Outlier Removal for Super-Resolution Problem Using QR-Decomposition

Mahmood Fathy **Naser Mozayani**
Mahmood Amintoosi
Computer Engineering Department,Iran University of Science and Technology
Narmak, Tehran, Iran

Abstract - *Super-resolution algorithms reconstruct a high resolution image from a set of low resolution images of a scene. Recent advances in Super-Resolution techniques show trends towards robustness against outliers. Even under normal lighting condition, images may have motion outliers due to independent moving objects. In this article a new method for dealing moving objects is proposed, which those frames containing moving objects are identified using QR-Decomposition, a known method in Linear Algebra. This detection procedure comes between the registration and the reconstruction steps of super-resolution techniques. The detected frames as outlier frames will be dropped and hence the reconstruction step only works on other frames. The simulation results show the better performance of the proposed algorithm in compare with some other super-resolution methods.*

Keywords: Super-Resolution, Matrix Decomposition, Segmentation, Background Detection

1. Introduction

The Super-Resolution (SR) techniques fuse a sequence of low-resolution images to produce a higher resolution image. The low resolution images may be noisy, blurred and have some displacement with each other. A common matrix notation which is used to formulate the super-resolution problem [3, 6] is as follows:

$$\underline{Y}_k = DHF_k\underline{X} + \underline{V}_k, \quad k = 1, \dots, N \tag{1}$$

where $[r^2M^2 \times r^2M^2]$ matrix F_k is the geometric motion operator between the high-resolution frame $\underline{X}$ (of size $[r^2M^2 \times 1]$) and the k^{th} low-resolution frame $\underline{Y}_k$ (of size $[M^2 \times 1]$) which are rearranged in lexicographic order and r is the resolution enhancement factor. The camera's point spread function (PSF) is modeled by the $[r^2M^2 \times r^2M^2]$ blur matrix H, and $[M^2 \times r^2M^2]$ matrix D represents the decimation operator. $[M^2 \times 1]$ vector V is the system noise and N is the number of available low-resolution frames. We assumed that decimation operator D and blur matrix H is same for all images. As in [6] we consider $\underline{Z} = H\underline{X}$, so $\underline{Z}$ is the blurred version of the ideal high-resolution image $\underline{X}$ and the SR problem is broken in two separate steps:
1) Finding a blurred high-resolution image from the low-resolution measurements ($\hat{\underline{Z}}$).
2) Estimating the de-blurred image $\hat{\underline{X}}$ from $\hat{\underline{Z}}$.
Hence the SR problem can be formulated as follows:

$$\hat{\underline{Z}} = \underset{\underline{Z}}{\text{Arg Min}} \Big[\sum_{k=1}^{N} ||DF_k\underline{Z} - \underline{Y}_k||_p^p \Big] \tag{2}$$

Figure 1 shows a schematic diagram of the overall framework of Super-Resolution techniques, which is a bit modified version of figure 1.2 of [15].

Existence of outliers in input images is one of the major challenges in super-resolution domain. Farsiu *et al*. in [6] considered three common sources of outliers in super-resolution system: 1. Error in motion estimation. 2. Inconsistent pixels: effect of an object which is only present in a few low-resolution frames (e.g. the effects of a flying bird in a static scene). 3. Salt and Pepper noise. In this paper we focus on the second type of outliers (moving objects). In the context of Super-resolution, trends toward robustness against outliers have been grown recently [4, 5, 6, 7, 15, 16, 19].
Eren *et al*. [4] proposed a robust, object-based approach for SR problem using POCS framework. Their proposed method employs a validity map and a segmentation map. The validity map disables projections based on observations with inaccurate motion information for robust reconstruction in the presence of motion estimation errors. For their approach an accurate motion segmentation must be available, which is difficult to obtain in the presence of aliasing and noise. They assumed that objects of interest are marked on a

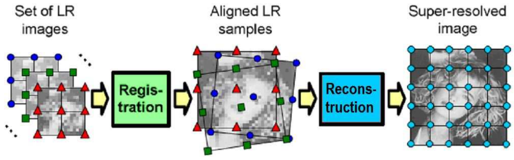

Figure 1. An overall view of Super-Resolution steps.

