Volume : 4 | Issue : 7 | July 2015 • ISSN No 2277 - 8179

Derivation of Optimum Image Data Base Size in Super Resolution using DCT with Local Ternary Pattern Operator

Engineering

KEYWORDS : Image Database , Discrete Cosine Transform(DCT), Local Ternary Pattern operator(LTP), Super Resolution(SR).

Kiran P. Patel | Lecturer, Department of EC, Govt. Polytechnic, Gandhinagar

ABSTRACT

In this paper, I present what should be an ideal size of image data base. There is always a enough database size so that we can have best possible Super Resolved image in less time . I use a new learning based technique to super-resolve a low resolution image using "Local Ternary Pattern Operator" and a database of low resolution (LR) images and corresponding high resolution (HR) versions. The local geometry of an image is conveyed by image features such as edges, corners and curves. We encode these features with local ternary pattern operator. The missing high resolution features of the low resolution observation are learnt in the form of discrete cosine transform coefficients from high resolution images in the training database. Experiments are conducted on real world natural images and results are compared with taking different sizes of database of low resolution (LR) images and corresponding high resolution (HR) versions. The results help us in deciding the optimum size of the database.

I. INTRODUCTION

In electronic imaging applications, high quality images with fine details are often desired. Images with high resolution (HR) offer more details. High Resolution means that pixel density within an image is high, and therefore an HR image can offer more details that may be critical in various applications. The learning based super resolution uses a database consisting of high resolution images to learn the missing high frequency details of the super-resolved image. Many researchers have attempted to solve the problem of learning based super resolution. Freeman et al. [8] were first to solve the super-resolution problem by learning the finer details from a database of high resolution images. Learning based algorithms are generally used when a single low resolution image is available as an observation. In learning-based SR, correspondences between low and high resolution image patches are learned from a database of low and high resolution image pairs (usually with a relative scale factor of 2), and then applied to a new low-resolution image to recover its most likely high-resolution version. In learning-based SR, this missing high-resolution information is assumed to be available in the high-resolution database patches, and learned from the low-res/high-res pairs of images in the database.. High resolution imaging could be helpful obtaining a better classification of regions in an image, in diagnosis a medical image or a more pleasing view in a HD television. In this paper, we have proposed a super-resolution approach to reconstruct a high resolution image using learning based technique. The missing high frequency details are learned in the form of discrete cosine transform coefficients from the database of LR-HR image pairs. While learning, I reconstruct the local structural information of the image so that the HR image contains sharp edges and corners. Local geometry of the image is captured by modeling the features with local ternary pattern (LTP) operator. The high frequency coefficients in the DCT transform domain are used to reconstruct the fine details pertaining to high resolution information. I have taken different sizes of image database & observed the resulted HR image & time taken to reconstruct the HR images. The rest of the paper is organized as follows. In section II, I describe the proposed approach with four clear steps: Discrete cosine transform, image feature modeling, searching similar features and learning HR feature coefficients. I illustrate the performance of the proposed approach in section III. Finally, we conclude in section IV.

II. THE PROPOSED APPROACH

THe block diagram of the proposed approach is shown in Fig. 1. The training database consists of 250 low resolution training images and their high resolution versions all captured using a real camera. Low resolution images are captured with 1-x zoom setting and high resolution images are captured with 2-x zoom setting of the camera. The flow of the algorithm goes through four steps. Initially, we upsample the test image and the LR training images with bicubic interpolation method by a factor of 2. Then we take a block based discrete cosine transform of test image and HR training images. In the second step we model image features using LTP. Next we search the training database for similar features that match to the features in the test image. We then learn the high frequency details in the form of discrete cosine transform coefficients and finally obtain the super resolved image by applying inverse DCT.

A. Discrete Cosine Transform

Like other transforms, the Discrete Cosine Transform (DCT) attempts to decorrelate the image data. After decorrelation each transform coefficient can be encoded independently. We use DCT transform domain here as DCT coefficients reflect the pixel intensities of the image into DC coefficients and AC coefficients with low and high frequency regions that can be utilized to recover image details in the region of interest.

