A Fast Image Registration Approach based on SIFT Key-points Applied to Super-Resolution

Mahmood Amintoosi Mahmood Fathy
Nasser Mozayani
Computer Engineering Department,
Iran University of Science and Technology,
Narmak, Tehran, Iran,
mamintoosi@ieee.org {mahFathy,Mozayani}@iust.ac.ir

Abstract

An accurate image registration is a fundamental stage in many image processing problems. In this paper a new and fast registration approach based on Scale Invariant Feature Transform key-points descriptors, under Euclidean transformation model is proposed. The core idea of the proposed method is estimation of rotation angle and vertical and horizontal shifts using averaging of differences of SIFT key-points pairs descriptors. The method is simple but requires some tuning modules for accurate estimation. Orientation modification and compensation and shift compensation are some of the proposed modules. The proposed method is fast, it is about 5 times faster than RANSAC method for model parameters estimation. The accuracy of the proposed method is compared with some popular registration methods. Various comparisons have been done with LIVE database images with known motion vectors. The experimental results show the high performance of the proposed algorithm in a super-resolution application.

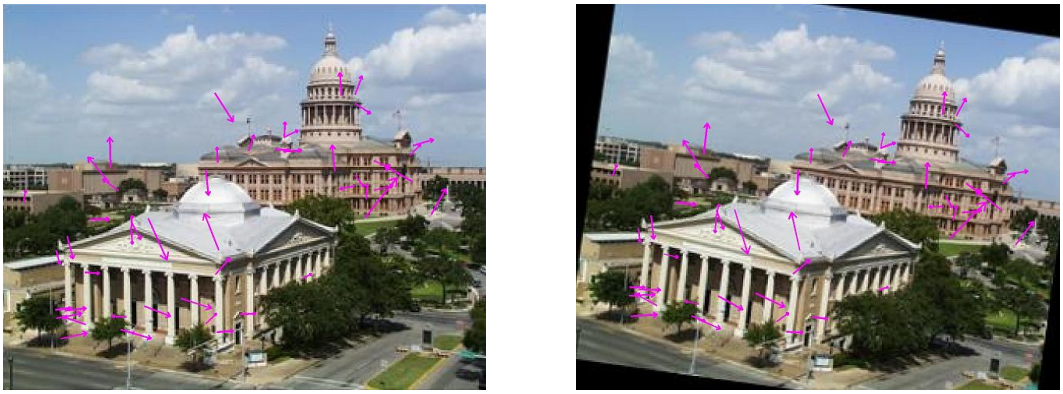
Keywords: Image Registration; Super-Resolution; SIFT Key-points.

1 Introduction

One of the most critical aspects of many applications in image processing and computer vision, including Super-Resolution, is the image registration problem. Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. It geometrically aligns two images, the reference and sensed images (Zitová & Flusser, 2003).

In image processing literatures, a variety of registration categories has been used. Regarding the transformation model among the images (such as translation, affine or projective) the registration method may be different. But, they can be categorized into two main approaches: area-based methods and feature-based methods. While the former uses the information from all pixels, the latter requires only a sparse set of feature correspondences to fit the motion model (Pham, 2006).

The Lucas-Kanade registration algorithm (Lucas & Kanade, 1981), is a famous area-based method, which is the basis of many other methods (Keren et al., 1988; Baker et al., 2004) . Their approach is based on using 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. As Taylor series only give a good approximation for small offsets, these registration methods are generally applied iteratively using a Gaussian pyramid. Vandewalle et al. (2006) 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(Vandewalle et al., 2006). Another fast image registration



(a) 50 selected keypoints from total 894 keypoints

(b) 50 selected keypoints from total 921 keypoints

Figure 1: Some of SIFT keypoints of an image and its transformed version (of size 252×316). Keypoints are displayed as vectors indicating scale, orientation, and location. The SNR=70 and distance ratio=.1 in the Lowe's algorithm. The transformation parameters of the right side image are: $t_x = 9, t_y = 1, \phi = 8°$.

which is used in image stabilization context is Gray Coded Bit Plane Matching (GC-BPM)(Ko et al., 1999). This method is very computationally efficient since it uses binary boolean operations, but its performance is lower than popular methods such as Keren et al. (1988).

In many image processing applications, such as some Super-Resolution problems (Fathy et al., 2008), global translational motion model is assumed, in which the low resolution input images have small rotation and translation differences with respect to each other. In this paper our goal is to find the registration parameters between two images, with the above mentioned assumption, directly from their SIFT key-points' descriptors.

Following introducing SIFT by Lowe (1999, 2004), various applications of it, including matching and registration reported by some researchers. Mikolajczyk & Schmid (2005) compared the performance of some descriptors computed for local interest regions and their results showed that the SIFT-based descriptors have the highest performance. Yi et al. (2008) used SIFT key-points for multi-spectral remote images. They proposed a matching method and called it SR-SIFT algorithm (SIFT matching with Scale Restriction) to reduce the incorrect matches. The famous RANSAC[1] algorithm (Fischler & Bolles, 1981) has been used many times for removing outliers (incorrect matches) from SIFT key-points pairs and estimating a homography matrix between two images (Amintoosi et al., 2008; Yuan et al., 2008; Amintoosi et al., 2009).

The original matching method proposed and implemented by Lowe (2004)[2] consists of a nearest neighbor search and a heuristic criteria suggested by him, which is the ratio of closest to the second closest neighbor (named 'distance ratio'). The method is very powerful in finding correct matches among putative key-points. Figure 1 shows two versions of an instance image from LIVE dataset (shown in figure 2) and some of their SIFT keypoints matches with the Low's program. The rotation angle of the second image with respect to the first one, was $8°$. The key-points' orientations differ from ϕ, the image rotation angle with respect to the reference image.

Our goal is to estimate registration parameters (t_x, t_y, ϕ) between two images, directly from SIFT key-points' descriptors. The angle resulting from difference of corresponding key-points' orientations is considered as an estimation of ϕ. The proper estimation of the rotation angle with this method has some limitations which is discussed in the next section. After estimating the rotation angle between image pairs, the key-points' locations of the second image are rotated in a proper manner. Computing the registration parameters has been done following an outlier removal stage.

The experimental results showed that the proposed method is about 5 times faster than RANSAC for obtaining registration parameters, while its estimation accuracy is competitive with RANSAC method.

[1] RANdom SAmple Consensus (RANSAC)

[2] The implementation is available online at: http://www.cs.ubc.ca/∼lowe/keypoints/

Figure 2: Some of LIVE database images (Sheikh, Wang, Cormack & Bovik, Sheikh et al.) which are used in the experiments of this paper.

Also its precision is compared with some famous registration method like Keren et al. (1988) method. The performance of the proposed method is compared with some other methods in a super-resolution application, while a high resolution (HR) image is achieved from the motion compensated low resolution(LR) images with a Super-Resolution reconstruction method. Our implementation results show the high performance of the proposed method.

The rest of the paper organized as follow: section 2 explains the proposed method. Section 3 provides the experimental results and section 4 is dedicated to the concluding remarks.

