RESOLUTION ENHANCEMENT OF VIDEO SEQUENCES WITH ADAPTIVELY WEIGHTED LOW-RESOLUTION IMAGES AND SIMULTANEOUS ESTIMATION OF THE REGULARIZATION PARAMETER

Hu He, Lisimachos P. Kondi

Dept. of Electrical Engineering, SUNY at Buffalo, Buffalo, NY 14260

ABSTRACT

In many imaging systems, the resolution of the detector array of the camera is not sufficiently high for a particular application. Furthermore, the capturing process introduces additive noise and the point spread function of the lens and the effects of the finite size of the photo-detectors further degrade the acquired video frames. The goal of resolution enhancement is to estimate a high-resolution image from a sequence of low-resolution images while also compensating for the above-mentioned degradations. In this paper, we propose a technique for image resolution enhancement with adaptively weighted low-resolution images (channels) and simultaneous estimation of the regularization parameter. The weight coefficients work as the cross-channel fidelity to each low-resolution image, while the regularization parameter acts as the within-channel balance between data and prior model for each channel. Experimental results are presented and conclusions are drawn.

1. INTRODUCTION

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. An example scenario is the case of a camera that is mounted on an aircraft and is imaging objects in the far field. The vibrations of the aircraft will generally provide the necessary motion between the focal plane array and the scene, thus yielding frames with subpixel motion between them and minimal occlusion effects.

In this paper, we extend our previous results in [1] and [2] by proposing a technique for the adaptive update of cross-channel weights and the simultaneous estimation of the regularization parameter of each channel. The rest of the paper is organized as follows. In section 2, a regularized cost function is addressed for image resolution enhancement. In section 3, we rewrite the cost function in multi-channel form to establish the relationship between the overall regularization parameter and the individual parameters for each channel. We then develop our technique for the estimation of both the cross-channel weights and the regularization parameter. In section 4, experimental results are presented. Finally, in section 5, conclusions are drawn.

2. MAP-BASED RESOLUTION ENHANCEMENT

The problem of resolution enhancement is an active research area. In this review, we concentrate on Bayesian methods. Maximum A Posteriori (MAP) estimation with an edge preserving Huber-Markov random field image prior is studied in [3], [4]. MAP based resolution enhancement with simultaneous estimation of registration parameters has been proposed in [1], [2], [5]. Resolution enhancement with estimation of the regularization parameter for each low-resolution image has been studied in [6]. In our recent work in [2], a method for simultaneous estimation of the regularization parameter was also proposed. In this paper, we extend the result in [2] to more general case considering a different noise level for each low-resolution image (channel).

The image degradation process is modeled by a linear blur, motion, subsampling by pixel averaging and an additive Gaussian noise process. 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 $N = L_1 N_1 L_2 N_2$ and L_1 and L_2 represent the down-sampling factors in the horizontal and vertical directions respectively. The kth low-resolution frame can be denoted as $\mathbf{y} = \left[\mathbf{y}_1^T, \mathbf{y}_2^T, \cdots, \mathbf{y}_p^T\right]^T$ for $k = 1,2,\cdots p$. The system can be modeled as

$$\mathbf{y} = \mathbf{W_s z} + \mathbf{n}, \tag{1}$$

where matrix $\mathbf{W_s} = \left[\mathbf{W}_{\mathbf{s},1}^T, \mathbf{W}_{\mathbf{s},2}^T, \cdots, \mathbf{W}_{\mathbf{s},p}^T\right]^T$ contains the operation of blur, motion, subsampling by pixel averaging. The vector $\mathbf{s}_k = \left[s_{k,1}, s_{k,2}, \cdots s_{k,K}\right]^T$ contains the K registration parameters for frame k measured in reference to a fixed high-resolution grid. In contrast with [1], [2] and [5], we assume that the noise samples $\mathbf{n} = \left[\mathbf{n}_1, \mathbf{n}_2, \cdots \mathbf{n}_p\right]^T$ are independent Gaussian, but with variances σ_k^2 that are not necessarily identical. We propose to determine **z** and **s** by minimizing the following regularized cost function:

$$L(\mathbf{z},\mathbf{s}) = \sum_{k=1}^{p}\left(\frac{1}{\sigma_k^2}\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2\right) + \frac{1}{\lambda}\left\|\mathbf{D}\mathbf{z}\right\|^2, \quad (2)$$

where λ is the tuning parameter and **D** represents a high-pass filter. The above cost function can be also obtained from MAP estimation using a Gaussian-Markov random field (GMRF) image prior model as in [1], [2], [5].

