Edge Model Based High Resolution Image Generation⋆

Malay Kumar Nema[1], Subrata Rakshit[1], and Subhasis Chaudhuri[2]

[1] Centre for Artificial Intelligence and Robotics, Bangalore
[2] VIP Lab, Department of Electrical Engineering, IIT Bombay, Mumbai

Abstract. The present paper proposes a new method for high resolution image generation from a single image. Generation of high resolution (HR) images from lower resolution image(s) is achieved by either reconstruction-based methods or by learning-based methods. Reconstruction based methods use multiple images of the same scene to gather the extra information needed for the HR. The learning-based methods rely on the learning of characteristics of a specific image set to inject the extra information for HR generation. The proposed method is a variation of this strategy. It uses a generative model for sharp edges in images as well as descriptive models for edge representation. This prior information is injected using the Symmetric Residue Pyramid scheme. The advantages of this scheme are that it generates sharp edges with no ringing artefacts in the HR and that the models are universal enough to allow usage on wide variety of images without requirement of training and/or adaptation. Results have been generated and compared to actual high resolution images.

Index terms: Super-Resolution, edge modelling, Laplacian pyramids.

1 Introduction

Generation of high resolution (HR) images from low resolution (LR) images have been attempted through reconstruction based approaches and learning based approaches. Reconstruction based approaches require multiple images. They make use of subpixel shifts between images to pool in the extra information needed to create the HR image. Methods employed include sub-pixel registration, nonuniform interpolation [1][2] and frequency domain approaches [3][4]. An exhaustive list of methods can be found in [5], [6]. Learning based approaches build a relation between LR and HR images, based on the imaging process and/or description of corresponding edges between LR and HR. Multiresolution based mehods are a natural choice for this problem. The multiresolution representations seperate the information in images by frequency. The generation of HR is essentially the problem of generating the missing (hypothetical) level(-1) subband. Solutions have been proposed based on zoom [7][8], wavelet [9] and contourlet [10] coefficients. A detailed discussion can be obtained from [11]. The problem may be

---

⋆ This work is supported by DRDO through project CAR-008.

decomposed into three parts: (i) formulating a model for predicting edges in HR based on edges in LR (ii) using the model for calculating the high frequency components to be added and (iii) injecting those (postulated) components in a manner consistent with the known lower frequency components as given in the LR image. This paper presents a new method for each of these, based on Laplacian pyramids and their variations. Laplacian pyramids [12] are chosen for the multiresolution representation as they provide the simplest and most regular representation of edges in the subbands. There is a single, non-directional subband at each level, unlike the three subbands for 2D wavelets. The 4/3 redundancy also leads to more regular structures in the subband.

Our model for predicting edges in the HR image is based on the observation that edges arising due to occlusion remain sharp at every resolution. Edges arising due to shading and surface patterns should become blurred when resolution is increased. Conversely, the majority of the sharp edges in an image may be assumed to be occlusion edges. For generating an HR image, these sharp edges must be identified in the LR and their sharpness must be restored in the HR image. Simple interpolation will blur all edges while increasing feature sizes. Our HR generation is limited to the objective of preserving sharpness of edges which are sharp in LR. (A single image method cannot, in any case, introduce new edges and features into the HR image.) The Laplacian pyramid, like all other multiresolution representations, creates a hierarchy of subbands encoding edges of decreasing sharpness. Thus the first subband, designated L_0, captures all the sharp edges in the image. This paper presents methods for interpolating L_{-1} from L_0, restoring the sharpness of the edge representations and ensuring consistency between this modified L_{-1} and the given LR image.

The current work makes use of two recent results related to Laplacian pyramids. It has been shown that Laplacian subbands can be represented using edge model elements [13]. (Note that these are descriptive models describing the patterns corresponding to edges of various sharpness and geometry, as opposed to the generative model for edges discussed earlier.) A model-based description is convenient for altering the sharpness of selected edges without affecting the resolution or frequency content of the image as a whole. Another result of relevance to the present work is the extension of Burt's Laplacian pyramids called Symmetric Residue pyramids [14] (SRP). It addresses the issue of consistency between pyramid subbands and the nature of independent information at each level. The methods given there are used in the present work to initially postulate an L_{-1} from given L_0 and then, after selective edge sharpening, to ensure consistency between modified L_{-1} and L_0.