2 The Proposed Method

Assume we have a continuous two-dimensional reference signal $f_0(X)$ and its shifted and rotated version $f_1(X)$:

$$f_1(X) = f_0(R(X + \Delta X)), \tag{1}$$

with

$$X = \begin{pmatrix} x \\ y \end{pmatrix}, \Delta X = \begin{pmatrix} t_x \\ t_y \end{pmatrix}, R = \begin{pmatrix} cos\phi & -sin\phi \\ sin\phi & cos\phi \end{pmatrix} \tag{2}$$

Our goal is to estimate (t_x, t_y, ϕ) between a pair of LR images, directly from SIFT key-points' descriptors. Among the various features used in feature based image registration methods, SIFT key-points of Lowe (1999) has gained a great attention in recent years. SIFT key-points are identified as the local maxima or minima of the difference-of-Gaussian filters across scales. To determine a key-point's orientation, a gradient orientation histogram is computed in the neighborhood of the key-point. Peaks in the histogram correspond to the dominant orientations. Each key-point is denoted by a vector $(x_i, y_i, \sigma_i, \theta_i)^T$, which (x_i, y_i), σ_i and θ_i denote location, scale and orientation for i^{th} key-point, respectively (Yi et al., 2008). The usual method for finding a match for each key-point is identifying its nearest neighbor. To ensure a correct match, Lowe (2004) suggests that the ratio of the closest to the second-closest neighbors must be less than a threshold.

Suppose that $(x_i^1, y_i^1, \theta_i^1)$ and $(x_i^2, y_i^2, \theta_i^2)$ are the locations and orientations of i^{th} key-point in image 1 -as the reference frame- and another image 2 (from total N matches found by D.Lowe's suggestion). Let $\Delta x_i = x_i^1 - x_i^2$, $\Delta y_i = y_i^1 - y_i^2$ and $\Delta\theta_i = \theta_i^1 - \theta_i^2$; the following notations for $\Delta x, \Delta y$ and$\Delta\theta$ is used in the following subsections for explaining the proposed method:

$$\begin{aligned} \Delta x &= \{\Delta x_1, \cdots, \Delta x_N\} \\ \Delta y &= \{\Delta y_1, \cdots, \Delta y_N\} \\ \Delta\theta &= \{\Delta\theta_1, \cdots, \Delta\theta_N\} \end{aligned} \tag{3}$$

Our approach for estimating t_x, t_y and ϕ, from $\Delta x, \Delta y$ and $\Delta\theta$ is simple. It is based on two assumptions: the normality of displacements and the separability of the shift and rotation estimations. At first based on the normality of displacements, the incorrect matches are reduced by removing those data points which are far from the mean. Then based on the second assumption[3], ϕ, Δx and Δy are estimated. The rotation angle, ϕ is approximated by averaging of $\Delta\theta$, in which some modifications are needed for accurate estimation. Then the key-points' locations of the second image are compensated according to the estimated rotation angle, ϕ. Finally t_x, t_y are approximated by averaging Δx and Δy. The details are discussed in the following sections.

2.1 Rotation Estimation

The rotation angle ϕ may be approximated by averaging $\Delta\theta$, i.e. $\phi = \overline{\Delta\theta}$. Because this method is based on θs, the rotation angles of SIFT key-points, in the following, we denote this method of estimation of rotation angle as T-SIFT. But as we will see later, this method does not operate well for negative rotation angles. There are some questions, related to averaging $\Delta\theta$:

- Whether each delta follows a known distribution, like Normal distribution?
- If yes is the average of samples, has meaningful difference with the true value of parameters?

In the following, at first the normality of the distribution of $\Delta\theta$ will be discussed. As we will see $\Delta\theta$, based on T-SIFT, does not follow the Normal distribution, when the rotation angle is negative, unless some modifications are applied.

2.1.1 On the Normality of $\Delta\theta$

Suppose that we have M images that should be aligned with respect to a reference image in which their rotation angle are equal ϕ and their horizontal and vertical shifts are unknown. Let ϕ_m be the estimated rotation angle of m^{th} image with respect to the reference image; $\Phi = \{\phi_1, \cdots, \phi_M\}$ can be considered as a random variable. Here we discuss whether we can consider the average value of Φ,$\overline{\Phi}$ as an estimation of rotation angle between two images (ϕ) or not.

A Z-test is any statistical test for which the distribution of the test statistic under the null hypothesis can be approximated by a normal distribution. According to the Central Limit Theorem (CLT), $\overline{\Phi}$ is approximately normally distributed for large samples, the next step is to determine whether the expected value ϕ under the null hypothesis, has meaningful difference with its true value or not.

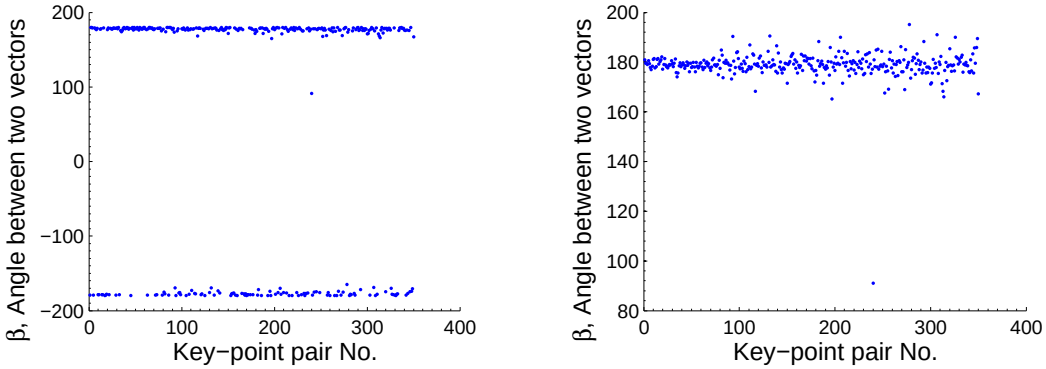
The experimental results over LIVE dataset (Sheikh, Wang, Cormack & Bovik, Sheikh et al.), verified that the estimated ϕ has not meaningful difference with true ϕ with the T-SIFT method when $0 \leq \phi < 180$, with 95% confidence; but it does not hold when $\phi < 0$. Table 4 shows the result of Z-test for some methods which will be explained in more details later. The reason for these disappointing result is calculating of dot product for computing angle in Lowe's matching procedure, which is used in T-SIFT.

In the following we first describe our solution for dealing with this problem and then explain our method for removing outliers, which leads to a better estimation of registration parameters.

2.1.2 The Orientation Modification

We used the implementation of SIFT key-points extraction and matching by Lowe (2004) which computes dot products between unit vectors ($v_1 = sin\theta_i^1, cos\theta_i^1$) and ($v_2 = sin\theta_i^2, cos\theta_i^2$) rather than Euclidean distances. It is computationally efficient, but it would not specify whether v_1 is ahead or behind v_2. In most math libraries $acos(.)$ will usually return a value between 0° and 180°. We used the following method for indicating whether v_1 is ahead or behind v_2.

