Image Resolution Enhancement using NSCT based Learning with LBP as Feature Model

Karishma V. Koladia[1] Prakash P. Gajjar[2] Parul V. Pithadia[3]
[1, 2, 3] E & C Department
[1]L. D.College of Engineering Ahmedabad, India [2] Government Polytechnic for Girls Surat,India
[3] Government Engineering College Surat,India

Abstract---In this paper, we propose a new technique for feature preserving spatial resolution enhancement of an image captured at low spatial resolution. We use a training database containing low resolution (LR) images and their high resolution (HR) versions. In an image, different features like edges, corners, curves and junctions are important to convey its local geometry. We use Local Binary Pattern (LBP) operator to represent texture of an image. The missing high resolution details of the low resolution observation are learnt in form of Non-subsampled Contourlet Transform (NSCT) coefficients of the high resolution images in the training database. We demonstrate the effectiveness of the proposed technique by conducting experiments on real world gray scale images. The results are compared with existing learning-based approaches. The proposed technique can be used in applications such as medical imaging, remote surveillance, wildlife sensor networks where the transmission bandwidth, the camera cost and the memory are main constraints.

Keywords: Non-Subsampled Contourlet Transform(NSCT) ,Local Binary Pattern (LBP), Super-Resolution (SR)

I. INTRODUCTION

In the field of photogrammetry, high resolution (HR) images are often desired. HR imaging offers more details that can be useful for better analysis, interpretation and classification of information in an image. One way to obtain HR images depends on hardware solution. High precision optics and charge coupled devices (CCDs) can be used to obtain HR images directly from the camera. But, this solution is not appropriate for general purpose commercial applications due to application specific limitations such as cost, memory, sensor dimensions and shot noise, transmission bandwidth etc. Therefore, many algorithmic approaches are designed to obtain HR images. Super-Resolution (SR) methods attempt to recover a high resolution image from one or more low resolution (LR) images. Methods for SR can be classified into two main categories: (i) Classical Multi-image SR and (ii) Example-based SR. The classical multi-image SR method uses a set of LR images of the same scene to obtain a superresolved image[14], [19], [15], [17]. This method is limited to only small increase in resolution. This limitation of the classical multi-image SR methods has led to the development of ’Example-based SR’ also termed as ’Learning-bases SR’ or ’Image Hallucination’ [18], [22]. Since then a number of example-based SR algorithms have been proposed [16], [17], [20], and [21]. This method learns the correspondences between low and high resolution image patches from a database of low and high resolution images pairs and recovers the high resolution version of the given LR test image.

The authors in [2] use a training database containing LR images and their HR counterparts all captured from a real camera and obtains initial HR estimate using local binary pattern (LBP) operator and discrete wavelet transform (DWT) based learning. DWT divides the high frequency only into three sub-bands at each level (i.e. horizontal, vertical and diagonal). It is suitable for describing Isotropic point singularity but is not able to detect line or curve singularity. Wavelets are blind to the smoothness along edges commonly found in natural images. So, wavelets cannot accurately express the direction of edges and contours in an image. The limitation of the approach is also due the inability of DWT to capture geometrical regularities and to detect contours. *Patel et al.* [4] approach introduces an algorithm to reconstruct HR version of an LR observation using local ternary pattern (LTP) and discrete cosine transform (DCT). The limitation of this approach is that DCT gives shift variant representation of an image and poor recovery of edge information. DCT is also unable to describe directional features of images efficiently because DCT coefficients reflect the pixel intensities of an image into DC and AC coefficients with low and high frequency regions. In [5], the authors propose an algorithm for image resolution enhancement using a training database and direct pixel based learning with the help of non-subsampled contourlet transform (NSCT). Because of direct pixel based learning method used in this approach, the local geometry and texture of an image is not well preserved in the resulting HR image. Considering this limitation of the aforementioned methods, we present a new technique for image resolution enhancement using NSCT based learning and LBP as a feature model. We use a non-subsampled contourlet transform (NSCT) to learn its coefficients for the missing high frequency details from the detailed sub-bands of high resolution training images in the database. NSCT is a flexible and efficient transform.