reference low resolution frame interactively. Moreover they have to generate a spatially varying threshold for their validity map structure. Zomet *et al*. [19] proposed a robust median estimator, which is combined in an iterative process to achieve a super resolution algorithm. Their method is interesting and is the basis or a part of some other methods [5, 6, 7, 15, 16]. But their method suffers from ghost artifacts, due to scene crowding, as was shown in figure 2(e and f) of [19]. Farsiu *et al*. [6] proposed a Shift-and-Add (S&A) method which is a simple fusion approach, but outlier effects are apparent in its output. Other variation and combination of Shift and Add method such as L1 + Bilateral TV which is also proposed by Farsiu [7] are good and robust techniques, but having many parameters, makes their usage not so easy. The normalized convolution method presented by Pham [15, 16] produces very good results, but it is a very high time consuming method. In the proposed method we detect and reject, those frames containing moving objects as outliers, which is based on work of Amintoosi *et al*. [1, 2] for background modeling. The key idea of the proposed method lies in identification of the outlier frames using QR factorization technique, a known method in Linear Algebra [10]. QR decomposition can be applied to decompose a given system and indicate the degree of the significance of the decomposed parts. Selection of those frames which are without any moving object, was conceptually obtained by choosing those parts which have weak contribution.

The reminder of this paper is organized as follows: section 2 explains the proposed method. Section 3 provides simulation results and section 4 discusses about pros and cons of the proposed method. The last section describes the conclusion and future works.

2. The Proposed Method

The problem of picking the most influential columns of a given matrix is known as subset selection [10, 17]. With considering some assumption we relate this problem to outlier detection. Suppose that (a) the camera is fixed, (b) the noise domain is very smaller than the signal domain and (c) the background can be seen totally in β percentage of input frames. If $\underline{Y} = \left[\underline{Y}_1, \underline{Y}_2, \ldots, \underline{Y}_N\right]$ is the matrix representation of N input frames,those columns of $\underline{Y}$ correspondence to background frames are almost similar to each other. Hence these columns of $\underline{Y}$ are *sufficiently dependent*. With an extra assumption, -which will be mentioned in section 2.2- those frames containing moving objects are *sufficiently independent*. The core of the proposed method is identification and removing this independent columns from the data set. But the input frames in SR problem don't satisfy assumption (a). In the following subsections we first review the QR-Decomposition as a tool for subset selection problem and then we will explain our solution.

2.1. Subset Selection and QR-Decomposition

The subset selection arises in many practical problems such as rule base reduction [13, 17], which the underlying model is represented by a matrix $P \in \mathbb{R}^{M\times N}$. One approach for selecting the most important columns of P is in short as follows [10]:

1. Estimate n, the effective rank of P.
2. Calculate a permutation matrix Π such that the columns of the matrix $P_n \in \mathbb{R}^{M\times n}$ in $P\Pi = \left[P_n, P_{N-n}\right]$ are sufficient independent.

The usual technique for estimating n is based on finding a well defined gap between singular values of P [10]. As we will see in the next subsection, at least in the current stage of our method, we will use a simple method for dealing with rank revealing.

Calculating the permutation matrix Π can be done with QR factorization. The QR decomposition of $P \in \mathbb{R}^{M\times N}$ is given by $P\Pi = QR$, where $\Pi \in \mathbb{R}^{N\times N}$ is a permutation matrix, $Q \in \mathbb{R}^{M\times N}$ has orthonormal columns, and $R \in \mathbb{R}^{N\times N}$ is upper triangular. The QR decomposition is uniquely determined by the permutation matrix Π , and many techniques have been proposed to compute it, The most well known is the column pivoting strategy [10].

2.2. Applying the QR-Decomposition for outlier detection in SR

Our outlier detection is based on the calculation of the permutation matrix Π that extracts an independent subset of columns $\underline{Y}_k$, assumed to correspond to those frames containing moving objects. Since existence of displacement between input images is necessary in SR methods, our assumption (a) in the previous section is violated. For satisfying this assumption, we made it after the registration step. Figure 2 shows the overall SR framework after adding the outlier removal stage to one was shown in figure 1.

