International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 24 Issue 9 – JUNE 2017.

A SELF-EXAMPLE BASED SUPER RESOLUTION ALGORITHM FOR SURVEILLANCE IMAGES

Divya Priya.T.A
PG Student,Dept. of ECE,
SSN College of Engineering,
Chennai, Tamil Nadu, India

Jino Hans.W
Assistant Professor,Dept. of ECE,
SSN College of Engineering,
Chennai, Tamil Nadu,India

***Abstract*—This paper presents an Example-based super-resolution for surveillance images in which high-resolution (HR) image is reconstructed from single input image without depending on external dataset of images. The main idea of our technique is that random patches selected in HR images match some of the patches in low-resolution image (LR). Self-examples are obtained by learning a dictionary which consists of LR input image generated by a double pyramid of recursively scaled and interpolated images. Moreover, a linear regression function is learned to map the correspondence between HR patch and its corresponding LR patch. Finally, objective metrics such as PSNR and SSIM has been evaluated. Comparison with state-of-the-art methods, based on external and internal dictionary shows that our algorithm produce clearer images with sharp edges.**

Keywords-Super-resolution, Surveillance images, Self-examples, double pyramid, regression

I. INTRODUCTION

Super-resolution (SR) is a technique in which high-resolution images are reconstructed from low-resolution images. SR mainly aims to increase the resolution of the images. The higher the resolution, the more the image details. SR images are required for many application particularly, for surveillance purposes in which enhancement of smaller, region of interest objects is very necessary.

SR methods are commonly classified into two major categories. a) Multi-frame super-resolution b) Single image super-resolution. Both the methods differ from each other with respect to the input images. In Multi-frame SR, more number of input images are used to recover the HR images whereas, in single-image SR only one input image is used for the process. Park et al. [1] has given a good description about multi-frame SR methods. Here, multiple images of a same scene with different views is used to estimate the super resolved output image. However, computational complexity arises due to more number of input images.

Single-image SR (SISR), on the other hand, aims to construct a SR image without the need of large set of external images i.e. a single unique input image is enough to reconstruct the images.Single-image SR predict the missing HR details that are not present in the given original image. SISR methods are classified into two main divisions: interpolation-based method [2]-[4], and learning method. Interpolation-method is the simplest method to implement. It determine the unknown pixel in HR image by using interpolation function. However, this method produces ringing effects and blurring artifacts particularly, in natural images with strong discontinuities. Due to the powerfulness and benefits of learning method numerous literature has been appeared in recent years.

Learning based algorithm follows pixel-based procedures, in which HR output value is inferred via statistical learning [5], [6], or patch-based procedures, where HR value is estimated from a dictionary which consists of HR and its corresponding LR patches. Example-based SR is a learning algorithm which follows patch-based procedure [7].

Example-based SR recover the HR images from LR image by learning the high frequency details from the training image pairs. An efficient prior must be learned to find the correspondence between the HR and LR image pairs. The image prior is of two types: Explicit prior, implicit prior.Explicit prior makes use of mathematical energy function of an image such as contour let, gradient profile etc. The training time of these explicit priori based algorithm is generally very large for big training set

Implicit priors are learned from training image pairs. It requires a collection of HR images and its corresponding LR images. The correspondence between the HR images and its LR images is learned as an implicit prior. The implicit prior can be learned either by direct mapping approach or indirect mapping approach. The indirect mapping approach includes algorithm such as K-nearest neighbor [8]-[10], which is computationally expensive for practical applications. The direct mapping approach learns the relation between LR and

HR images as a regression function and it is also computationally efficient.

In this paper, an example-based SR algorithm which uses internal dictionary to recover HR images is presented. First, a “double pyramid” is build where traditional image pyramid [11] is connected with a pyramid of interpolated images. Finally, HR patches are reconstructed through direct mapping approach. A matrix-based implicit prior is used to learn the relationship between HR patches and LR interpolated patches. This matrix-based implicit prior also preserves the image level information and thus will reduce artifacts.

The objective of this paper is to estimate the high frequency details from low-resolution surveillance images and to preserve the image level information and also to evaluate the effectiveness of the algorithm with state-of-the-art methodologies.

