Super Resolution Using Graph-cut

Uma Mudenagudi, Ram Singla, Prem Kalra, and Subhashis Banerjee

Department of Computer Science and Engineering
Indian Institute of Technology Delhi
Hauz Khas, New Delhi, 110016, India
{uma,pkalra,suban}@cse.iitd.ernet.in
ram.singla@gmail.com

Abstract. This paper addresses the problem of super resolution - obtaining a single high-resolution image given a set of low resolution images which are related by small displacements. We employ a reconstruction based approach using MRF-MAP formalism, and use approximate optimization using graph cuts to carry out the reconstruction. We also use the same formalism to investigate high resolution expansions from single images by deconvolution assuming that the point spread function is known. We present a method for the estimation of the point spread function for a given camera. Our results demonstrate that it is possible to obtain super-resolution preserving high frequency details well beyond the predicted limits of magnification.

1 Introduction

In this paper we investigate the problem of obtaining a single high-resolution image given a set of low resolution images which are related by small displacements. We pose super-resolution as a reconstruction problem using the MRF-MAP formalism [1], and use approximate optimization using graph cuts [2] to carry out the reconstruction. We also use the same formalism to investigate high resolution expansions from single images by deconvolution assuming that the point spread function is known. We present a method for the estimation of the point spread function (PSF) for a given camera.

There have been several different approaches to super-resolution, with estimation of high-resolution (HR) images from multiple low resolution (LR) observations related by small motions being by far the most common one. Most of these methods are based on accurate registration and solve the super resolution reconstruction using variants of gradient descent with or without a smoothness prior [3–5]. Super-resolution has also been tried from multiple defocused images [6], varying zoom [7] and photometric cues [8]. Reconstruction based approaches to super-resolution model the low resolution image formation process to establish a relation between the unknown high resolution image and the low resolution observations, and use the relationship to derive algorithms to estimate the high resolution image essentially by an inversion process [6–10]. The inversion process is typically ill-conditioned and it often necessitates the use of smoothness or other priors [9, 11, 12] to obtain reasonable solutions. In [13], Baker and Kanade examine the limits of such processes and derive that for most point spread functions and blur kernels the estimation process is non-invertible or ill-conditioned. Further, the number of possible solutions grow at least quadratically with the desired magnification factor. They also show that this large growth in the number of solutions makes super-resolution difficult even with smoothness priors and the resulting solutions often fail to recover the high frequency details. In [14] the authors attempt to derive exact bounds on magnification factors based on a perturbation analysis. Their results indicate that under practical situations the magnification bound is only 1.6 for effective super-resolution. These results are obtained under the assumption of box PSF and local translations.

A large number of super-resolution algorithms have been based on the MAP-MRF formulation [6–8, 15] which indeed is a powerful framework for modeling the super-resolution problem. However, traditional algorithms for obtaining the MAP estimate, which in most cases result in non-convex optimization problems, have been based on simulated annealing or Iterated Conditional Mode (ICM) which provide no guarantee on the quality of the solution. Recently, Boykov et al. [2] have proposed a new algorithm based on graph-cut optimization which, under mild conditions on the nature of the objective function, can provide such guarantees. In this paper, we investigate whether with use of suitable priors, obtaining a good solution near the global optimum using an MRF-MAP formulation can indeed provide acceptable quality of reconstruction even beyond the derived limits. Unlike some methods in the literature [10–12] we do not learn the prior from examples of high resolution images, because such priors can then only be used to reconstruct similar high-resolution images. Instead, we use generic smoothness priors which are suitable for most situations.

The main contributions of this paper are as follows:

1. We give a formulation of super-resolution reconstruction from multiple displaced images using the MRF-MAP framework and solve using a graph-cut optimization. Typical graph-cut applications [16, 17] assume that the data term is a function of a single pixel. However, in the case of super-resolution, the intensity observed at a pixel is affected by neighboring pixels through a convolution representing blurring with the PSF. We formulate how such convolution based data terms can be approximated in the graph-cut formalism so that the resulting model of neighborhood interaction is regular which is necessary for graph-cut optimization [2]. We also present a method of estimation of the point spread function (PSF) of a camera as an off-line calibration process. We model the combined effects of the lens and the sensor and assume that all images are obtained using a camera whose PSF model is available.
2. We use the same formulation to deal with high-resolution expansion of single images using deconvolution.
3. Our results demonstrate that it is possible to obtain super-resolution preserving high frequency details well beyond the predicted limits of magnification of 1.6 in [14]. Note, that the limits in [14] are derived assuming box PSF and

local translations and our conditions are more generic. Our results demonstrate that even in presence of noise the super-resolved reconstruction is close to ground truth. In fact, in case of black and white images containing printed characters, we obtain high quality super-resolved images even without using a smoothness prior.