[3]The separability of the shift and rotation estimations was shown by Vandewalle (2006, chap. 3).



(a) β values, demonstrate two clusters, near to -180° and 180°

(b) Resolving the problem by the proposed method in 2.1.3.

Figure 3: Scatter plot for β in eq. (4) over $\Delta\theta$ between two instance images with rotation angle equal 179°.

The orientation of each key-point is in the range of [-180,180], hence the difference between angles of two vectors will be in the range of [-360,360]. Suppose that $\alpha = \Delta\theta_i$ is the angle between i^{th} key-points pair, $\beta(\alpha)$, the modified orientation of α is defined as follows:

$$\beta(\alpha) = \begin{cases} \alpha & \alpha \in [-180, 180] \\ \alpha - 360 & \alpha \in (180, 360] \\ \alpha + 360 & \alpha \in [-360, -180) \end{cases} \tag{4}$$

Note that the modification is applied on the differences of SIFT key-points' orientations and the resulting angle may be positive or negative. The investigation of 4 is left to the reader. The advantage of this modification is that the difference angle between two key-points, indicates the direction of rotation angle.

2.1.3 The problem of rotation angles near 180°

It should be mentioned that with the above orientation modification, we have some problems for estimating of rotation angles, when it is close to 180 or -180 degree. Consider two key-point pairs shown in table 1.

Table 1: An example demonstrating a problem encountered in large angles.

Key-point Pair	θ_i^1	θ_i^2	$\alpha = \Delta\theta_i$	$\beta(\alpha)$
1	179	2	177	177
2	178	-3	181	-179

For both of these key-point pairs, the rotation angle (α) is close to 180 degrees, but β is 177 and -179 degrees. In this situation some parts of the correspondences show positive angle (close to 180°) and some of them show negative angle (near -180); which is not suited case for our algorithm that is based on averaging. Figure 3(a) shows this case for two images with the rotation angle of 179 degree. As can be seen many of the angles of the corresponding key-points' pairs have been clustered in two groups: one close to 180 and the other close to -180. For solving this problem, it is sufficient to replace each $\beta(\alpha)$ with its corresponding α for angles close to 180 or -180, if the maximum of PDF of β is close enough to 180 or -180. In this paper 20° has been chosen as a closing threshold. The resulting scatter plot after this process has been shown in figure 3(b).

Even after the above procedure, some outliers may exist. So those points which are far from the mean value more than $2.5 * \sigma_\phi$ as outliers; where $\sigma_\phi = std(\Delta\theta)$. The new mean value of $\Delta\theta$ is our estimation

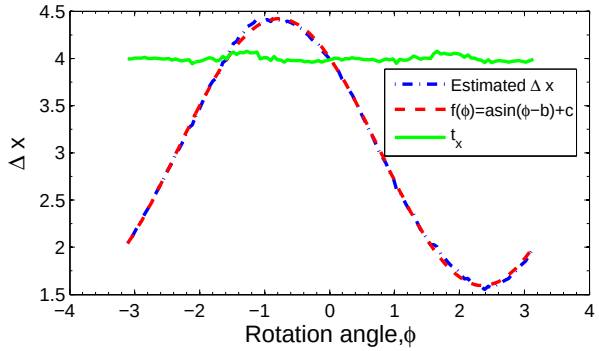


Figure 4: The fitting model for $\overline{\Delta x}$ over rotation angle ϕ, $-\pi \leq \phi \leq \pi$.

of ϕ. The experimental results verified that the estimated rotation angle, ϕ has not meaningful difference with true ϕ with 95% confidence (see section 3, Table 4); hence the estimated rotation angle is reliable.

2.2 Shift Estimation

The overall method for estimating t_x and t_y is based on averaging of Δx and Δy, but there is some notes which is discussed in the following section.

2.2.1 Orientation Compensation

It is obvious that using the average of Δx and Δy, the horizontal and vertical shifts have not been approximated correctly; unless the second image rotated based on ϕ and then t_x, t_y are estimated. If the image rotated by ϕ degree, we have to re-find the SIFT keypoints and rerun the matching procedure. Instead of this time consuming method, we just rotate the locations of key-points, which were found formerly, based on the estimated ϕ. The new $\Delta x, \Delta y$ based on the rotated key-points of the second image are computed. The t_x, t_y are approximated by averaging of $\Delta x, \Delta y$ after an outlier removal based on σ_x, σ_y from $\Delta x, \Delta y$.

2.2.2 Shift Compensation

Our experimental results with fixed known t_x and variable t_y over various values of ϕ ($-2\pi < \phi \leq 2\pi$), show that the estimated t_x has a sinus shape function. It is why the estimated value is not accurate as enough. The following sinusoidal function is fitted on the estimated Δx over ϕ:

$$\overline{\Delta x}(\phi) \approx f(\phi) = a.sin(\phi - b) + c \tag{5}$$

The result of an experience with $t_x = 4$ over 3480 images is shown in figure 4. The dashdot line in figure 4 is the estimated horizontal shift ($\overline{\Delta x}$). For each of 29 LIVE database images, 120 random image with $t_x = 4$ pixel, random vertical shifts and 120 rotation angles (-178°:3:180°) were generated. Hence each point in the figure 4 demonstrate the average of $\overline{\Delta x}$ over 29 images with an specified rotation angle. This process is repeated for $t_x = 2, 6$; the parameters of the fitted functions for the mentioned t_x are shown in table 2.

Table 2: Estimated parameters of fitted function of $f(\phi) = a.sin(\phi - b) + c$ for horizontal shift estimation.

t_x	a	b	c
2	-1.42	0.78	1
4	-1.42	0.78	3
6	-1.42	0.78	5

Algorithm 1 Registration based on SIFT key-points' descriptors.

Input: The pair of images: Image1 and Image2

Output:Registration parameters (t_x, t_y, ϕ).

1: Extract SIFT key-points
2: Find correspondence pairs of key-points, based on D.Lowe's suggestion.
3: Compute the difference of key-point pairs descriptors: $\Delta\theta, \Delta x$ and Δy according to (3).
4: Estimate rotation angle, ϕ based on averaging of $\Delta\theta$ using the method described in section 2.1.
5: **for** i=1 to 2 **do**
6: Estimate horizontal and vertical shifts, t_x, t_y from $\Delta x, \Delta y$ using the method described in section 2.2.1.
7: Remove outliers from $\Delta\theta$ and re-estimate ϕ.
8: **end for**
9: Compute $\phi = \overline{\Delta\theta}$,
compute t_x, t_y according to equations (7) and (8).

As can be seen in table 2, a and b are fixed over various t_x and $c = t_x - 1$, regardless of t_x. Substituting $c = t_x - 1$ in (5) yields:

$$\begin{aligned} &\overline{\Delta x}(\phi) = a.sin(\phi - b) + t_x - 1 \\ &\Rightarrow \\ &t_x = \overline{\Delta x}(\phi) - a.sin(\phi - b) + 1 \end{aligned} \tag{6}$$

The dashed line in figure 4 is the fitting function, $f(\phi) = -1.42.sin(\phi - .78) + 3$, and the solid line is t_x, compensated based on (6). We performed a z-test of the null hypothesis that the estimated t_x's are a random sample from a normal distribution with mean 4, against the alternative that the mean is not 4. The result indicated a failure to reject the null hypothesis at the 5% significance level.