But there is another problem yet, we have black boundaries in many frames, resulting of motion compensation of registration step. This problem prohibits us to use the QR-Decomposition. An easy method to circulate this difficulty is considering an inner sub-matrix of each frame, so that the assumption (a) is fulfilled. This sub-image selection is temporary accomplished only on this stage. The reconstruction

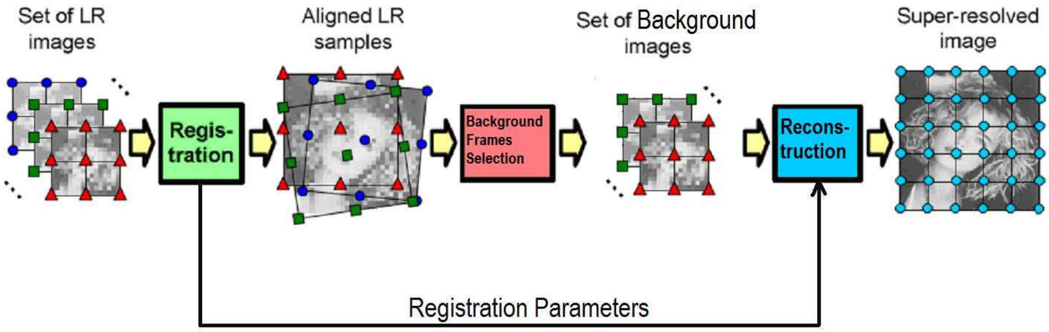

Figure 2. The schematic diagram of the overall SR method with the proposed outlier rejection step.

stage will act on the whole image. We have to find the coordinates $(x_1, y_1), (x_2, y_2)$of each image so that the resulting sub-images represent an equal portion of the scene. If $(\Delta x_k, \Delta y_k)$ be the motion vector between the k^{th} frame and a reference image, row coordinates can be computed as follows:

$$y_1 = Max\{1, \underset{k}{Max}\{\Delta y_k\}\} \tag{3}$$

$$y_2 = M + Min\{0, \underset{k}{Min}\{\Delta y_k\}\} \tag{4}$$

Column coordinates x_1, x_2 can be computed in a similar way.

Now we need a sub-matrix of the motion compensated form of each low-resolution input frame. Suppose that F'_k (of size $[M^2 \times M^2]$) similar to F_k, is the geometric motion operator for k^{th} low resolution image, then $\underline{Y'}_k$ will be the motion compensated form of $\underline{Y}_k$ as follows:

$$\underline{Y'}_k = \underline{F'}_k \underline{Y}_k \tag{5}$$

Let Y'_k (of size $[M \times M]$) be the 2-dimentional representation of $\underline{Y'}_k$, we use a notation used in Linear Algebra for representing a sub-matrix of Y'_k by removing some rows and/or columns of it.

Definition 1 *(Sub-matrix) Suppose that A is an $M \times N$ matrix, $I \subseteq \{1, \ldots, M\}$ and $J \subseteq \{1, \ldots, N\}$ are two sets with m and n elements, respectively. Then the* sub-matrix $A(I|J)$ [1] *(of size $(M-m) \times (N-n)$) obtained from A by removing those rows and columns of A which their indices are in I and J, respectively.*

Remark. If I $(J)$is empty set, none of the rows(columns) of A will be removed. Note that the above definition can be extended to higher dimensions.

Now let $I = \{1, \ldots, y_1 - 1, y_2 + 1, \ldots, M\}$ and $J = \{1, \ldots, x_1 - 1, x_2 + 1, \ldots, N\}$, hence $Y'_k(I|J)$ is our desired portion of compensated low resolution frame k.

[1]This notation should not be confused with the conditional probability

An issue which is not addressed in [2, 1] is that the pivoting algorithm favors columns of R with a large norm. related to the norm of the columns of P. The trick which is used here is to utilize the gradient of $\underline{Y'}$ in horizontal direction. let x axis denotes horizontal direction, we define $\underline{Y''}$ as:

$$\underline{Y''} = \frac{\partial \underline{Y'}}{\partial x} \tag{6}$$

Where: $\underline{Y'} = \left[\underline{Y'}_1(I|J), \underline{Y'}_2(I|J), \ldots, \underline{Y'}_N(I|J)\right]$. Our simulation results showed the better performance of the gradient with QR Decomposition. Hence the QR-Decomposition can be apply on $\underline{Y''}$.

