Superresolution using perceptually significant side information

Vikas Ramachandra* and Truong Q. Nguyen *Fellow, IEEE*

***Abstract*—We investigate the problem of super-resolution of images in the presence of side information. In some situations, when some information of the original image is available to the sender, it can be embedded into the low resolution images, either in the pixels themselves or in the headers. This information can be later used when required to reconstruct the superresolved image. For this, a novel multiresolution histogram matching based superresolution procedure is outlined. The proposed technique gives better results compared to contemporary resolution enhancement algorithms.**

I. Introduction

Superresolution (SR) is the problem of reconstructing a high resolution image from a single low resolution image or set of low resolution images, each of which contributes some unique information. SR is an ill-posed problem since there might be many high resolution images which give the same low resolution image set. We need to resort to regularization by imposing prior knowledge about the high resolution image, to restrict its solution space to a visually plausible set [1]. In ordinary SR at the receiver, one would impose a prior like the Markov random Field [20] or the constrained Total Variation norm model [13], [14] in conjunction with an estimation procedure to get an estimate of the high resolution image. These priors are inspired by general image statistics. However, if we had more specific 'side' information about the actual image we are trying to reconstruct, our estimate would be much better.

We believe that our method can perform better than conventional resolution enhancement techniques in the following two scenarios:

- Superresolution of transmitted low resolution images over limited bandwidth.
- Deblurring of print media (text and images on paper) captured by mobile phone cameras etc.

In the above scenarios, the chosen side information can either be incorporated in the image headers or embedded into the low resolution image itself (before transmission or printing) using data embedding schemes such as Quantization Index Modulation [2]. In this paper, we look at what 'side' information is perceptually well suited for SR. An algorithm that actually uses that information for high resolution image estimation is also explored.

The authors are with the University of California, San Diego, 9500 Gilman Drive, San Diego CA 92093. Email: {vikas,nguyent}@ucsd.edu. Phone: 858 534 5669. * Corresponding author.

A. Previous work and our contributions

Previous related work which uses data embedding for image processing (non-security) applications includes improving coding efficiency by embedding color information in the image itself [4], using data embedding to enable good error concealment for transmitted images [5], and for transmitted videos over lossy networks [6], and embedding information to help selectively filter transmitted compressed image regions [17]. However, none of the previous works use partial information embedding for superresolution and estimation of the original image like we do. Also, in papers like [12] and [16], one uses side information to enhance transmitted images, but these methods do not make use of the side information motivated by perceptually significant parameters like we propose.

More recently, [3] embedded statistics of the original image into transmitted images, which could be used as an objective quality metric. Our work is inspired by recent advances in texture synthesis. It is discussed in [7], [10] and [11] that many textures can be reconstructed by matching the histograms of the filter responses of a set of well-selected bandpass filters. In [7], the authors proposed an iterative projection method with constraints imposed on the multiscale oriented pyramid coefficients, and were able to construct meaningful textures from random initial images. However, it was concluded that this method does not work for images.

Our contributions in this paper are twofold. Firstly, we have modified the above mentioned texture synthesis method, turning it into an image deblurring scheme, making it suitable for superresolution. Our proposed modifications are explained in the sections which follow. Secondly, we have presented a framework for using side information for image resolution enhancement.

Our technique is general in that it can be used to further augment the performance of any existing popular SR or image enhancement technique like [8] and [9] which make use of regularization of this ill posed inverse problem. Our method can be viewed as a method to enhance prior knowledge. We believe that this can also help do away with simplistic but unrealistic assumptions being made at present.

The proposed algorithm follows these steps:

- Side information embedding- At the image print output device or transmitter, embed a compact representation of the side information into the image(s) being stored/ sent, or, if possible, encode the side information in the header.
- Extraction and decoding of the side information- At the receiver or image capture device, when required, decode the embedded/encoded information in the set of image(s)

received or stored after capture.

- The SR technique- Use the decoded information to aid the SR image estimation procedure, in the form of prior knowledge and original image statistics.

B. Organization

This paper is organized as follows. Section 2 explains the proposed SR method, with a detailed discussion on the proposed changes to the existing texture synthesis scheme in section 2B. Also, the gradient projection based estimation scheme to obtain the high resolution image is briefly outlined in section 2C. Results for each of the examples are presented in section 3, and section 4 concludes the paper.

II. The proposed superresolution technique

Side information is incorporated into the low resolution image, either using QIM or in the header. This side information is used in the resolution enhancement process. Two important questions are addressed below,

- What is the perceptually significant side information to be embedded.
- How can the received side information be used to enhance the received image quality.

