}$$

performs the operation of motion, blur and subsampling. Therefore $\mathbf{W_z}$ for frame k can be written as:

$$\mathbf{W}_{\mathbf{z},k} = \mathbf{S}\mathbf{B}_k\mathbf{M}_k, \tag{5}$$

where $\mathbf{S}$ is the $N_1N_2 \times N$ subsampling matrix, $\mathbf{B}_k$ is the $N \times N$ blurring matrix, and $\mathbf{M}_k$ is the motion matrix. The PSF of blurring is assumed to be space-invariant, normalized and having non-negative elements within a 2-D rectangular support $l_1 \times l_2$. In this paper, for the special case under consideration, we have same space invariant blur for each frame and global translation shift of the subpixel motion among them. Under these assumptions, $\mathbf{B}_k$ is a Tikhonov matrix and approximately semiblock circulant (SBC). Also, we can exchange the order of the motion and blur in (5) such that $\mathbf{W}_{\mathbf{z},k} = \mathbf{S}\mathbf{M}_k\mathbf{B}_k$ Thus, each frame can be modeled as

$$\mathbf{y}_k = \mathbf{W}_{\mathbf{z},k}\mathbf{z} + \mathbf{n}_k = \mathbf{S}\mathbf{M}_k\mathbf{B}\mathbf{z} + \mathbf{n}_k. \tag{6}$$

We drop the subindex k from $\mathbf{B}_k$ from above equation. MAP based super-resolutions have studied in our previous work [2], [3], with the assumption that the PSF is well known. However, in many applications, the blur due to the PSF is only partially known. In this paper, for the over-determined system which has enough low-resolution frames available, $p \geq L_1 \times L_1$, we propose to first reorganize and interlace the low-resolution frames to construct an interlaced image using the registration parameters. This process can be summarized as:

(i) stabilize the low-resolution frames. One frame, usually the first frame, is used as the reference frame, which has motion [0, 0] on the high-resolution grid. Other frames are stabilized to have motion vector within the Cartesian product set $\{0, 1, \ldots, L_1 - 1\} \times \{0, 1, \ldots, L_2 - 1\}$ on the high-resolution grid, which is equivalent to subpixel motion on the low-resolution grid with resolution $\frac{1}{L_1}$ and $\frac{1}{L_2}$ on the two axes.

(ii) If enough frames exist and all the combinations in the above Cartesian product set exist, we can select those $L_1 \times L_2$ stabilized frames bearing unique motion from the Cartesian product set. These frames all together bear the information of the PSF blurred original high-resolution image, $\mathbf{Bz}$, with different translation shift among them.

(iii) Motion compensate and interlace the above $L_1 \times L_2$ stabilized frames into the high-resolution grid. The result, an interlaced high-resolution image, can be formed from the PSF blurred original high-resolution image, $\mathbf{Bz}$, via subpixel shift and averaging, which is also equivalent to the uniform blur convolution with support size $L_1 \times L_2$. Note here we also use the symmetric property of the uniform blur: 2-D flip of the uniform blur remain unchanged.

(iv) Now we can rewrite the system model as

$$\bar{\mathbf{y}} = \mathbf{B}_u\mathbf{B}\mathbf{z} + \bar{\mathbf{n}}. \tag{7}$$

Here $\mathbf{B}_u$ stands for the uniform blur matrix with size $N \times N$, which is squared and approximately semiblock circulant (SBC). The multiplication of two SBC matrices, $\mathbf{B}_u\mathbf{B}$, is also a SBC matrix [5]. The interlacing process will not change the variance of the interlaced AWGN noise, $\bar{\mathbf{n}}$. The interlaced image $\bar{\mathbf{y}}$ now has size $N \times 1$, same size as the high-resolution image $\mathbf{z}$. In case there are not $L_1 \times L_2$ stabilized frames formed in step (ii) available, a local interpolation can be used in step (iii) for a smoothed solution. Now the resolution enhancement problem reduces to a restoration problem with two low-pass filters to deblur: one is the blur due to the point spread function (PSF) of the optical lens, and the other is the uniform blur due to the decimation (subsampling by pixel averaging).