Similar experiments was done for fixed t_y. Hence the following functions are used for shifts compensation:

$$t_x = \overline{\Delta x}(\phi) - a_x.sin(\phi - b_x) + 1, \qquad a_x = -1.42, \quad b_x = 0.78 \tag{7}$$

$$t_y = \overline{\Delta y}(\phi) - a_y.sin(\phi - b_y) + 1, \qquad a_y = 1.42, \quad b_y = -0.78 \tag{8}$$

2.3 The Overall Algorithm

The repetitive patterns in the images, produce some mismatches in the matching stage of Lowe's algorithm. Here, those points which are far from the mean more than $2.5 * \sigma$ are removed as outliers. The overall framework based on the previous stages and outlier removal is shown in algorithm 1. The experimental results showed better performance when ϕ re-estimated after computing shift parameters. Hence in the algorithm we have a for loop.

We named our proposed method OXYT-SIFT, which each letter is described in table 3. Based on table 3, other variations of the proposed method can be named easily, for example OT-SIFT stands for the proposed method, where we have only Orientation modification and rotation estimation based on $\Delta\theta$, without shift estimation. In the next section we will see the experimental results of the proposed method in image registration and an application to Super-Resolution.

3 Experimental Results

Our experiments has been done over LIVE dataset images (Sheikh, Wang, Cormack & Bovik, Sheikh et al.) in which some of them were shown in figure 2.

We mentioned in section 2.1 that the estimated ϕ has not meaningful difference with true ϕ with the proposed method in section 2.1.1 with 95% confidence. For each image of LIVE dataset, 120 image

Table 3: Describing the letters in OXYT-SIFT, the name of the proposed method.

Letter	Explanation
O	Stands for Orientation Modification described in section 2.1.2.
X	Stands for shift estimation along X axis (section 2.2)
Y	Stands for shift estimation along Y axis (section 2.2)
T	Stands for estimation of rotation angles based on θ, without any modifications

were generated with random $t_x, t_y (\in [-10, 10])$ and over various $\phi \in [-180, 180]$. Table 4 shows the result of MATLAB ZTEST function for $\Delta\theta$ with the mentioned T-SIFT method and its modified version based on orientation modification, named as OT-SIFT described in section 2.1. As can be seen, the Null hypothesis about normal distribution of $\Delta\theta$ can not rejected with T-SIFT method only for $\phi > 0$; but the Null hypothesis can not rejected for all values of rotation angle by OT-SIFT method.

Table 4: Z-test over some rotation angles (ϕ)

	ϕ=-178		ϕ=-127		ϕ=-76		ϕ=-25		ϕ=26		ϕ=77		ϕ=128		ϕ=179	
Image	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT
bikes	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
building2	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
buildings	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
caps	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
carnivaldolls	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
cemetry	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
churchandcapitol	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
coinsinfountain	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
dancers	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
flowersonih35	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
house	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
lighthouse	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
lighthouse2	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
manfishing	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
monarch	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
ocean	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
paintedhouse	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
parrots	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
plane	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
rapids	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
sailing1	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
sailing2	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
sailing3	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
sailing4	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
statue	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
stream	1	0	1	0	1	0	1	0	0	0	0	0	0	0	1	0
studentsculpture	1	0	1	0	1	0	1	0	0	0	0	0	0	0	1	0
woman	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
womanhat	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0

Also the effect of noise for normality of $\Delta\theta$ has been tested here. Table 5 shows the result of z-test over various SNRs with fixed value of $\phi = -45°$. The SNR equal 100, means without noise. These experiments ensures that the estimated rotation angle with the proposed method does not have significance difference with its true value.

Table 5: Z-test over various SNRs for $\phi = -45°$

	SNR=20		SNR=45		SNR=70		SNR=100	
Image	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT	T-SIFT	OT-SIFT
bikes	1	0	1	0	1	0	1	0
building2	1	0	1	0	1	0	1	0
buildings	1	0	1	0	1	0	1	0
caps	1	0	1	0	1	0	1	0
carnivaldolls	1	0	1	0	1	0	1	0
cemetry	1	0	1	0	1	0	1	0
churchandcapitol	1	0	1	0	1	0	1	0
coinsinfountain	1	0	1	0	1	0	1	0
dancers	1	0	1	0	1	0	1	0
flowersonih35	1	0	1	0	1	0	1	0
house	1	0	1	0	1	0	1	0
lighthouse	1	0	1	0	1	0	1	0
lighthouse2	1	0	1	0	1	0	1	0
manfishing	1	0	1	0	1	0	1	0
monarch	1	0	1	0	1	0	1	0
ocean	1	0	1	0	1	0	1	0
paintedhouse	1	0	1	0	1	0	1	0
parrots	1	0	1	0	1	0	1	0
plane	1	0	1	0	1	0	1	0
rapids	1	0	1	0	1	0	1	0
sailing1	1	0	1	0	1	0	1	0
sailing2	1	0	1	0	1	0	1	0
sailing3	1	0	1	0	1	0	1	0
sailing4	1	0	1	0	1	0	1	0
statue	1	0	1	0	1	0	1	0
stream	1	0	1	0	1	0	1	0
studentsculpture	1	0	1	0	1	0	1	0
woman	1	0	1	0	1	0	1	0
womanhat	1	0	1	0	1	0	1	0

3.1 Run Times

Table 6 shows the average run times of different methods for estimation of registration parameters over 3480 images. The proposed OXYT-SIFT method is about 5 times faster than RANSAC in average. For each of 29 LIVE images, 120 random images with known motion vectors was created. Their SIFT key-points were extracted and saved before computing running times.

The initial SIFT matching time which is common in the methods is discarded. The mean run time for RANSAC was 16.78^{ms}, against to 3.5^{ms} for OXYT-SIFT, i.e. the OXYT-SIFT is faster than RANSAC about 5 times.

Distance Ratio for matching SIFT key-points was set to 0.8, and distance threshold for deciding outliers in RANSAC homography was set to 0.01.

Table 6: Average run times over 3480 images for estimating of motion parameters.

Method:	T-SIFT	OXYT-SIFT	RANSAC
Time (milisecond):	2.81	3.50	16.78

3.2 Registration Comparison

In addition to RANSAC, we applied three registration methods: frequency method(Vandewalle et al., 2006), GC-BPM(Ko et al., 1999) and Keren et al. (1988) for comparing to the proposed registration approach.

For every image in LIVE dataset, 4 distorted image was generated by resizing each image by a factor of .5, then the image was shifted by known values among the X and/or Y axis (in the range of [-10,10] pixels) , and rotated by a specified angle (in the range of [-10,10] degrees). Adding noise to the image, so that the SNR of the produced image was 70dB and JPEG compression were the last steps. It should be mentioned that the motion parameters for the first distorted image was set to zero as it is a reference frame. The images were generated in a manner to be used in registration comparison and in a super-resolution application.

