SuperResolution Image Reconstruction Using a Hybrid Bayesian Approach

Tao Wang, Yan Zhang, and Yong Sheng Zhang

Zhengzhou Institute of Surveying and Mapping, No. 66 Longhai Middle Road, Zhengzhou 450052, China
wangtaoyn1@163.com

Abstract. There are increasing demands for high-resolution (HR) images in various applications. Image superresolution (SR) reconstruction refers to methods that increase image spatial resolution by fusing information from either a sequence of temporal adjacent images or multi-source images from different sensors. In the paper we propose a hybrid Bayesian method for image reconstruction, which firstly estimates the unknown point spread function(PSF) and an approximation for the original ideal image, and then sets up the HMRF image prior model and assesses its tuning parameter using maximum likelihood estimator, finally computes the regularized solution automatically. Hybrid Bayesian estimates computed on actual satellite images and video sequence show dramatic visual and quantitative improvements in comparison with the bilinear interpolation result, the projection onto convex sets (POCS) estimate and Maximum A Posteriori (MAP) estimate.

1 Introduction

There are increasing demands for high-resolution (HR) images in various applications. Although the most direct way to increase spatial resolution is to use a HR image acquisition system, fabrication limitations and high cost for high precision optics and image sensors are always prohibitive concerns in many commercial applications. Therefore, the new image SR reconstruction approach, which is capable of generating a HR image from multiple low-resolution (LR) images, has been a hot research topic recently[1].Since Tsai and Huang's work[2], many work has been reported in the literature, including the weighted least-squares algorithm[1], the nonuniform interpolation approach[1], the POCS method[3-4] and MAP Bayesian approach[5-7]. Among these algorithms, the Bayesian approach is most notable for it robustness and flexibility in modeling noise characteristics and a priori knowledge about the solution. Assuming that the noise process is white Gaussian, the Bayesian estimation with convex energy functions ensures the uniqueness of the solution. But existing Bayesian reconstruction methods suffer from several impractical assumptions. Previous research often assumes that PSF is definitely known during reconstruction, which is impossible for actual images reconstruction as many uncertain blurring factors are involved during imaging process. Further, the image prior model founded upon the upsampled LR image greatly affects the quality of the reconstruction result as the LR images are

I. King et al. (Eds.): ICONIP 2006, Part II, LNCS 4233, pp. 412 – 419, 2006.

already contaminated and the resulted prior model is not robust to noise. Finally, the edge threshold parameter of the image prior model needs to be adjusted by experienced experts empirically, which limits the wide usage of the Bayesian estimator.

Therefore, we propose a novel hybrid Bayesian estimator for SR image reconstruction. Under the Bayesian framework, it deconvolutes the upsampled LR image to access PSF and approximation value for the ideal HR image with APEX algorithm first, and then models the HMRF image prior model and assesses its edge threshold parameter through maximum likelihood (ML) estimation, finally regularizes the ill-posed reconstruction process automatically.

2 Statement on Hybrid Bayesian Reconstruction Algorithm

Above all we formulate an observation model that relates the original HR image to the observed LR image. Consider the desired HR image $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$, $N = L_1N_1 \times L_2N_2$, which is sampled at or above the Nyquist rate from a hypothetically bandlimited continuous scene. L_1 and L_2 are the horizontal and vertical down-sampling factors, respectively. Let the kth LR image be denoted as $\mathbf{y}^{(k)} = [y^{(k)}_1, y^{(k)}_2, \ldots, y^{(k)}_M]^T$, $M = N_1 \times N_2$. During the imaging process, the observed LR image result from warping, blurring, and subsampling operators performed on $\mathbf{x}$ and is also corrupted by additive noise, we can then represent the observation model as

$$\mathbf{y}_k = \mathbf{DHT}_k\mathbf{x} + \mathbf{n}_k = \mathbf{W}_k\mathbf{x} + \mathbf{n}_k \quad \text{for } 1 \le k \le \mathrm{p}. \tag{1}$$

where $\mathbf{T}_k$ is a warp matrix, $\mathbf{H}$ represents a blur matrix **,** $\mathbf{D}$ is a subsampling matrix and $\mathbf{n}_k$ represents a noise vector, assumed to be Gaussian, white and stationary, p is the number of images.