3. ADAPTIVE UPDATE OF CHANNEL WEIGHT AND ESTIMATION OF THE REGULARIZATION PARAMETER

Alternatively, we can rewrite the cost function in Equation (2) as the sum of individual smoothing functionals for each of the *p* low-resolution images as:

$$\begin{aligned} L(\mathbf{z},\mathbf{s}) &= \sum_{k=1}^{p}\left\{\frac{1}{\sigma_k^2}\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2 + \frac{1}{\lambda_k}\left\|\mathbf{D}\mathbf{z}\right\|^2\right\} \\ &= \sum_{k=1}^{p}\frac{1}{\sigma_k^2}\left\{\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2 + \frac{\sigma_k^2}{\lambda_k}\left\|\mathbf{D}\mathbf{z}\right\|^2\right\}, \quad (3) \\ &= \sum_{k=1}^{p}\frac{1}{\sigma_k^2}\left\{\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2 + \alpha_k\left\|\mathbf{D}\mathbf{z}\right\|^2\right\} \end{aligned}$$

where α_k is the regularization parameter for each channel defined as

$$\alpha_k = \frac{\sigma_k^{\,2}}{\lambda_k}, \quad (4)$$

and λ_k is the tuning parameter of each channel. If the noise samples have different variance for each channel, we can not drop the factor $1/\sigma_k^2$ from Equation (3). Instead, we should give different fidelity to each channel. As in [2], we assume no prior information of the noise variance σ_k^2 and we propose the following method to adaptively update the weight assigned to each channel and the regularization parameter:

1. Introduce the cross-channel weight coefficient c_k and rewrite the cost function as:

$$L(\mathbf{z},\mathbf{s}) = \frac{1}{2}\sum_{k=1}^{p} c_k\left\{\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2 + \alpha_k\left\|\mathbf{D}\mathbf{z}\right\|^2\right\}, \quad (5)$$

where coefficient c_k satisfies: (a) c_k is proportional to $1/\sigma_k^2$, or equivalently, inversely proportional to the residual norm $\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2$; (b) $\sum_{k=1}^{p} c_k = c$, where c is a positive constant. In this paper, we choose $c = p$, the number of low-resolution images. This constraint can avoid the trivial solution of the cost function. The solution for criteria (a) and (b) with choice $c = p$ is:

$$c_k = \frac{R_{ave}}{\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2}, \quad (6)$$

where R_{ave} is the average residue norm defined as

$$R_{ave} = \frac{p}{\sum_{k=1}^{p}\left\{\frac{1}{\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2}\right\}}. \quad (7)$$

2. An iterative algorithm can be used to reconstruct the high-resolution image, i.e., **z** can be updated iteratively as

$$\hat{z}_r^{n+1} = \hat{z}_r^{n} - \varepsilon^n g_r(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n) \quad (8)$$

until convergence is reached. *n* is the iteration number starting from 0. The cost function can be minimized using the coordinate-descent method. This iterative method starts with an initial estimate of **z** obtained using interpolation from a low resolution frame. Then, for a fixed **z**, the cost function is minimized with respect to **s**. Thus, the motion of each frame is estimated. Then, for fixed **s**, a new estimate for **z** is obtained. This procedure continues until convergence is reached, i.e., **z** and **s** are updated in a cyclic fashion.

The gradient $g_r(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n)$ can be obtained from

$$g_r(\mathbf{z},\mathbf{s}) = \frac{\partial L(\mathbf{z},\mathbf{s})}{\partial z_r} = \sum_{k=1}^{p} c_k\left\{\left(\mathbf{W}_{\mathbf{s},k}^T\mathbf{W}_{\mathbf{s},k}\right)\mathbf{z} - \mathbf{W}_{\mathbf{s},k}^T\mathbf{y}_k + \alpha_k\left(\mathbf{D}^T\mathbf{D}\right)\mathbf{z}\right\}. \quad (9)$$

3. Following the same procedure as in [2], [7], we impose the following requirements for each α_k in the resolution enhancement scenario: it should be a function of the regularized noise power of the data and its choice should yield a convex functional whose minimization would give the high-resolution image. The imposed properties on α_k require a linear function between α_k and each term of the cost function:

$$\alpha_k = f\left(L_k(\alpha_k, \mathbf{z})\right) = \gamma_k\left\{\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2 + \alpha_k\left\|\mathbf{D}\mathbf{z}\right\|^2\right\}. \quad (10)$$

Thus, the choice of regularization parameter α_k for the regularization functional is given by

$$\alpha_k = \frac{\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2}{\frac{1}{\gamma_k} - \left\|\mathbf{D}\mathbf{z}\right\|^2}. \quad (11)$$

