Super Resolution on The Centroid Reference Grid

Hossein Rezayi

Engineering department of Ferdowsi University of Mashhad, Iran

hrezayi2001@yahoo.com

Seyed Alireza Seyedin

Engineering department of Ferdowsi University of Mashhad, Iran

Seyedin@um.ac.ir

Abstract— **In the multiframe image super-resolution (SR) methods a sequence of low-resolution (LR) images is used to create a high-resolution (HR) image. In all aforegoing methods, all of the LR images were registered with respect to one of existing images (called reference image) and their samples were mapped on an up-scaled grid of this reference image (called reference grid (RG)). However, the use of this RG may cause the samples appears with a highly non-uniform distribution. Additionally, final precision and quality may be dependent on the selection of reference image. In this paper a new RG which reduces the non-uniformity of sample distribution is proposed. This RG is produced by incorporating all of the images; therefore the final result is not dependent on the selection of reference image. We also propose a method to find this RG (will be referred as centroid RG). Moreover, this RG is unique.**

Keywords—Super resolution; Reference grid; Image interpolation; Homography; Sample density function

I. Introduction

Multiframe image super resolution (SR) is the process of fusing several low-resolution (LR) images to create a high-resolution (HR) image with better quality. Since 1984 when Tsai and Huang [1] proposed the first SR method, several SR techniques have been emerged. In a large amount of literature, these techniques have been reviewed and compared (e.g. [2]-[5]). Almost all of the aforegoing SR methods comprise three processes: image registration, image interpolation and image restoration (Fig.1) [2]. In the image registration phase, all of the LR images are registered with respect to one of them (called reference image) and their samples are mapped on an up-scaled grid of this reference image (called reference grid (RG)) (Fig.1). However, such selection of the RG may cause the samples appears with a highly non-uniform distribution especially when the motion model is homography (Fig.2). In the image interpolation phase the pixel values of RG are calculated using the existing samples have been mapped on the RG. Non-uniform distribution of the samples should be considered in the interpolation phase. The best non-uniform interpolation methods consider an adaptive coverage of neighboring samples so that the number of samples incorporate in the interpolation of each pixel is approximately the same [6]-[8]. But these methods have some disadvantages such as space-variant precision and quality, high computational cost and non-optimal usage of samples in some regions of images. Additionally, if a motion model which is more complex than the translational and Euclidean is used, the increasing factor of resolution has not an obvious interpretation, because different selections of reference image

may result in various final HR image with dissimilar details. In this paper to reduce the non-uniformity of the sample distribution on the RG, we propose to use a different RG based on the homographies between all existing LR images. All the LR image samples are mapped on this RG and then interpolated. Since the average motion of all LR image samples during mapping on this RG is less than the other RGs, the non-uniformity of samples on this RG is less than the other RGs. The reminder of this paper is organized as follows. In Section II problem formulation for SR is discussed. Section III is devoted to the development of centroid RG. Implementation details will be brought in Section IV and finally experiments and conclusion is appeared in Sections V and VI, respectively.

II. Problem formulation for SR

A sequence of K discrete LR images g_k ($M_g \times N_g$ pixels) in which $1 \le k \le K$, is considered. The column vectors $\mathbf{g}_k$ are lexicographically ordered LR images and all are stacked in one column vector, i.e. $\mathbf{g} = [\mathbf{g}_1^T, ..., \mathbf{g}_K^T]^T$ [9]. The aim of the SR process is to reconstruct the original HR image f of size $M_f \times N_f$ pixels using existing LR images as well as some known information about the original HR image such high frequency limitation [9]. The column vector $\mathbf{f}$ is the lexicographically ordered original HR image. The decimation factor (increasing factor of resolution) is assumed to be the same in vertical and horizontal directions (i.e. $\rho = M_f / M_g = N_f / N_g$). By applying the three principal operations to the original HR image $\mathbf{f}$ the LR image $\mathbf{g}_k$ is obtained as follows [9]:

$$\mathbf{g}_k = \mathbf{DB}_k \mathbf{S}(\boldsymbol{\alpha}_k)\mathbf{f} + \mathbf{n}_k \qquad (1)$$