The SR image reconstruction problem is ill-posed. A well-posed problem can be formulated under the MAP stochastic framework by introducing a priori constraint,

$$\mathbf{x} = \arg\max\{\log P(\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_\mathrm{p} | \mathbf{x}) + \log P(\mathbf{x})\}. \tag{2}$$

Both the priori image model $P(\mathbf{x})$ and the conditional density $P(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_\mathrm{p} | \mathbf{x})$ will be defined by a priori knowledge concerning $\mathbf{x}$ and the statistical information of noise. If the motion estimation error between images is assumed to be independent and noise is assumed to be an independent identically distributed zero mean Gaussian distribution, the conditional density can be expressed in the compact form

$$P(\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_\mathrm{p} | \mathbf{x}) = \prod_{k=0}^{\mathrm{p}} P(\mathbf{y}_k | \mathbf{x}) = \prod_{k=0}^{\mathrm{p}} \left\{ \frac{1}{(2\pi)^{\frac{N}{2}} \sigma^N} \exp\left\{ -\frac{1}{2\sigma^2} \|\mathbf{y}_k - \mathbf{W}_k\mathbf{x}\|^2 \right\} \right\}. \tag{3}$$

where σ^2 is error variance, $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_\mathrm{p}]^T$, $\mathbf{W} = [\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_\mathrm{p}]^T$.

In order to reconstruct the high-frequency information lost through imaging, we take the HMRF prior model, which represents piecewise smooth image data[5],

$$P(\mathbf{x}) = \frac{1}{Z}\exp\{-\frac{1}{2\beta}\sum_{c \in S}\rho(d_l^t\mathbf{x}, \alpha)\} = \frac{1}{Z}\exp\{-\frac{1}{2\beta}\sum_{m,n}\sum_{l=1}^{4}\rho(d_l^t\mathbf{x}, \alpha)\}. \tag{4}$$

where Z is a normalizing constraint, β is the temperature parameter, c is a local group of pixels contained within the image cliques S, α is the edge threshold parameter separating the quadratic and linear regions. The quantity $d_l^t\mathbf{x}$ measures the second-order finite differences in four directions at each pixel in the HR image, small in smooth locations and large at edges[5].The likelihood of edges in the data is controlled by the Huber penalty function

$$\rho(x,\alpha)=\begin{cases} x^2, & |x|\le\alpha, \\ 2\alpha|x|-\alpha^2, & |x|>\alpha, \end{cases}. \tag{5}$$

The regularized solution is then equivalent to minimizing the cost function

$$U(\mathbf{x})=\|\mathbf{y}-\mathbf{W}\mathbf{x}\|^2+\sum_{m,n}\sum_{l=1}^{4}\rho(d_l^t\mathbf{x},\alpha)\cdot \tag{6}$$

The HMRF prior model should be founded on the ideal HR image and parameter α should also be decided upon it. But in the existing MAP research, the upsampling LR image is usually taken as substitute for the ideal HR image and α is set empirically. However, the ideal HR image can't be approximated by its degraded version because the LR image is blurred and noisy. Parameter α estimated on a blurred image has too high a value and leads to over-smoothed solutions. Parameter α estimated on a noisy image is too low, and provides insufficient regularization, leading to noisy solutions. A bad initialization for the prior model often leads to degenerated solutions[7]. The Bayesian estimator is only significant and supplies good regularized estimate in the case of ideal HR image Therefore, an approximation of **x** has to be accurately determined before reconstruction.

We choose APEX algorithm to compute the approximation of **x** as the deconvoluted result produced by APEX algorithm is sufficiently close to the original image to enable us to set up an accurate HRMF prior model. Moreover, the unknown PSF can also be determined. In the following, we detail how to get an approximation of **x** with APEX algorithm, how to estimate parameter α from the approximation image, and how to generate a reconstruction estimate automatically.

