

Image Registration for Super-Resolution using SIFT Key-points

Mahmood Amintoosi Mahmood Fathy
Nasser Mozayani

Computer Engineering Department, Iran University of Science and Technology
Narmak, Tehran, Iran

mamintoosi@ieec.org, {mahFathy, Mozayani}@iust.ac.ir

Abstract: Super-Resolution algorithms reconstruct a high resolution image from a set of low resolution images of a scene. An accurate image registration is a fundamental stage in all Super-Resolution methods. In this paper a new restriction criteria for matching SIFT key points, for registration under the condition of Super-Resolution problems is proposed. This is based on the assumption of normal distribution for location and orientation differences of correct matches under global translational motion model. The accuracy of the proposed method is compared with some famous registration methods. A quantitative comparison has been done with a collective of images with known motion vectors. Also a subjective comparison has been done with the results of various registration methods including the proposed method for Super-Resolution. The experimental results show the better performance of the proposed algorithm in compare with some other registration methods.

Keywords: Image Registration, Super-Resolution, SIFT Key-points

1. Introduction

One of the most critical aspects of the many applications in image processing and computer vision is the accurate estimation of motion, also known as image registration. Image registration also plays a central role in the context of super-resolution. The Super-Resolution (SR) techniques fuse a sequence of low-resolution images to produce a higher resolution image. The low resolution (LR) images may be noisy, blurred and have some displacement with each other. These methods utilize information from multiple observed images to achieve restoration at resolutions higher than that of the original data. Super-Resolution restoration methods register the observed images to a common reference frame in order to reconstruct the high resolution (HR) image. The image registration process thus requires knowledge of the visual motion occurring in the observed image sequence. Since this typi-

cally is not known, the motion information must be estimated from the observed image sequence in order to effect the restoration process [1]. It is widely recognized [2] that the accuracy of the motion estimates is arguably the limiting factor in Super-Resolution restoration performance, so any fruitful consideration of this problem promises significant returns[1].

In SR literatures a variety of registration approaches has been used. Regarding the transformation model between images (such as translation, affine or projective) the registration method may be different. But they can be classified into two main approaches: feature-based methods and area-based methods. While the former requires only a sparse set of point correspondences to fit the motion model, the latter uses the information from all pixels [3].

The registration step of Irani and Peleg [4] is based on the pioneering works of Keren *et al.*[5] and Lucas and Kanade [6]. This is an area-based method which produces good results under global translational motion. Their approach is the use of a Taylor series approximation of the images. The motion parameters are the unknowns in the approximation, and they can be computed from the set of equations that can be derived from this approximation. Because Taylor series only give a good approximation for small offsets, these registration methods are generally applied iteratively using a Gaussian pyramid. Vandewalle *et al.*[7] used a frequency-based registration method, which at first, the rotation parameters are estimated from a radial projection of the absolute values of the Fourier transform image. A simple one-dimensional correlation can be performed to compute the rotation angle from the projections for two images. Then Shifts estimated from the linear phase difference between the rotation corrected images. This method performs well if the images have some directionality[7]. Another fast image registration which uses in image stabilization context is Gray Coded Bit Plane Matching (GC-BPM)[8]. This method is very computationally efficient since it uses binary boolean operations, but its performance is lower than area-based methods such as [5].

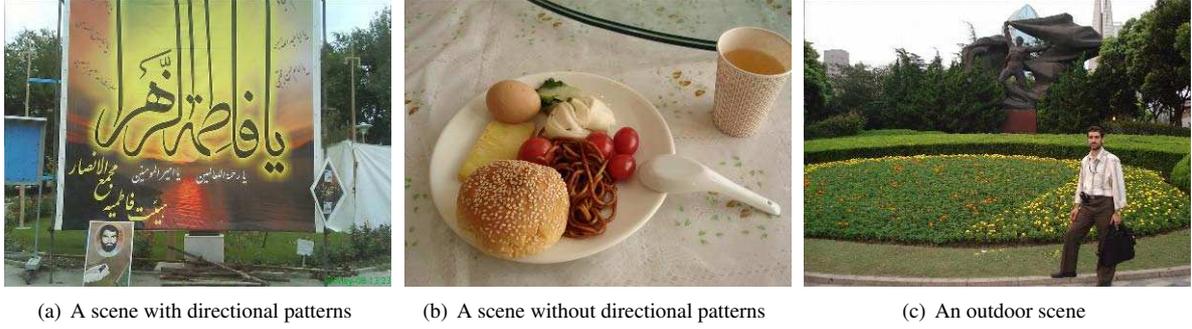


Figure 1. Images with different patterns used in this paper.