The paper is organized as follows: the section 2deals with the self-example based SR and Matrix-based implicit prior. The proposed SR algorithm for surveillance images is discussed in section 3, in which construction of double-pyramid, matrix-value regression operator and the reconstruction phase are explained. The performance of the proposed algorithm is evaluated and compared with other state-of-the-art methods in section 4 and finally the conclusion of the paper is given in section 5.

II. SELF-EXAMPLE BASED SR

Single image SR is a process of reconstructing the HR image by observing only one unique LR image. We consider the following equation (1) to generate LR image from HR Image,

$$I_L = (I_H * G) \downarrow_S \quad (1)$$

Where, I_L and I_H are the LR and HR image respectively, G is the Gaussian blur kernel and $\downarrow_S$ represents down-sampling by a scale factor of s.

Example-based SR recover the HR image from the above generated LR image by the use of a dictionary. The dictionary consists of HR image and its LR image counterparts. The LR image is the down-sampled version of HR image. The LR and HR patches are thus collected in a dictionary. Example-based SR consists of two phase.

1) A training phase, where the dictionary with LR and HR patches are constructed.

2) The Super resolution phase, where the HR image is reconstructed from the LR image.

A. Self-examples

In traditional methods, to obtain a relationship between low resolution images and high resolution images, external training dataset is needed. This is simply to achieve a similar quality of the internal representation of high resolution image. Moreover, this approach requires hundreds of external images, which makes the system inefficient. To tackle this problem, a pyramidal structure for a single image is obtained as it contains more internal representation of the image at different scales. The input image is either conveniently, down-scaled or up-scaled to different scale factors. The correspondence between these LR and HR patches is learned to obtain the super resolved output image. The patches learned in this way, is called as the self-examples. Thus, a single image alone is sufficient to obtain the super-resolution image using a combination of matrix and statistical techniques.

Two kinds of learning schemes are defined in the literature: one step schemes [12], [13] and [14], and image pyramidal schemes in which recursively scaled pyramids are constructed. In one step schemes, only one training image pair is used for the construction of dictionary, thus reducing the computational time of the algorithm. The technique depicted in [11] made a way for many self-example based algorithm using image pyramids. In contrast to one-step scheme, many training image pair obtained by recursively down-sampling the LR input image The later method is used in this paper to build the dictionary.

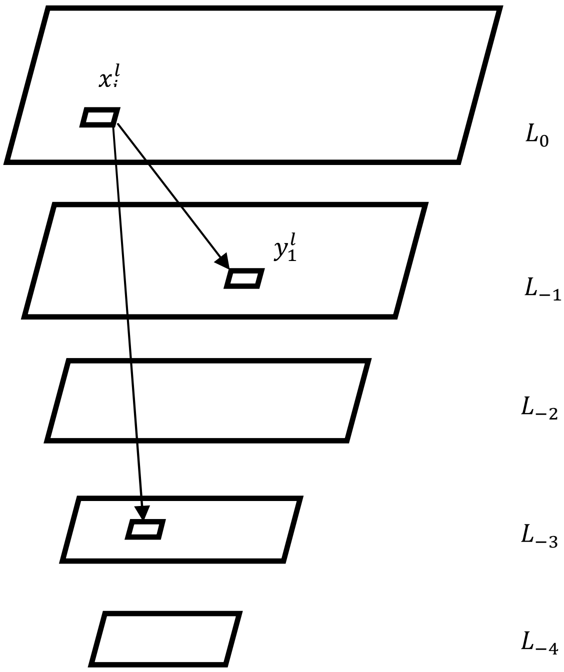


Fig. 1 Construction of single pyramid

Pyramid construction and searching of self-examples are depicted in Fig. 1.$L_o$represents the low resolution input image andL_{-1}, L_{-2}, L_{-3}, L_{-4} are the four sub-levels. Here, x_i^l represents the given input low resolution patch. Having x_i^l , neighbors (y_1^l, y_2^l ....) can be found in pyramid at any level

B. Matrix-based implicit prior

The correspondence between LR and HR patch is learned by a linear matrix-based regression operator. Let I= $(x_i, y_i)_{i=1}^n$ be the training image pair set, where x and y are LR and HR patch respectively of size p x p. An optimal matrix regression (x,y)

$$y=R.x \quad (2)$$

If the patches are assumed to be full rank matrices, then the regression operator can be deduced from (2)

$$R = yx^{-1} \quad (3)$$

where, x^{-1} is the Moore-Penrose inverse matrix of x . But most of the natural images are not full-rank matrices. Hence, the regression operator cannot be determined as described in (3). A least square regression is solved to find the optimal matrix value operator.