The paper is organised as follows. We provide relevant details of edge modelling of Laplacian subbands in Section 2. Subsequently we briefly review SRP in Section 3. Section 4 discusses generation of L_{-1} using edge-model representation, including effect of subsampling on the modelling process. Section 4 ends with the algorithm for generation of L_{-1}. Section 5 provides results and performance evaluation and the paper concludes in Section 6.

2 Edge-Modelling

For purposes of interpolation for HR, we need a mechanism to model the dominant, sharp edges in the L_0 subband. As such, it is important to have a representation that is edge based rather than pixel or frequency based since it will allow edge specific modifications. As the SRP based interpolation (see Sec 3) can fill in the mid- and lower-frequency components later, the MSE of the modelling process is not that critical. The Laplacian subband edge-modelling process described in [13] provides such a model. The modelling elements shown in the Figure 1 are the ones which are called Primitive Set(PS)-28.
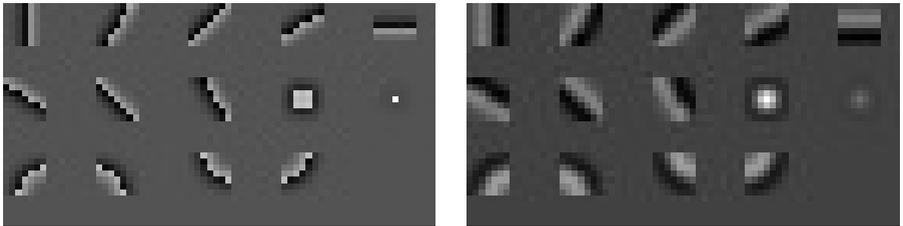
(a) The 14 elements based on sharp edges (b)The 14 elements based on smoother edges

Fig. 1. The PS28 primitive set elements. The 7×7 elements, scaled for display.

The PS-28 elements are defined as 7×7 images. The PS is chosen to be a set of sharp and blur edge elements as represented in Laplacian subbands. Due to its band-pass nature, only edges of certain thickness can be dominant in any subband. As mentioned in [13], the problem simplifies if one opts for the task of selecting 7×7 blocks that can be represented using individual PS elements (the modelling approach) instead of representing arbitrary 7×7 blocks as a superposition of a given set (projection onto basis approach). Denote the PS elements as $\mathbf{p_i}$ and the block to be modelled as $\mathbf{x}$. For computing a representation using a basis set, the procedure would be to compute a set of coefficients, α_i, to best represent any given $\mathbf{x}$. For edge modelling in [13], the procedure adopted is to take each element $\mathbf{p_i}$ and see which part of the image it can best model. Various model fit criteria are used to determine in which order the Laplacian gets modelled by the various model elements. The objective is to find an $\mathbf{x}$, model element $\mathbf{p_i}$ and associated scalar α that minimises

$$J = \frac{\|(\mathbf{x} - \alpha.\mathbf{p_i})\|}{\|(\mathbf{x})\|}. \tag{1}$$

The edge-model element for a particular location was picked on the basis of energy and modelling error. For each 7×7 block extracted from the Laplacian image, the following were computed. The energy of the extract determines the amount of signal present. Laplacians, like all subbands, are zero-mean and sparse. The energy is concentrated in only a few areas. Only blocks having energy above a threshold were considered for modelling. The threshold was initially set high

and reduced with each iteration till it reached a lower cut off. This cut off determined the termination of the iterative process, as errors below this threshold were not modelled further. Only blocks whose energy crosses the current threshold, were considered for modelling by elements of the primitive set. A sample result of the accuracy of modelling process is shown in Figure 2. It is seen that the modelling of Laplacian subbands can be done to a reasonable accuracy.
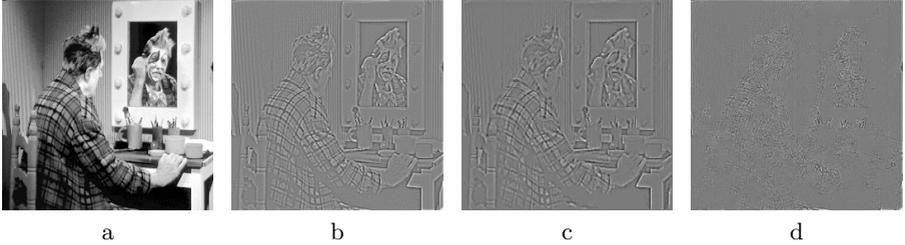

a   b   c   d

Fig. 2. Demonstration of modelling accuracy: (a) `Clown` image, (b) Laplacian of `Clown` image, (c) Modelled Laplacian (d) Modelling error