Because of the image size restriction in GC-BPM method, 6 images of dataset was dropped and the comparisons was done on remaining 23 images. Since for each reference image we had 3 distorted images, the total number of tested images is 69. Table 7 shows the estimated parameters (t_x, t_y, ϕ) for 23 of 69 distored images, along with their ground truth values, with various methods (instead of GC-BPM for table size limitation).

Table 7: Estimated parameters of some images with various methods.

	t_x					t_y					ϕ				
Image	**Grand Truth**	Vandewalle	Keren	OXYT SIFT	RANSAC	**Grand Truth**	Vandewalle	Keren	OXYT SIFT	RANSAC	**Grand Truth**	Vandewalle	Keren	OXYT SIFT	RANSAC
bikes	**-2**	-1.95	-2.04	-2.01	-2.02	**-3**	-2.87	-3.04	-3.00	-2.99	**1**	1.00	0.97	0.95	0.99
building2	**-7**	0.38	-7.02	-7.01	-6.87	**6**	-0.15	5.83	6.01	5.84	**-8**	-8.10	-7.82	-8.10	-8.00
buildings	**-3**	0.72	-2.86	-2.98	-3.18	**7**	0.38	6.40	6.96	7.14	**10**	9.80	10.22	10.12	10.01
caps	**-3**	-1.88	-3.09	-2.95	-3.09	**2**	1.09	1.02	2.00	2.18	**4**	3.10	3.66	3.84	3.96
carnivaldolls	**-1**	0.03	-0.91	-1.15	-1.08	**-9**	0.89	-8.82	-8.97	-8.93	**2**	1.70	1.77	1.91	2.03
cemetry	**0**	-0.01	-0.09	0.00	0.07	**10**	-0.79	9.93	10.03	9.93	**-4**	-4.00	-3.87	-4.15	-4.01
churchandcapitol	**-3**	-2.30	-2.86	-3.08	-3.15	**-3**	-1.65	-2.97	-2.99	-3.01	**1**	0.20	1.07	0.98	0.99
coinsinfountain	**5**	1.02	4.93	4.98	4.93	**4**	0.74	4.16	4.01	4.03	**3**	2.90	2.56	2.87	3.02
dancers	**-5**	-0.13	-5.20	-5.01	-5.01	**9**	-0.18	8.40	9.01	9.02	**0**	-0.10	0.39	-0.01	-0.01
flowersonih35	**-7**	-1.53	-6.79	-6.99	-7.08	**2**	-0.47	1.57	2.00	2.04	**2**	1.40	1.86	2.04	1.99
house	**-2**	0.69	3.53	-2.00	-2.20	**7**	0.30	-2.63	6.99	7.14	**10**	9.60	3.77	9.93	10.00
lighthouse2	**-1**	-0.69	-1.12	-1.05	-1.02	**-4**	-3.01	-4.22	-4.03	-3.94	**3**	2.60	2.82	3.01	2.99
manfishing	**-5**	0.54	-4.94	-5.04	-5.23	**-8**	0.39	-8.27	-8.01	-7.76	**10**	10.00	9.88	10.03	9.97
monarch	**-2**	-1.54	-1.98	-2.00	-2.08	**-4**	-3.35	-3.97	-3.97	-3.96	**-1**	-1.20	-0.92	-0.88	-0.99
ocean	**-10**	-0.36	-8.36	-10.02	-10.08	**-5**	0.70	-5.45	-5.03	-4.95	**6**	5.80	5.81	6.00	5.98
paintedhouse	**-1**	0.44	-1.69	-1.03	-1.07	**-7**	0.14	-7.37	-7.01	-6.89	**5**	4.90	4.75	4.82	5.01
parrots	**-1**	-0.99	-1.58	-1.03	-1.17	**-1**	-0.87	-2.31	-1.04	-0.86	**7**	6.70	5.12	6.85	7.03
plane	**-10**	-0.17	-7.39	-10.09	-10.11	**10**	0.13	1.35	10.01	9.92	**0**	-0.10	0.52	0.09	-0.08
rapids	**-1**	-0.97	-1.15	-1.03	-1.19	**-3**	-2.91	-3.13	-2.99	-2.84	**8**	7.70	7.76	8.03	8.00
sailing1	**-4**	1.00	-3.99	-3.89	-4.04	**-8**	0.30	-9.57	-7.79	-7.88	**7**	6.70	6.28	7.07	7.03
sailing4	**-4**	0.62	-4.02	-3.98	-3.85	**-9**	0.70	-8.80	-9.03	-9.21	**-9**	-8.70	-8.66	-9.06	-9.03
stream	**-10**	-0.51	-9.66	-10.00	-10.05	**10**	-0.19	9.37	10.00	10.06	**3**	0.00	2.70	3.13	2.99
studentsculpture	**-5**	0.15	-4.75	-5.01	-5.04	**5**	0.42	4.89	5.00	5.06	**2**	1.90	1.98	1.94	2.01

For better comparison, in figure 5 the Mean Square Error (MSE) between motion parameters esti-