III. PROPOSED SR ALGORITHM

In this section, we presented our proposed SR algorithm based on self-examples using double pyramidal approach. The flow chart representing the proposed methodology is shown in Fig. 2

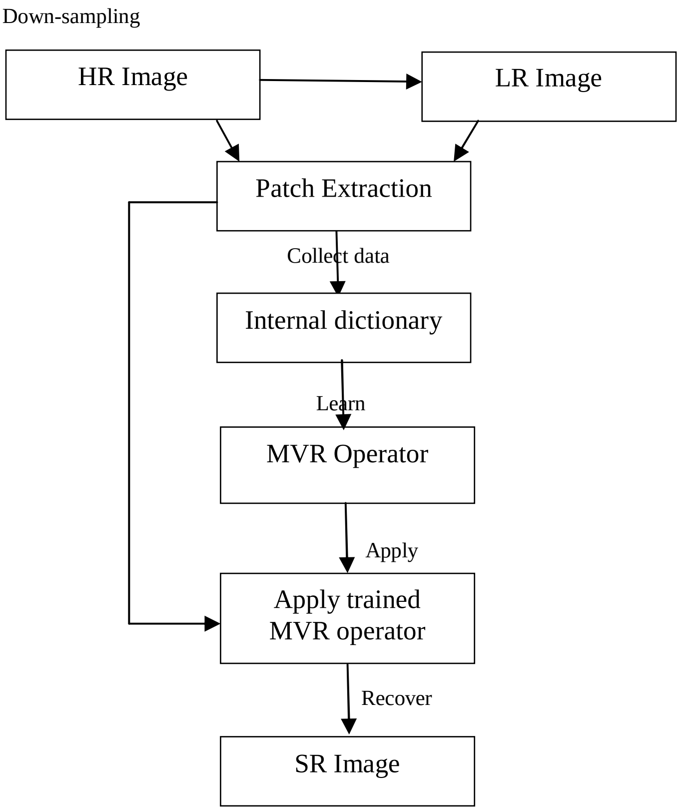


Fig. 2 Methodology of the proposed algorithm

A. *Construction of Double Pyramid*

The objective of SR algorithm is to recover HR image from the degraded version of LR image as given in image generation model (1). The blur kernel G is chosen to be Gaussian kernel ofvariance σ_G^s, where s is the scale factor. If the LR input image is of size M x N, then the final super-resolved output image is of size *s*M x *s*N.

As a beginning stage for our internal dictionary learning method, we first take the single pyramid described in Fig. 1. A certain number of sub-levels are constructed from the LR input image itself (L_0), by using the following relation.

$$L_{-n} = (L_0 * G_n) \downarrow_{p^n} \quad (4)$$

where, p is the scale-factor of the pyramid. The sub-level image $L_{-n}$is the down-scaled version of the input imageL_0. The Gaussian kernelG_n is then computed using the formula (5) as described in [16]

$$\sigma_{G_n}^2 = n.\sigma_G^2 \ . \log(p)/\log(s) \quad (5)$$

After creating a single pyramid, each sub-level L_{-n} is interpolated by a factor p. The interpolated level thus obtained H(L_{-n}), where H is the up-scaling operator. The image size of the interpolated level H(L_{-n}) is same as the non-interpolated level present above in the pyramid L_{-n+1} .Thus H(L_{-n}) and L_{-n+1} forms the HR and LR training image pair respectively. The LR and HR internal dictionary is formed by using the pair sets [H(L_{-n}), L_{-n+1}] , where n=1,2….N, where N is the number of sub-levels, LR and HR patches of equal size are formed. Fig. 3 describes the above scheme, with the “double pyramid”.