Here we have to find the effective rank of $\underline{Y''}$. Assumption (c) leads us that $(1-\beta)N$ can be a good estimation for the effective rank, with the following extra assumption:
Assumption (d): Non-background frames are not identical. This assumption will be false for example when a moving object remains on a fixed location for a few frames and no other moving objects exist in these frames. [2]

If $\{f_1, \ldots, f_N\}$ represent the frame numbers, resulting of the permutation matrix Π of QR-Decomposition, we reject the first $n = (1-\beta)N$ frames as outliers. Based on the input matrix of QR-Decomposition, β and the aforementioned set, we define operators *OFI* (Outlier Frames Indices) and *NOFI* (Non Outlier Frames Indices) as follows:

Definition 2 *(*OFI *and* NOFI *Operators) Suppose that $P \in \mathbb{R}^{M \times N}$ and β are given, and $\{f_1, \ldots, f_N\}$ represents the column numbers of P, according to the permutation matrix Π of QR-Decomposition of P, and $M \geq N$, we define OFI (Outlier Frames Indices) and $NOFI$ (Non Outlier Frames Indices) as follows:*

$$OFI(P, \beta) = \{f_1, \ldots, f_{(1-\beta)N}\}$$
$$NOFI(P, \beta) = \{f_{(1-\beta)N+1}, f_{(1-\beta)N+2}, \ldots, f_N\}$$

[2]This assumption may be seems as a limiting constraint, but we can deal with this difficulty by applying the QR-decomposition until the remaining matrix is not near rank deficient, which is not implemented in the current version of our simulations.

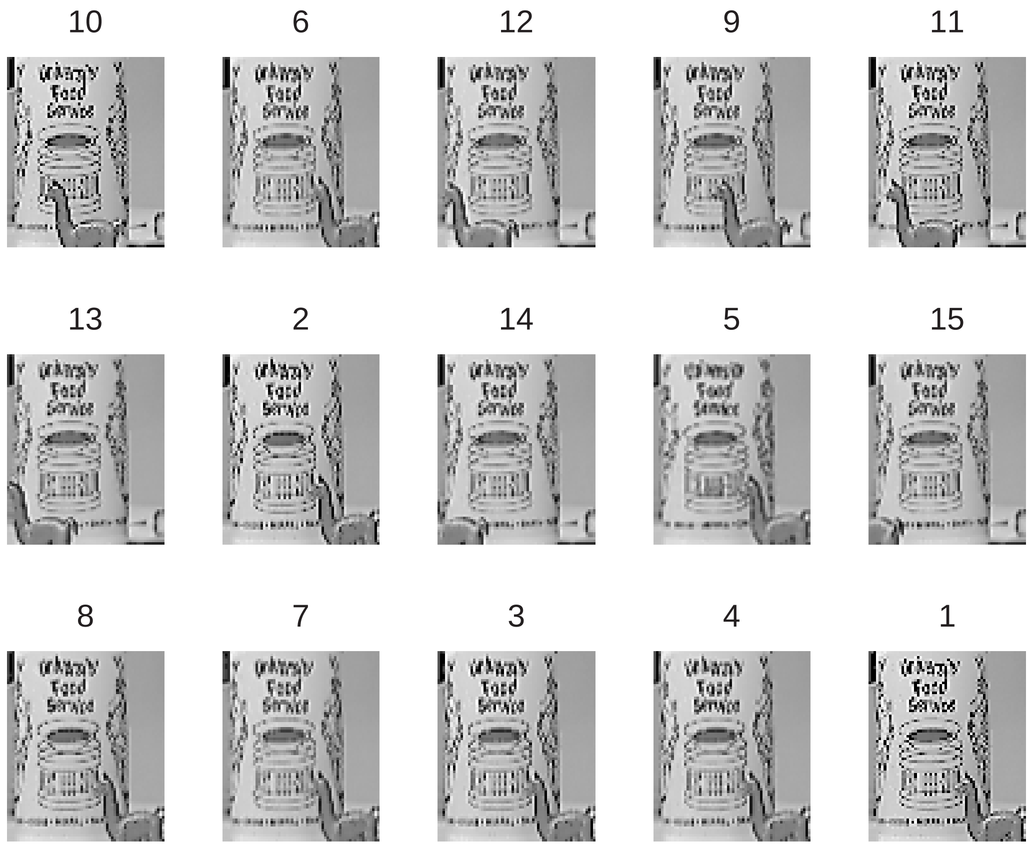

Figure 3. Rearrangement of test frames based on permutation matrix Π. The title of each image, represents its order number in the original sequence.

Restriction of $\underline{Y}$ to columns indicated by the set $NOFI(\underline{Y''}, \beta)$, leads us to a new set of frames, which does not contain any outlier frame and is our desire for super-resolution problem. In other words, $\underline{Y}\big(\varnothing | OFI(\underline{Y''}, \beta)\big)$ eliminates outlier columns of $\underline{Y}$, which their indices are indicated by $OFI(\underline{Y''}, \beta)$.