B. Image Feature Representation using LTP

B..I. Local Ternary Pattern

The Local Ternary Pattern (LTP)[14] at a point c is defined by:

$$\mathrm{LTP}(C) = \sum_{k=0}^{7} 3^k s(I_k - I_c) \qquad (1)$$

$$h(u, i) = \begin{cases} -1, & (u-i) < -t \\ 1, & (u-i) > t \\ 0, & else \end{cases} \qquad (2)$$

The LTP encoding process is illustrated in fig.2(b).However, in LTP, calculations becomes huge. So LTP can be converted into two Local Binary Patterns,namely Up LTP and Down LTP. Mathematically, let:

$$LTP(C) = \sum_{k=0}^{7} 2^k s(I_k - I_c) \qquad (3)$$

$$h_{up}(u, i) = \begin{cases} 1, (u-i) > t \\ 0, \quad else \end{cases} \qquad (4)$$

$$h_{dw}(u, i) = \begin{cases} 1, (u-i) < -t \\ 0, \quad else \end{cases} \qquad (5)$$

Then, the LTP vector is the vector of these ternary values over a given neighborhood . Thus, the ith bit of the LTP vector for the point c is given by:

$$LTP_c(i) = h(I_{ki}, I_c) \qquad (6)$$

Where ki is the ith neighborhood point for c.Tan and Triggs[14].

C. Searching similar features

In this step our algorithm searches the training database to find LR training images with similar features to that of the test image. For this we divide the LTP encoded test image and LR training images

into 8×8 image patches. We compare 8×8 = 64 pixels in each patch of the test image with the corresponding image patches of all the LR training images in the database and select the LR training image with maximum number of matching pixels in the patch under consideration as the best match for that patch. This indicates that there are maximum matching features such as edges, corners and curves in the patch of the test image and the best matching image patch of the LR training image. Thus we have a matching LR training image with similar features corresponding to all the 8×8 image patches of the test image.

We apply this procedure to each image patch in the test image and find best matching LR training image for the corresponding patch.

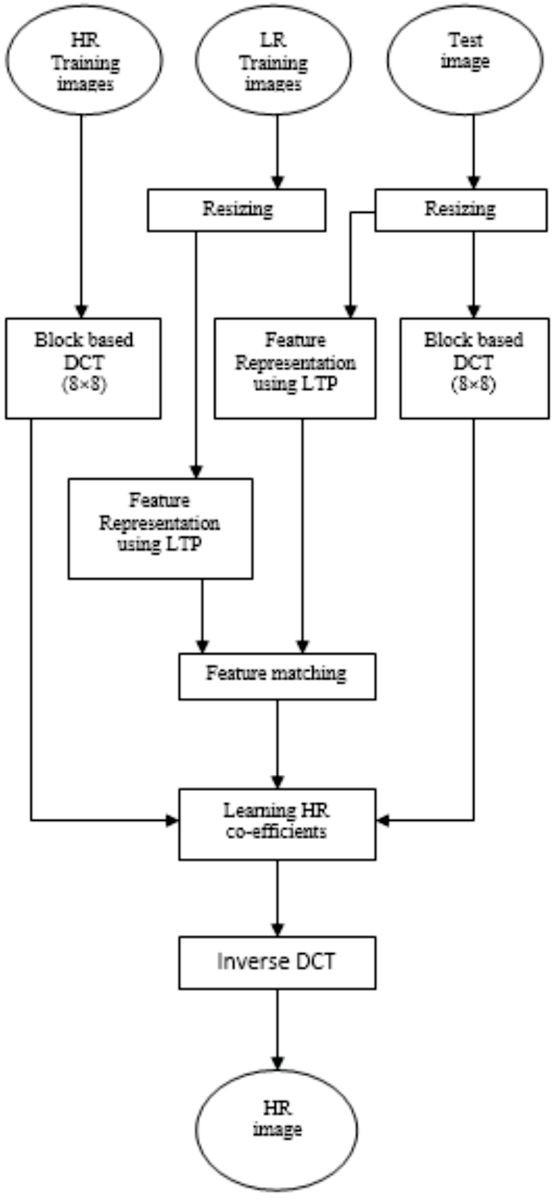


Fig. 1. Block diagram of the proposed approach.

D. Learning HR Feature Coefficients

Once we get the best matching LR training image corresponding to each patch of the test image, we learn the high resolution information in the form of detailed DCT coefficients from the HR training database. Since HR training image in the database is the true high resolution image, we extract the high frequency details from it for improving the resolution of the test image. We extract the DCT coefficients pertaining to the high frequency details from the HR image that corresponds to the best matching LR image and copy these DCT coefficients to appropriate locations in the block of test image. We repeat this process for each 8 x 8 blocks of the test image. Once all the DCT coefficients are learned, we apply inverse DCT to the learned image to obtain high resolution version of the test image.