where $\mathbf{n}_k$ is additive white Gaussian noise (AWGN), $\mathbf{D}$ is the down-sampling operator (appears practically as a $M_g N_g \times M_f N_f$ matrix), $\mathbf{B}_k$ is the blurring operator (appears practically as a $M_f N_f \times M_f N_f$ matrix) and $\mathbf{S}(\boldsymbol{\alpha}_k)$ is the warping operator (appears practically as a $M_f N_f \times M_f N_f$ matrix) [9]. $\boldsymbol{\alpha}_k$ is the motion vector for warping the grid of original HR image $\mathbf{f}$ (RG), $\overline{\mathbf{X}}_{ref}$, onto the grid of kth LR image, $\mathbf{X}_k$. The matrices $\overline{\mathbf{X}}_{ref}$ and $\mathbf{X}_k$ are homogenous representation of grid points [10]. In the current SR methods,
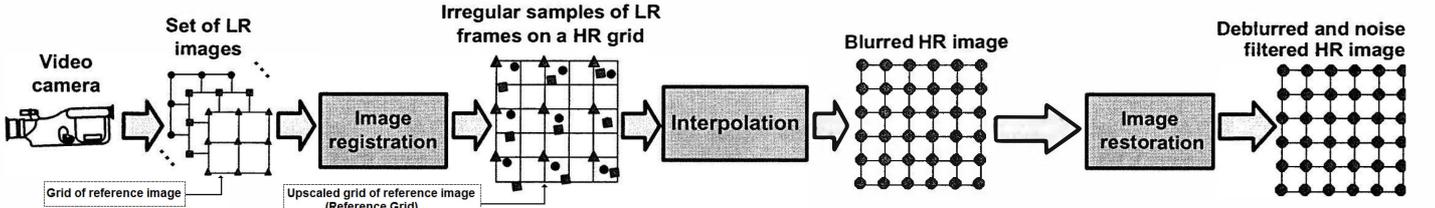
Fig.1 Schematic presentation of usually common three phases of the SR technique, partially adopted from [11].

one of the existing LR images is considered as the reference image and its up-scaled grid is taken as the RG.

The combination of the three mentioned principal operations can be considered as a unit combinational operation which is realized by a matrix $\mathbf{W}_k$ as follows [9]:

$$\mathbf{g}_k = \mathbf{W}_k \mathbf{f} + \mathbf{n}_k \qquad (2)$$

The homography motion model has eight degrees of freedom. Hence, $\boldsymbol{\alpha}_k = [h_{1k},...,h_{8k}]$ where $h_{1k},...,h_{8k}$ are the main elements of the 3x3 homogenous matrix [10] as follows:

$$\mathbf{Q}_k = \begin{bmatrix} h_{1k} & h_{4k} & h_{7k} \\ h_{2k} & h_{5k} & h_{8k} \\ h_{3k} & h_{6k} & 1 \end{bmatrix} \qquad (3)$$

Therefore, we can write

$$\mathbf{X}_k = \mathbf{Q}_k \overline{\mathbf{X}}_{ref} \qquad (4)$$

Equation (2) presents the relations between the original HR image and LR images. These relations can be merged into one equation as:

$$\mathbf{g} = \mathbf{W}\mathbf{f} + \mathbf{n} \qquad (5)$$

where $\mathbf{W} = [\mathbf{W}^1(\boldsymbol{\alpha}_1)^T,...,\mathbf{W}^K(\boldsymbol{\alpha}_K)^T]^T$ and we call it combinational coefficient matrix and $\mathbf{n} = [\mathbf{n}_1^T,...,\mathbf{n}_K^T]^T$.