III. EM Based Super-resolution

Expectation Maximization (EM) based image restoration has been studied in previous work [6], [7]. The EM algorithm is an iterative approach for computing maximum-likelihood (ML) estimates of the unknown parameters. Restoration from partially known blur using the EM algorithm was proposed in [7]. For a system modeled as

$$\mathbf{y} = \mathbf{Hz} + \mathbf{n}, \tag{8}$$

the conditional mean of the restored image is updated in the frequency domain in a LMMSE fashion E-step as [7], [8]:

$$M_{\mathbf{z}|\bar{\mathbf{y}}} = \frac{\bar{H}^*(i)S_{\mathbf{z}}(i)Y(i)}{|\bar{H}(i)|^2S_{\mathbf{z}}(i) + NS_{\Delta h}(i)S_{\mathbf{z}}(i) + S_{\Delta \mathbf{y}}(i)}, \tag{9}$$

for $i = 1, \ldots, N$. Here $S_{\mathbf{z}}$, $S_{\Delta h}$ and $S_{\Delta \mathbf{y}}$ are the DFT of the covariance matrices $R_{\mathbf{z}}$, $R_{\Delta h}$ and $R_{\Delta \mathbf{y}}$, respectively. The EM algorithm is useful for its capability to identify these unknown covariance matrices, while simultaneously restoring the degraded image.

However, the results from image restoration can not be directly applied to the resolution enhancement scenario, because of the subsampling (decimation) process in the acquisition model. This leads to a non-square degradation matrix which is not circulant as in the restoration case and hence cannot be diagonalized and operated on by the EM algorithm in the frequency domain [9]. In previous work by Woods et. al. [9], an interlaced observation is also formed and the EM algorithm is applied, however, their model limits the unknown parameters in the PSF and registration and the expression of these parameters doesn't have a closed form. Gradient desecent is used to calculate these parameter, while it will bring some computation cost, and possibility of instability. In our approach, after the process from above section, we have overcome this difficulty. Comparing equations (7) and (8), we apply transform $\bar{\mathbf{H}} = \overline{\mathbf{B}_u\mathbf{B}} = \mathbf{B}_u\bar{\mathbf{B}}$ and get the E-step for super-resolution as

$$M_{\mathbf{z}|\bar{\mathbf{y}}} = \frac{B_u^*(i)\bar{B}^*(i)S_{\mathbf{z}}(i)Y(i)}{|B_u(i)\bar{B}(i)|^2S_{\mathbf{z}}(i) + NS_{\Delta h}(i)S_{\mathbf{z}}(i) + S_{\Delta \mathbf{y}}(i)}, \tag{10}$$

TABLE I

THREE CASES OF SYNTHETIC TEST FOR "LENA".

	σ^2	L_1, L_2	p	$\mathbf{s}_k^T = [s_{k,1}, s_{k,2}]$
Case I	1.5	1	1	$\{0\} \times \{0\}$
Case II	1.5	2	4	$\{0,1\} \times \{0,1\}$
Case III	1.5	4	16	$\{0,1,2,3\} \times \{0,1,2,3\}$

TABLE II

RESULTS OF "LENA" USING THE THREE METHODS.

PSNR (dB)	Bilinear interpolation	EM	LMMSE
Case I	26.4019	27.6598	29.0723
Case II	25.5224	27.8427	29.1170
Case III	23.4162	26.6088	27.4551

Similar transform can be made to other equations in the EM procedure in [7] to extend the application from restoration to super-resolution.

IV. EXPERIMENTAL RESULTS

A number of experiments were conducted, some of which are presented here. To test the performance of our algorithms, we first use the 256x256 "Lena" test image for a synthetic test. The PSF is a Gaussian PSF with support size 7×7 and variance $\sigma^2 = 1.5$. Three cases, Case I-III, as listed in Table I, are tested, with Case I the EM based restoration. The PSF noise and the AWGN noise are both gaussian type with both $SNR_{\mathbf{y}}$ and SNR_h fixed at 30dB. The global shift $\mathbf{s}_k^T$ belongs to the vectors generated from the given Cartesian product listed in the table. The PSNR of the reconstructed images for "Lena" using three methods (Bilinear interpolation, LMMSE, EM) are listed in Table 2. Here, $PSNR_{HR}$ is $10log_{10}\frac{255^2}{MSE_{HR}}$, where MSE stands for the mean squared error between the original high-resolution image and the estimated high-resolution image.

The original frame and the interlaced image are shown in Fig. 1, 2. The reconstructed "Lena" image from bilinear interpolation (BI) of the first low-resolution frame, EM, LMMSE in case II are shown in Fig. 3, 4 and 5. The PSNRs of the reconstructed images using these three methods are listed in Table II. From the results, we can see that our algorithm is better than the bilinear interpolation and close to the LMMSE results. The advantage over LMMSE is that our algorithm does not need the power spectrum information of the original image and the additive noise.