III. EXPERIMENTAL RESULTS

The training database consists of a set of 250 images of various scenes. The LR and HR images are of sizes 64×64 and 128×128, respectively. The test images are of size 64×64 and the super-resolved images are of size 128×128. I started with database of 100 images & increased in the number of 25 , i.e , 125 , 150 , 175 ,200 etc & observed the time taken to reconstruct the HR image in each case .A few of the LR training images from the database are selected as test images. Thus their true high resolution versions are available for comparison of the reconstructed images. LR-HR pairs of the test images are removed from the database.The quantitative comparison of the results is presented using mean squared error (MSE) and peak signal to noise ratio (PSNR) defined by,

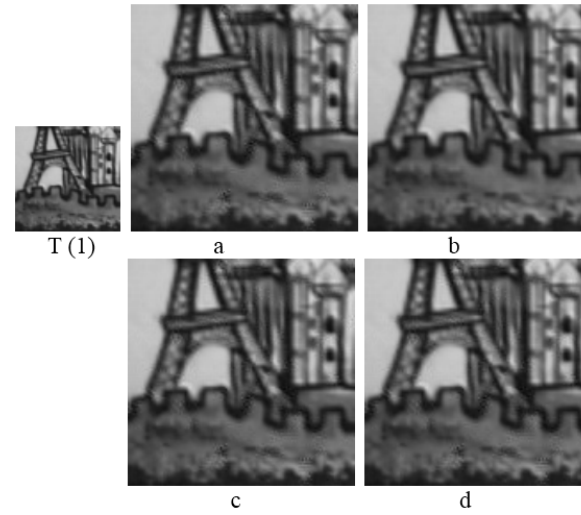
$$MSE = \sqrt{\frac{\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}[f(i,j)-f'(i,j)]^2}{M\times N}} \quad (7)$$

$$\text{PSNR} = 10\log 10\left[\frac{255\times 255}{\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{(f(i,j)-f'(i,j))^2}{M\times N}}\right] \quad (8)$$

where f(I; j) is the original HR version of the reconstructed image *f* (i,j) and M×N is the size of the images. The structural similarity (SSIM) index is a method for measuring the similarity between two images. The local similarity S(x; y) between image patches x and y taken from the same locations of the two images under comparison is calculated as,

$$S(x,y) = \left(\frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}\right)\left(\frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}\right)\left(\frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}\right) \quad (9)$$

Where ,μx and μy are local sample means of image patches x and y, respectively. Similarly σ_x and σ_y are the local sample standard deviations of x and y, r espectively. Here σ_{xy} is the sample cross correlation of x and y after removing their means. C1, C2 and C3 are small positive constants preventing the numerical instability. The SSIM score of the entire image is computed by averaging the SSIM values of the patches across the image [5], [6]. Higher value of S(x; y) indicates better performance respectively. All the experiments were performed on computer having Intel®Core™ i3-2330m processor working at 2.2GHz and having 3GB RAM.



Volume : 4 | Issue : 7 | July 2015 • ISSN No 2277 - 8179

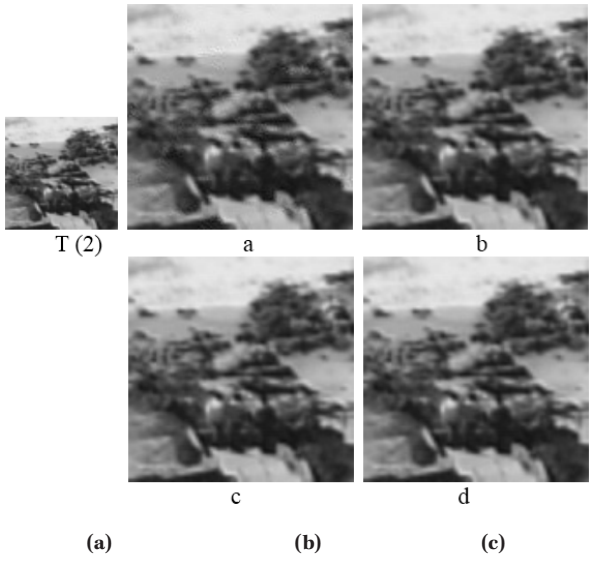


Fig. 2 Results. (T) Low resolution observations (64 × 64),(a) super-resolved images using 100 images Database (128 × 128), (b) 125 images Database (128 × 128), (c)150 images Database (128 × 128) (d) 175 images Database (128 × 128)(e) respectively.