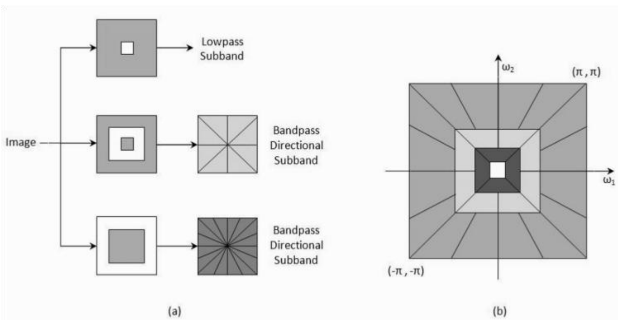

Fig. 1: Non-subsampled contourlet transform. (a)NSFB structure that implements the NSCT.(b)Idealized frequency partitioning obtained with the NSCT.

This fully shift-invariant, multiscale, and multi direction expansion transform is effective for images with smooth contours which allows different angular resolution at different scale and direction. NSCT is efficient in image denoising and image enhancement. In the proposed approach, NSCT is used as the multiscale transform (MST) tool to provide a better representation of the contours and overcome limitations of DWT and DCT.

The remainder of the paper is organized as follows.We briefly describe non-subsampled contourlet transform(NSCT)and filter banks used in it in section II. We then discuss about local binary pattern (LBP) operator in section III. We describe different steps of the proposed algorithm in Section IV. It is followed by Section V, where we describe our experimental setup, training database and present our results for real world images. We finally conclude the paper in Section VI with a summary of the obtained results.

II. NON-SUBSAMPLED CONTOURLET TRANSFORM

The non-subsampled contourlet transform (NSCT) is a shift-invariant version of contourlet transform. NSCT has excellent properties in the process of image decomposition including shift invariance, multiscale and multidirection [25]. This efficient and flexible transform is effective for images with smooth contours which allows different angular resolution at different scales and directions [8]. Fig. 1(a) displays an overview of the NSCT. Fig. 1(b) illustrates the structure consisting of a bank of filters which splits the 2-D frequency plane in the sub-bands.The NSCT is divided into two shiftinvariant parts: 1) a non-subsampled pyramid structure that gives the multiscale property and 2) a non-subsampled DFB structure that ensures directionality.

A. *Non-subsampled Pyramid (NSP)*

The multiscale property of the NSCT is obtained by NSP that achieves a sub-band decomposition similar to that of the Laplacian pyramid. It uses two-channel non-subsampled 2-D filter banks. Fig. 2(a) illustrates the non-subsampled pyramid (NSP) decomposition. For each subsequent stage, the filters are obtained by upsampling the filters of the first stage. The multiscale property is achieved without using additional filter design. A similar decomposition can be obtained by removing the downsamplers and upsamplers in the Laplacian pyramid and then upsampling the filters accordingly. Thus, certain parts of the noise spectrum in the processed pyramid coefficients has been filtered[9], [10].

B. *Non-subsampled Directional Filter Bank (NSDFB)*

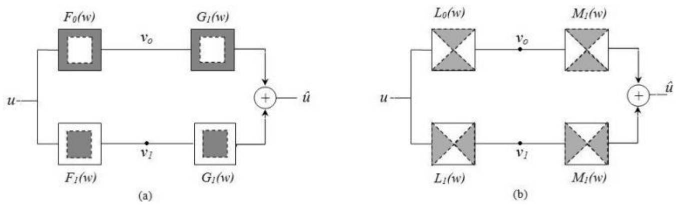

Fig. 2: Non-subsampled contourlets and filter banks (a)Pyramid NSFB. (b)Directional NSFB

The critically-sampled two-channel fan filter banks and resampling operations have been combined to construct NSDFB. This results in a tree-structured filter bank that splits the 2D frequency plane into directional wedges. A non-subsampled DFB (NSDFB) obtains a shift-invariant directional expansion.

