CHOICE OF THRESHOLD OF THE HUBER-MARKOV PRIOR IN MAP-BASED VIDEO RESOLUTION ENHANCEMENT

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

Abstract

MAP (Maximum A Posteriori) -based resolution enhancement technique with Huber-Markov random field (HMRF) as the image prior has been proposed in the literature, and better preserves image discontinuities when compared with a Gaussian prior model. The reconstruction relies on the choice of Huber function parameter, or threshold T. There is no explicit selection of T in the previous studies. In this paper, we propose a method for choosing the threshold of the HMRF image prior in MAP based resolution enhancement. The method is based on the fact that the threshold T of the Huber function in the HMRF image priors is physically the trade-off between high-frequency components and low-frequency components for imagery data. High-pass filtering using the discrete Laplacian kernel along with the Huber function is used as the smoothness measure. When the high-passed value is less than T, the measure is a parabola function, while when the value is larger than T, the smoothness measure becomes a linear function. We hence define two different sets and derive the MAP estimator as a function of T. Experimental results are presented and conclusions are drawn.

Keywords: *Resolution enhancement, Huber-Markov random field, threshold, MAP estimation.*

1. INTRODUCTION

The goal of resolution enhancement 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 this paper, we propose a method for the selection of the threshold of the HMRF image prior in MAP-based resolution enhancement. The rest of the paper is organized as follows. In section 2, the MAP-based resolution enhancement algorithm is presented along with the HMRF-based prior model. In section 3, experimental results are presented. Finally, in section 4, conclusions are drawn.

2. MAP-BASED RESOLUTION ENHANCEMENT TECHNIQUE WITH HMRF IMAGE PRIOR

2.1 Review of MAP Based Technique

Several resolution enhancement techniques have been proposed in the literature. In this review, we concentrate on Bayesian methods. Maximum A Posteriori (MAP) estimation with an edge preserving Huber-Markov random field (HMRF) image prior is studied in [1], [2]. In our previous work [3], we proposed a joint MAP registration (estimation of the relative motion between frames) and high-resolution image estimation algorithm using an HMRF prior model. In [3], two approaches were proposed, which differ in the selection of image smoothness measure. The first approach employs a measure that is based on a discrete Laplacian kernel, while the second approach uses a finite difference approximation of second order derivatives at each pixel of the high-resolution image estimate. There are two important parameters in this technique, the tuning parameter λ and the threshold T of the Huber function. The former one is further studied in [4], leaving the threshold T as the focus of this paper.

2.2 Observation Model

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

CCECE 2004- CCGEI 2004, Niagara Falls, May/mai 2004
0-7803-8253-6/04/$17.00 ©2004 IEEE

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 image through blurring, motion shift and subsampling. Let the *k*th low-resolution frame be denoted as $\mathbf{y}_k = [y_{k,1}, y_{k,2}, \cdots y_{k,M}]^t$ for $k = 1,2,\cdots p$ and where $M = N_1 N_2$. The full set of *p* observed low-resolution images can be denoted as

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

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 $m = 1,2,\cdots M$ and $k = 1,2,\cdots p$. The weight $w_{k,m,r}(\mathbf{s}_k)$ represents the "contribution" of the *r*th high-resolution pixel to the *m*th low resolution observed pixel of the *k*th frame. The vector $\mathbf{s}_k = [s_{k,1}, s_{k,2}, \cdots s_{k,K}]^t$, 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 will be 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

$$\mathbf{y} = \mathbf{W}_\mathbf{s} \mathbf{z} + \mathbf{n}, \quad (3)$$

where matrix $\mathbf{W}_\mathbf{s} = [\mathbf{W}_{\mathbf{s},1}^T, \mathbf{W}_{\mathbf{s},2}^T, \cdots, \mathbf{W}_{\mathbf{s},p}^T]^T$ contains the operation of blur, motion, subsampling by pixel averaging.