Next, we use the same setup as above, with upsampling ratio $L_1 = L_2 = 2$, but fix $SNR_{\mathbf{y}}$ at SNR=30dB and vary SNR_h. The plot of the PSNR for the reconstructed image versus SNR_h is shown in Fig. 6. Also, we fix SNR_h at SNR=30dB and vary $SNR_{\mathbf{y}}$. The plot of the PSNR for the reconstructed image versus $SNR_{\mathbf{y}}$ is shown in Fig. 7.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a resolution enhancement algorithm based on the EM framework. we propose a two-stage process: first to reorganize and interlace the low-resolution frames to construct an interlaced image. This interlaced image is equivalent to a uniform blur process of the PSF blurred image. In the next stage, the resolution enhancement problem reduces to a restoration problem with two low-pass filters to deblur: one is the blur due to the point spread function (PSF) of the optical lens, and the other is the uniform blur. The EM based restoration algorithm can be efficiently computed in the frequency domain to reduce the computational cost, with partially know PSF blur information and unknown power spectrum of the AWGN noise. Experimental results show that our algorithm provides visually satisfying reconstructions. Future work will take the more complicated case with inaccurate registration, i.e., registration noise into consideration.

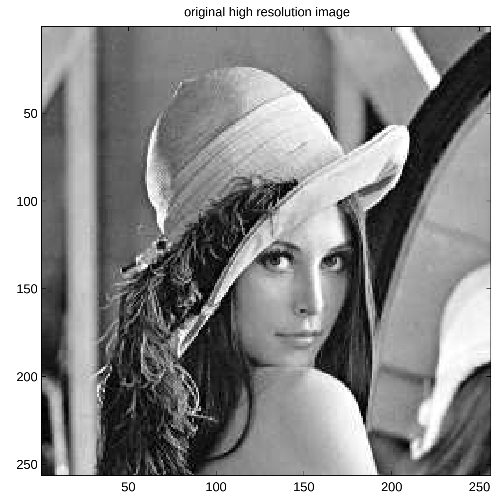

Fig. 1. Original image.

REFERENCES

[1] D. Kundur, D. Hatzinakos, "Blind image deconvolution revisited", *IEEE Signal Processing Magazine*, vol. 13, pp.61 - 63, Nov. 1996.

[2] H. He, L.P. Kondi, "MAP based resolution enhancement of video sequences using a Huber-markov random field image prior model", in *Proc. IEEE International Conference on Image Processing*, Vol. 2, pp. 933-936, 2003.

[3] H. He, L.P. Kondi, "Resolution Enhancement of Video Sequences with Simultaneous Estimation of the Regularization Parameter", *Journal of Electronic Imaging*, vol.13, pp. 586-596, July 2004.

[4] R. C. Hardie, K. J. Barnard, E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images", *IEEE Trans. Image Processing*, pp. 1621-1633, Dec. 1997.

[5] A. K. Katsaggelos, K. T. Lay, N. P. Galatsanos, "A general framework for frequency domain multi-channel signal processing", *IEEE Trans. Image Processing*, vol. 2, pp. 417-420, July 1993.

[6] T. J. Hebert, K. Lu, "Expectation-maximization algorithms, null spaces, and MAP image restoration", *IEEE Trans. Image Processing*, vol. 4, pp. 1084-1095, Aug. 1995.

[7] V. Z. Mesarovic, N. P. Galatsanos, M. N. Wernick, "Restoration from partially-known blur using an expectation-maximization algorithm", *1996 Conference Record of the Thirtieth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 95-99, 3-6 Nov. 1996.

[8] N. P. Galatsanos, V. Z. Mesarovic, R. Molina, A. K. Katsaggelos, "Hierarchical Bayesian image restoration from partially known blurs", *IEEE Trans. Image Processing*, vol. 9, pp. 1784-1797, Oct. 2000.

[9] N. A. Woods, N. P. Galatsanos, A. K. Katsaggelos, "EM-based simultaneous registration, restoration, and interpolation of super-resolved images", *2003 International Conference on Image Processing*, vol. 2, pp. 303-306, 14-17 Sept. 2003.

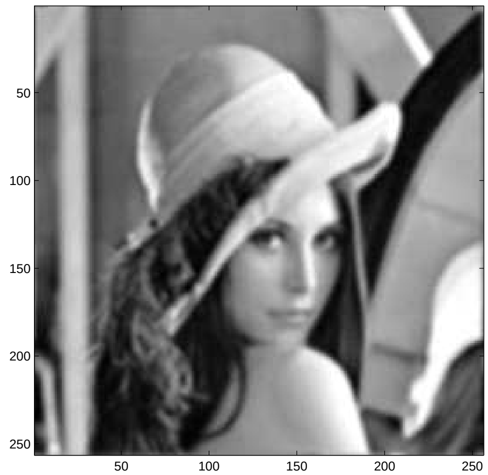

Fig. 2. Interlaced image in high-resolution grid.