Hence the SR problem formulated in 2 can be reformulated as follows:

$$\hat{\underline{Z}} = \underset{\underline{Z}}{\text{Arg Min}} \Big[\sum_{k=1}^{N-n} ||DG_k\underline{Z} - \underline{W}_k||_p^p \Big] \qquad (7)$$

where

$$\underline{W} = \underline{Y}\big(\emptyset | OFI(\underline{Y''}, \beta)\big), \\ G = F\big(\emptyset | \emptyset | OFI(\underline{Y''}, \beta)\big), \\ n = (1-\beta)N$$

$\underline{W}, G$ in the above formula indicate that, those frames, which their indices is in $OFI(\underline{Y''}, \beta)$ will be removed from the input data set. Figure 2 shows the overall framework after adding the proposed method between registration and reconstruction steps of figure 1. As it can be seen in figure 2, and obvious from $G = F\big(\emptyset | \emptyset | OFI(\underline{Y''}, \beta)\big)$, another separate registration step is not required for reconstruction stage. We will discuss about the pros and cons of the proposed method after the next section.

3. Simulation Results

We applied our proposed method on a test video images shown in figure 3 and compared its performance with some other algorithms. The used sequence can be downloaded from Prof. Milanfar's web site (http://www.ee.ucsc.edu/ milanfar). In this sequence, two separate sources of motion were present. First, by shaking the camera a global motion