3   Symmetric Residue Pyramids

Symmetric Residue Pyramids [14] were proposed as an extension to Burt Laplacian pyramids. Initial aim of the SRP was to achieve better signal compaction. It makes use of the fact that the Laplacian subband at the level i, L_i and gaussian subband at next level $i+1$, G_{i+1} are related due to the redundancy in the Laplacian pyramids. The set of all possible L_i can be divided into equal sized cosets corresponding to each possible G_{i+1}. The process of generating an acceptable L_i given a G_i starts with an initial guess, which may even be a blank (zero) image. An iterative process is deployed to get one of the acceptable L_i: (exp is the expand/interpolation operation, ss is subsampling and lpf is low-pass filtering)

1. $L_i[0] =$ Initial guess (may even be 0 image)
2. $G_i[k] = exp(G_{i+1}) + L_i[k]$ (usual pyramid reconstruction)
3. $L_i[k+1] = G_i[k] - exp(ss(lpf(G_i[k])))$

The above process can be used to guess L_i to within an element in the correct coset. The reconstruction process only needs the difference between this element and the actual L_i, denoted as SL_i. The SRP is defined by $SL_i, i = 0...n$, G_{n+1}. Decomposition of `Lena` image using Burt Laplacian pyramid and SRP is shown in Figure 3(a) and 3(b) respectively.

The above processes are of relevance to the current work for two reasons. The iterative scheme, initialzed with a blank L_{-1}, is a good way of interpolating an L_{-1} from L_0. It is certainly better than simple interpolation (essentially $L_{-1} = exp(L_0)$). More importantly, it highlights the need for, and the method of utilization, of additional information. The SRP subbands show that the missing information is with regards to high frequency informaion about the sharp edges.
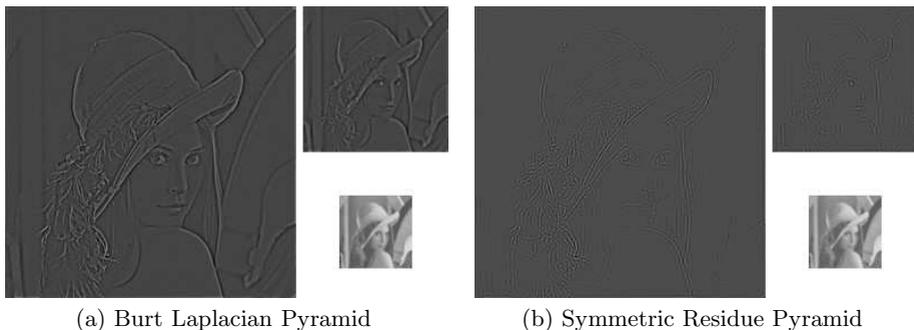
(a) Burt Laplacian Pyramid   (b) Symmetric Residue Pyramid

Fig. 3. Decompositions of `Lena` image

This is just the kind of information that could be provided by the edge model based interpolation or sharpenning of a Laplacian. Moreover, any information fed as a prior $(L_{-1}[0])$, can be considered as an addition of two components: one that is part of the L_{-1} coset corresponding to G_0 and one that is not. The former constitutes the part of the prior that is consistent with the LR and latter the inconsistent part. Once the iteration process is run, the latter is eliminated. This is of importance because it means that our generation of the prior need not be very accurate. A certain amount of error correction can be performed.

Having reviewed the necessary building blocks, the generation of L_{-1} is addressed next.

4   Generation of L_{-1} Using Edge-Model Representation

Given the edge-model based representation of the Laplacian subbands, it is intuitive to use direct method of placing them at calculated position in $2X$ (2 times) to get L_{-1}. We have done experiments which reveal that the modelling of subsampled images does suffer from errors while modelling any of the even or odd location. The even or odd location depends on the choice of pixels in subsampling process. The error becomes prominent when a $2X$ Laplacian image i.e. L_{-1} is generated using the description. Though the error could be suppressed by a choice of higher energy threshold for modelling, it would affect the modelling accuracy at the L_0 level itself. This is shown in Figure 4. The image consists of four 64×64 blocks. The left top pixel of the boxes is placed at even-even (block at left top), odd-even (block at right top), even-odd (box at bottom left), and odd-odd (block at bottom right) locations. Figure 4(d) shows the modelling error at LR only. It is clearly visible that the modelling error is more in case of the blocks placed at the location where one of the starting positions happens to be odd. The bottom right block suffers the most as it is placed at the odd-odd location.