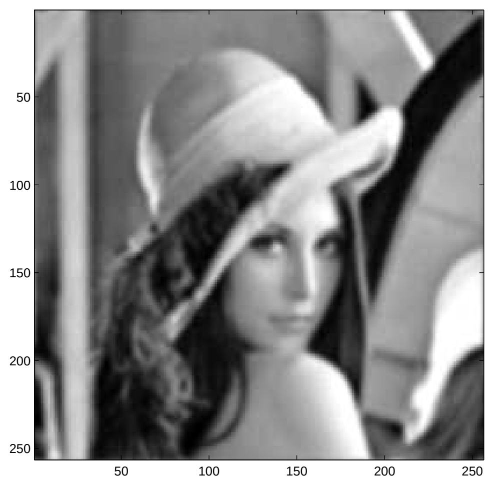

Fig. 3. Bilinear interpolation of first low-resolution frame with upsampling ratio 2.

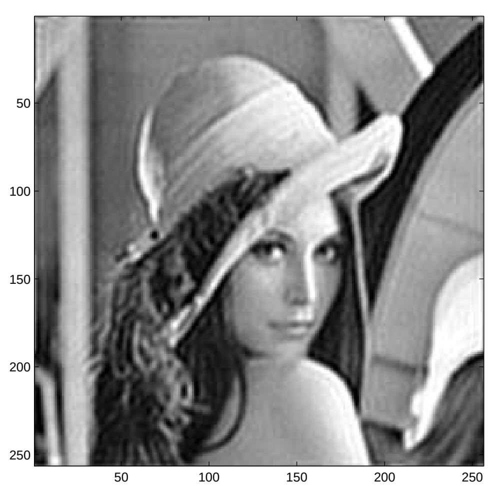

Fig. 4. Reconstruction of Lena image using EM with upsampling ratio 2.

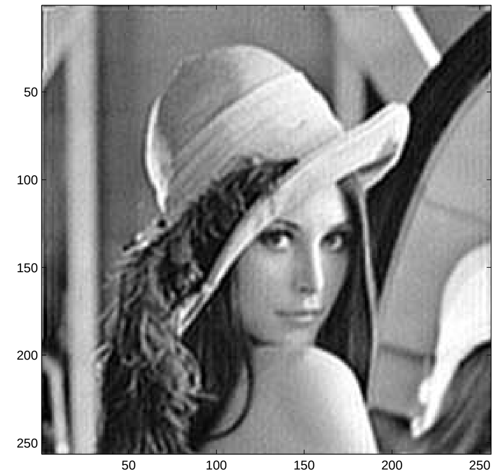

Fig. 5. Reconstruction of Lena image using LMMSE with upsampling ratio 2.

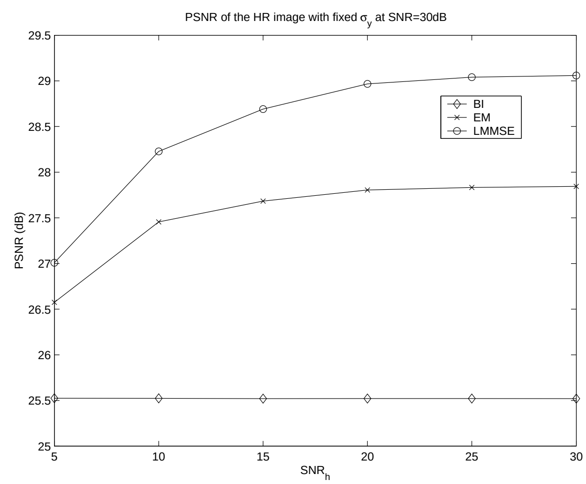

Fig. 6. Plot of PSNR of the reconstructed high-resolution image vs. SNR of the PSF noise.

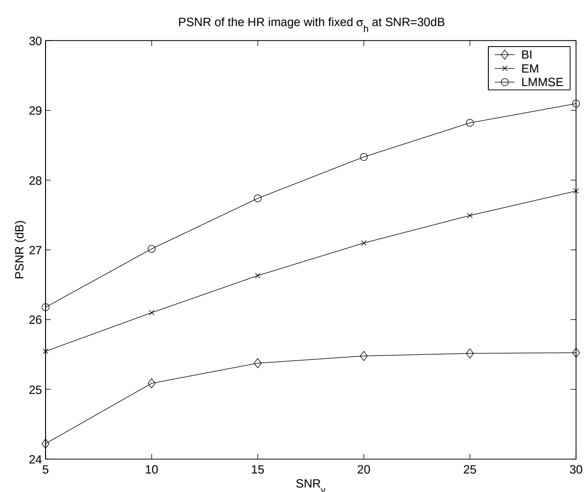

Fig. 7. Plot of PSNR of the reconstructed high-resolution image vs. SNR of the AWGN noise.