After introducing Scale Invariant Feature Transform (SIFT) by D. Lowe [9, 10], various application of it, including matching and registration reported by researchers. Mikolajzyk and Schmid [11] compare the performance of descriptors computed for local interest regions and the results show the SIFT-based descriptors are the best. Amintoosi *et al.*[12] used these features for registration of images under projective model in a SR problem. Yi *et al.* used SIFT key-points for multi-spectral remote images[13]. They proposed a scale restriction criteria for SIFT match to reduce the incorrect matches.

In many Super-Resolution approaches such as [14, 15, 4, 16], global translational motion is assumed, which the low resolution input images have a bit differences with each other about rotation and vertical and horizontal shifts. In this paper we propose a displacement restriction criteria for removing the incorrect matches with the assumption of Gaussian Probability Distribution Function (PDF) for the mentioned displacements. A HR image is achieved from these motion compensated LR images with a Super-Resolution reconstruction method.

Section 2 explains the proposed method. Section 3 provides experimental results and section 4 describes the concluding Remarks and future works.

2. The Proposed Method

Interest points (called key points in the SIFT framework) are identified as local maxima or minima of the difference-of-Gaussian filters across scales. To determine the key point orientation, a gradient orientation histogram is computed in the neighborhood of the key point. Peaks in the histogram correspond to dominant orientations. The key points are denoted by a vector $(x_i, y_i, \sigma_i, \theta_i)^T$, which denote location, scale and orientation. The usual method for finding a match for each key point is identifying its nearest neighbor. To ensure correct match, Lowe[9] suggests that the ratio of closest to second-closest neighbors must be less than a threshold [13]. In our proposed method, this is done as a preprocess stage for finding corresponding matches between two images, but the resulting matches may be contain incorrect matches as outliers, yet.

In many SR applications, global translational motion with small displacements is supposed. In registration step of SR methods, usually an specified image considered as

reference frame and other images, are aligned with respect to this reference frame. With regarding a Normal distribution for difference of locations of corresponding key points between two images, we can distinguish and remove outliers.

Suppose that $(x_i^1, y_i^1, \theta_i^1)$ and $(x_i^2, y_i^2, \theta_i^2)$ are the location and orientation of i^{th} key point in image 1 -as reference frame- and another image 2 (from total N matches found after preprocess stage). let $\Delta x_i = x_i^1 - x_i^2$, our empirical results showed that $\Delta x = \{\Delta x_1, \dots, \Delta x_N\}$ as a random variable has a Gaussian Probability Distribution Function (PDF) with mean μ_x and variance σ_x^2 . Figure 3 shows the histogram and estimated Gaussian PDF for Δx and $\Delta \theta$ between figures 2(a) and 2(d). Δy as Δx also has a similar bell shape. But we need a better method for checking the similarity of our data to the normal distribution. The normal probability plot [17] is a graphical technique for assessing whether or not a data set is approximately normally distributed. The data are plotted against a theoretical normal distribution in such a way that the points should form an approximate straight line. Departures from this straight line indicate departures from normality. The normal probability plot for Δx and $\Delta \theta$ has been shown in figure 4. As it can be seen from figures 3(a) and 4(a), normal distribution is a good model for Δx ; but as figures 3(b) and 4(b) show, $\Delta \theta$ does not approximated well with normal distribution. At least at this stage of our research we assume a normal distribution for it, with accepting a bit error for estimation of $\Delta \theta$. Investigation of figure 3(b) shows that we have a bit shift in approximating μ_θ with a Normal PDF, against the mean of the plotted histogram.

Hence we can remove those points which are far from the mean more than $2.5 * \sigma_x$ as outliers. Repeating this process for Δy and $\Delta \theta$ will remove some other outliers. After removing outliers, with recalculating the means of Δx , Δy and $\Delta \theta$ the registration parameters will be in hand. It should be mentioned that the difference between key points orientations (θ_i^1, θ_i^2) has been taken by computing dot products of vectors $(\sin \theta_i^1, \cos \theta_i^1)$ and $(\sin \theta_i^2, \cos \theta_i^2)$.

In the next section we will see the result of the mentioned outlier removal for image registration and Super-Resolution.



(a) LR1-A downsampled version of 1(a) without any displacement (b) LR2- A downsampled version of 1(a) with -2.5 pixel horizontal shift (c) LR3- A downsampled version of 1(a) with -4 pixel vertical shift (d) LR4- A downsampled version of 1(a) with -2.5,-4 pixel shift along X,Y axis, respectively, plus rotation with 3 degree

Figure 2. Produced low resolution images from figure 1(a).