The image formation process can be assumed to be modelled by the expression [15][16] $\mathbf{y} = \mathbf{DBMx} + \mathbf{n}$. Here $\mathbf{M}$ is a warp matrix, $\mathbf{B}$ represent a blur matrix and $\mathbf{D}$ is a subsampling matrix, and $\mathbf{n}$ is noise vector. $\mathbf{y}$ and $\mathbf{x}$ are the
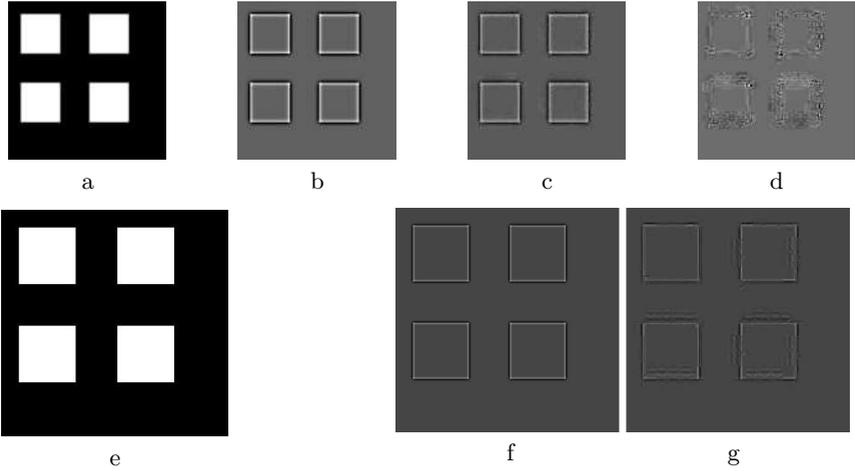
Fig. 4. Subsampling and grid positions. Top: (a) Low resolution (LR) image (b) L_0 of LR (c) Modelled L_0 of LR. (d) modelling error. Modelling errors after subsampling are more for odd grid positions. Bottom: (e) Original (HR) image (f) True L_0 of HR image (g) Estimate of HR's L_0, based on modelling description of L_0 of LR. The interpolation error also depends on position.

LR and HR respectively. **D** breaks the equivalence between even and odd grid points in the HR image. In the above example the modelling elements generated happened to be the ones which correspond to edges at even positions, *i.e.,* the positive peak corresponds to even row or column location in the image. It is not desirable to circumvent this problem by incorporating modelling elements for both even and odd location edges. It doubles the number of elements from PS-28 to PS-56 and increases the complexity of the modelling process. In practise, this approach does not lead to reduction in modelling error. For the specific case of interpolation and edge sharpening, an indirect solution is devised based on post-interpolation modelling.

4.1 Dealing with Modelling Error Due to Subsampling

Our final objective is to define an L_{-1} based on some model reconstruction. As such a model for L_0 itself is not a necessity. One solution is to interpolate L_0 to $\hat{L}_{-1}$ *before* performing the edge modelling. The quality of the interpolation method is clearly important. In our case we have used the symmetric residue process to get the $\hat{L}_{i-1}$, as explained earlier. After interpolation, we generate the edge-model description of the interpolated image's Laplacian subband. The variation in edge patterns due to even-odd positions is attenuated because the equivalence between various grid points gets restored when they are upsampled. As edge representation gets blurred during interpolation, the model is dominated by the elements corresponding to the thicker edges. These modelling elements are now replaced by sharp edge modelling elements. By doing so we undo the
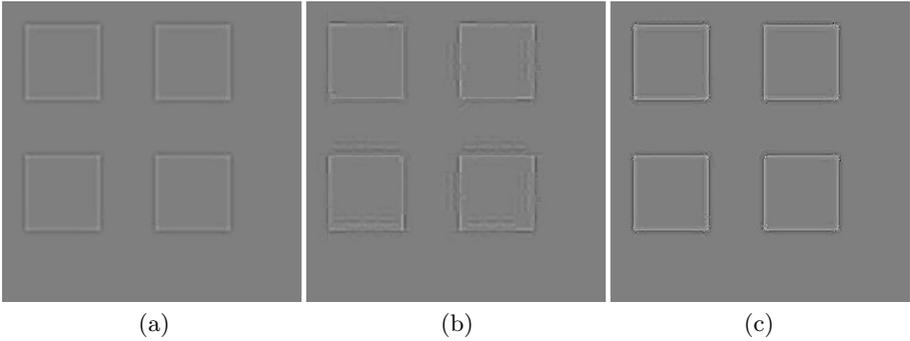
(a)   (b)   (c)