2.3 HMRF Image Prior and Laplacian Smoothness Kernel

A general form of the HMRF density is

$$P_r(\mathbf{z}) = A \exp\left\{ -\frac{1}{\lambda} \sum_{c \in C} \rho\left(\mathbf{d}_c^t \mathbf{z}\right) \right\}, \quad (4)$$

where *A* is a constant, λ is the temperature or tuning parameter of the density, *c* is a local group of points called clique and *C* denotes the set of all cliques in the image. $\mathbf{d}_c$ is a coefficient vector for clique *c* and $\rho(\bullet)$ is the Huber function:

$$\rho(x) = \begin{cases} x^2, & |x| \le T \\ T^2 + 2T(|x| - T), & |x| > T \end{cases}, \quad (5)$$

where *T* is the threshold of the Huber function.

In this paper, we use the second smooth measure in [3], measuring the image smoothness at a given pixel using a Laplacian kernel, i.e.,

$$\mathbf{d}_i^t \mathbf{z} = \sum_{j=1}^{N} d_{i,j} z_j, \quad (6)$$

with $d_{i,j}$ a Laplacian smoothness kernel defined as

$$d_{i,j} = \begin{cases} 1 & for\ i = j \\ -1/4 & for\ i, j : z_j\ is\ a\ cardinal\ neighbor\ of\ z_i \end{cases}. \quad (7)$$

The corresponding cost function is:

$$L(\mathbf{z}, \mathbf{s}) = \frac{1}{2\sigma_\eta^2} \sum_{m=1}^{pM} \left(y_m - \sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r \right)^2 + \frac{1}{2\lambda} \sum_{i \in N_T} \left(\sum_{j=1}^{N} d_{i,j} z_j \right)^2 + \frac{1}{2\lambda} \sum_{i \in N_T^C} \left\{ T^2 + 2T\left(\left| \mathbf{d}_i^t \mathbf{z} \right| - T \right) \right\}, \quad (8)$$

where N_T is set $\{ i \mid i \in (1,2,\cdots,N), |\mathbf{d}_i^t \mathbf{z}| \le T \}$ and N_T^C is the complement set of N_T.

2.4 Joint MAP Registration Algorithm

The above cost functions 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. In order to update the estimate **z**, we first estimate

$$\hat{\mathbf{s}}_k^n = \arg\min_{\mathbf{s}_k} \left\{ \sum_{m=1}^{M} \left(y_m - \sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r \right)^2 \right\}. \quad (9)$$

Also, the gradient can be obtained from

$$g_k(\mathbf{z}, \mathbf{s}) = \frac{\partial L(\mathbf{z}, \mathbf{s})}{\partial z_k}. \quad (10)$$

The step size ε^n can be found by solving

$$\frac{\partial L(\hat{\mathbf{z}}^{n+1}, \hat{\mathbf{s}}^n)}{\partial \varepsilon^n} = 0. \quad (11)$$

Then **z** can be updated recursively as

$$\hat{z}_k^{n+1} = \hat{z}_k^n - \varepsilon^n g_k(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n) \quad (12)$$

until convergence is reached. *n* is the iteration number starting from 0.

2.5 Choice of Threshold *T* of HMRF Prior

Threshold *T* of HMRF image priors is physically the trade-off between high-frequency components and low-frequency components for imagery data. For the synthetic test, the original high-resolution image is known, threshold T can be set such that

$$\frac{f_T(\mathbf{z})}{f(\mathbf{z})} = \rho \,. \qquad (13)$$

Here $f(\mathbf{z})$ is the discrete smoothing norm from the 2-D filtering using Laplacian kernel, and $f_T(\mathbf{z})$ is the smoothness norm when threshold T is taken into consider (any value lower than T is set to zero). ρ is a predetermined cutoff ratio, roughly corresponding to the percentage of high-frequency components in the image. Therefore, T can be determined in the synthetic test for a specific ρ . And ρ is usually chosen within $(0, 0.5]$ under the assumption than there is more energy for the low-frequency components than that for the high-frequency components. If no prior information of the energy distribution is available, ρ can be set as 0.5 to allow enough high-frequency components appear in the constructed high-resolution image, and we name the corresponding T as $T_{0.5}$. In Fig. 1, we show the plot of threshold versus ρ for "Cameraman" and "Lena" image, with $T_{0.5}$=2 and $T_{0.5}$=2.25 respectively. For real data, the threshold T can be selected within a reasonable range, which is usually chosen to cover the range of the threshold T found from the synthetic tests.