3 Hybrid Bayesian Reconstruction Solution

3.1 APEX Prior Blind Deconvolution

The APEX[8] method is a FFT-based direct blind deconvolution technique, which is applicable to a restricted two-dimensional radially symmetric shift-invariant G class blurs. The OTF (Optical Transfer Function) form of G class blur $h(x, y)$ is defined as

$$H(\varepsilon,\eta)=\int_{R^2}h(x,y)e^{-i2\pi(\varepsilon x+\eta y)}dxdy \;=\; e^{-a\left(\varepsilon^2+\eta^2\right)^b}. \tag{7}$$

where (a>0, 0<b<1). When just blurring factor considered, the relationship between the HR image $x(x,y)$ and the LR image $y(x,y)$ in the frequency domain is as follows,

$$Y(\varepsilon,\eta)=H(\varepsilon,\eta)X(\varepsilon,\eta)+N(\varepsilon,\eta). \tag{8}$$

where $Y(\varepsilon, \eta)$, $X(\varepsilon, \eta)$ and $N(\varepsilon, \eta)$ are Fourier transforms of $x(x, y)$, $y(x, y)$ and $n(x, y)$, respectively. We may surely assume that the noise n(x, y) satisfies $\int_{R^2} |n(x, y)| dxdy \le f(x, y)dxdy = \sigma > 0$ (σ is a normalizing constant), so that we can ignore N(ε, η) and further normalize (8) into (9), assuming Y(ε, η), X(ε, η) and the OTF keep the following relation in a region Ω in the frequency domain

$$\log|Y(\varepsilon,\eta)| \approx -a(\varepsilon^2+\eta^2)^b + \log|X(\varepsilon,\eta)|. \tag{9}$$

We replace $\log|X(\varepsilon,\eta)|$ by negative constant $-A$ and solve (a, b) in (9) with nonlinear least squares algorithms. Putting (a, b) into (10), we can get the optimal approximation value for ideal HR image after inverse Fourier transform. $\overline{H}$ is the conjugate of H, K and s are adjustable parameters

$$X(\varepsilon,\eta) = \frac{\overline{H}(\varepsilon,\eta)Y(\varepsilon,\eta)}{|H(\varepsilon,\eta)|^2 + K^{-2}|1-H^s(\varepsilon,\eta)|^2}. \tag{10}$$

3.2 Maximum Likelihood Estimation on HMRF Parameter

The ML estimation of the edge threshold parameter α based on the approximation value provided by APEX deconvolution is calculated as

$$\hat{\alpha} = \arg\max P(\mathbf{x}|\alpha). \tag{11}$$

Parameter α can be assessed according to a predetermined cutoff ratio T ($T = f_\alpha(d_l^t\mathbf{x}) / f(d_l^t\mathbf{x})$), which corresponds to the percentage of high-frequency components in the image. $f(d_l^t\mathbf{x})$is the norm from($\|\ \|$) of the second order derivative, $f_\alpha(d_l^t\mathbf{x})$ is the norm when α is taken into consider (any value lower than α is set to zero). Since the approximation of the original image is known, T can be chosen according to the available information of energy distribution in the HR image. After ratio T is set, the estimation on α consists of solving the system

$$\partial \log P(\mathbf{x}|\alpha) / \partial\alpha = \partial\left[\sum_r \sum_{l=1}^{4} \rho(d_l^t\mathbf{x}, \alpha)\right] \Big/ \partial\alpha = 0. \tag{12}$$

where r is the component within the high frequency components set. Thus it gives $\hat{\alpha} = \sum_{r\in R} (|d_l^t\mathbf{x}|) / n$, n is the number of high-frequency components.

3.3 Gradient Projection Solution

We select the improved Newton gradient optimization technique to compute the unique minimum solution, which searches the global minimum of the objective function along the Newton direction. Any starting point $\mathbf{x}_0$ that satisfies (1) is valid. We use APEX restored image as the initial value $\mathbf{x}_0$. Suppose the gradient matrix of the cost function $U(\mathbf{x}_i)$ is $\mathbf{g}_i = \nabla U(\mathbf{x}_i)$ and the Hessian matrix is $\mathbf{G}_i = \nabla^2 U(\mathbf{x}_i)$ (i=0,…,K) , in each iteration the Newton direction $\mathbf{p}_i$ is calculated as

$$\mathbf{p}_i = -\mathbf{G}_i^{-1}\mathbf{g}_i . \quad (13)$$

And $\hat{\mathbf{x}}$ moves in the Newton direction $\mathbf{p}_i$ with step size τ_i to minimize $U(\mathbf{x}_i)$.

$$\tau_i = -\frac{\mathbf{p}_i{}^T \mathbf{p}_i}{\mathbf{p}_i^T (\mathbf{G}_i)\mathbf{p}_i} . \quad (14)$$

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \tau_i \mathbf{p}_i . \quad (15)$$

A sequence of iterates $\{\mathbf{x}_i\}_{i=0}^K$, more closely to $\hat{\mathbf{x}}$, are generated. The convergence is achieved until the relative state change for a single iteration has fallen below a predetermined threshold ε , such that $\|\mathbf{x}_{i+1} - \mathbf{x}_i\| / \|\mathbf{x}_i\| \le \varepsilon$.