Also, following the same procedure for convergence requirement as in [7], we get

$$\frac{1}{\gamma_k} > \frac{\varepsilon p\left\|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\right\|^2 \phi_{\max}(\mathbf{D}^T\mathbf{D})}{2 - \varepsilon p\,\phi_{\max}(\mathbf{W}_{\mathbf{s},k}^T\mathbf{W}_{\mathbf{s},k})} + \left\|\mathbf{D}\mathbf{z}\right\|^2. \quad (12)$$

where $\phi_{\max}(\cdot)$ stands for the maximum eigenvalue of a matrix. We show in [2] that

$$\phi_{\max}(\mathbf{W}_{\mathbf{s},k}^T \mathbf{W}_{\mathbf{s},k}) = \frac{1}{(L_1 L_2)^2}. \quad (13)$$

Therefore, inequality (12) becomes

$$\frac{1}{\gamma_k} > \frac{\varepsilon p \phi_{\max}(\mathbf{D}^T \mathbf{D})}{2 - \left(\varepsilon p / (L_1 L_2)^2\right)} \left\| \mathbf{y}_k - \mathbf{W}_{\mathbf{s},k} \mathbf{z} \right\|^2 + \left\| \mathbf{Dz} \right\|^2. \quad (14)$$

If we select step size ε as

$$\varepsilon = \frac{2}{p} \left(\frac{(L_1 L_2)^2}{(L_1 L_2)^2 \phi_{\max}(\mathbf{D}^T \mathbf{D}) + 1} \right), \quad (15)$$

inequality (12) will become

$$\frac{1}{\gamma_k} > \left\| \mathbf{y}_k - \mathbf{W}_{\mathbf{s},k} \mathbf{z} \right\|^2 + \left\| \mathbf{Dz} \right\|^2. \quad (16)$$

As in [2], a simple fixed choice $1/\gamma_k = 2\left\| \mathbf{y}_k \right\|^2$ is used in this paper to satisfy convergence. Thus, we can obtain the simultaneous estimation of regularization parameter for each channel as

$$\alpha_k = \frac{\left\| \mathbf{y}_k - \mathbf{W}_{\mathbf{s},k} \mathbf{z} \right\|^2}{2\left\| \mathbf{y}_k \right\|^2 - \left\| \mathbf{Dz} \right\|^2}. \quad (17)$$

Now, the weight coefficients c_k in Equation (6) work as the cross-channel fidelity, while the regularization parameter α_k in Equation (17) acts as the within-channel balance between data and prior model for each channel.

4. EXPERIMENTAL RESULTS

To test the validity of our algorithm, we used the 256x256 "cameraman" image for a synthetic test in this paper. The original high-resolution image was blurred with a Point Spread Function (PSF), globally shifted, subsampled and added with AWGN noise to produce a sequence of 16 low-resolution images. The PSF was generated from a Gaussian blur with variance 1.7. The down/up sampling ratio was $L_1= L_2=4$. Global shift $\mathbf{s}_k^T$ belongs to the set generated from the Cartesian product of $\{0, 1, 2, 3\} \times \{0, 1, 2, 3\}$. Four cases, Case I-IV, as listed in Table 1, were tested. In case IV, eight unrelated low-resolution images were generated from the "Lena" image. The first frame was selected as reference frame and bilinear interpolation of the first frame was chosen as the first estimate of high-resolution image **z**. The algorithm was carried out for 20 iterations or until convergence was reached when $\left\| \hat{\mathbf{z}}^{n+1} - \hat{\mathbf{z}}^n \right\|^2 / \left\| \hat{\mathbf{z}}^n \right\|^2 < 10^{-6}$.

Table 1. Four cases of synthetic test for "cameraman"

	p=16: Number of low-resolution frames	σ_k^2
Case I	16 frames of "Camerman"	1
Case II	16 frames of "Camerman"	k
Case III	16 frames of "Camerman"	k^2
Case IV	8 frames of "Camerman" followed by 8 unrelated frames of "Lena"	1

The PSNR of the reconstructed image for "cameraman" using the three methods (Bilinear, Proposed Method, Non-channel-weighted Simultaneous Method) are listed in Table 2. The bilinear interpolation of the reference frame is shown in Fig. 1. The reconstructed images from Proposed Method and Non-channel-weighted Simultaneous Method of "cameraman" in Case III, IV are listed in Fig. 2~5, respectively.

Table 2 Results of "cameraman" using the three methods

PSNR (dB)	Bilinear	Proposed Method	Non-channel-weighted Simultaneous Method
Case I	21.26	24.07	24.07
Case II	21.26	23.96	23.94
Case III	21.26	23.70	22.78
Case IV	21.26	24.00	16.07

5. CONCLUSION

We have proposed a technique for adaptively weight update of each low-resolution image (channel), and simultaneous estimation of the regularization parameter for digital image resolution enhancement. The weight coefficients work as the cross-channel fidelity to each low-resolution image, while the regularization parameter acts as the within-channel balance between data and prior model for each channel. Our experimental results demonstrate the performance of the proposed algorithm, which can be easily applied to real data. In all the cases considered, the proposed algorithm gives a good reconstruction.