From Equation (5), we can easily see that if the threshold T is large enough, ρ becomes one and the HMRF prior reduces to a GMRF prior.

3. EXPERIMENTAL RESULTS

A number of experiments were conducted, some of which are presented here. To test the performance of our algorithm, we first use the 256x256 "Cameraman" and "Lena" images for synthetic tests. 16 low-resolution frames were generated via Gaussian blurring (variance 1.7), global motion shift of $\{0,1,2,3\}\times\{0,1,2,3\}$ in the high-resolution grid, subsampling by L_1 = L_2 =4 and addition of AWGN noise (variance 1). The threshold T is set to $T_{0.5}$. The first frame is selected as reference frame and bilinear interpolation plus CLS (Constraint Least Squared) filtering of the first frame is chosen as the first estimate of the high-resolution image $\mathbf{z}$. Algorithm is carried out for 20 iterations or until convergence is reached when $\left\|\hat{\mathbf{z}}^{n+1}-\hat{\mathbf{z}}^{n}\right\|^2 / \left\|\hat{\mathbf{z}}^{n}\right\|^2 < 10^{-6}$.

The PSNR of the reconstructed images for "Cameraman" and "Lena" using three methods (Bilinear, GMRF and HMRF/$T_{0.5}$) are listed in Table 1.

The bilinear interpolation of the first low-resolution frame and the reconstructed "Cameraman" and "Lena" images from GMRF and HMRF/$T_{0.5}$ are shown in Fig. 2(a)-(c), 3(a)-(c), respectively.

Table 1 Results of "Cameraman" using the three methods

PSNR (dB)	Bilinear Interpolation	GMRF	HMRF/$T_{0.5}$
Cameraman	21.26	23.87	23.94
Lena	23.20	26.55	26.75

4. CONCLUSION

As expected, HMRF gave a good reconstruction result and preserve more information at the edges than when using Gaussian prior model. In this paper, we proposed a method for the choice of threshold T for HMRF prior, which is the physical trade-off between high-frequency components and low-frequency components for imagery data. Synthetic test verify the validity of out algorithm.

References

[1] 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.

[2] 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.

[3] 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, pp933 – 936, 2003

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

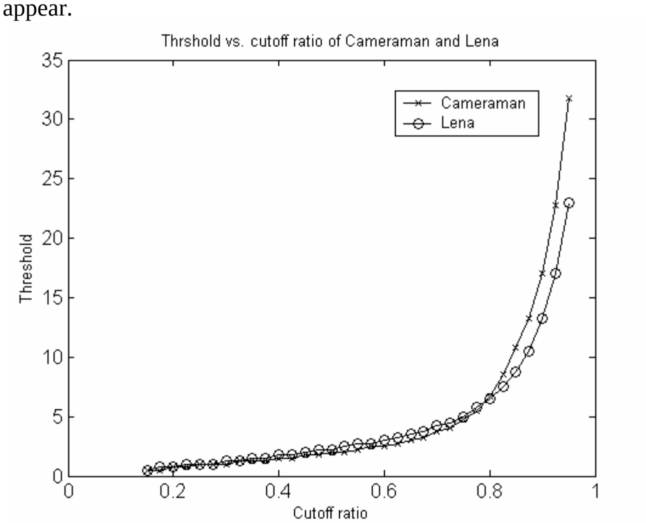

Fig. 1. Threshold vs. cutoff ratio of "Cameraman" and "Lena" images, with $T_{0.5}$=2 and $T_{0.5}$=2.25 respectively.