Fig. 5. Generation of sharp edges by post-interpolation model switch. (a) L_{-1} by interpolating L_0 has thick edges. (b) L_{-1} obtained from edge-model description of L_0 has errors (cf Fig 4 g). (c) L_{-1} obtained by modelling the left image and switching model elements before reconstruction has sharp edges with no ringing artefacts.

blurring caused by interpolation and restore the sharpness of edges at $2X$ also. The effect is shown in the Figure 5.

4.2 L_{-1} Generation Algorithm

The process of generating HR images from LR images, as developed in previous sections, is summarized here. The standard pyramid notations are used: initial image is G_0, its first subband is L_0, the HR image is denoted G_{-1} and its first subband is L_{-1}. Edges in Laplacians refer to the their representations as coupled positive-negetive linear structures.

1. Given a G_0 (LR image), generate $\hat{L}_{-1}$ using the iterative process used in SRP with the null prior (**0** image). This gives an estimate of L_{-1} with edges at correct locations but with wrong widths.
2. Generate an edge model description of $\hat{L}_{-1}$.
3. Switch the blur elements to corresponding sharp elements and reconstruct to get $\tilde{L}_{i-1}$. This version of L_{-1} should have sharp edges but may not have lower frequency details due to modelling errors.
4. Using G_0 (LR image), again generate L_{-1} using the iterative process used in SRP, but with $\tilde{L}_{-1}$ as the prior.
5. Genrate the HR image as $G_{-1} = exp(G_0) + L_{-1}$.

The above algorithm uses the SRP and Laplacian edge modelling processes in a way that allows them to complement each other. The SRP based interpolation is able to insert the correct mid- and low-frequency components in $\hat{L}_{-1}$, consistent with G_0. However, it can neither insert nor alter any high-frequency components $(\pi/2 \ldots \pi)$. This leads to blurring of sharp edges. By invoking the assumption that sharp edges arise due to occlusion, we proceed to sharpen all edges in the image that are sharp enough to be in L_0. This is done by modelling the initial estimate $(\hat{L}_{-1})$ and replacing thick edge models with sharp edge models. This revised estimate, $\tilde{L}_{-1}$, has high frequencies injected at the right places.

In order to minimize any errors introduced by the model-switch process, it is regularized by once again passing through the SRP process.

5   Results

In order to evaluate the proposed method, its performance was tested on a diverse set of five images and compared with the results for bicubic interpolation. The representative set is shown in Figure 6, consisting of an artificial image, a texture dominated image, a linear edge dominated image and two portraits. The available images were taken as true HR (THR) and the LR images were generated using blurring and subsampling of these images. The HR images generated from the LR for the proposed and bicubic methods were compared to the THR. The result is shown in Figure 7. The HR images generated by the proposed method has sharper edges. An analysis of the errors indicate that the proposed method does best for isolated edges. In regions having dense edges (as in Barbara texture regions), the errors are largest. This is mainly due to Laplacian edge modelling limitations. However, the errors are no larger than for bicubic interpolation. Thus the modelling step does not introduce instabilities at these regions.
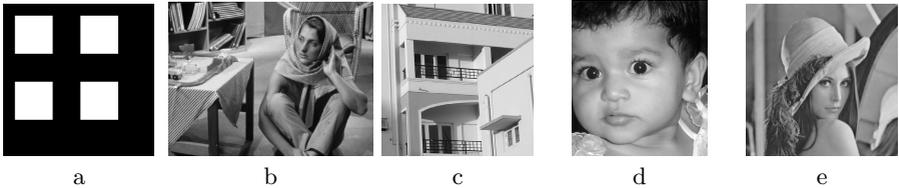