6. REFERENCES

[1] L.P. Kondi, D. Scribner, J. Schuler, "A Comparison of Digital Image Resolution Enhancement Techniques", *SPIE AeroSense*, Orlando, FL, April 2002.

[2] H. He, L.P. Kondi, "Resolution Enhancement of Video Sequences with Simultaneous Estimation of the Regularization Parameter", *Journal of Electronic Imaging*, to appear.

[3] R. R. 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] R. R. Schultz and R. L. Stevenson, “A Bayesian approach to image expansion for improved definition,” *IEEE Trans. Image Processing*, vol. 3, pp. 233–242, May 1994.

[5] R. C. Hardie, K. J. Barnard, “Joint MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images,” *IEEE Trans. Image Processing*, Vol. 6, No. 12, pp. 621–1633, December 1997.

[6] E. S. Lee, M. G. Kang, “Regularized Adaptive High-Resolution Image Reconstruction Considering Inaccurate Subpixel Registration”, *IEEE Trans. Image Processing*, Vol. 12, No. 7, pp. 826–837, July 2003.

[7] M. G. Kang, A. K. Katsaggelos, "Simultaneous Multichannel Image Restoration and Estimation of the Regularization parameters", *IEEE Trans. Image Processing*, May 1997, pp. 774-778.

Fig. 1. Bilinear interpolation of reference frame

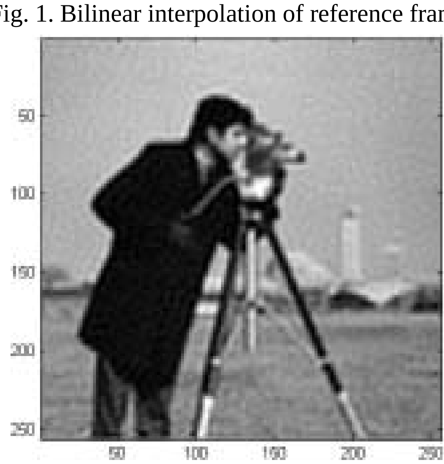

Fig.2 Reconstructed image of “cameraman” from Proposed Method (Case III)

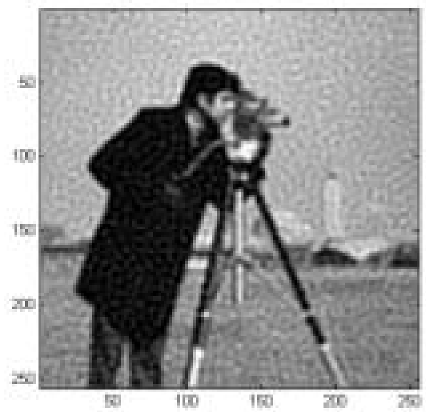

Fig. 3 Reconstructed image using non-channel-weighted Simultaneous Method (Case III)

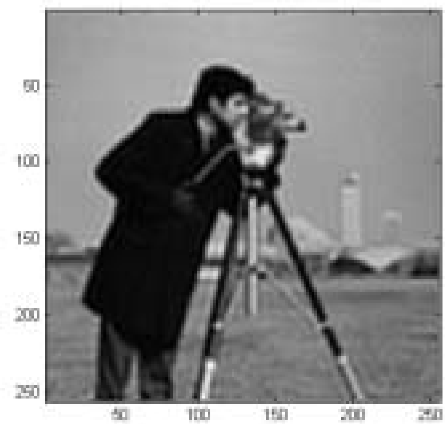

Fig.4 Reconstructed image of “cameraman” from Proposed Method (Case IV)

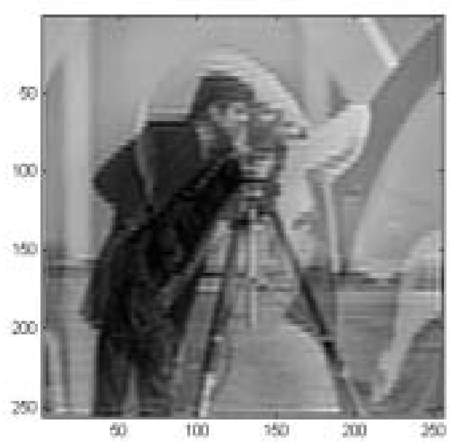

Fig. 5 Reconstructed image using non-channel-weighted Simultaneous Method (Case IV)