The NSDFB is constructed by eliminating the downsamplers and upsamplers in the DFB. This is done by switching off the downsamplers/upsamplers in each two-channel filter bank in the DFB tree structure and upsampling the filters accordingly. Thus, a tree composed of two-channel non-subsampled filter banks (NSFBs) is obtained. Fig. 2(b) illustrates the non-subsampled DFB decomposition[11].

The NSCT is constructed by combining the NSP and the NSDFB as shown in Fig. 1(a). The frame elements are localized in space and oriented along a discrete set of directions. The NSCT offers flexibility because it allows any number of directions in each scale. We observe that there are three classes of pixels in the transform coefficients: strong edges, weak edges and noise.The strong edges correspond to the pixels having large magnitude coefficients in all sub-bands. The weak edges correspond to the pixels having large magnitude coefficients in some directional sub-bands but small magnitude coefficients in other directional sub-bands within the same scale. The noise corresponds to those pixels with small magnitude coefficients in all sub-bands[8]. The NSCT is shift-invariant so that each pixel of the transform sub-bands corresponds to that of the original image in the same spatial location. So, we can gather the geometrical information pixel by pixel from the NSCT coefficients.

III. LOCAL BINARY PATTERN

A local binary pattern operator was first introduced as a descriptor summerizing the local gray-level structure for texture classification by *Ojala et al.*[1]. LBP has been adapted for various applications such as smoke detection, object detection, dynamic texture recognition, face recognition and shape localization.The LBP operator transforms an image into an array or image of integer labels describing small-scale appearance of the image. These labels are known as LBP codes which can be used for further image analysis.Generally, the LBP operator works in a 3x3 pixel block of an image. The pixels in this block are thresholded with respect to the value of its center pixel and we get an unique binary code for each image patch. This LBP code is converted into decimal value that gives the label to the center pixel of each block. By encoding the entire image using LBP operator, the local geometry and texture can be represented. As the neighborhood consists of 8 pixels, a total of $2^8 = 256$ labels can be obtained depending on the relative gray values of the center and the pixels in the neighborhood [27].

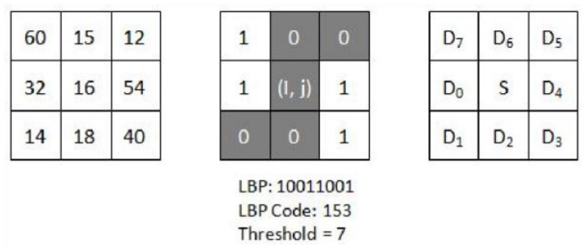

Fig. 3: Binary Pattern and LBP coding of a 3x3 image patch

Consider a 3x3 image block. Let Us be the pixel intensity at location *(i,j)* and D_k be the pixel intensities of the neighboring pixels of *(i,j)*. For a center pixel having location *(i,j)*, the LBP code can be computed using the following equation.

$$L(i,j) = \sum_{k=0}^{7} f(D_k - S)2^k$$

where,

$$t(a) = \begin{cases} 1 & if \quad |a| \quad > \theta \\ 0 & if \quad |a| \quad \leq \theta \end{cases}$$

Here,θ is a threshold value.An example of a 3x3 image patch, its binary pattern and LBP coding are given Fig.3. The local binary patterns are exceedingly informative in conveying reliable local structural information about the image. In this algorithm, we use LBP as image feature model for analyzing local geometric structures and comparing them for finding matching HR images from the database.

IV. THE PROPOSED APPROACH

The block diagram of the proposed approach is shown in Fig.4. The training database consists of low resolution training images and their high resolution versions all captured using a real camera. These images in the database cover a wide range of scenes. Low resolution training images are captured with 1 x zoom setting and high resolution training images are captured with 2x zoom setting of camera.