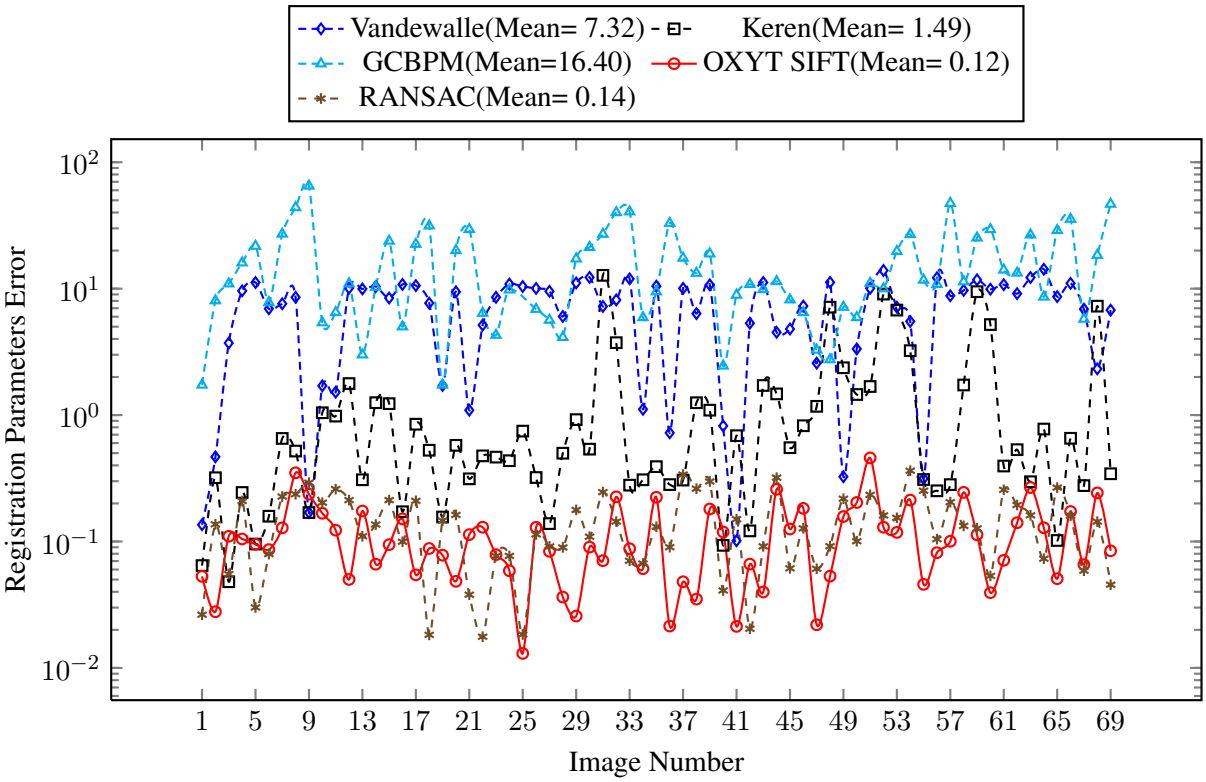


Figure 5: Mean Square Error between estimated registration parameters and their grand truth values, over 69 distorted images.

mated by all mentioned registration methods for all 69 images is demonstrated. The average value of each method is demonstrated beside its legend. As it can be seen the proposed method produced better results against the others in average.

Since 3 registration parameters have different units, the following error measure is a better criterion than the errors in parameters. For each estimated model, the MSE is computed over 30 random points in the image coordinate frame of the distance between their current and correct transformed locations. Figure 6 shows comparison of the MSE between the estimated locations and true locations by mentioned registration methods. As can be seen, the proposed method after RANSAC method produced better results against the others in average.

3.3 Dealing with Large Angles

In the above experiments, the range of rotation angles was [-10,10] degrees but the proposed method can be used to a wider range ([-180,180]). Figure 7 shows the result of running the mentioned methods on an image rotated by various angles in the range of [0,180]. The aforementioned criterion, the MSE between the estimated locations and true locations of 30 pixels is chosen for comparison.

As can be seen the proposed method produced the best result in average. Its performance does not decrease with increasing the rotation angle, in contrast to other methods. The first 3 registration methods in figure 7 are not suitable for estimation of large rotation angles. The performance of RANSAC method, decreases when ϕ increases.

3.4 Experimental Results for SR problems

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 displacements

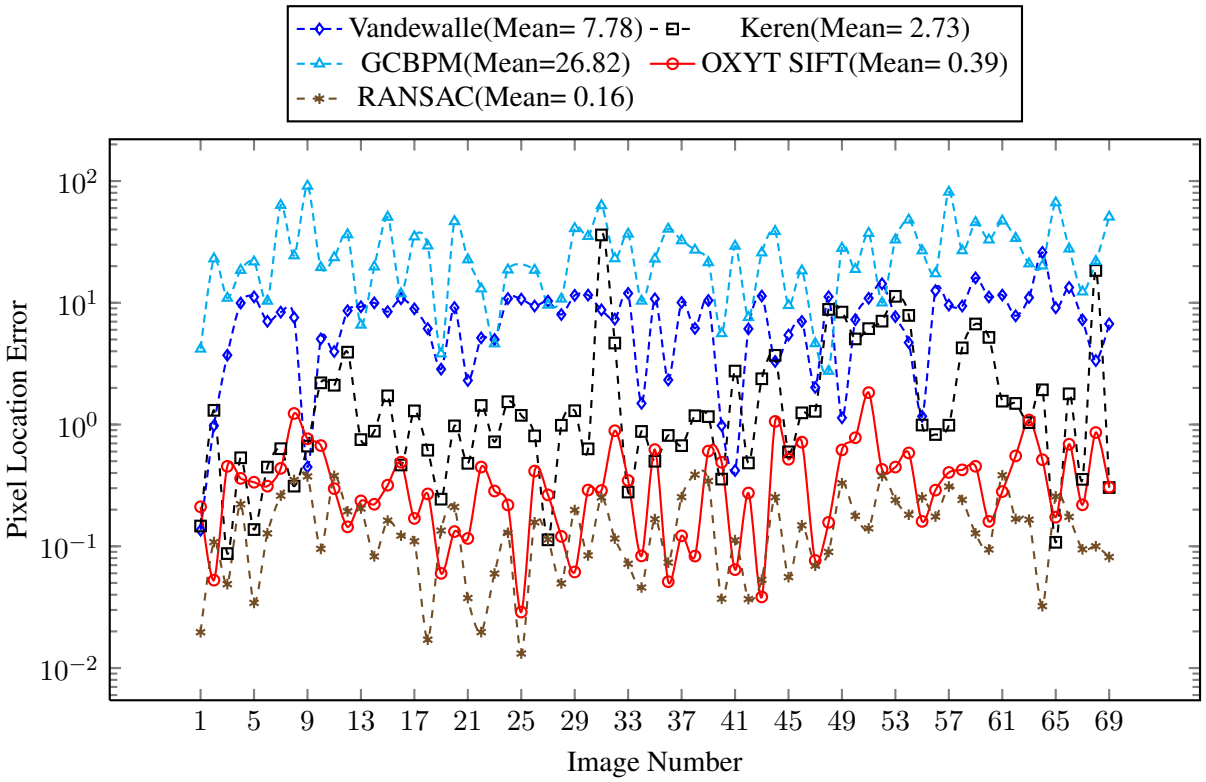


Figure 6: Mean Square Error between estimated registration pixel location error and their grand truth values, over 69 distorted images