A. Choosing the required side information

It was argued in [7] that texture images with identical higher order marginal statistics and joint statistics across subbands are indistinguishable to the human eye. It was also also shown in [3] that the marginal distribution of the wavelet coefficients changes in different ways for different types of image distortions. Based on these discussions, we choose the side information to be the histograms and joint subband statistics in the (multiresolution) steerable pyramid decomposition of the original high resolution image. We first briefly review the texture synthesis scheme from [7] below, and then explain our proposed changes to this scheme.

The texture synthesis method builds textures from random initial images. In the first step of texture synthesis, the steerable pyramid is implemented by recursively splitting an example image into a set of oriented subbands and a lowpass residual band. The filters used in this decomposition are polar separable in the Fourier domain. The recursive procedure is initialized by splitting the input image into lowpass and highpass portions, using the following filters:

$$L_0(r,\theta) = L(r/2,\theta)/2$$
$$H_0(r,\theta) = H(r/2,\theta)$$

where r, θ are polar coordinates. In this paper, we have used K = 4 orientation bands, and N = 4 pyramid levels (scales), for a total of 18 subbands (16 oriented, plus highpass and lowpass residuals). Figure 1 shows the pyramid decomposition structure, where $B_k(w)$'s are bandpass oriented filters. Figure 2 shows the actual subbands (real and imaginary parts) for the 3 scale 4 orientation decomposition of a disk image.

In the model in [7], the constraint functions are defined on the coefficients of these subbands of the original image

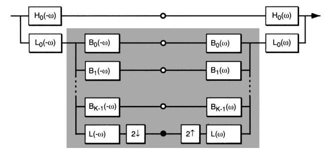


Fig. 1. Pyramidal decomposition- The image is split into lowpass and highpass subbands, the lowpass subband is further split recursively. The reconstruction structure is to the right of the circles.

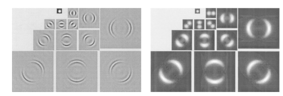


Fig. 2. 3 scale 4 orientation pyramidal decomposition of the disk image. Left: real parts of oriented bandpass images at each scale. Right: Magnitudes of these subbands.

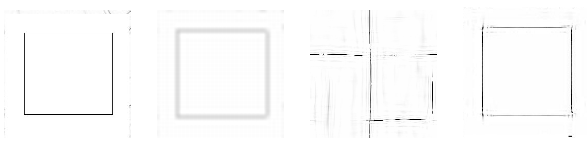
(available to the sender). The chosen perceptually significant parameter set includes three normalized sample moments (variance, skewness and kurtosis), together with the range (minimum and maximum intensities) for the pixel statistics, the skewness and kurtosis of the lowpass subband images, and the variance of the highpass subband images are computed at each level of the recursive pyramid decomposition. Along with this, the scheme also requires both raw coefficient as well as magnitude cross-correlation terms across orientations and scales, as well as phase cross-correlation terms across subbands. The texture image is constructed by an iterative sequential gradient projection method (starting from a random image) which imposes each of the above constraints (extracted from a reference image or texture patch) on the present estimate of the texture being constructed. The algorithm works well in practice for texture synthesis. However, this method was noted to fail for synthesizing natural images from random initializations.

B. Proposed modifications to texture synthesis

Here, we outline our proposed modifications to the texture synthesis scheme, converting it into a deblurring method. We also outline the reasons for these changes through example images.

- The initial image estimate is not random, instead a bicubic interpolated image is used.
- We eliminate the constraints on the raw coefficient cross-correlation, as well as relative local phase constraints.
- A deringing stage is introduced.

1) Random versus bicubic initialization for superresolution: One of the reasons why the texture synthesis algorithm fails for natural image synthesis is because of the random initialization. The set of multiscale constraints are not sufficient to capture all the higher order statistical information in *natural images*, especially global structures. This is different from *textures* where the local regions reflect the global structure, since the texture image consists of a repetition of those local patches. However, for natural images, a low resolution version of the actual image captures vital higher order information and global structure which cannot be enforced using just the multiscale constraints listed before. Below, we present a simple example image of a square of size 256X256 pixels (figure 3(a)), which was reconstructed using (a) a random initial guess and multiscale constraints (figure 3(c)), and (b) low resolution image of size 32X32 (bicubic upsampled by 8) (see figure 3(b)) as the initial guess with multiscale constraints (figure 3(d)). In both cases, constrained gradient projection was used (as explained in the next subsection). It is clear that the random estimate with the constraints is unable to fully capture the global structure, whereas (even a very) low resolution image does well at this.