III. CENTROID REFERENCE GRID

As mentioned in the Section I, if a motion model which is more complex than the translational and Euclidean is used, the non-uniformity of samples on a RG which is associated to one of existing LR images is considerable. For this reason we propose a new RG constructed by contributing the grids of all LR images and is not the up-scaled form of one of existing LR grids. In the animation field, projective interpolation is an issue concerns to transition states between two or more image grids especially when their relative motions is homography [12]. There is several methods for projective interpolation include linear interpolation of four vertices of quadrilateral grids, direct interpolation of homography matrices $((1-\alpha)\mathbf{I}+\alpha\mathbf{Q}$ where $\mathbf{I}$ is identity matrix), dividing the quadrilateral grid into two triangle grids and then affinely interpolate their triangle grids correspondingly and finally (the best one) decomposing the homography matrices into their pure component matrices [10], i.e. translation, rotation, scaling, shearing and pure homography, and then linear interpolating the pure component matrices [13]. Some of these methods are not acceptable for our purpose (because, for example, they transform the quadrilateral grid into approximately a segment line during their transition) and some others are suitable only for couple of images. Hence, the RGs provided by these methods during their transition states are not as advantageous as our proposed
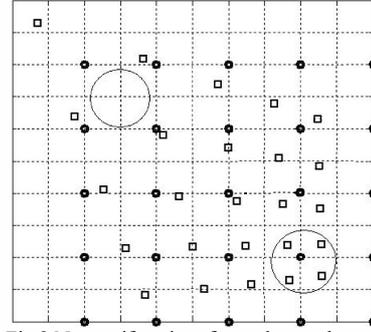
Fig.2 Non-uniformity of samples on the RG when the motion model is homography

one. Since the new RG is similar to the centroid of a polygon in which the vertices are corresponding to existing LR image grids, we call it centroid RG (CRG). Fig.3 represents how this RG is derived for 2, 3 and 4 image grids. This can be easily generalized for the larger number of
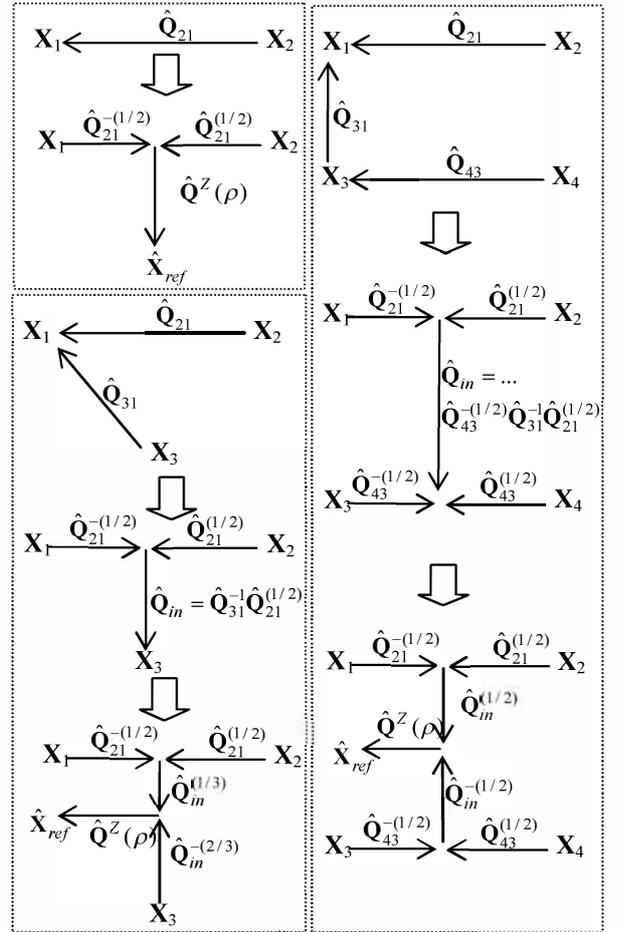
Fig.3 Calculating the CRG for 2, 3 and 4 LR image grids.