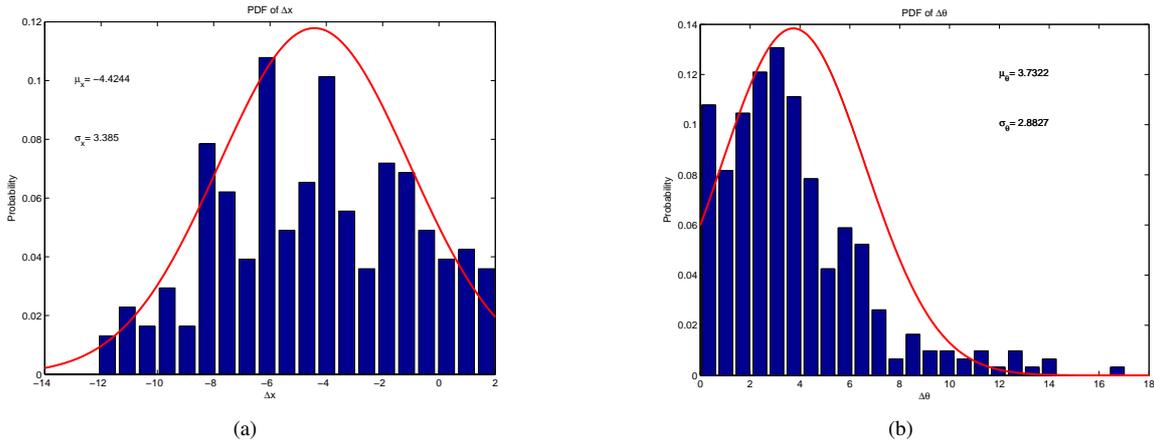


Figure 3. The histogram and estimated Gaussian PDF for Δx and $\Delta \theta$ between 2(a) and 2(d).

3. Experimental Results

We applied our proposed method on a set of test images shown in figure 1 and compared the proposed registration performance with some other methods. Because for having objective comparison we need some images with known displacement, our implementations was done under controlled conditions. In the experiments we create 4 LR images from each of 3 images shown in figure 1. These images have some different patterns. Figure 1(a) is an image captured from our university campus which has many directional patterns. Figure 1(b) is an image which does not have directional patterns. Figure 1(c) is an outdoor scene having different patterns. For every high resolution image, we considered 4 set of known motion parameters. For every set, we first shifted the HR image by these known values among the X and/or Y axis, rotated it by the specified angle and then resized it by a factor of .5. Hence when the shift value be an odd number, the resulting LR image would have a motion vector with non integer value. The produced LR images from image 1(a) is shown in figure 2. It should be mentioned that the motion parameters for first LR image as our reference frame is set to zero, hence it is only a down sampled version of HR image. Every HR image with its corresponding LR images are one data set. This approach with different motion vectors in vertical and horizontal directions (in the range of [0,9] pixels) and

different rotation angles (in the range of [0,5] degree) was used to produce 4 LR images from each image shown in figure1.

We applied 3 registration methods: frequency method [7], GC-BPM[8] and Keren *et al.* method [4, 5] for comparing with our proposed registration approach. Because we have the ground truth data, the objective comparison is possible. Figure 5(a) shows comparison of the MSE between motion parameters (x, y, θ) estimated by various mentioned registration methods and the known parameters. We had 3 HR images in figure1 and from each image we produced 4 LR images. The estimated motion parameters are about to LR images 2,3 and 4 with respect to the first LR image in each data set. Hence the x-axis has 9 points in figure5(a). As it can be seen the proposed method and Keren method produced better results against the others in average. Also as expected the frequency method, was not good in the second data set (figure1(b)), which does not have directional patterns.