Section 2 describes the image formation process, modeling of the high resolution image as MRF. This section also gives MRF-MAP solution using graph-cut optimization for both single image expansion and SR reconstruction using multiple images. In Section 3 we present results to demonstrate the effectiveness of our method. The conclusions are given in Section 4.

2 Super-resolution with MRF-MAP

2.1 Image Observation Model

The image observation model is given by Equation 1, as in [3, 18]

$$\mathbf{g_k} = DH_kT_k\mathbf{f} + \eta_k \qquad 1 \leq k \leq n \tag{1}$$

where $\mathbf{f}$ is the HR image, $\mathbf{g_k}$ is the k^{th} observed LR image, D is the sub-sampling matrix, T_k is the affine transformation that maps the HR image to k^{th} LR image, H_k is the space invariant PSF of the camera for the k^{th} LR image and η_k is the observation noise.

Figure 1 summarizes the observation model.

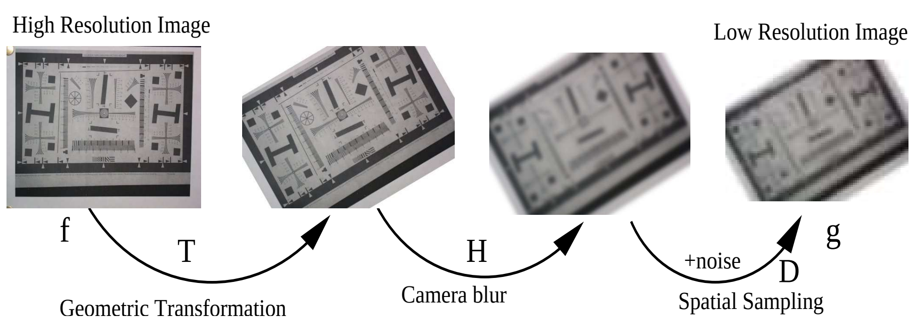


Fig. 1. Low Resolution Image Observation Model

Given the LR image and a magnification factor, the decimation matrix D is fixed. T_k can be estimated using any image registration technique, we use Hierarchical Model based Motion Estimation by Bergen et al.[19]. The estimation of PSF (H) is discussed in Section 2.2. We model the HR image as an MRF and use a maximum *a posteriori* (MAP) estimate as the final solution. The problem of SR reconstruction can be posed as a labeling problem where each pixel is assigned a label.

In the case of SR reconstruction, the posterior energy is given by

$$E(f|g) = \sum_k \|DH_kT_k\mathbf{f} - \mathbf{g_k}\|^2 + \sum_{p,q\in\mathcal{N}} V_{p,q}(f_p, f_q) \tag{2}$$

where, $V_{p,q}(f_p, f_q)$ are clique potentials which act as a smoothness prior and $\mathcal{N}$ is a neighborhood system. We minimize the posterior energy using the graph-cut technique proposed by Boykov et al. [2].

2.2 PSF Estimation

Assuming no motion blur, the edge spread function (ESF) captures the blurring effect of an ideal step edge by the image formation process. This includes both blurring due to the lens and the camera sensor. Under the Gaussian PSF assumption, the ESF $s(x)$ for a normalized edge is given by

$$s(x) = \frac{1}{2}(1 + erf(\frac{x}{\sigma\sqrt{2}})) \tag{3}$$

Given a calibration pattern with a set of ideal step edges, we estimate parameters of ESF by fitting the Equation 3 to a normalized edge in the least square sense. Note that in our MRF formalism we do not require shift invariant PSF, however, our estimates with a modern digital camera (Olympus, C-4000 zoom) indicate that the PSF is approximately space invariant. Typically the estimates of σ are around 0.4, hence we approximate the Gaussian PSF with a 3×3 mask.