The whole procedure of the hybrid Bayesian estimator is summarized as follows.

1) Upsample the LR images according to the enhancement factor q using bilinear interpolation, construct matrix **D** according to q, construct the geometric distortion matrix **T** using the hierarchical block matching[5].
2) Deconvolute the reference upsampling image with APEX algorithm to obtain the optimal approximation value for HR image and PSF.
3) Calculate the Newton direction $\mathbf{p}_i$.
4) Compute the step size τ_i and update the state according to (14) and (15).
5) If convergence criterion is satisfied, the estimate is given as $\hat{\mathbf{x}} = \mathbf{x}_{i+1}$. Otherwise, increment $\mathbf{x}_{i+1} = \mathbf{x}_i + \tau_i \mathbf{p}_i$ and return step 3.

4 Results

In order to demonstrate the performance of the proposed algorithm, two groups of experiment results are presented here, which involve actual satellite remote sensing images and actual video sequence grabbed from a digital video film during play back. The enhancement factor is set to be 2. The bilinear interpolation scheme, the POCS algorithm[3], the Huber-MAP algorithm[5] and our proposed hybrid Bayesian estimator (HBE) are applied in each group of test.

In the first group of test, we try to generate a HR satellite image from a sequence of five 5.0m resolution SPOT 5 satellite images. Fig.1 (a) is the reference 5.0m resolution LR image. The bilinear interpolation of the reference image, the POCS, Huber-MAP and HBE estimates are shown in shown in Fig. 1(b), 1(c), 1(d) and 1(e) respectively. Fig. 1(f) is the 2.5m resolution SPOT 5 image. The PSNR (Peak Signal-to-Noise Ratio) of the bilinear interpolation is 20.1, those of POCS, Huber-MAP and HBE estimates are 24.3, 25.2 and 26.7 respectively. Obviously, the HBE method achieves a significant improvement in PSNR, with considerably much higher resolution than the bilinear interpolation, POCS and Huber-MAP estimates.