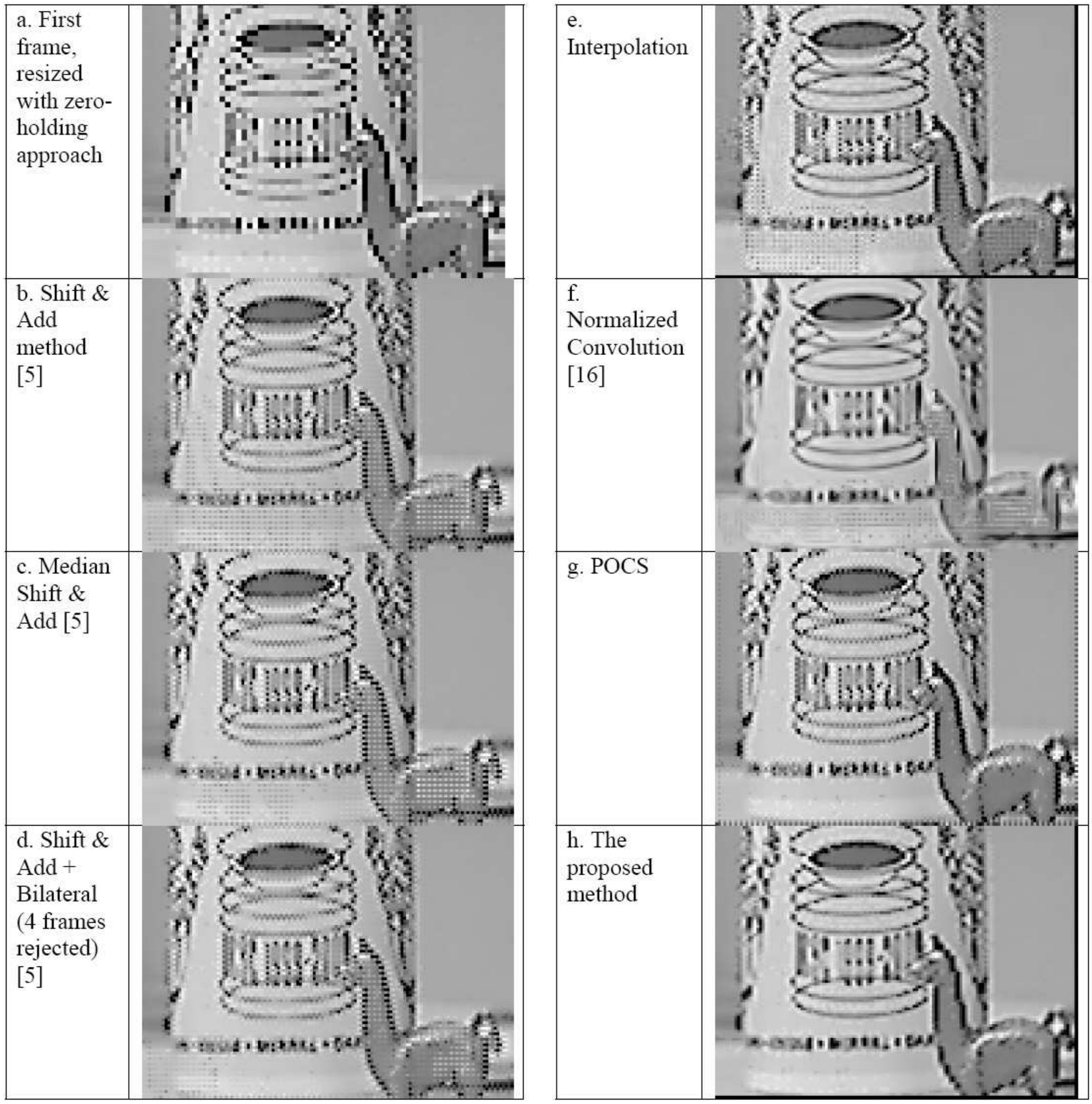

Figure 4. Results of different super resolution methods applied to the Cropped Alpaca sequence (b-h). Outlier effects are apparent in the reconstruction methods(b-f). The shadow of the Alpaca is removed in the POCS (g) and proposed methods(h).

was created for each individual frame. Second, an Alpaca statue was independently moved in to ten frames out of the total 55 input frames [5]. We used only 15 frames (frame numbers: 38-46,50-55) in our simulations, which are shown in figure 3. Frame numbers 47-49 are dropped, because in these frames, the movement of Alpaca statue is paused and our assumption (d) is violated. The shown images are motion-compensated form of the original frames. We used the proposed method by Irani and Peleg [12] for image registration, which is based on the work of Luacs and Kanade [14].

We renumbered the selected frames from 1 to 15, that is shown by the image's title in figure 3. In the last 7 frames (frames numbered from 9-15), Alpaca statue is moving (47 percent of the sequence length).

Figure 3 (from top to bottom and left to right) shows the order of the frames based on the permutation matrix Π, resulting of QR Decomposition. As it can be seen, the outlier frames, shifted to the beginning of the list. The parameter β is set to 1/3, hence the images shown in the first two rows of figure 3 are rejected and the reconstruction step is done only by the last 5 images. In implementation of the proposed method, the interpolation approach is used for reconstruction step. The test results are shown in figure 4. Note the relative improvement in the quality of the proposed method as a result of detecting those frames containing moving objects and not using them in the reconstruction process. Inspection of the results indicates that other approaches -except that POCS method- suffer from artifacts.

4. Discussion

This section is devoted to pros and cons of the proposed method and future works.

- The big question about the proposed method is that removing the entire frame, containing a moving object is not rational.If one frame in SR has a moving object, this object will be visible in many other frames. The better idea is not to remove entire frames, but to select for each region all frames where the desired region has no moving objects. In our previous works about background modeling [1, 2] we divided each frame into some smaller blocks. But at the current stage of this work, the main purpose was introducing QR-Decomposition as a tool for outlier detection in Super-Resolution context. Dealing with the mentioned question is our future work by acting on small blocks instead of the whole frame.
- Super-resolution under photometric diversity is another challenge in SR domain[9]. Since some kind of frame brightness changing is multiplying the pixels' intensities with a specified factor and these two frames are linearly dependent. In contrast with the usual background modeling, the proposed method is promising for dealing such situations.
- One may be wondered why the famous RANSAC method[8] was not used for outlier detection. RANSAC method is used when we have some data points (inliers and outliers) and we have to estimate a model for our data; such as Homography matrix [11] in stereo correspondence or image registration. But here we have not a model for data fitting and therefor we cannot use RANSAC method for outlier removal.
- An intuitive result leads us that, a breakdown point[3] smaller than .5 - which is indicated in [5] as the lowest breakdown point with median estimator- may be achieved under some assumptions, by the proposed method.
- Another issue related to the proposed method is initializing β, which is related to the effective rank of $\underline{Y''}$. The effective rank of $\underline{Y''}$ identified by a (relative) gap in the singular values. The gap represents a point for partitioning frames into background and non-background ones in our method. However, often, the singular values tend to decrease smoothly without any clear gap. In such cases one can guess β by visual inspection. Automating the finding of effective rank is in progress.
- Separating the SR for background and foreground has been done before by Eren *et al*. [4]. The proposed method should motivates new approaches for the SR problem from this point of view.