Flow of the algorithm goes through four main steps. Initially, we upsample the test image using bicubic interpolation method by a factor of 2. Then we decompose the test image and HR training images by a non-subsampled contourlet transform (NSCT). In the second step, we model the image features present in the sub-band *0* of the test image and HR training images using LBP operator. The third step searches HR training images with identical LBP code at each pixel in sub-band *0* to that of the test image. In the last step, we learn the missing high frequency details in the form of NSCT coefficients. Finally, we reconstruct the super-resolved image by applying inverse NSCT. The following subsections give detailed description of the proposed approach.

A. Image Feature Modeling Using LBP

The local structural variations of an image are represented by various image features in terms of edges, corners, curves and junctions. The second step of the proposed algorithm is to represent those features using local binary pattern (LBP).The local binary patterns can represent fundamental properties oftexture, providing the vast majority of 3x3 patterns present in images.This 8-neighborhood system allows us to represent eachfeature with an 8-bit LBP code.

The LBP is a highly discriminative operator in the sense that it uniquely records the influence of neighboring pixels on the center pixel and the occurrences of various patterns in the neighborhood of each pixel. This helps us to model different features such as edges, corners, curves, and junctions in an image. Some of the observed features and their LBP codes are shown in Fig.5. The LBP code formed after thresholding the 8neighboring pixels dictate the underlying geometric structure. The value of the center pixel is modified according to this binary pattern, thus the LBP code represents the feature enclosed in an image patch under consideration. We capture the local texture patterns of the image in the form of LBP codes by encoding the entire image using LBP operator. We apply LBP coding to the sub-band *0* of NSCT decomposition of bicubic version of the test image and HR training images.

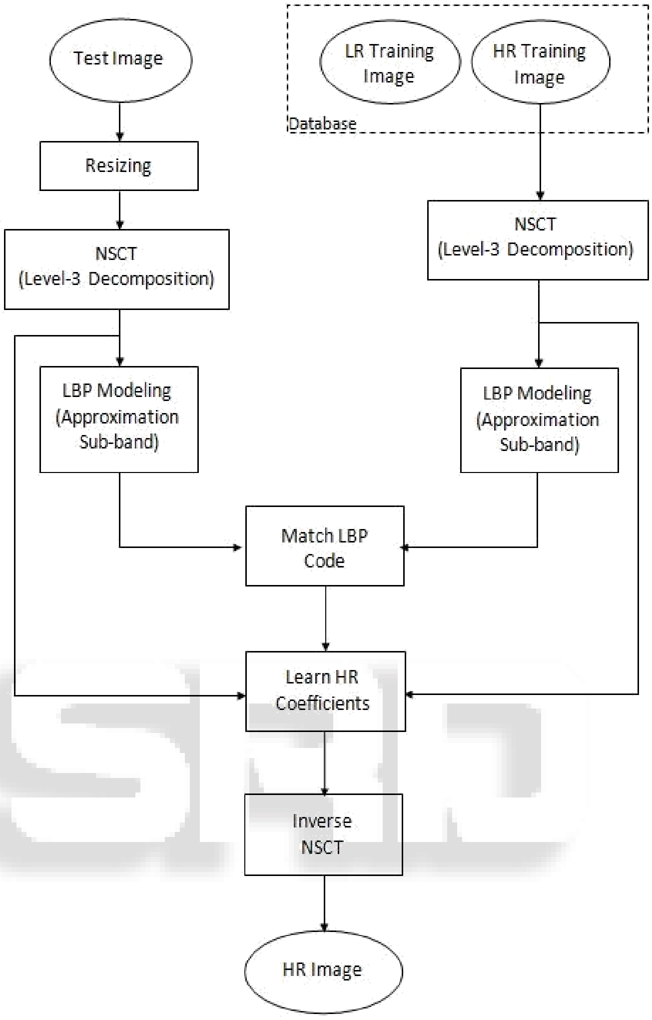

Fig. 4: Schematic representation of the proposed approach for image resolution enhancement

B. Searching Matched HR Training Images

In this step of the proposed algorithm, we search the training database to identify the HR training images withsimilar structure for each pixel in the sub-band *0* at level3 NSCT to that of the given test image.