a   b   c   d   e

Fig. 6. Representative set (a) Box (b) Barbara (c) Building (d) Apoorva (e) Lena

5.1   Numerical Evaluation

Aside from visual inspection, it is desirable to have a numerical evaluation of performance. In the present case, availbility of ground truth allows for computation of error based SNR metrics. However, the standard SNR for images would be dominated by the energy present in low frequencies. In order to emphasize the accuracy of the finer details for HR generation, modified measures are used. We have defined HF PSNR (*High Frequency Proportional Signal to Noise Ratio*) where we consider only the HF noise and a proportional amount of the signal power (as actual HF signal is often very small and peak HF power unrealistic). The HF PSNR is thus defined as

$$\text{HF PSNR} = 10 \log \left[\frac{\text{NA} \times \sum |FT_{orig}[i,j]|^2}{\hat{\sum} |FT_{est}[i,j] - FT_{orig}[i,j]|^2} \right] \tag{2}$$

where $\sum$ is summation over all components,
$\hat{\sum}$ is summation over the HF components $[0.25\pi, \pi]$,
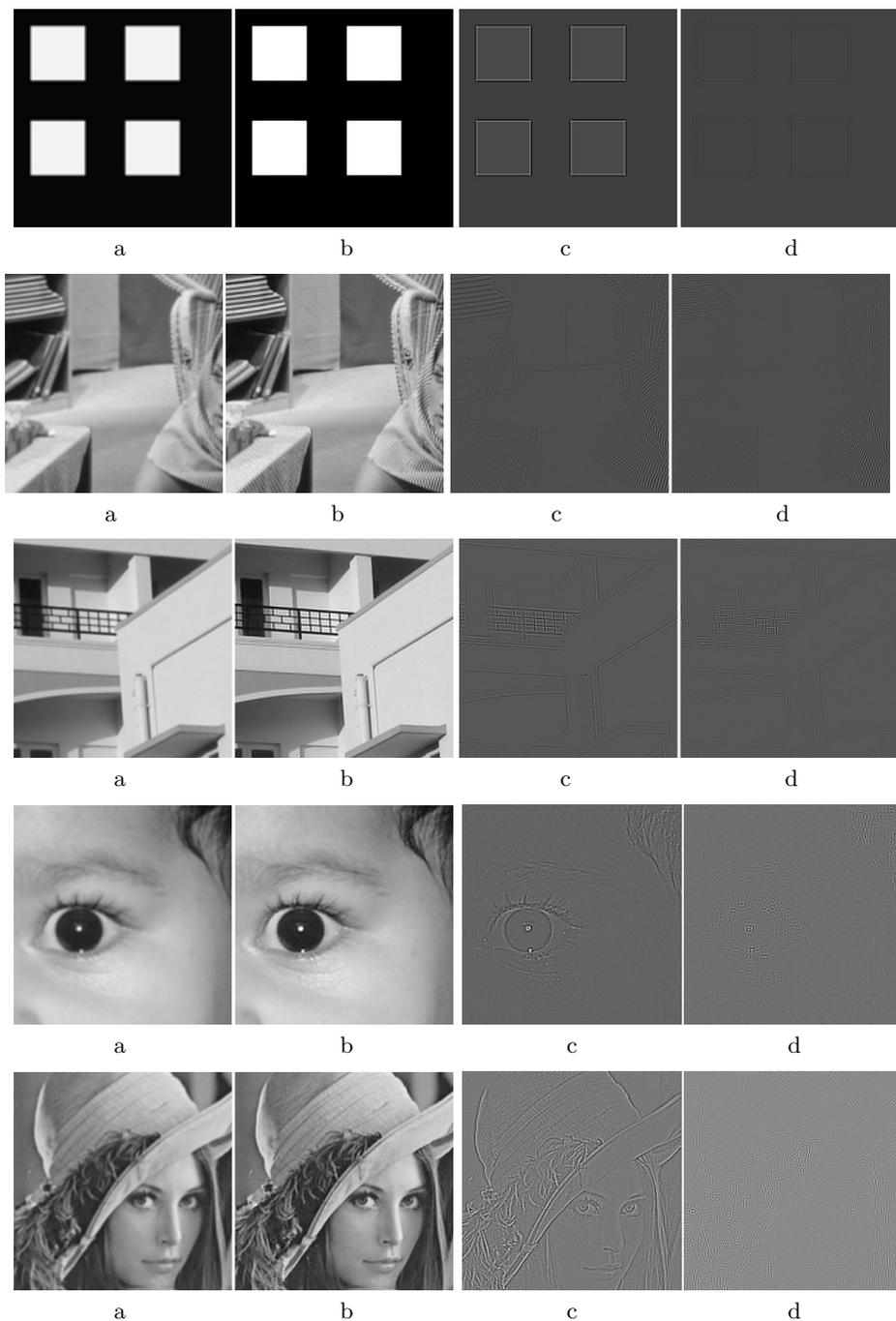NA = fraction of spectral components summed over in the denominator,