We used the estimated parameters with every registration methods of each data set for producing a high resolution image with a magnifying factor of 2. Among the SR reconstruction methods the interpolation approach is used here. Quantitative comparison has been done with the produced HR image and the original HR image (figure5(b)). The produced HR images are shown in figure 6 for subjective comparison. Note the relative improvement in the

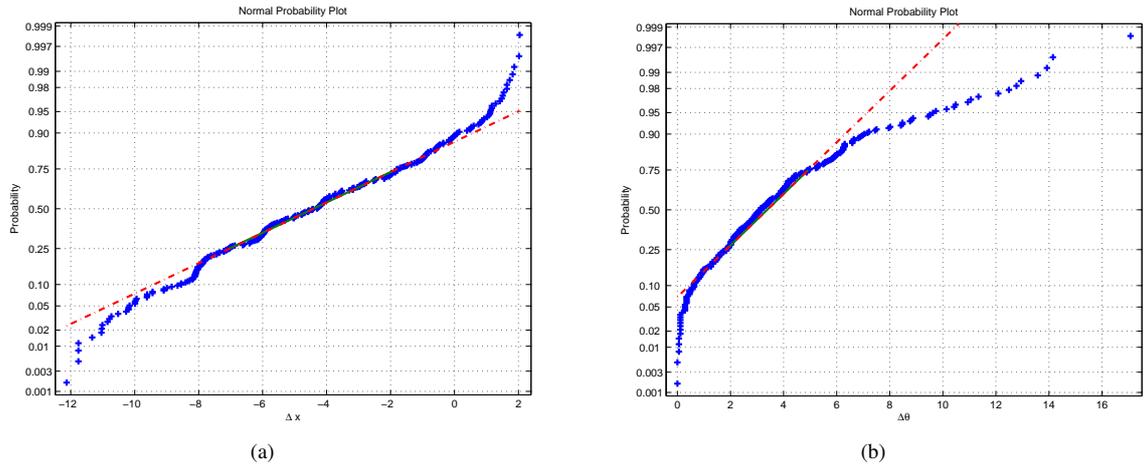
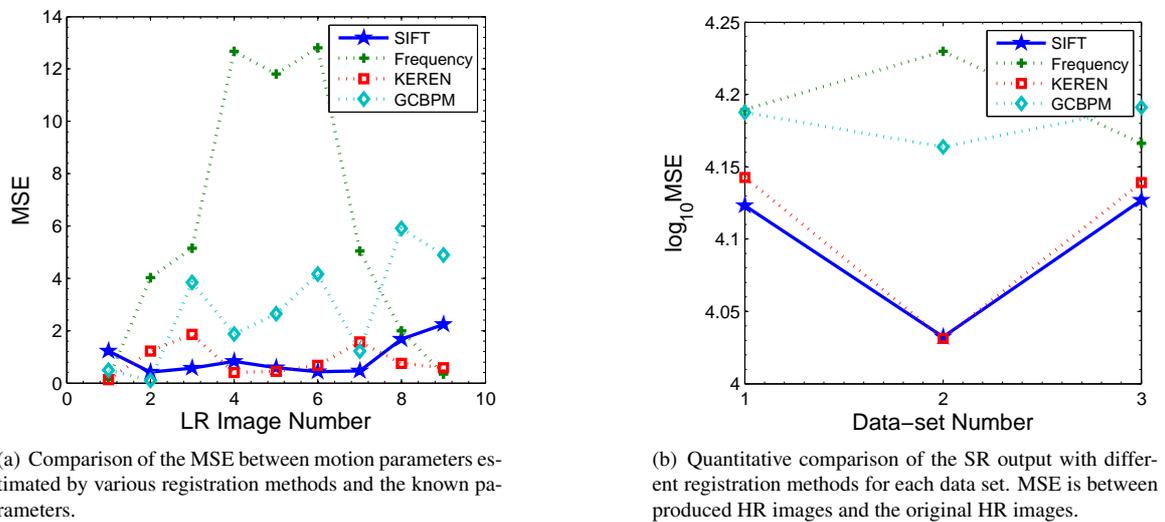


Figure 4. The normal probability plot for Δx and $\Delta\theta$ between 2(a) and 2(d).



(a) Comparison of the MSE between motion parameters estimated by various registration methods and the known parameters.

(b) Quantitative comparison of the SR output with different registration methods for each data set. MSE is between produced HR images and the original HR images.

Figure 5. Quantitative Comparison

quality of the proposed method (figure 6(h)) as a result of the better image registration. Inspection of the results indicates that other approaches -except that Keren method (figure 6(g))- suffer from bad registration. For better visual comparison a small region containing an Arabic text has been enlarged.

4. Conclusion and Future Works

Precise sub-pixel image registration is a basic requirement for a good reconstruction in Super-Resolution methods. If the images are inaccurately registered, the high resolution image is reconstructed from incorrect data and is not a good representation of the original signal. Registration using SIFT key points are very powerful and promising approach. In this paper a new restriction criteria for removing outliers from matched key points of Lowe[9] was proposed. In summary the innovations of this paper are as follows:

- Justification that displacements between corresponding key points, under global translational motion

model, is approximately normally distributed (by normal probability plot[17]).