Here, we take the advantage of the fact that the pixels with similar LBP codes represent similar geometric structures. We use such similarities for learning HR structures. After allocating the LBP code to each pixel location in the sub-band *0* of the test image and all HR training images, we search for the HR training images with identical LBP code to that of the test image for each pixel location. Ultimately, we proceed to search for the HR training images having features same as the features around the pixel of the test image under consideration. To find the best matching HR training image from the group of all matched HR training images, we need to compare NSCT coefficients as described in the next section.

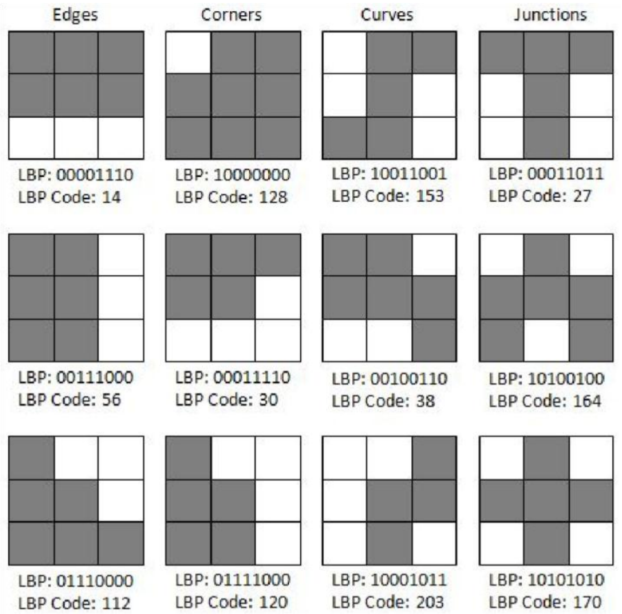

Fig. 5: Samples of image features and corresponding LBP codes. Each column represents three different samples for various image feature.

C. *Mapping HR coefficients*

This section describes the method to learn NSCT coefficients from the matched HR training images and to map those HR coefficients to sub-band *III* of NSCT decomposition the test image. In order to obtain the best matching HR training image, we refine the matching criterion. This is achieved by comparing all 6 NSCT coefficients in sub-bands *I* and *II* of the test image with the corresponding coefficients of each HR training image included in the group of matched HR training images. Finally, we map the remaining 16 HR coefficients of sub-band *III* from the best matching HR training image to corresponding pixel locations of the test image.

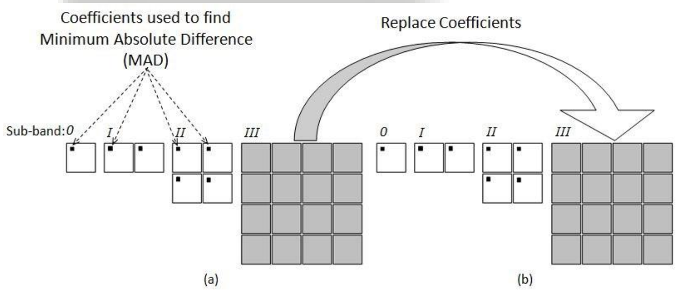

Fig. 6: Replacing HR coefficients(a)NSCT coefficients of best matching HR training image (b)NSCT coefficients of upsized test image