2.3 Energy Minimization Using Graph-Cuts

The typical energy functions using MRF formulation are of the following form:

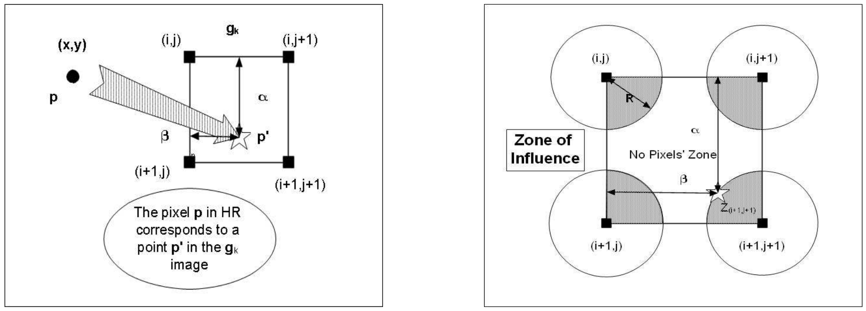
$$E\left(f\right) = \sum_{p\in\mathcal{S}} Data_p\left(f_p\right) + \sum_{p,q\in\mathcal{N}} V_{p,q}\left(f_p, f_q\right) \tag{4}$$

$Data_p\left(f_p\right)$ is a function derived from the observed data that measures the cost of assigning the label f_p to the pixel p. $V_{p,q}\left(f_p, f_q\right)$ measures the cost of assigning labels f_p, f_q to adjacent pixels p, q and is used to impose spatial smoothness. In order to minimize E using graph cuts a specialized graph is created. The form of the graph depends on the exact form of V and on the number of sites. The minimum cut on the graph minimizes the energy E either locally or globally [2].

Graph cuts can minimize only *graph-representable* energy functions. An energy function is graph-representable iff each term $V_{p,q}$ satisfies the regularity constraint [20].

For our problem the MRF sites are pixels of the HR image and labels are possible intensity values. The energy function presented in Equation 2 is in the coordinates of the LR images. In what follows, we re-write energy function in the coordinate system of the HR image.

The label at any site p is influenced by its neighbor due to the blur convolution. For a particular image site p, in the HR image, mapping to site p' in some



(a) (b)

Fig. 2. (a) Mapping of the Pixel from HR grid to LR grid (b) Circle of Influence of the pixel in LR grid

LR image potentially influences $four$ pixels $(i, j), (i, j + 1)(i + 1, j)(i + 1, j + 1)$ as shown in Fig 2(a).

We divide the space bounded by the $four$ pixels into $five$ zones − one each for the $four$ LR pixels and one *No Pixels' Zone* as shown in Fig. 2(b). The zone of each pixel is called *Zone of Influence* of that pixel. If the site p maps into one of the *Zone of Influence* of a pixel in the LR image then the expected label at the site p must be the label of that pixel. A site p mapping to a *No Pixels' Zone* of an LR image indicates that the LR image does not have any useful information about the expected label at the site p. A site p mapping to *No Pixels' Zone* for all LR images indicates that the sub-pixel displacement for this site p is not available in any of the LR images. For such a site p, we estimate the expected value of the label by interpolating the result from all the $four$ pixels in the reference LR image. From the above observation, we see that the data term for site p may be given as:

$$Data_p(f_p) = \sum_{k=1}^{n} \beta_k ||h_p * f_p - g^k_{DT_k p}||^2 \tag{5}$$

where h_p is the blur kernel at site p and $g^k_{DT_k p}$ is the expected label at the site p with precision β_k for the k^{th} LR image. $\beta_k = 0$ for a site p that is mapped into *No Pixel Zone*. If $\beta_k = 0$ for all LR images then $\beta_1 = 1$ and $g^1_{DT_1 p}$ is set to the interpolated value from the reference LR image. Further, $\sum_{k=0}^{n} \beta_k = 1$.