[3] The breakdown point is the smallest percentage of outlier contamination that may force the value of the estimate outside some range [5].

5. Concluding Remarks

Existence of outliers in the input low resolution images is one of the major challenges in super resolution domain. In this article a new method for dealing moving objects have proposed. In the proposed method, those frames containing moving objects were identified via QR-Decomposition, which is used frequently for subset selection. The proposed outlier removal routine lies prior to reconstruction process and it is very susceptible for combination with other super-resolution techniques. Although in simulation of the proposed method, the interpolation approach is used for reconstruction step, however every other reconstruction method can be used instead of interpolation. In contrast with [5] our method does not need many parameters setting, its only parameter is estimation of the percentage of the sequence length, which the scene is empty; that can be estimated by visual inspection. In contrast with [15] which is a heavy and time consuming method, our method is very light.

As future works, we plan to apply the proposed method to test on different lighting conditions and to acting on small blocks instead of the whole frame. Moreover, mathematically justification of its breakdown point and considering it inside the reconstruction stage is in progress.

6. Acknowledgment

The authors wish to thank to Prof. Milanfar, Dr. Farsiu and Dr. Vandewalle [18] for providing us with their Super-Resolution packages.

References

[1] M. Amintoosi, F. Farbiz, and M. Fathy. A QR Decomposition based mixture model algorithm for background modeling. In *ICICS2007, Sixth International Conference on Information, Communication and Signal Processing*, pages 1–5, Singapore, December 2007.

[2] M. Amintoosi, F. Farbiz, M. Fathy, M. Analoui, and N. Mozayani. QR-Decomposition-based Algorithm for Background Subtraction. In *ICASSP2007, IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1093–1096, USA, 2007.

[3] M. Elad and A. Feuer. Restoration of single super-resolution image from several blurred, noisy and downsampled measured images. *IEEE Trans. Image Processing*, 6:1646–1658, Dec 1997.

[4] P. E. Eren, M. I. Sezan, and A. Tekalp. Robust, object- based high-resolution image reconstruction from low-resolution video. *IEEE Trans. Image Processing*, 6(10):1446–1451, 1997.

[5] S. Farsiu. *A Fast and Robust Framework for Image Fusion and Enhancement*. PhD thesis, Electrical Engineeting, UC Santa Cruz, December 2005.

[6] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar. Robust shift and add approach to super-resolution. In *Proc. of the 2003 SPIE Conf. on Applications of Digital Signal and Image Processing*, pages 121–130, Aug 2003.

[7] S. Farsiu, M. D. Robinson, M. Elad, , and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, 2004.

[8] 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*, 24(6):381–395, 1981.

[9] M. Gevrekci and B. K. Gunturk. Superresolution under photometric diversity of images. *EURASIP J. Appl. Signal Process.*, 2007(1):205–205, 2007.

[10] G. H. Golub and C. F. van Loan. *Matrix Computations*. Johns Hopkins Univ. Press, Baltimore, 3 edition, 1996.

[11] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[12] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP: Graph. Models Image Process.*, 53(3):231–239, 1991.

[13] O. Kaynak, K. Jezernik, and A. Szeghergyi. Complexity reduction of rule based models: a survey. In *Proceedings of the 2002 IEEE International Conference on Fuzzy Systems*, pages 1216–1221, 2002.

[14] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI81*, pages 674–679, 1981.

[15] T. Q. Pham. *Spatiotonal Adaptivity in Super-Resolution of Under-sampled Image Sequences*. PhD thesis, aan de Technische Universiteit Delft, 2006.

[16] T. Q. Pham, L. J. van Vliet, and K. Schutte. Robust fusion of irregularly sampled data using adaptive normalized convolution. *EURASIP Journal on Applied Signal Processing*, 2006.

[17] M. Setnes and R. Babuska. Rule base reduction: Some comments on the use of orthogonal transforms. *IEEE Trans. Syst., Man, Cybern, Part C*, 31(2):199–206, May 2001.

[18] P. Vandewalle, S. Ssstrunk, 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)*, 2006:Article ID 71459, 14 pages, 2006.

[19] A. Zomet, A. Rav-Acha, and S. Peleg. Robust super resolution. In *Proceedings of the Int. Conf. on Computer Vision and Patern Recognition (CVPR)*, pages 645–650, Dec 2001.