Let $N_0(i,j)$ be the NSCT coefficient at a location (i,j) in the sub-band 0 of the test image. Similarly, $N_1(i,j)$,$N_2(i,j)$ and $N_3(i, j)$ to $N_6(i,j)$ correspond to sub-band I and sub-band II respectively. We find the HR training image having the best matching geometrical features with the test image in terms ofminimum absolute difference (MAD) by comparing $N_1(i,j)$ to $N_6(i,j)$. Then we replace the corresponding 16 coefficients of sub-band III of the test image by that of the best matching HR training image for a given location (i,j). Here, we use the following equation to obtain MAD.

$$\hat{C}(i,j) = \operatorname{argmin}[|N_1(i,j) - N_1^{(m)}(i,j)| + |N_2(i,j) - N_2^{(m)}(i,j)| + |N_3(i,j) - N_3^{(m)}(i,j)| + |N_4(i,j) - N_4^{(m)}(i,j)| + |N_5(i,j) - N_5^{(m)}(i,j)| + |N_6(i,j) - N_6^{(m)}(i,j)|]$$

Where,$\hat{C}(i, j)$ is an index of best matching HR training image from database corresponding to the location (i,j) and $1 \leq \hat{C}(i,j) \leq L$. Here, m is the index of the HR training image and k is the index of NSCT coefficient. m=1,2....L where, L is the number of matched HR training images in the group. Fig. 6 illustrates representation of NSCT coefficients in different sub-bands and mapping of HR coefficients from the sub-band III of the best matching HR training image to that of the given test image.

V. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of the proposed technique to enhance spatial resolution of a low resolution test image using a training database consisting of LR and HR image pairs. All the experiments are conducted on real world images. The test images are of size 64x64 and the super-resolution is achived for an upsampling factor of 2. The size of the super-resolved image is 128x128. Our training database consists of 250 sets of images all of real world scenes. Each set consist of an LR image and corresponding HR versions. A few of the LR training images from the database are selected as test images. Thus their true high resolutions are available for comparison of the reconstructed images. LR and HR pairs of the test images are removed from the database. The quantitative comparison of the results is presented using structural similarity (SSIM) and peak signal to noise ratio (PSNR).The SSIM score of the entire image is computed by averaging the SSIM values of the patches across the image[6], [7]. Higher values of SSIM and PSNR indicate better performance respectively. All the experiments were performed on computer having IntelCore i3 processor working at 2.20GHz and having 4GB RAM.

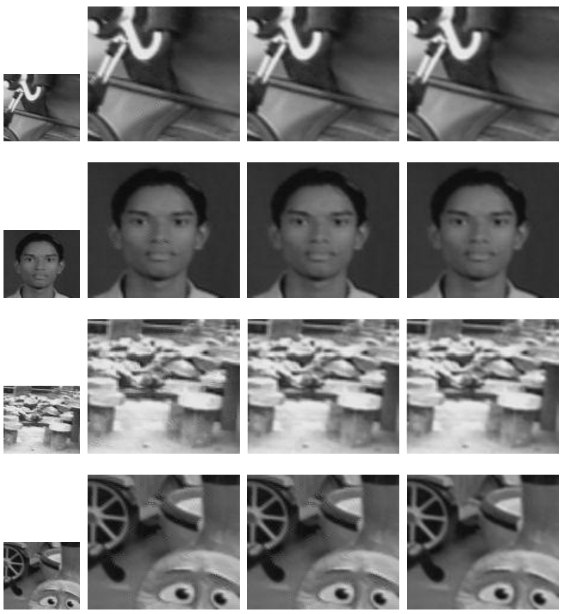
Fig. 7: (a)LR test image of size 64x64 (b)Resultant image of Patel et al. approach[4] (c)Resultant image of Koladia et al. approach[5] (d)Resultant image of proposed approach

Results of the experiments conducted on the real world images are shown in Fig.7. Fig.7(a) shows the observed low resolution test images. Fig.7(b) shows the images obtained using *Patel et al*.[4] approach. Fig.7(c) includes the images obtained using *Koladia et al*.[5] approach. Fig.7(d) shows the super resolved images obtained using the proposed approach.The test images used for experiments contain different image details. Image 1 contains horizontal and vertical edges. Image 2 shows a human face with various features. Image 3 is a scene consisting of various like junctions and curves. Image 4 contains different features like sharp edges, curves and contours as well as smooth areas.