(a) original (b) Low resolution image (c) SR with random initialization (d) SR with bicubic initialization

Fig. 3. SR with random versus bicubic initialization for a square

2) Problems with raw subband coefficient correlations and relative local phase: Raw subband coefficient correlations and relative local phase are two constraints which are important for texture synthesis,as noted in [7] since they enforce local constraints. However, for superresolution, we are more interested in edges and larger details as well, and these two constraints actually yield local non-linear distortions. Therefore, we propose to do away with these constraints. Below, we present an example image in figure 4(a) which was downsampled by 4 first (see figure 4(b)), the low resolution image was superresolved with and without the above constraints (figures 4(c) and 4(d) respectively).We see that imposing the two extra constraints noted above leads to non-linear distortions. It can be seen that the estimated high resolution image still contains some ringing artifacts which we address below.

3) Deringing using non-linear fuzzyfilter: As mentioned before, the proposed superresolution method suffers from ringing artifacts. Ringing was also observed in other applications of the multiscale multioriented transform [19]. To combat this, we use a nonlinear fuzzy filter proposed by Vo et al. in [18]. Ringing artifacts are directional, and this method proposes to design non-linear fuzzy filters which can capture and filter out these distortions. Here, we present an example of the superresolved image alone (figure 4(d)), and compare it to superresolution with deringing, which is the final output of our proposed framework (figure 4(e)). It is clear that the fuzzyfilter is able to almost fully remove the ringing.

C. Constrained gradient projection based superresolution

All the modified parameter values thus chosen (which corresponds to about a kilobyte of side information) are embedded into the low resolution image(s) or encoded into the image header. The image is then either transmitted or put on print paper. At the receiver or camera software, the side information is first extracted as mentioned above. We then upsample the LR image to have the same dimensions as the required SR image using bicubic interpolation. This image serves as the starting input to the constrained gradient projection algorithm.

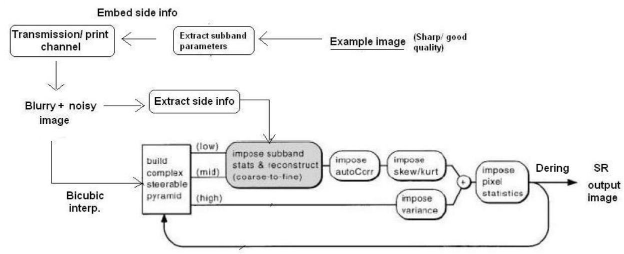


Fig. 5. The iterative SR reconstruction method

The constraints imposed are based on the statistics of the original image in the side information. The bicubic upsampled low res image is decomposed into a complex steerable pyramid. An iterative coarse-to-fine procedure imposes the statistical constraints on the lowpass and bandpass subbands, while simultaneously reconstructing a lowpass image. The autocorrelation of the reconstructed lowpass image is then adjusted, along with the skew and kurtosis, and the result is added to the variance-adjusted highpass band to obtain the SR image. The marginal statistics are imposed on the pixels of this image, and the entire process is repeated. The procedure is outlined in Figure 5.

Let x_0 be the initial estimate of a subband of the high resolution image. Then, for each constraint c_k, we have a projection P relating the subband coefficients obtained at iteration x_n to the the subband coefficients from the previous iteration x_{n-1} onto the set imposed by p, i.e., $x_n = P_{c_k}(x_{n-1})$. Let $\phi(x)$ be one of the parameters which is being constrained (like mean, variance etc.). A straightforward way of performing these projections is to modify the current estimate along the gradient of the constrained parameter under consideration, i.e.

$$x_n = x_{n-1} + \lambda_k \nabla \phi_k(x) \tag{1}$$

where λ_k is chosen such that the parameter meets its constraint, i.e., $\phi(x) = c_k$. The reader is referred to [7] for details of these projection operators.

III. Results

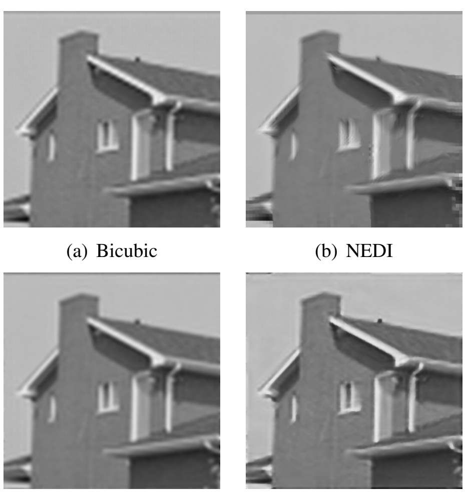
We take the high resolution images and *downsample them by 4* to get low res images. Then bicubic interpolation is used to get an intermediate image of the same size of the HR image, which is input as the initial guess to the proposed algorithm. Here, one could consider bicubic interpolation as an algorithm which does not use any side information. From figure 6, we see that for the 'Text' image, the proposed method makes the text readable and sharp, unlike only bicubic interpolation. As with other images, our method improves upon the PSNR too. For the 'House' image in figure 7, we compare the results of our method with bicubic interpolation, a Markov random field prior based Bayesian MAP SR estimation algorithm [20] and new edge directed interpolation (NEDI) [15] which are other popular interpolation methods. The MRF based algorithm is representative of techniques which use generic image statistics to help regularize the SR image solution space. NEDI uses second order statistics only. Our method adds

(a) Original (b) Bicubic (c) SR with raw coeff and relative phase constraints (d) SR without those two constraints (e) SR with deringing: Final output

Fig. 4. Estimated high res image at various stages of the superresolution process

visually significant details which could neither be inferred by the generic prior model nor only second order statistics based modeling.