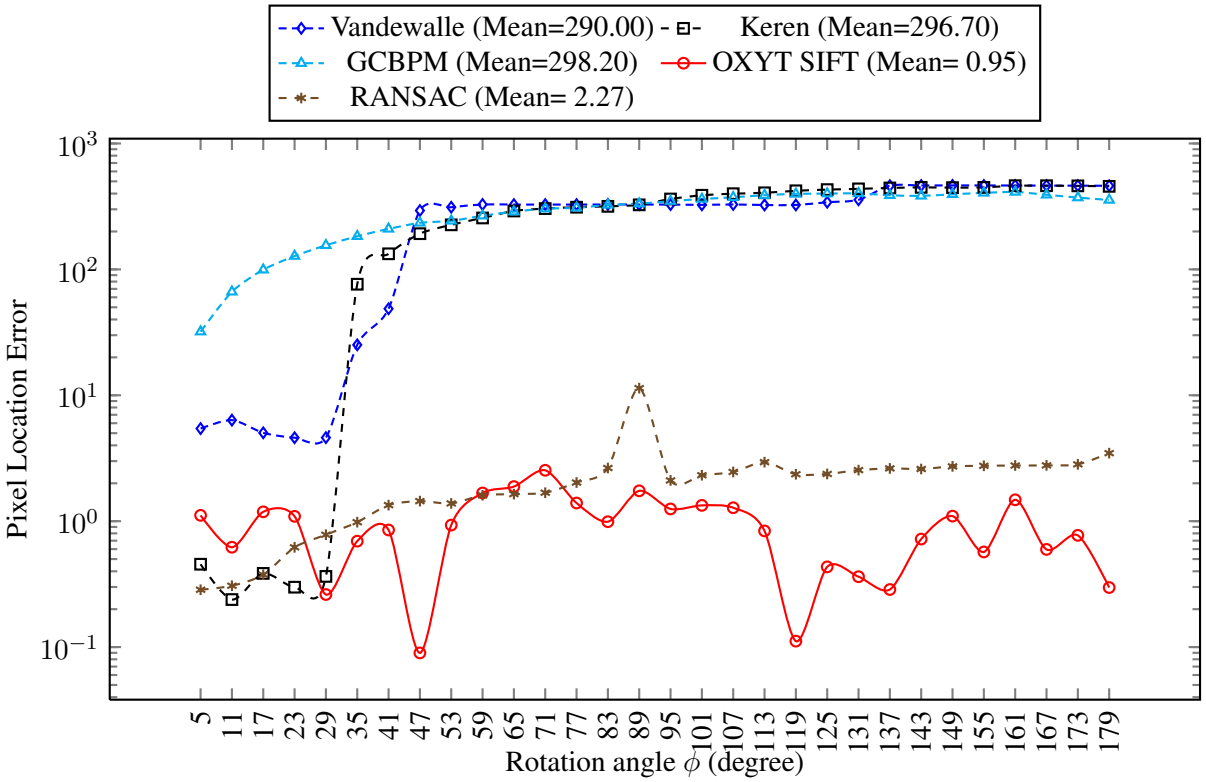


Figure 7: Registration pixel location error for 'Bulidings' image of size 384×256 over various values of rotation angles.

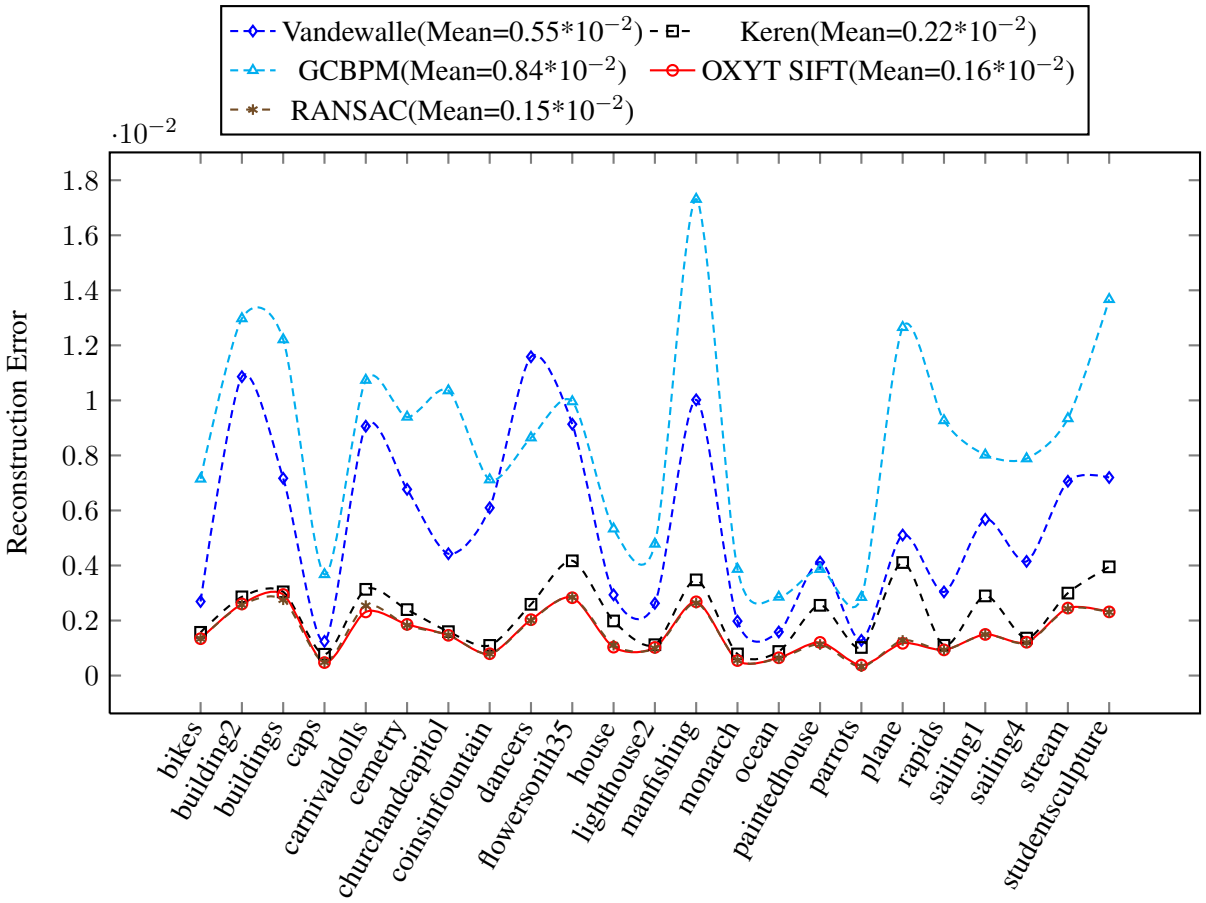


Figure 8: MSE between reconstruction HR image with the mentioned registration methods as the first stage of super-resolution, over 4 LR images corresponding to each LIVE image.

with each other. These methods utilize information from multiple observed images to achieve restoration at resolutions higher than that of the original data. The Super-Resolution restoration methods register the observed images to a common reference frame in order to reconstruct the high Resolution (HR) image. Since Euclidean transformation model is a common assumption in multi frame super-resolution literatures, here it is chosen as an application of the proposed method.

As mentioned earlier, for each image in the dataset, 4 LR image was generated with random motion vectors. Hence for each image of LIVE, as a HR image we have 4 LR images, in which the first LR image is considered as the reference image. The motion parameters are estimated with various registration methods to produce a high resolution image with a magnifying factor of 2. Among the SR reconstruction methods the interpolation approach is used here.

Figure 8 show the MSE between the produced HR images with each registration method against the real HR image. As can be seen the proposed method and RANSAC method produced the better results. Note that in this experiment the rotation angles was belong to [-10°,10°], based on the results of the previous section, the better result of the proposed method in larger rotation angles is expected.

3.5 Visual Comparison

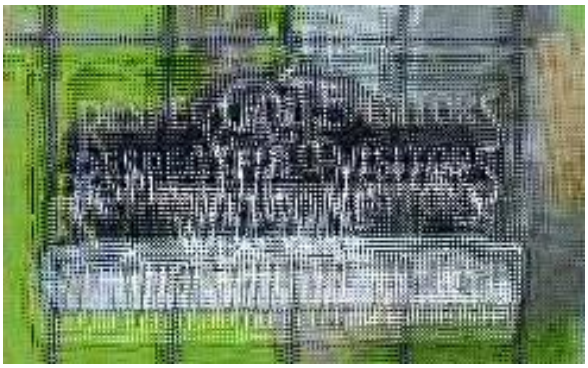
For visual comparison the super-resolution results, when various methods has been used as registration stage of Interpolation reconstruction method, has been used. The ‘cemetry’ image, in which contains a text area has been chosen and the results are shown in figure 9. For a better visual comparison a small region, containing the text, has been enlarged and is shown in the figure.