Image	Patel et.al approach[4]	Koladia et.al approach[4]	Proposed approach
SSIM			
1	0.9050	0.9157	0.9187
2	0.9307	0.9340	0.9364
3	0.8201	0.8287	0.8292
4	0.8724	0.8872	0.8897
PSNR			
1	64.7408	64.8852	64.9381
2	77.1294	77.9424	78.0328
3	53.2624	53.2925	53.2938
4	63.4814	63.7623	63.7967

Table 1: Results Comparison

The quantitative performance using SSIM and PSNR as measurement indices is shown in Table I. It shows considerable improvement in the figures of the proposed approach as compared to the Patel et al. approach and Koladia et al. approach. Higher value of PSNR implies that the super resolved images are less noisy. Use of NSCT helps to extract edges and contours and leads to better structure preservation thus improving SSIM score of the resultant images.

VI. CONCLUSION

In this paper, we have presented a new technique for image resolution enhancement. We use non-subsampled contourlet transform (NSCT) coefficients to learn the missing high frequency information of the test image from high resolution training images contained in the database. We use LBP coding to detect the important features and texture of an image sothat they can be preserved in the resulting HR image. The experimental results show that proposed algorithm outperforms the other resolution enhancement techniques. The quantitative comparison shows a considerable improvement in PSNR and fairly good SSIM scores. The super resolved images are better in terms of quantity and appearance because local geometric structure information like directional edge details and contour like structures are reconstructed well. The use of NSCT in the proposed algorithm helps to describe line singularities and different directional information more accurately. The use of LBP helps in image modeling to preserve geometric structures of an image when it is reconstructed with higher resolution with a scale factor of 2.

REFERENCES

[1] T. Ojala and M. Pietikainen and T. T. Maenpaa, Multiresolution GrayScale and Rotation Invariant Texture Classification with Local BinaryPatterns, IEEE transactions on pattern analysis and machine intelligence. vol. 24, no. 7, pp.971-987, 2002.

[2] P. V.Pithadia and P. P. Gajjar, ’Feature Preserving Super-Resolution: Use of LBP and DWT’, IEEE International conference on Devices, Circuits and Systems(ICDCS) vol. 41, no. 12, pp. 34453462, 2012.

[3] P. V.Pithadia and P. P. Gajjar, ’Super-resolution Using DCT Based Learning With LBP as Feature Model’, IEEE International conference on Computing, Communication and Networking Technologies(ICCCNT). vol. 41, 2012.

[4] K. P.Patel and P. P. Gajjar, ’Learning based Image Super Resolution using DCT with Local Ternary Pattern operator’, IEEE International conference. vol. 41, no. 12, pp. 34453462, 2012.

[5] K. V. Koladia and P. V. Pithadia, ’Image Resolution Enhancement using Non-Subsampled Contourlet Transform’,International Journal ofMultidisciplinary Educational Research(IJMER) pp.2277 - 7881 , 2014.

[6] Z. Wang and A. C.Bovik and H. R.Sheikh and E. P.Simoncelli, Image quality assessment : From error visibility to structural similarity,IEEE Transaction on Image Processing. vol. 13, no. 4, pp. 600-11, 2004.

[7] Z. Wang and A. C.Bovik, Mean squared error: love it or leave it ?,IEEE Signal Processing Magazine. ,pp. 98-117, 2009.

[8] Arthur L. da Cunha, Jianping Zhou and Minh N. Do, The Nonsubsampled Contourlet Transform:Theory, Design, and Applications, IEEE Transactions on Image Processing. vol. 15, no. 10, 2005.

[9] P. J. Burt and E. H. Adelson, ’The Laplacian pyramid as a compact image code’, IEEE Transactions on Communication., vol. 31, no. 4 , pp. 532540, 1983.