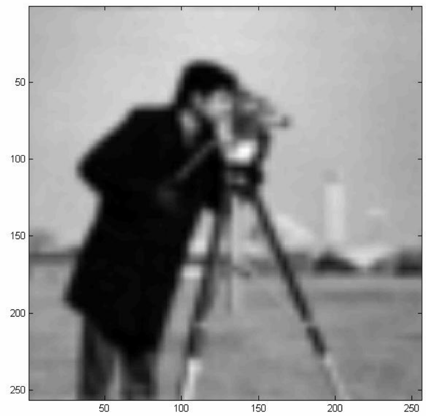

Fig. 2(a). Bilinear interpolation of first low-resolution frame of "Cameraman"

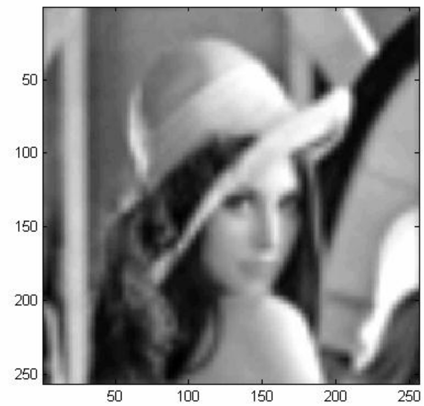

Fig. 3(a). Bilinear interpolation of first low-resolution frame of "Lena"

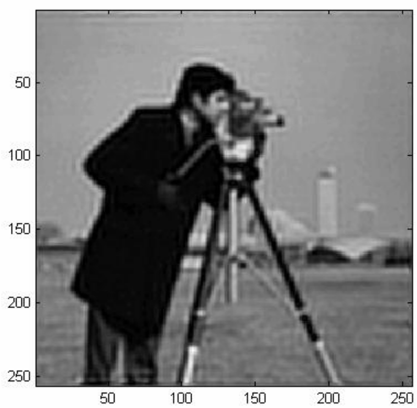

Fig. 2(b). Reconstruction of "Cameraman" using GMRF as the image prior

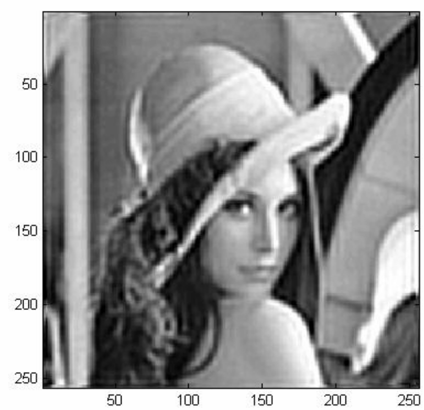

Fig. 3(b). Reconstruction of "Lena" using GMRF as the image prior

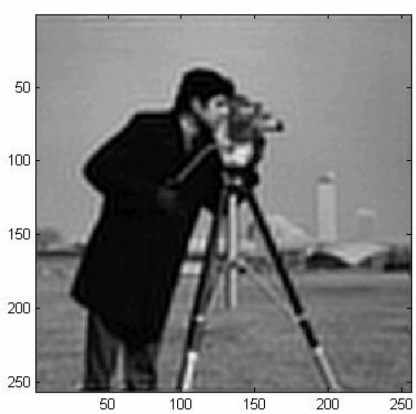

Fig. 2(c). Reconstruction of "Cameraman" using HMRF as the image prior and threshold of Huber function set to $T_{0.5}$

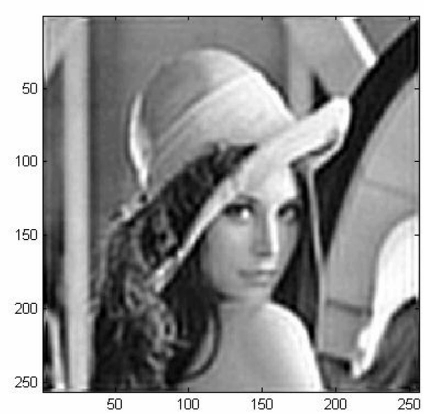

Fig. 3(c). Reconstruction of "Lena" using HMRF as the image prior and threshold of Huber function set to $T_{0.5}$