The energy function is still not in the standard form for energy minimization using graph-cuts. The data term of site p also depends on the neighbors of p due to blurring operator. We now approximate the data term as a sum of two terms − (i) a term that depends only on the observed data at site p (ii) a term that depends on the neighbors of p. In following equations we assume h_p is a blur kernel with values w_{pp} at the center and w_{pq} at the neighbor q of p. Expanding we obtain the following:

$$Data_p(f_p) = D_p^*(f_p) + \sum_{k=1}^{n} \beta_k [\sum_{q \in \mathcal{N}_p} (2(w_{pp} f_p - g_{DT_k p}^k) w_{pq} f_q + (w_{pq} f_q)^2)] + 2 \sum_{q,r \in \mathcal{N}_p} w_{pq} w_{pr} f_q f_p$$

$$where, \qquad D_p^*(f_p) = \sum_{k=1}^{n} \beta_k [(w_{pp} f_p - g_{DT_k p}^k)^2] \tag{6}$$

Hence, the energy is

$$E(f) \approx \sum_{p \in \mathcal{C}} D_p^*(f_p) + \sum_{p,q \in \mathcal{N}} \phi_{p,q}(f_p, f_q) + \sum_{p,q,r \in \mathcal{N}} \psi_{p,q,r}(f_p, f_q, f_r) \tag{7}$$

where ϕ_{pq} is the interaction associated with two neighboring pixels and ψ_{pqr} is with three neighboring pixels, which is ignored due to first-order neighborhood approximation. For energy E to be *graph representable* the function ϕ_{pq} must be regular. But due to the term $f_p f_q$ in ϕ_{pq} the regularity condition breaks. We eliminate this dependency by further approximating $(w_{pp} f_p - g_{DT_k p}^k) = \Delta_p^k$ and $(w_{qq} f_q - g_{DT_k q}^k) = \Delta_q^k$, where Δ_p^k and Δ_q^k are fixed for a particular α-expansion move during the minimization using graph-cuts.

The equations for ϕ_{pq} after approximation is:

$$\phi_{pq}(f_p, f_q) = \sum_{k=1}^{n} \beta_k [2\Delta_p w_{pq} f_q + (w_{pq} f_q)^2] + \sum_{k=1}^{n} \beta_k [2\Delta_q w_{qp} f_p + (w_{qp} f_p)^2] + V_{pq}(f_p, f_q) \tag{8}$$

The single image expansion is the special case with $n = 1$ and T_1 as the identity transformation. We have used the graph-cut library provided by Kolmogrov [21] for our implementation.

3 Results

In this section we present SR reconstruction results on both synthetic and real images. In all cases the attempted magnification is $4 \times pixel - zoom$ (four in each dimension). We compare our results of single image expansion and multiple image SR reconstruction using graph-cut with bilinear interpolation and iterative back projection (IBP) method proposed by Peleg and Irani [3] (Since it is a popular method and we have its implementation). All our real images are obtained with an Olympus digital camera (C-4000 zoom).

In Figure 3, we show a synthetic example of SR reconstruction for noisy observations with a calibration image. For both SR reconstruction using multiple images and single image expansion we have used the linear truncated smoothness prior, given by $V_{p,q} = \lambda(min(8, |f_p - f_q|))$. We generate input images with affine transformation with additive uniform noise of SNR=2 for multiple SR reconstruction. We register 24 LR images and carry out SR reconstruction as described in 2.3. Note that with multiple observation images the restoration method can effectively remove the noise and yet preserve the high frequency details. The results are more smooth because of registration errors in the presence of noise. Even the single image expansion can handle noise to a certain extent and even the resolution is better in the boxes in the lower left corner.

In Figures 4, 5 and 6 we show SR reconstruction using multiple images and single image expansion results for some real images. In each case for the single image expansion we use the same smoothness prior as above. For SR reconstruction using multiple images we use the same smoothness prior for results in Figure 4(d). For results in Figures 5(d) and 6(d) we do not use any smoothness prior. In each of these cases we use the same estimate of σ =0.473, since $f - number$ was fixed at 2.8.

It is evident from the results that the super resolution reconstruction using MAP-MRF using graph-cuts, both for multiple images and single image expansion, preserves the high frequency details.

The results can be downloaded from our site at
http://www.cse.iitd.ac.in/˜uma/publication.html.