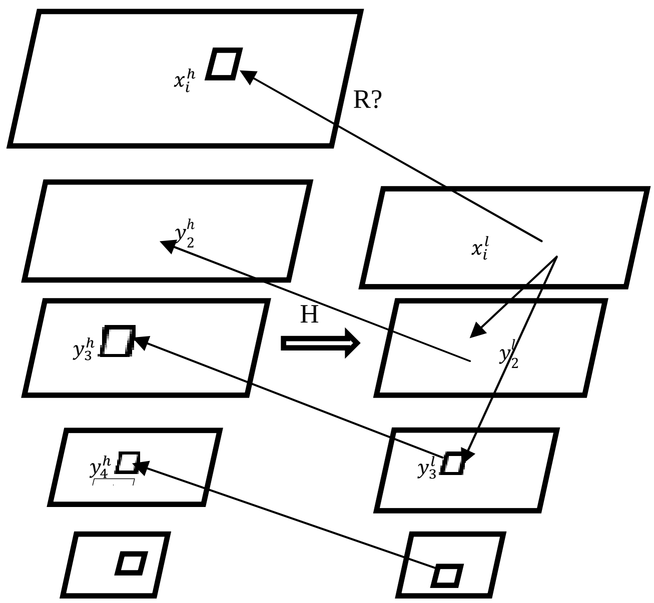


Fig. 3 Searching of self-examples in double-pyramid

B. Matrix-Value Regression Operator

Let T = {(x_1, y_1),……(x_n, y_n)} denotes the training image patch-pair set. Let R be the operator that maps LR patch with HR patch. Here, x_i represents the low resolution patch and $y_i$represents the high resolution patch. To learn the optimal matrix R^*, a least square regression technique is used from the training set T.

R^*=argmin$_R \sum_{i=1}^{n} \|y_i - Rx_i\|_F^2$ (6)

where, $\|.\|_F$ represents the Frobenius norm. Let *F(R)*= $\|y_i - Rx_i\|_F^2$. Taking partial fraction and equating it to zero gives,

$$\frac{\partial}{\partial R}F(R) = 0 \qquad (7)$$

Now, equation (7) becomes,

$$\frac{\partial}{\partial R}\sum_{i=1}^{n}(<y_i, y_i>_F - 2<y_i,\ Rx_i>_F + <Rx_i, Rx_i>_F) = 0$$
$$\frac{\partial}{\partial R}\sum_{i=1}^{n}(<y_i, y_i>_F - 2<y_i x_i^T,\ R>_F + <Rx_i x_i^T, R>_F) = 0$$

$$\Sigma_i^n Rx_i x_i^T = \Sigma_i^n y_i x_i^T$$

$$\text{R} = \Sigma_i^n y_i x_i^T (\Sigma_i^n x_i x_i^T)^{-1}$$

The optimal matrix regression value operator is given by,

$$R^* = S_1 S_2^{-1} \qquad (8)$$

where, $S_1 = \Sigma_i^n y_i x_i^T$ and S_2 = $(\Sigma_i^n x_i x_i^T)$ are auxiliary matrices. The inverse of S_2 is found by factorizing using Singular Value Decomposition (SVD). Hence, S_2= UΣV^T where, Σ is a diagonal matrix with singular values and V and U are orthogonal matrices. Therefore,

$$R^* = S_1\ (V\ \Sigma^{-1} U^T)\ (9)$$

The regression operator in (9) gives the relation between the low resolution patch and high resolution patch from which fine details can be reconstructed to produce super-resolved output image.

C. Reconstruction phase.

In this phase, the given input test image I is interpolated with a scale factor s. Patches of size N X N is obtained from both the input HR image and the LR image, which is obtained after interpolation. The correspondence between the collection of LR and HR patches are learned by the optimal matrix value regression operator obtained in the equation (). This regression operator super-resolve the LR patch, resulting in the super-resolved image output.

$$I_{hr} = R^* * I_{lr}\ \ (10)$$

where, I_{lr} and I_{hr} are LR and HR patch, respectively and R^* is the optimal matrix regression operator. Merging the super-resolved patches gives the super-resolution HR image.