Fig. 7. Comparative results (selected areas magnified for display) : (a) Output of Bicubic interpolation (b) Output of our method (c) Error in bicubic interpolation, (d) Error in our method. Dynamic range of residual images are stretched to suit display.

Table 1. The HF PSNR and L$_{-1}$PSNR values for the representative image set. The calculation is done for SRP based method, our method and bicubic interpolation. The HF PSNR is calculated over 0.25π to π.

image	HF PSNR for bicubic	HF PSNR for SRP	HF PSNR for our method	L$_{-1}$ PSNR for bicubic	L$_{-1}$ PSNR for SRP	L$_{-1}$ PSNR for our method
Box	19.11	33.65	36.13	30.98	44.20	46.11
Barbara	24.16	25.76	25.52	33.45	34.46	34.22
Building	24.86	26.58	26.64	32.20	33.42	33.52
Apoorva	34.95	34.94	34.66	37.41	37.20	37.05
Lena	30.68	31.84	31.45	34.17	34.87	34.50

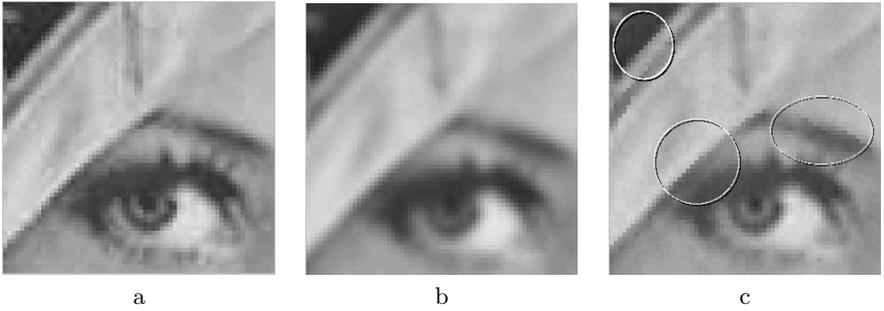|		
a	b	c

Fig. 8. Sharpening of edges (a) Original image (b) Bicubic interpolated (c) Proposed method. Note that some edges have been rendered even sharper than in the original HR image (marked by ovals). This happens due to our underlying model that all sharp edges are occlusion edges that should be rendered as step edges. These enhancements are visually acceptable though they degrade the SNR measures.

$FT_{est}[i,j]$ = Fourier Transform of the generated HR image and
$FT_{orig}[i,j]$ = Fourier Transform of the original HR image.

The HF-PSNR considers all high frequencies, whether or not they play a role in defining edges. Conversely, it ignores the role played by lower frequency components in edges. As the generation of HR was critically dependent on generation of the L_{-1}, we define L_{-1}-PSNR [1] where the PSNR is computed on the L_{-1} image rather than the G_{-1} image.

$$\text{L}_{-1}\text{PSNR} = 10\log\left[\frac{(2^B-1)^2}{\text{MSE}_{\text{L}_{-1}}}\right]. \tag{3}$$

where $\text{MSE}_{\text{L}_{-1}}$ is mean squared error defined for L_{-1} subband as

$$\text{MSE}_{\text{L}_{-1}} = \frac{1}{m \times n}\sum_m \sum_n (L_{-1}[m,n] - \tilde{L_{-1}}[m,n])^2. \tag{4}$$

---

[1] For consistency with earlier sections, we continue to refer to the THR and HR as G_{-1} and the corresponding Laplacian level as L_{-1}.