- Outlier removal based on the assumption of normal PDF of displacement and removing those points which are far from the mean as enough,
- Using SIFT key points orientation directly for image registration, in contrast to other methods such as RANSAC¹ [18] which produce a fitting model,
- Using SIFT key points with this manner for image registration in the Super-Resolution context.

As mentioned earlier $\Delta\theta$ does not approximated well with a normal distribution. As future works we plan to (i) find a better distribution function for $\Delta\theta$ and (ii) comparing the result with another outlier removal approach such as RANSAC.

¹RANdom SAmples Consensus (RANSAC)



(a) Resized of LR image 2(a)



(b) A magnified region of 6(a)



(c) Frequency method



(d) GC-BPM



(e) A magnified region of 6(c)



(f) A magnified region of 6(d)



(g) Keren



(h) SIFT (this paper)



(i) A magnified region of 6(g)



(j) A magnified region of 6(h)

Figure 6. Super-resolution results with different registration methods. The effect of not precise registration by frequency and GC-BPM methods is seen in (c),(d). For better visual comparison a small region of (a),(c),(d),(g),(h), containing an Arabic text has been enlarged and is shown in (b),(e),(f),(i),(j),respectively.

Acknowledgment

The authors wish to thank Prof. Milanfar and Dr. Farsiu [14] for providing us with their Super-Resolution package. We also thank to Dr. D. Lowe for his SIFT key-points program [10].

References

- [1] S. Borman, "Topics in multiframe superresolution restoration," Ph.D. dissertation, University of Notre Dame, Notre Dame, IN, May 2004.
- [2] R. Schultz and R. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. on Image Processing*, vol. 5, no. 6, pp. 996–1011, June 1996.
- [3] T. Q. Pham, "Spatiotonal adaptivity in super-resolution of under-sampled image sequences," Ph.D. dissertation, aan de Technische Universiteit Delft, 2006.
- [4] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graph. Models Image Process.*, vol. 53, no. 3, pp. 231–239, 1991.
- [5] D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using sub-pixel displacement," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 1988, pp. 742–746.
- [6] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *IJCAI81*, 1981, pp. 674–679.
- [7] P. Vandewalle, S. Süsstrunk, and M. Vetterli, "A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution," *EURASIP Journal on Applied Signal Processing (special issue on Super-resolution)*, vol. 2006, pp. Article ID 71459, 14 pages, 2006.
- [8] S. Ko, S. Lee, S. Jeon, and E. Kang, "Fast digital image stabilizer based on gray-coded bit-plane matching," *IEEE Transactions on Consumer Electronics*, vol. 45, no. 3, pp. 598–603, Aug 1999.
- [9] D. G. Lowe, "Object recognition from local scale-invariant features," in *ICCV '99: Proceedings of the International Conference on Computer Vision-Volume 2*. Washington, DC, USA: IEEE Computer Society, 1999, p. 1150.
- [10] —, "Distinctive image features from scale-invariant key-points," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, 2004, The implementation is available online at: <http://www.cs.ubc.ca/~lowe/keypoints/>.
- [11] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [12] M. Amintoosi, M. Fathy, and N. Mozayani, "Reconstruction+synthesis: A hybrid method for multi-frame super-resolution," in *(MVIP08) 2008 Iranian Conference on Machine Vision and Image Processing*, University of Tabriz, Tabriz, Iran, Nov. 4-7 2008. [Online]. Available: webpages.iust.ac.ir/mamintoosi/papers/Fustruction.pdf
- [13] Z. Yi, C. Zhiguo, and X. Yang, "Multi-spectral remote image registration based on SIFT," *Electronics Letters*, vol. 44, no. 2, pp. 107–108, January 2008.
- [14] S. Farsiu, "A fast and robust framework for image fusion and enhancement," Ph.D. dissertation, Electrical Engineering, UC Santa Cruz, December 2005. [Online]. Available: <http://www.ee.ucsc.edu/milanfar>
- [15] M. Fathy, N. Mozayani, and M. Amintoosi, "Outlier removal for super-resolution problem using QR-Decomposition," in *IPC08: Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition*, USA, July 14-17 2008, pp. 271–277. [Online]. Available: <http://webpages.iust.ac.ir/mamintoosi/papers/IPC08.pdf>
- [16] P. Vandewalle, "Super-resolution from unregistered aliased images," Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, 2006.
- [17] J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey, *Graphical Methods for Data Analysis*. New York: Chapman and Hall, 1983.
- [18] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.