image grids. In this figure, $\hat{\mathbf{Q}}_{jk}$ is the homogenous matrix for transforming the jth LR grid to the kth LR grid and $\mathbf{Q}^Z(\rho)=[\rho\ 0$
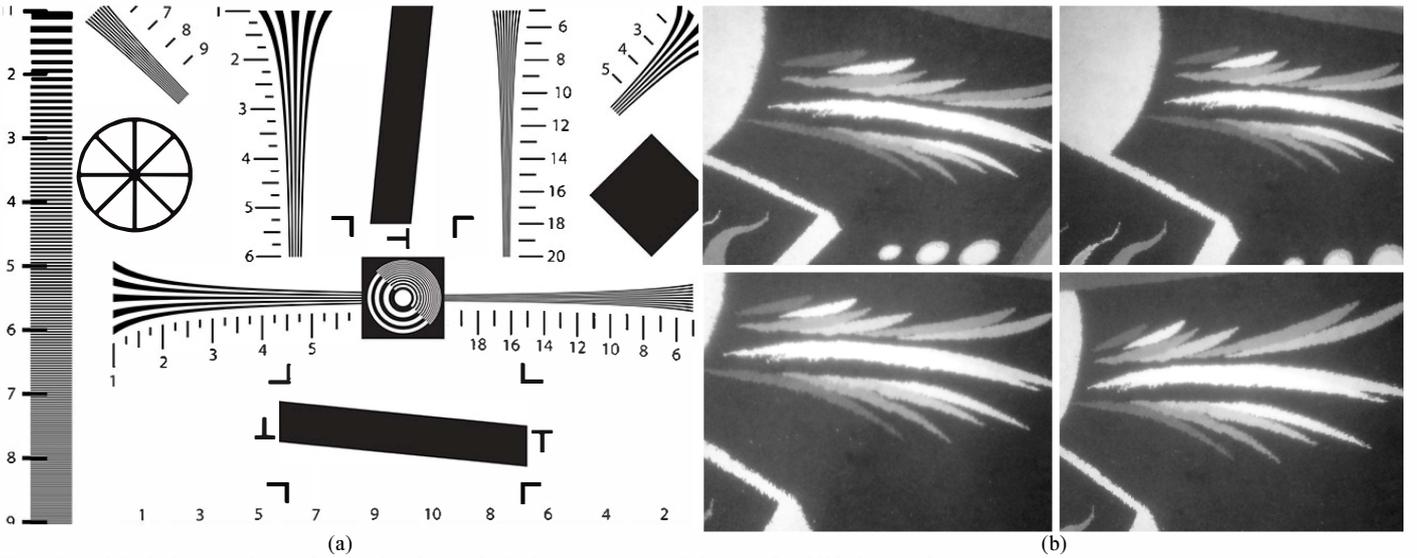
Fig.4 The original 'chart' test image for creating the synthetic dataset (a).Four LR image of real life dataset 'Carpet'(b)

0; 0 ρ 0; 0 0 1] is pure zooming homogenous matrix transforms a LR grid to its corresponding HR one.

It is not surprising to say that the proposed RG is unique because the centroid of a polygon is unique too. Nevertheless, it can be proved readily mathematically for simple motion models (like pure translation, rotation, scaling and homography), and practically for complex motion models (like general homography) that the proposed RG is unique. But none of RG created by other methods is unique. These RGs are completely path dependent.

A. Sample Density Function

The amount of non-uniformity of samples on the different RGs can be presented by the sample density function (SDF) on RG. If the homography matrix which maps the RG on k^{th} LR image grid is $\overline{\mathbf{Q}}_k$ with elements $\overline{h}_{1k},...,\overline{h}_{8k}$, it can be readily proved that SDF will be as follows:

$$SDF(x,y) = \sum_{k=1}^{K} \frac{\det(\overline{\mathbf{Q}}_k)}{(\overline{h}_{3k}x + \overline{h}_{6k}y + 1)^3} \qquad (6)$$

This function is positive in the feasible region of RG plane.

IV. IMPLEMENTATION DETAILS

It should be noted that the use CRG is not an alternative for the adaptive interpolation methods used in SR techniques. Rather it can be even used along with these methods to further improvement of the final performance of the SR technique. Although by applying the CRG, it can be unnecessary to use the heavy techniques such as adaptive normalized convolution [7],[8]. We will implement two SR methods that can handle homography motion model. The first method is the Delaunay Triangulation interpolation based method [11]. The second method is a maximum a posterior (MAP) estimator based SR method used by several authors (e.g. [9], [14] and [15]). In the MAP estimator method we used bilateral total variation (BTV) [9] as a regularizer. Since the motion model is homography, we applied the same method used in [14] for calculating the combinational coefficient matrix, $\mathbf{W}$. Another interesting

advantage of using CRG is its impact on the sparsity of the combinational coefficient matrix $\mathbf{W}$. The blur kernels have smaller coverage on the CRG rather than RG which are up-scaled versions of their corresponding LR image grids.