Fig.2 shows the results of the experiments conducted for the up sampling factor $f = 2$ and threshold 2. The test images are shown in fig 2(T). The image expanded using 100 images database are shown in fig 2 (a) . Images obtained using 125 images database are shown in fig. 2 (b). & 150 & 175 images database in fig. 2 (c) & fig. 2 (d) respectively. Test images used in experiment are different in terms of picture details.

Table I shows the quantitative performance using MSE, PSNR and SSIM as measurement indices.

Table 1
Performance comparision

Image database size	MSE	PSNR	SSIM	Time taken to reconstruct the HR image (in Minutes)
Image T (1)				
100	237.3135	55.4431	0.8343	11
125	237.3215	55.4428	0.8352	14
150	237.3242	55.4513	0.8361	17
175	237.3251	55.4527	0.8372	20
200	237.3254	55.4531	0.8376	23
Image T(2)				
100	222.3590	56.2424	0.8053	10
125	222.3611	56.2435	0.8057	13
150	222.3624	56.2444	0.8067	16
175	222.3638	56.2458	0.8078	19
200	222.3641	56.2460	0.8082	22

IV. Conclusion

I have presented a approach to trade off between time & quality of a super-resolved image using single observation. Image features are recovered by representing them with local ternary pattern operator that encodes the local structural variations into LTP codes. The missing high frequency details are learned from a database in form of discrete cosine transform coefficients. As we can see after some point the time taken to reconstruct the image is much more than the improvement in the quality of the reconstructed image. One can choose the appropriate database size which can yield good results in minimum possible time depending on one's application & requirement.

REFERENCE

[1] Jing Tian and Kai-Kuang Ma, 'A survey on super-resolution imaging ,signal ,Image and Video processing' ,(5)(3) (2010) 329-342. | [2] S.C.Park,M.K.Park and M.G.Kang , 'Super-Resolution Image Rcognition: A Technical Overview' , IEEE Signal Processing Magazine,(20)(3)(2003) 21-36. | [3] S.Borman and R.L. Stevension, 'Spatial Resolution Enhancement of Low- Resolution Image Sequences:A Comprehensive Review with Directions for Future Research,Technical Report', University of Notre Dame,1998. | [4] 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,(13)(4)(2004) 600-11 | [5] Z.Wang and A.C.Bovik,' Mean squared error: love it or leave it ?',IEEE Signal Processing Magazine, (2009) 98-117 | [6] T. Ojala and M. Pietikainen and T. T. Maenpaa ,' Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local BinaryPatterns', IEEE transactions on pattern analysis and machine intelligence.(24)(7) (2002) 971-987 | [7] Changhyun Kim, Kyuha Choi, Ho-young Lee, Kyuyoung Hwang and Jong Beom Ra, 'Robust learning-based super-resolution', IEEE International Conference on Image Processing (ICIP), (2010) 2017 - 2020 | [8] S. Rajaram and M. D. Gupta and N. Petrovic and T. S. Huang, 'Learning based nonparametric image super resolution', EURASIP Journal of Appllied Signal Processing. (2006)(2) (2006) 1 – 11. [9] C. V. Jiji and M. V. Joshi and S. Chaudhuri, 'Single-frame image superresolution using learned wavelet coefficients', International Journal of Imaging Systems and Technology. (14)(3) (2004) 105-112. | [10] W. T. Freeman and E. C. Pasztor and O. T. Carmichael 'Learning lowlevel vision', International Journal of Computer Vision.(40) (1) (2000)25-47. | [11] K.P.Patel ,P.V.Pithadia,P.P.Gajjar , 'Learning based Image Super Resolution using DCT with Local Ternary Pattern operator' .IEEE International conference on Computing, Communicat ion & Networking Technologies(ICCC 2013). | [12] X.Tan and B.Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in IEEE InternationalWorkshop on Analysis and Modeling of Faces and Gestures, 2007, pp.168–182. | [13] Saeed Dabbaghchian, Masoumeh P. Ghaemmaghami and Ali Aghagolzadeh,' Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology', Pattern Recognition, (43)(4) (2010) 14311440. | [14] Nuno Roma and Leonel Sousa, A tutorial overview on the properties of the discrete cosine transform for encoded image and video processing, Signal Processing, (91)(11) (2011) 24432464. | [15] P. P. Gajjar and M. V. Joshi, 'New Learning based Super-Resolution: Use of DWT and IGMRF Prior', IEEE Transactions on Image Processing. (19)(5) (2010) 1201-1213