Fig. 6. Text: L to R (PSNRs in braces)- Original high-res, bicubic (25.63 dB), proposed (26.34 dB).



(a) Bicubic (b) NEDI (c) SR with general MRF prior (d) Proposed SR

Fig. 7. Comparison of the proposed SR method with bicubic, general prior based SR and NEDI for 'House' image

`Additional results are at http://videoprocessing.ucsd.edu/~vikas/sr_results.html`

IV. Conclusion

A superresolution method which uses side information is proposed. We embed information about the original high res image into the low res images or send it in the header. We have provided an example SR algorithm in the form of constrained gradient projection. One can use side information, when available, to regularize the SR problem. This solution is suited for improving the quality of transmitted low res images, as well as deblurring print media captured using cell phone cameras.

References

[1] S Borman and RL Stevenson, "Super-resolution from image sequences-a review." *Midwest Symposium on Circuits and Systems,* 9-12 Aug. 1998.

[2] B Chen and GW Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Trans. on Information Theory*, Vol.47,Iss 4, 2001.

[3] Z Wang, G Wu, HR Sheikh, EP Simoncelli, E Yang and AC Bovik, "Quality-aware images," *IEEE Trans. on Image Processing,* Vol. 15, no. 6, June 2006, pp. 1680-1689.

[4] P Campisi, D Kundur, D Hatzinakos and A Neri. "Compressive data hiding: an unconventional approach for improved color image coding." *EURASIP J. Appl. Signal Process.* 2002, pp. 152-163.

[5] Y Liu and Y Li, "Error Concealment of Digital Images Using Data Hiding," *Proceedings of the Ninth DSP Workshop,* Oct. 15 - 18, 2000.

[6] DL Robie and RM Mersereau, "Video error correction using steganography." *EURASIP J. Appl. Signal Process.* 2002, pp. 164-173

[7] J Portilla and EP Simoncelli, "Parametric texture model based on joint statistics of complex wavelet coefficients," *IJCV,* Vol. 40, no. 1, Oct. 2000

[8] M Elad and A Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Trans. on Image Processing,* Vol.6, Iss.12, Dec 1997, pp. 1646-1658.

[9] AJ Patti, I Sezan and M Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time." *IEEE Trans. on Image Processing,* Vol.6, Iss.8, Aug 1997, pp. 1064-1076

[10] DJ Heeger and JR Bergen, "Pyramid-based texture analysis/synthesis," *ACM SIGGRAPH 95,* 1995, pp. 229-238

[11] SC Zhu, Y Wu and D Mumford, "Filters, Random fields and Maximum entropy (FRAME) : Towards a unified theory for texture modeling," *IJCV,* 1998, vol. 27, no. 2, pp. 107-126.

[12] D Barreto, L Alvarez, R Molina, A Katsaggelos and GM Callico, "Region-based super-resolution for compression," *Multidimensional Systems and Signal Processing,* Vol. 18, No. 2-3, Sep. 2007.

[13] D Capel and A Zisserman, "Computer vision applied to super resolution." *IEEE Signal Proc. Magazine,* Vol.20, Iss.3, May 2003.

[14] S Farsiu, D Robinson, M Elad and P Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. on Image Processing,* Vol.13, Iss.10, Oct. 2004, pp. 1327- 1344

[15] X Li and M Orchard, "New edge-directed interpolation," *IEEE Trans. on Image Processing,* Vol.10, Iss.10, Oct 2001, pp. 1521-1527

[16] S Wittmann and T Wedi, "Transmission Of Post-Filter Hints For Video Coding Schemes," *ICIP,* 2007.

[17] D Silverstein and S Klein, "Precomputing and encoding compressed image enhancement instructions," US Patent 5822458.

[18] DT Vo, S Yea and A Vetro, "Spatio-temporal Fuzzy Filtering for Coding Artifacts Reduction", *SPIE VCIP,* San Jose, Jan. 2008.

[19] EP Simoncelli, WT Freeman, EH Adelson and DJ Heeger, "Shiftable multiscale transforms," *IEEE Trans. on Information Theory*, 1992.

[20] RR Schultz and RL Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. on Image Processing,* vol.5, no.6, Jun 1996, pp.996-1011.