4 Conclusions

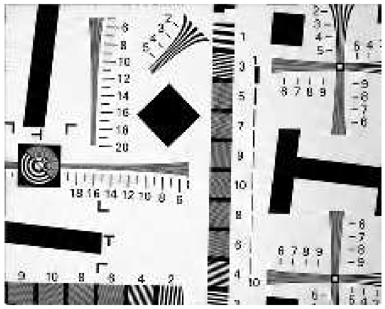
We have formulated the SR reconstruction problem in the framework of MRF-MAP and have proposed a solution using graph-cuts. We also carry out single image expansion using the same framework. The results demonstrate that the proposed framework for SR reconstruction using multiple images preserves the high frequency details.

The results may improve if we estimate registration parameters in the same graph-cut formulation.

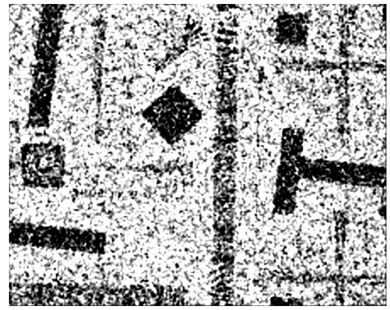
Our method is not real time. We are currently exploring ways to make it fast.

References

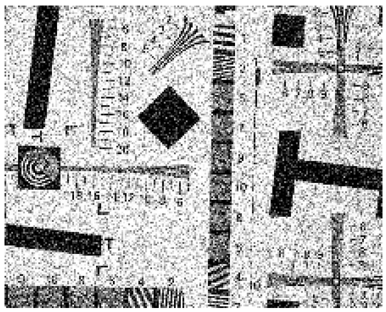
1. S.Geman, Geman, D.: Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. IEEE Trans. on Pattern Analysis and Machine Intelligence **6** (1984) 721–741
2. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Transactions on PAMI **23** (2001) 1222–1239