V. EXPERIMENTS

In this section, the CRG is compared to the other RGs which are the up-scaled versions of the existing LR image grids. Then the impact of selecting the CRG and other RGs on the performance of two mentioned SR techniques is discussed. Due to limited space of paper, only two datasets including a synthetic dataset and a real-life dataset will be used for this aim. SDF is used as a measure for comparison RGs. The peak signal to noise ratio (PSNR) of the reconstructed HR images as a measure for presenting the impact of selecting these RGs on the SR techniques is defined as follows [9]:

$$PSNR(\widetilde{\mathbf{f}}) = 10\log_{10}(M_f N_f / \| \mathbf{f} - \widetilde{\mathbf{f}} \|^2) \qquad (7)$$

where $\widetilde{\mathbf{f}}$ and $\mathbf{f}$ are reconstructed and real HR images, respectively.

A. Experiment on Synthetic Dataset

This synthetic dataset is a sequence of four images which are created using the 654x490 pixel 'chart' test image (Fig. 4(a)). The LR images are produced using (2) the same as method has been used in [14]. Two number of the LR images are shown in Fig.5. Four homography matrices have been created such that the CRG of created LR images grids is the grid of current test image. The standard deviation of Gaussian blur is set to 0.2 LR pixel. Other required parameters are set as [9]. The SDF of CRG and one of other RGs are displayed in Fig. 6. It is observed that SDF of CRG has smaller variations. The reconstructed HR image using the two SR techniques with CRG and one of other RGs are shown in Fig. 7. The PSNR of reconstructed HR images are expressed in Table 1.

B. Experiment on Real Life Dataset

A sequence of four 654x490 pixel images from a Carpet (Fig.4 (b)) is used for this experiment. The reconstructed HR
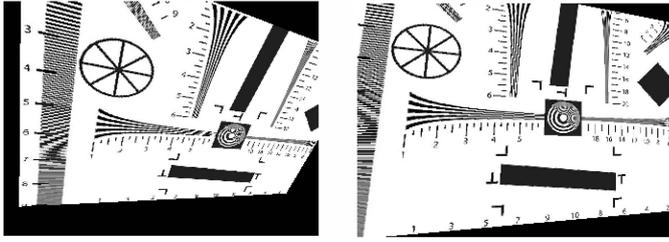
Fig.5 Two degraded LR images of 'Chart' test image

(a)   (b)
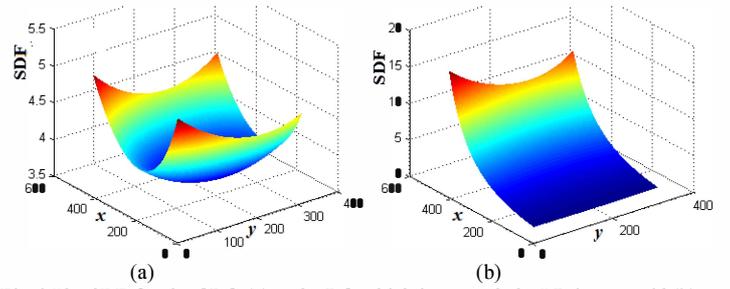
Fig.6 The SDF for the CRG (a) and a RG which is up-scaled a LR image grid (b)