IV. RESULTS AND PERFORMANCE METRICS

In this section, the proposed algorithm is tested on some images shownin Fig.4. These images are used as input from which super-resolved images are obtained and compared with other state-of-art example-based SR algorithm. We have considered bi-cubic interpolation, Sparse SR algorithm to analyze the efficiency.Surveillance images shown in Fig. 4 are downloaded from Caltech webpage dataset is used to evaluate the efficiency by simulating using MATLAB.The single-input image is downscaled to obtain the LR images from which HR and LR patches are extracted and super-resolved using a scale-factor of 3. The original code are obtained from respective author’s page for the other methods in Table 1. Performance metrics such as Peak to Signal Noise Ratio (PSNR) and Structural Similarity Index (SSIM) is calculated and summarized in Table 1. PSNR is the measure of quality between original input image and scaled up image.

a) Image 1

b) Image 2

c) Image 3

Fig. 4 Input Images

It describes the variation in pixel values. If the PSNR value is high, then the quality of the image will be higher. PSNR is computed using the mean square error which is given below

$$PSNR(x,y) = 10 log_{10} (\frac{255^2}{MSE_{xy}})$$

where,

$MSExy = ||x - y||^2 / W*H$,

W is the width of both images, and H is the height of both images patches.

SSIM is used to measure the similarity between the input image and the reconstructed image.

$$SSIM_{(x,y)} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

where,

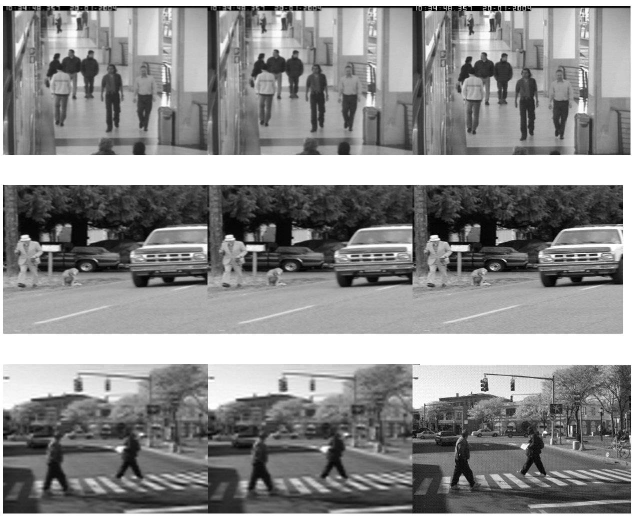
$$\mu_x = \frac{1}{W * H \sum_{i=1}^{W*H} x_i} \; ; \; \mu_y = \frac{1}{W * H \sum_{i=1}^{W*H} y_i}$$

$$\sigma_x = \frac{1}{W * H - 1 \sum_{i=1}^{W*H} ((x_i - \mu_x)^2)^{\frac{1}{2}}}$$

$$\sigma_y = \frac{1}{W * H - 1 \sum_{i=1}^{W*H} ((y_i - \mu_y)^2)^{\frac{1}{2}}}$$

Input Image	Bi-cubic		Sparse SR		Proposed	
	PSNR	**SSIM**	**PSNR**	**SSIM**	**PSNR**	**SSIM**
Image1	26.17	0.7389	28.84	0.8154	**29.91**	**0.8795**
Image2	31.62	0.6572	32.02	0.7547	**33.63**	**0.7991**
Image3	25.82	0.7003	26.92	0.7396	**27.71**	**0.7957**

Tab. 1 Performance comparison of PSNR and SSIM with other state-of-the-art methods



(a) (b) (c)

Fig. 5. Comparative results of the super-resolved surveillance images. The methods considered are: (a) Bi-cubic interpolation, (b) Sparse SR, (c) Proposed method

Subjective assessment of SR relies on some properties of the recovered SR image. The image is outwardly reviewed for its instinctive nature and sharpness to evaluate the quality of the recovered image. The high-frequency details in the image is used to evaluate the sharpness of an image. The quality of the image is affected if the fine-details of the image is not protected, resulting in ringing effects.These characteristics in the images can be assessed by visually comparing the images.