The HF PSNR and L_{-1}-PSNR values calculated for the sample set is given in Table 1. In addition to bicubic, we have also compared the proposed method to interpolation by SRP using the **0** prior.

In Table 1 it is observed that the proposed method gives good results for all types of images, but its comparative advantage varies from image to image. For images with sharp straight edges (`Box, Building`) it is the best. For soft images (`Apoorva`) it scores less, though the output looks sharp in Figure 7. This discrepency is explained on detailed examination of the errors. The true high resolution image itself may not have perfectly sharp edges due to optics or processing. The proposed method makes them sharp by opting for the model element change. This is shown in Figure 8. While such sharpenning may be visually acceptable (or even be desirable), they degrade THR based SNR measures.

6   Conclusion and Future Work

A method of achieving HR image from a single LR image has been proposed. It is based on a generic generative edge model that removes the requirement for any training set and makes the method widely applicable. It exploits the model based description of Laplacian subbands and the Symmetric Residue pyramid techniques to generate a putative Laplacian subband corresponding to the desired HR image. The results of this method are good. At some places the edges are rendered sharper than in the original HR image. As no ringing artefacts are created by this over-compensation, it is not a major concern as far as Super-Resolution is concerned. Other errors are introduced due to modelling failures. Both the modelling and the SRP processes may need to be optimized for this method. Future direction of work will also focus on using this single image approach in association with reconstruction-based approaches to exploit multiple LR images optimally.

Acknowledgements

References

1. Ur, H., Gross, D.:   Improved resolution from sub-pixel shifted pictures. CVGIP:Graphical Models and Image Processing **54** (1992) 181–186
2. Nguyen, N., Milanfar, P.: An efficient wavelet-based algorithm for imaqe superresolution. In: Proc. Int. Conf. Image Processing. Volume 2. (2000) 351–354
3. Tsai, R., Huang, T.: Multiple frame image restoration and registration. In: Advances in Computer and Image Processing, CT: JAI Press Inc. (1984) 317–339
4. Rhee, S., Kang, M.: Discrete cosine transform based regularized high-resolution image reconstruction algorithm. In: Opt. Eng. Volume 38. (1999) 1348–1356

5. Chaudhuri, S., ed.: Super-Resolution Imaging. Norwell,MA: Kluwer Academic (2001)
6. Park, S.C., Park, M.K., Kang, M.G.: Super-resolution image reconstruction: A technical overview. In: IEEE Signal Processing Magazine. (2003) 21–36
7. Joshi, M., Chaudhuri, S.: Super-resolution imaging: Use of zoom as a cue. In: Proc. ICVGIP, Ahmedabad, India. (2002)
8. Joshi, M., Chaudhuri, S.: Zoom based super-resolution through sar model fitting. In: Proc.Intl Conf. on Image Processing (ICIP), Singapore,. (Oct. 2004)
9. Jiji, C.V., Joshi, M.V., Chaudhuri, S.: Single frame image super-resolution using learnt wavelet coefficients. Intl. J. Imaging Science & Tech. (special issue on high resolution image reconstruction) **14** (Sep. 2004) 105–112
10. Jiji, C.V., Joshi, M.V., Chaudhuri, S.: Single frame image super-resolution through contourlet learning. EURASIP J. Applied Signal Processing (2006) 1–11
11. Chaudhuri, S., Joshi, M.: Motion-Free Super-Resolution. (Springer)
12. Burt, P., Adelson, E.: The laplacian pyramid as a compact image code. IEEE Trans. Commuication **31** (1983) 532–540
13. Nema, M.K., Rakshit, S.: Edge-model based representation of laplacian subbands. In: Proc. Seventh Asian Conf. on Computer Vision (ACCV 7), Hyderabad, India,. (Jan. 2006) 80–89
14. Rakshit, S., Nema, M.K.: Symmetric residue pyramids: An extension to burt laplacian pyramids. In: Proc. IEEE ICASSP, Hong Kong,. (2003) III–317–III–320
15. Elad, M., Feuer, A.: Restoration of a single superresolution image from several blurred, noisy and undersampled measured images. IEEE Trans. Image Processing **6** (1997) 1646–1658
16. Naguyen, N., Milanfar, P., Golub, G.: Efficient generalizd cross-validation with applications to parameteric image restortion and resolution enhancement. IEEE Trans. Image Processing **10** (2001) 1299–1308