(a)   (b)   (c)   (d)
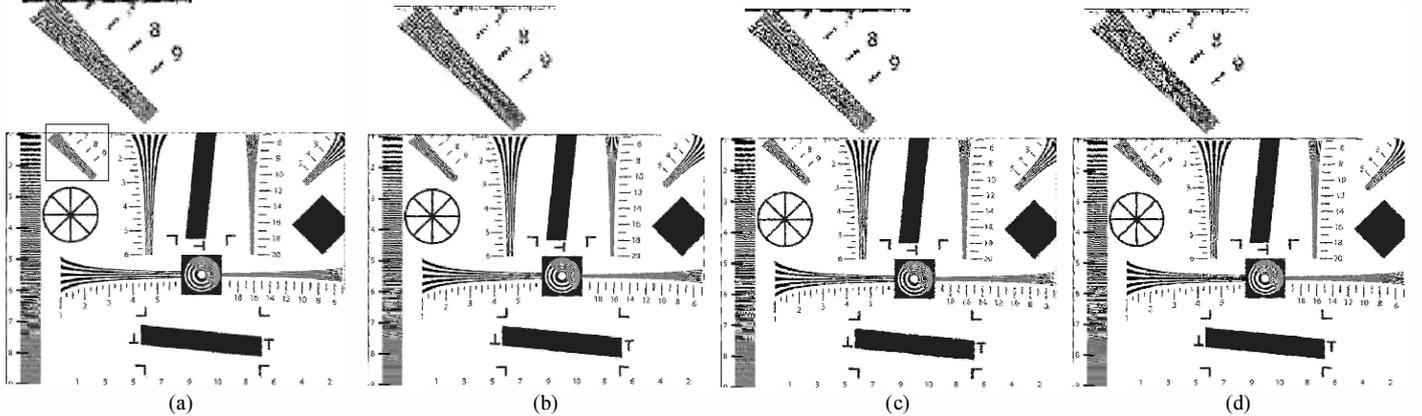
Fig.7 The reconstructed HR image using the interpolation based method (a,b) and MAP estimator based method (c,d) with CRG (a,c) and without CRG (b, d).

(a)   (b)   (c)   (d)
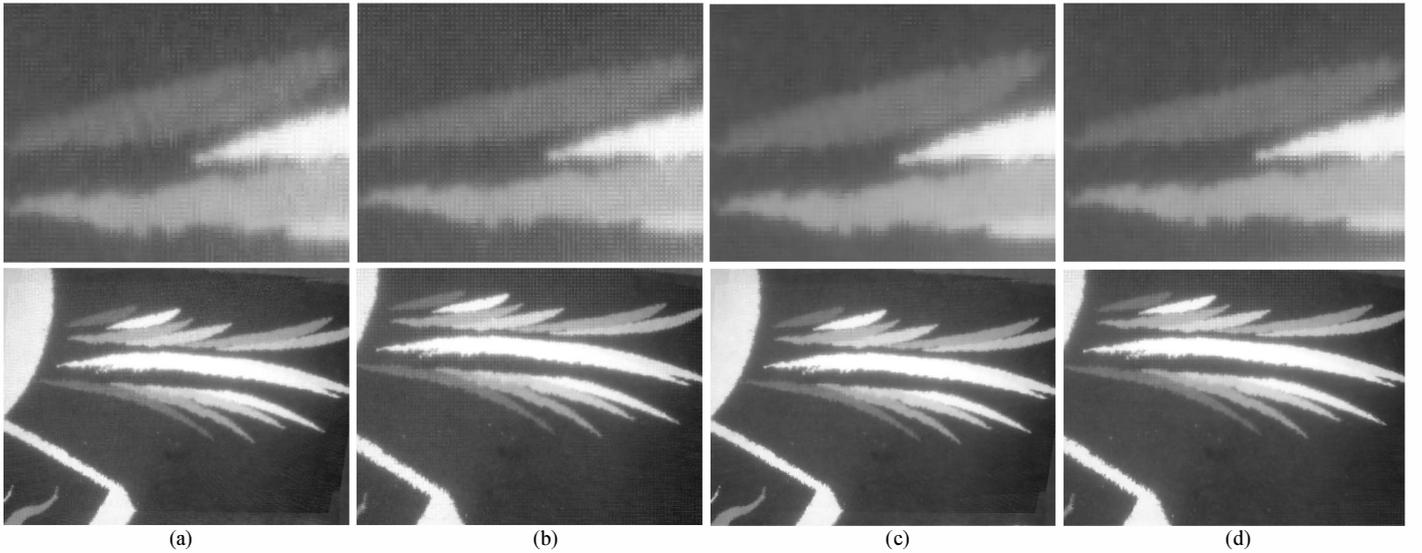
Fig.8 The reconstructed HR image using the interpolation based method (a,b) and MAP estimator based method (c,d) with CRG (a,c) and without CRG (b, d).

images using two SR techniques with and without CRG (the same as previous experiment) are indicated in Fig.8. The highlighted parts of these reconstructed HR images show the positive impact of selecting the CRG.