The visual comparison of the surveillance images with the other state-of-the art methods is shown in Fig. 5. The super-resolved output for all the methods clearly explains that our proposed algorithm produces clearer images with less artifacts than the other methods.

Tab.1 depicts the quantitative measure of the images. It clearly shows that the proposed method outperforms other state-of-the-art algorithm. SSIM value indicates that image-level information is preserved and shows the similarity between the ground truth and SR output image and also high PSNR value shows that the output image is recovered with less distortions.

V. CONCLUSION

In this paper, an example-based super-resolution algorithm which uses a single-image for reconstruction is presented. The algorithm initially, constructs a “double pyramid” to extract patches from a single-image and it is used to build an internal dictionary consisting of both LR and HR patches. A linear regression operator is learned to find the relation between both these patches and finally the output SR image is reconstructed. PSNR and SSIM has considerably high gain compared to other existing methods. The output super-resolved image is also visually cleaner with less artifacts. The algorithm makes use of only one single-image, without relying on external training database making it computationally efficient.

REFERENCES

[1] S. C. Park, M. K. Park, and M. G. Kang,“Super-resolution image reconstruction: A technical overview,” IEEE Signal Processing, vol. 20, no. 3, pp. 21-36, May 2003.

[2] X. Li and M. T. Orchard, “New edge-directed interpolation,” IEEE Trans. Image processing., vol. 10, no. 10, pp. 1521-1527, Oct. 2001.

[3] M. F.Tappen, B. C. Russell, and W. T. Freeman, “Exploiting the sparse derivation prior for super-resolution and image demosaicing”, in Proc. 3rd IEEE Int. Workshop. Comput. Theories Vis., 2003, pp. 1-4.

[4] R. Fattal, “Image upsampling via imposed edge statistics,” ACM Trans. Graph., vol. 26, no. 3, pp. 1-8, Jul. 2007, Art. ID 95.

[5] H. He and W. C. Siu, "Single image super-resolution using Gaussian process regression," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Jun 2011, pp. 449-456.
[6] K. Zang, X. Gao, D. Tao, and X. Li, "Single image super-resolution with non-local means and steering kernel regression," IEEE Trans. Image processing., vol. 21, no. 11, pp. 4544-4556, Nov. 2012.
[7] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based Super-resolution," IEEE. Comput. Graph. Appl., vol. 22, no. 2, pp. 56-65, Mar./Apr. 2002.
[8] W. Fan and D. Y. Yeung, "Image hallucination using neighbour embedding over visual primitive manifolds," in Proc. IEEE Conf. Comput. Vis. Pattern Recognition, Jul. 2007, pp. 1-7
[9] H. Chang, D. Y. Yeung, and Y. Xiong, "Super-resolution through neighbour embedding," in Proc. IEEE Conf. Comput. Vis. Pattern Recognition, vol. 1, Jul. 2004, pp. 275-282.
[10] T. M. Chang, J. Zhang, J. Pu, and H. Huang, "Neighbor embedding based super-resolution algorithm through edge detection and feature selection," Pattern Recognit. Lett., vol. 30, no. 5, pp. 494-502, Apr. 2009
[11] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single-image," in Proc. IEEE 12th Int. Conf. Comput. Vis. Oct. 2009, pp. 349-356
[12] J. Yang, Z. Lin, and S. Cohen, "Fast image super-resolution based on in-place regression," in Proc. IEEE Conf. Comput. Vis. Pattern Recognition, vol. 1, Jun. 2013, pp. 1059-1066.
[13] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," ACM Trans. Graph., vol. 28, no. 3, pp. 1-10, 2010.
[14] K. Zang, X. Gao, D. Tao, and X. Li, "Single-image super-resolution with multiscale similarity learning," IEEE Trans. Neural Netw. Learn. Syst, vol. 24, no. 10, pp. 1648-1659, Oct. 2013.
[15] C. Y. Yang, J. B. Huang, and M. H. Yang, "Exploiting self-similarities for singleframe super-resolution," in Proc. 10th Asian Conf. Comput, Nov. 2010, pp. 497-510