[10] M. N. Do and M. Vetterli,’Framing pyramids’, IEEE Transactions on Signal Processing, vol. 51, no. 9, pp. 23292342, 2003.

[11] R. H. Bamberger and M. J. T. Smith, ’A filter bank for the directional decomposition of images: Theory and design’, IEEE Transactions Signal Processing, vol. 40, no. 4, pp. 882893, 1992.

[12] E. Mohan, K. B. Jayarraman, U. Maheswaran, D. Sathiyaraj and G.Dhakshanamoorthi, ’A Novel Approach for Satellite Image Resolution Enhancement’, International Journal of Engineering and Advanced Technology (IJEAT), vol. 2, no. 4, pp. 2249 8958, 2013.

[13] Mr. G. M. Khaire, Prof. R. P.Shelkikar, ’Resolution Enhancement of images with Interpolation and DWT-SWT Wavelet Domain Components’, International Journal of Application or Innovation in Engineering and Management (IJAIEM) vol. 2, no. 9, pp. 2319 - 4847, 2013.

[14] K. V. Suresh, G. Mahesh and A. N. Rajagopalan, ’Super-resolution of license plates in real traffic videos’, IEEE Transactions on Intelligent Transportation Systems vol. 8, no. 2, pp. 321331, 2007.

[15] S. Farsiu, M. Robinson, M. Elad and P. Milanfar,’Fast and robust multiframe super resolution’, Transactions on Image Processing vol. 13, no. 10, pp. 13271344, 2004.

[16] P. P. Gajjar and M. V. Joshi, New Learning based Super-Resolution: Use of DWT and IGMRF Prior, IEEE Transactions on Image Processing vol. 19, no. 5, pp. 1201-1213, 2010.

[17] T. A. Stephenson and T. Chen, 'Adaptive markov random fields for example-based super-resolution of faces', EURASIP Journal of Applied Signal Processing vol. 2006, no. 2, pp. 1-11, 2006.

[18] W. T. Freeman, T. R. Jones and E. C. Pasztor,'Example-based superresolution', IEEE Computer Graphics and Applications vol. 22 , no. 2, pp. 56-65, 2002.

[19] D. Capel, 'Image Mosaicing and Super-Resolution', SpringerVerlag 2004.

[20] D. Capel and A. Zisserman,'Super-resolution from multiple views using learnt image models', Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition pp. 627-634, 2001.

[21] C. V. Jiji, M. V. Joshi and S. Chaudhuri,'Single-frame image superresolution using learned wavelet coefficients', International Journal of Imaging Systems and Technology vol. 14, no. 3, pp. 105-112, 2004.

[22] W. T. Freeman, E. C. Pasztor and O. T. Carmichael,'Learning low level vision', International Journal of Computer Vision vol. 40, no. 1 , pp. 25-47, 2000.

[23] S. S. Sisina,'Multiresolution Imagefusion using Non-subsampled Contourlet Transformation',International Journal of Engineering Research and Technology (IJERT) vol. 2, no. 9, pp. 2278-0181, 2013.

[24] Peyman Milanfar, 'Super-Resolution Imaging', Taylor and Francis/CRC Press. Series: Digital Imaging and Computer Vision. 2010.

[25] Liang Xu, Junping Du, and Qingping Li, 'Image Fusion Based on Nonsubsampled Contourlet Transform and Saliency-Motivated Pulse Coupled Neural Networks', Mathematical Problems in Engineering, Hindawi Publishing Corporation, id 135182, 2013.

[26] Mayank Agrawal and Ratnakar Dash, A thesis report on 'Development of Image Super-Resolution Algorithms', National Institute of Technology, Rourkela, 2013.

[27] M. Pietikainen, A. Hadid, G. Zaho, T. Ahonon, 'Computer Vision using Local Binary Patterns',Springer. pp.978-0-85729-747-1, 2011.