It is worth noting that inserting the CRG to SR techniques

has not additional complexity because its calculated is limited to few 3x3 matrix multiplications. It may be interesting to know that the number of non-zero elements for sparse matrix $\mathbf{W}$ is approximately 1.29e6 and 1.43e6 with and without CRG, respectively, for MAP estimator based SR technique.

VI. CONCLUSION

In this paper we proposed a new framework for SR problem especially when the motion model is more complex than Euclidean. Although this framework can be helpful for other motion model to contribute the maximum overlap regions in the SR technique. By Inserting the CRG into SR, it can be

TABLE I.   PSNR FOR TWO SR METHODS

	With CRG	Without CRG
Interpolation based method SR	*13.79*	*13.19*
MAP estimator based method SR	12	11.2

unnecessary to use the heavy techniques such as adaptive normalized convolution. CRG can affect the selection of increasing factor, indirectly. Since, when the CRG is not used in SR, a large amount of samples may be mapped on a very small region of RG and to avoid the shrinking of these samples, a larger value for increasing factor should be selected. This leads to an increase in computational cost and required memory.

References

[1] R. Tsai and T. Huang, "Multiframe image restoration and registration," Advances in Computer Vision and Image Processing, vol. 1, pp. 317–339, 1984.

[2] S. Park, M. Park and M. Kang. "Super-Resolution Image Reconstruction: A Technical Overview," IEEE Signal Processing Magazine, vol. 20, no. 3, pp.21– 36, 2003.

[3] S. Borman and R.L. Stevenson. "Spatial resolution enhancement of low-resolution image sequences. A comprehensive review with directions for future research," Laboratory for Image and Signal Analysis (LISA), University of Notre Dame, Notre Dame, Ind, USA, Tech. Rep., 1998.

[4] K. Katsaggelos, R. Molina and J. Mateos. Super Resolution of Images and Video, Morgan and Claypool Publishers, San Rafael, 2007.

[5] P. Milanfar. Super-Resolution Imaging, CRC Press, 2011.

[6] F. Xu, H. Wang, L. Xu and C. Huang. "A new framework of normalized convolution for superresolution using robust certainty," IEEE International Conference on Computer and Automation Engineering, vol. 2, pp. 144-148, 2010.

[7] T.Q. Pham, L.J.V. Vliet and K. Schutte. "Robust Fusion of Irregularly Sampled Data Using Adaptive Normalized Convolution," EURASIP Journal on Applied Signal Processing, vol. 2006, pp. 1-12, 2006.

[8] K. Zhang, G. Mu, Y. Yuan, X. Gao and D. Tao, "Video super-resolution with 3D adaptive normalized convolution," Elsevier Journal of Neurocomputing, vol. 94, pp. 140-151, 2012.

[9] H. Rezayi and S.A.R Seyedin, "Joint Image Registration and Super-Resolution Based on Combinational Coefficient Matrix," 2nd International Conference on Pattern Recognition and Image Analysis, Rasht, Iran, 2015.

[10] R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision, 2nd ed., Ed. Cambridge, United Kingdom: Cambridge University Press, 2003.

[11] S. Lertrattanapanich and N. Bose. "High Resolution Image Formation from Low Resolution Frames Using Delaunay Triangulation," IEEE Transaction on Image Processing, vol. 11, no. 12, pp. 1427–1441, 2002.

[12] T.Y. Wong, P. Kovesi and A. Datta. "Projective transformations for image transition animations," IEEE Conference on Image Analysis and Processing, pp. 493-500, 2007.

[13] K. Shoemake and T. Duff. "Matrix animation and polar decomposition," ACM International Conference on Graphics Interface, pp. 258–264, 1992.

[14] L. C. Pickup, D. P. Capel, S. J. Roberts, and A. Zisserman. "Bayesian Methods for Image Super-Resolution," Computer Journal, 2007.

[15] .Capel, D. Image Mosaicing and Super-resolution (Distinguished Dissertations). Springer, ISBN: 1852337710, 2004.