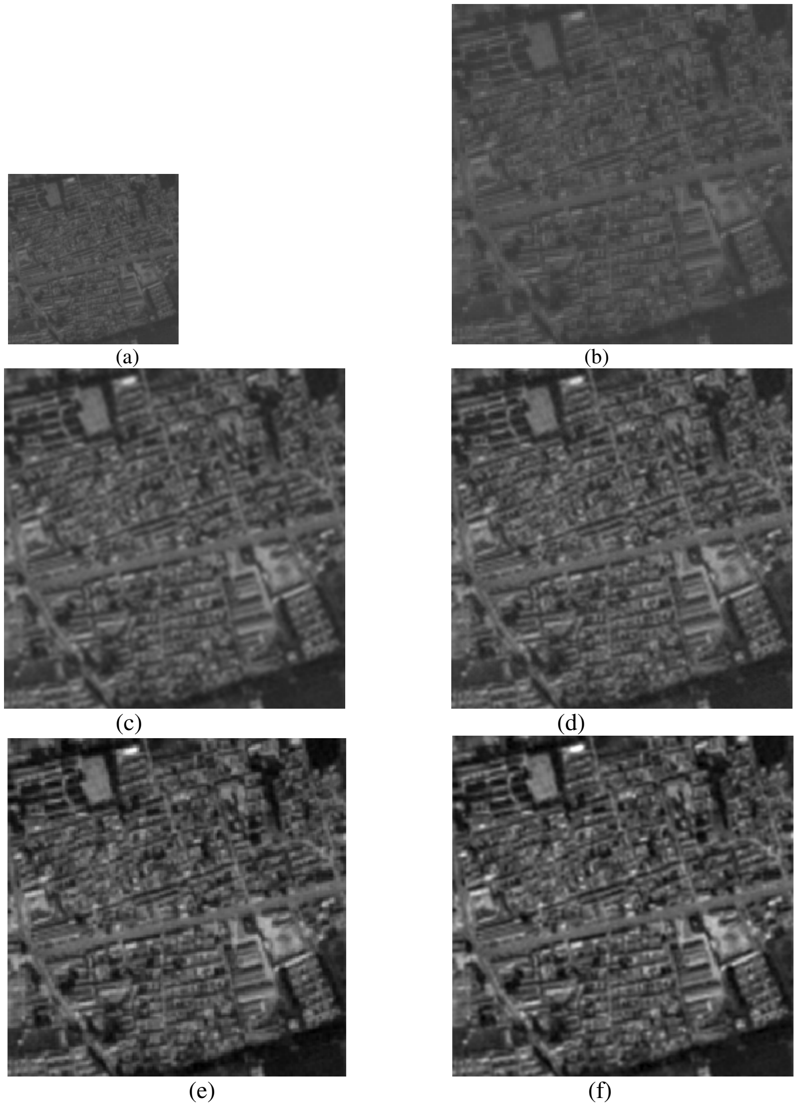

Fig. 1. Actual Satellite Image Sequence. (a) the reference 5.0m image. (b) Bilinear interpolation result. (c) POCS estimate. (d) Huber-MAP estimate. (e) HBE result. (f) the 2.5m HR image.

In the second group of test, nine frames are grabbed from the video sequence during playback. The frame shown in Fig. 2(b) is the bilinear interpolation of the reference frame in Fig. 2(a). The POCS result after 20 iterations is shown Fig. 2(c). The Huber-MAP result after 20 iterations is shown Fig. 2(d) and the HBE result after 16 iterations is shown Fig. 2(e). Fig. 2(f) is the original HR image.

Fig. 2. Actual Video Sequence. (a) the reference LR image. (b) Bilinear interpolation result. (c) POCS estimate. (d) Huber-MAP estimate. (e) HBE result. (f) the HR image.

The PSNR values of the bilinear interpolation, POCS, Huber-MAP and HBE estimates are 22.3, 25.2, 25.7 and 27.1 respectively. Experimental result shows that the image generated by the HBE approach outperforms those produced by bilinear interpolation, POCS and Huber-MAP estimators, especially in the areas of man's face and the bars far behind the man.

5 Conclusion

In the paper a novel hybrid Bayesian algorithm is proposed for HR image reconstruction from actual LR images or video sequence. The proposed approach firstly gets a good approximation of the ideal HR image, then estimate the edge threshold parameter from approximation data by ML estimation, and finally obtains a regularized reconstruction estimate automatically. Its main contributions are setting up an accurate HMRF image prior model, which enables the reconstruction processing to be carried out automatically and ensures the robustness of the estimate. Experimental results demonstrate this new technique is robust and gives very excellent reconstruction result in actual satellite data and video data. The resulted images exhibit much sharper and clearer details than images reconstructed by the bilinear interpolation, the POCS estimator and the Huber-MAP estimator.

References

1. Park, S. C., Park, M. K., Kang, M. G..: Super-Resolution Image Reconstruction: A Technical Overview. IEEE Signal Processing Magazine. 5 (2003) 21-36
2. Tsai R. Y., Huang, T.S.: Multiframe image restoration and registration. In: in Huang, T.S.(Ed.): Advances in computer vision and image processing, JAI Press, (1984) 317-339
3. Patti, A.J., Sezan, M. I., Tekalp, A. M.: Superresolution Video Reconstruction with Arbitrary Sampling Lattices and Nonzereo Aperture Time. IEEE Trans. Image Processing. 8 (1997) 1064-1997
4. Patti, A.J., Altunbasak, Y.: Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. IEEE Trans. Image Processing. 1(2001) 179-186
5. Schulz, R.R., Stevenson, R. L.: Extraction of High-Resolution Frames from Video Sequences. IEEE Trans. Image Processing. 6 (1996) 996-1011
6. Hardie, R.C., Barnard ,K.J., Armstrong E.E.: Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. IEEE Trans. Image Processing. 12 (1997) 1621-1633
7. Jalobeanu, A., Blanc-Féraud, L. et al.: An Adaptive Gaussian Model for Satellite Image deblurring. IEEE Trans. Image Processing, 4(2004) 613-621
8. CARASSO, A.S.,: THE APEX Method in Image Sharpening and the use of low exponent Lévy Stable Laws. SIAM J. APPL. MATH., 2(2002) 593-618