(a) Original HR frame

(b) Close-up of text area of HR frame

(c) Resized of the first LR frame (Nearest)

(d) GCBPM



(e) Vandewalle

(f) Keren

(g) RANSAC

(h) OXYT SIFT

Figure 9: Close-up views of the original HR image, replication (nearest) resized version of the first LR image, super-resolution results, when GCBPM, Vandewalle, Keren, RANSAC and the proposed 'OXYT SIFT' methods has been used as registration stage.

The LR images were produced with the way explained in section 3.2. Horizontal and vertical shifts of the 3 LR images with respect to the first LR image were {1,4,-1} and {-7,-5,10} pixels, respectively and the rotation angles were {-5,1,2} degrees. The first LR image is shown in figure 9(c); note that the lower text (NO THROUGHFARE PLEASE), is almost unreadable in the LR image. The result of GCBPM and Vandewalle methods (9(d), 9(e)) suffer from bad registration. Although the results of Keren, RANSAC and OXYT SIFT seems equal, but reffering to figure 8, indicates that the MSE of the two later methods are a bit smaller than the Keren method.

4 Conclusion

In this paper a new registration method with the assumption of Normal distribution of displacements of SIFT key-points' descriptors, after some modifications was proposed. The main idea is averaging of differences of key-points' orientations for rotation estimation. The key-points' locations of the second image are rotated based on the estimated rotation angle and vertical and horizontal displacements are approximated by averaging of differences of key-points' locations. Some modification and compensation has been done for accurate estimation of registration parameters.

In contrast to Keren et al. (1988) registration method, which is a repetitive method and can not handle large rotation angles between images, the proposed approach is a one step approach and can handle large rotation angles. In contrast to the frequency method of Vandewalle et al. (2006), the proposed approach does not need strong directionality in images. The GC-BPM (Ko et al., 1999) is a fast registration algorithm but produced poor results with respect to the proposed method. Finally the proposed method is faster than famous RANSAC algorithm for registration parameters estimation. Moreover the proposed method have some variations which can be used, when there are only vertical or horizontal shifts, rotation or combinations of them.

The main strength of the proposed method is in the situations where SIFT key-points of the images are known as a priory, and at the same time the registration parameters are requested; for example in an object recognition and tracking application based on SIFT key-points . In this case the proposed method is faster than RANSAC about 5 times for parameters estimation, while its accuracy is competitive to RANSAC.

The only limitation of the proposed method is that it can be used only for Euclidean transformation model (Translation+Rotation); although it is a usual assumption such as many super-resolution applications. As future work we plan to extend the proposed method for other transformation model such as similarity model (Euclidean + isotropic scaling) with the aid of Yi et al. (2008).

The various comparisons, showed that the proposed registration method outperforms some other popular methods. The experimental results showed the high performance of the proposed method in superresolution problem.

In summery the innovations of this paper are as follows:

- Using SIFT key-points' orientations directly for image registration, in contrast to other methods such as RANSAC which use only the location of SIFT key-points.
- Using SIFT key-points with this manner for image registration in the Super-Resolution context.
- Justification that the displacements between corresponding SIFT key-points' descriptors, under global translational motion model, is approximately normally distributed, after some modifications.

Acknowledgements

The authors wish to thank Dr. Patrick Vandewalle [4], for providing us with his Super-Resolution package, Dr. Alan Brooks [5] for his stabilization MATLAB codes and Prof. David Lowe for his SIFT program.

[4] Laboratoire de communications audiovisuelles, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland.
[5] EECS Dept., Northwestern Univ., United States, http://dailyburrito.com/m-files/.

Also the authors are indebted to Dr. Hamid Sheikh for providing LIVE dataset (Sheikh, Wang, Cormack & Bovik, Sheikh et al.) and A. Ghodsi [6] for his contribution.

References

Amintoosi, M., Fathy, M., & Mozayani, N. (2008). Reconstruction+synthesis: A hybrid method for multi-frame super-resolution. In *(MVIP08) 2008 Iranian Conference on Machine Vision and Image Processing*, (pp. 179–184)., Tabriz, Iran. Tabriz University. 2

Amintoosi, M., Fathy, M., & Mozayani, N. (2009). Regional varying image super-resolution. In *IEEE International Joint Conference on Computational Sciences and Optimization*, volume 1, (pp. 913–917)., Sanya, China. 2

Baker, S., Gross, R., & Matthews, I. (2004). Lucas-kanade 20 years on: A unifying framework. *International Journal of Computer Vision*, *56*, 221–255. 1

Fathy, M., Mozayani, N., & Amintoosi, M. (2008). Outlier removal for super-resolution problem using QR-Decomposition. In *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV08)*, (pp. 271–277)., USA. 2

Fischler, M. A. & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, *24*(6), 381–395. 2

Keren, D., Peleg, S., & Brada, R. (1988). Image sequence enhancement using sub-pixel displacement. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, (pp. 742–746). 1, 2, 3, 10, 15

Ko, S., Lee, S., Jeon, S., & Kang, E. (1999). Fast digital image stabilizer based on gray-coded bit-plane matching. *IEEE Transactions on Consumer Electronics*, *45*(3), 598–603. 2, 10, 15

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *ICCV '99: Proceedings of the International Conference on Computer Vision-Volume 2*, (pp. 1150)., Washington, DC, USA. IEEE Computer Society. 2, 3

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, *60*(2), 91–110. 2, 3, 4

Lucas, B. & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *IJCAI81*, (pp. 674–679). 1

Mikolajczyk, K. & Schmid, C. (2005). A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *27*(10), 1615–1630. 2

Pham, T. Q. (2006). *Spatiotonal Adaptivity in Super-Resolution of Under-sampled Image Sequences*. PhD thesis, aan de Technische Universiteit Delft. 1

Sheikh, H., Wang, Z., Cormack, L., & Bovik, A. Live image quality assessment database release 2. http://live.ece.utexas.edu/research/quality. 3, 4, 7, 16

Vandewalle, P. (2006). *Super-resolution from unregistered aliased images*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne. 4

Vandewalle, P., Süsstrunk, S., & Vetterli, M. (2006). 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. 1, 10, 15

[6] Ph.D Student in Spatial Processes, University Putra Malaysia

Yi, Z., Zhiguo, C., & Yang, X. (2008). Multi-spectral remote image registration based on SIFT. *Electronics Letters*, *44*(2), 107–108. 2, 3, 15

Yuan, Z., Yan, P., & Li, S. (2008). Super resolution based on scale invariant feature transform. In *Rroceedings of International Conference on Audio, Language and Image Processing (ICALIP2008)*, (pp. 1550–1554)., Shanghai, China. 2

Zitová, B. & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, *21*, 977–1000. 1