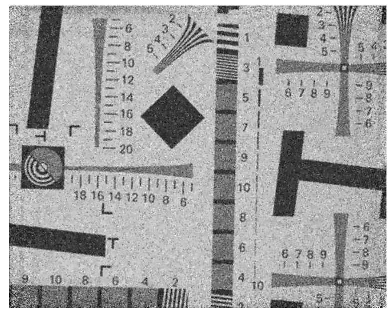
(a) Ground truth



(b) Noisy Bilinear Interpolated



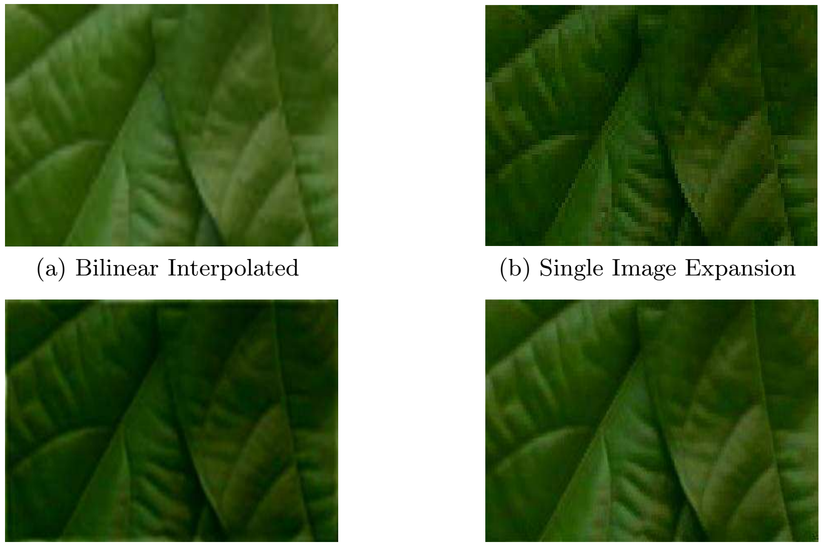
(c) Single Image Expansion



(d) SR method: Multiple Images

Fig. 3. Effect of Noise on HR Image

3. M.Irani, Peleg, S.: Improving resolution by image registration. CVGIP:Graphical Models and Image Processing **53** (1991) 231–239
4. M.Irani, Peleg, S.: Motion analysis for image enhancement: Resolution, occlusion, and transparency,. Journal of Visual Communication and Image Representation **4** (1993) 324–335
5. Borman, S., Stevenson, R.: Linear models for multi-frame super-resolution restoration under non-affine registration and spatially varying PSF. In Bouman, C., Miller, E., eds.: Computational Imaging II. Volume 5299 of Proceedings of the SPIE., San Jose, CA, USA (2004) 234–245
6. D.Rajan, Chaudhuri, S.: Simultaneous estimation of super-resolved scene and depth map for low resolution defocused observations. IEEE Trans. on Pattern Analysis and Machine Intelligence **25** (2003) 1102–1117
7. Joshi, M., Chaudhuri, S.: A learning based method for image super-resolution from zoomed observations. in Proc. fifth Int. Conf. on Advances in Pattern Recognition (2003) 179–182
8. Chaudhuri, S., Joshi, M.V.: Motion-Free Super-Resolution. Springer (2004)
9. Schultz, R.R., Stevenson, R.L.: A bayesian approach to image expansion for improved definition. IEEE transctions on Image processing **3** (1994) 233–242
10. Capel, D., Zisserman, A.: Computer vision applied to super resolution. IEEE Signal Processing Magazine (2003) 75–86
11. Pickup, L.C., Roberts, S.J., Zisserman, A.: A sampled texture prior for image super-resolution. Advances in Neural Information Processing Systems (2003)



(a) Bilinear Interpolated
(b) Single Image Expansion
(c) IBP Method
(d) SR Method: Multiple Images

Fig. 4. HR Image of leaves with $\lambda = 0.06$

12. Freeman, W.T., Jones, T.R., Pasztor, E.C.: Example-based super-resolution. IEEE computer Graphics and Applications (2002) 56–65
13. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. in Proceedings of the 2000 IEEE Conference on Computer Vision and Pattern Recognition (2000)
14. Lin, Z., Shum, H.Y.: Fundamental limits of reconstruction-based superresolution algorithms under local translation. IEEE Trans. on Pattern Analysis and Machine Intelligence **26** (2004) 83–97
15. Borman, S., Stevenson, R.: Super-resolution from image sequences – A review. In: Proceedings of the 1998 Midwest Symposium on Circuits and Systems, Notre Dame, IN, USA (1998) 374–378
16. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on PAMI **26** (2004) 1124–1137
17. Kolmogorov, V., Zabih, R.: Multi-camera scene reconstruction via graph cuts. European Conference on Computer Vision (2002)
18. Elad, M., Feuer, A.: Super resolution restoration of an image sequence: Adaptive filtering approach. IEEE Trans. on Image Processing **8** (1999) 387–395
19. Bergen, J.R., Anandan, P., Hanna, K.J., Hingorani, R.: Hierarchical model-based motion estimation. ECCV (1992) 237–252
20. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? IEEE Transactions on PAMI **26** (2004)
21. Kolmogorov, V.: http://research.microsoft.com/ vnk/. (2000)

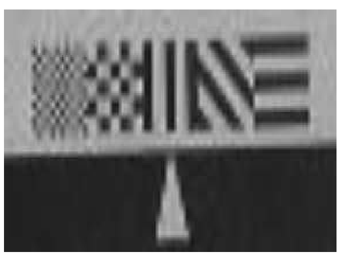
(a) Bilinear Interpolated

(b) Single Image Expansion

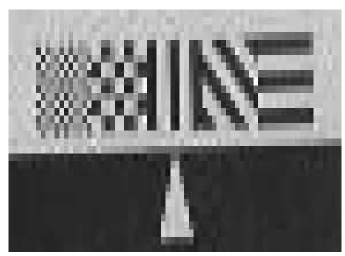
(c) IBP Method

(d) SR Method: Multiple Images

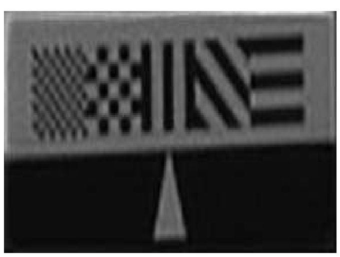
Fig. 5. HR Image of Text



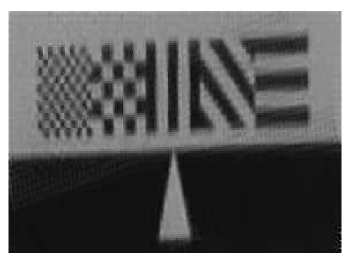
(a) Bilinear Interpolation



(b) Single Image Expansion



(c) IBP Method



(d) SR Method: Multiple Images

Fig. 6. HR Image of Pattern