Journal of Data Science 22 (4), 558–574
October 2024
DOI: 10.6339/24-JDS1139
Statistical Data Science

Image Registration for Zooming Using Similarity Matching

Sujay Das[1] and Partha Sarathi Mukherjee[1,*]
[1]*Interdisciplinary Statistical Research Unit, Indian Statistical Institute, Kolkata, India*

Abstract

Image registration techniques are used for mapping two images of the same scene or image objects to one another. There are several image registration techniques available in the literature for registering rigid body as well as non-rigid body transformations. A very important image transformation is zooming in or out which also called scaling. Very few research articles address this particular problem except a number of feature-based approaches. This paper proposes a method to register two images of the same image object where one is a zoomed-in version of the other. In the proposed intensity-based method, we consider a circular neighborhood around an image pixel of the zoomed-in image, and search for the pixel in the reference image whose circular neighborhood is most similar to that of the neighborhood in the zoomed-in image with respect to various similarity measures. We perform this procedure for all pixels in the zoomed-in image. On images where the features are small in number, our proposed method works better than the state-of-the-art feature-based methods. We provide several numerical examples as well as a mathematical justification in this paper which support our statement that this method performs reasonably well in many situations.

Keywords *local similarity; non-rigid body transformation; scaling transformation; similarity measure*

1 Introduction

In image processing literature, image registration is a method that is widely used for mapping one image to another. When we capture two pictures of the same image object, they are usually not aligned perfectly, i.e., they do not match perfectly pixel by pixel. Because of the difference in the position of the camera, and many other practical reasons, this occurs often. Therefore, it is meaningless to compare these images unless we align them properly. Typically, it is the process of estimating a function, which when applied to one image yields the other image. Now-a-days, imaging techniques are common in many fields such as medical diagnosis (Oliveira and Tavares, 2014; Sotiras et al., 2013; Pluim et al., 2003), monitoring of health and manufacturing processes (Matabosch et al., 2005; Špiclin et al., 2011), and so on. In all these applications, we are interested in comparing two or more images. To accomplish this in a reasonable way, we need to register relevant images effectively. Hence, with widely growing applications of imaging techniques, image registration has become an important research problem.

In literature, various methods of image registration (Brown, 1992; Wyawahare et al., 2009; Zitova and Flusser, 2003) are discussed. Depending on the nature of transformation, there are mainly two types of registration: registration of rigid body transformation (Hill et al., 2001; Ash-

*Corresponding author. Email: psmukherjee.statistics@gmail.com.

Received March 31, 2024; Accepted April 26, 2024

(a) (b) (c)

Figure 1: (a) 'Cameraman' reference image, (b) image rotated clockwise by 10^o, (c) zoomed version of (a).

burner and Friston, 2007; Pataky et al., 2008; Eddyy et al., 1996) and registration of non-rigid body transformation (Crum et al., 2004; Gaens et al., 1998; D'agostino et al., 2003). Rigid-body transformation includes translation (De Castro and Morandi, 1987; Wolberg and Zokai, 2000), rotation (De Castro and Morandi, 1987; Reddy and Chatterji, 1996; Alpert et al., 1990), reflection etc., whereas non-rigid body transformation includes shearing, elastic transformation etc. Among these, rigid-body transformation is most commonly used in many applications such as medical diagnostics. Depending on the nature of mapping used in registration, there are various types of registration procedure, c.f., registration based on image features (Qiu and Xing, 2013), intensity-based registration (Xing and Qiu, 2011), and many more. However, most methods in literature do not consider the fact that the two images may not be zoomed to the same extent.

It may happen sometimes that two images are zoomed unequally due to varying distance of the camera lens from the same image object or artificially zooming of the image object in one image to get a closer look of the image object. Therefore, if we have two images where one is a zoomed-in version of the other, then it poses before us the problem of registration which will enable us to compare the images in future. The images in Figure 1 are examples of various transformations such as translation, rotation and zooming-in of images.

Although there is a number of feature-based methods in the literature addressing the issue of zooming transformations, but direct intensity-based image registration approaches for addressing zooming are not easy to find. Image features are important characteristics of images, such as edges, corners, texture patterns, intensity gradients, and so forth. Image features are visually very prominent in most real-life images. There are several state-of-the-art feature matching algorithms in the literature, viz., Scale Invariant Feature Transform (SIFT) (Lowe, 2004), Speeded Up Robust Features (SURF) (Bay et al., 2008), Oriented Fast and Rotated Brief (ORB) (Rublee et al., 2011), Binary Robust Invariant Scalable points (BRISK) (Leutenegger et al., 2011), KAZE (Alcantarilla et al., 2012), Accelerated-KAZE (AKAZE) (Alcantarilla and Solutions, 2011) to name a few. In these feature-based registration techniques, one first matches the features of the image using the above-mentioned algorithms and then create the registered image using RANSAC (Derpanis, 2010) algorithm. Recently, Das et al. (2024) propose an edge-based image registration method for zooming transformation which requires explicit detection of edges at the beginning of the image registration procedure.

In this paper, we consider two images of equal resolution where one is a zoomed version of the other, and propose an intensity based method to register them. The central idea of this procedure is that, if we select a pixel in the zoomed-in image and consider a neighborhood around it, then we might find a similar neighborhood in the reference image. But, that neighborhood would be at a distant place compared to the zoomed-in image. Therefore, we construct our method in such a way that, corresponding to a certain neighborhood of a pixel in the zoomed-in image, it can find a similar neighborhood and corresponding pixel in the reference image. One major advantage of the proposed method is that it is very simple in its construction, and it does not require explicit detection of image features, such as edges, unlike the recent method proposed by Das et al. (2024), and many other popular feature-based techniques in the literature.

Primary contributions of this paper:

- It proposes an intensity-based image registration procedure for zooming transformation. Local similarity matching is the central idea behind the proposed method.
- The proposed method works on nearly all types of images that we encounter in practice, whereas the popular feature-based registration methods fail to work if the number of prominent image features are low.
- The proposed idea of local similarity matching is simple to understand and implement.
- Adapted versions of the central idea of local similarity matching can be used to develop many other techniques in data science.

In this paper, we describe our proposed registration method in Section 2, and provide its theoretical justification in Section 3. Section 4 presents our numerical studies. In Section 5, we compare our method with some state-of-the-art methods. We conclude by providing a few remarks on this method in Section 6.

2 Proposed Image Registration Procedure

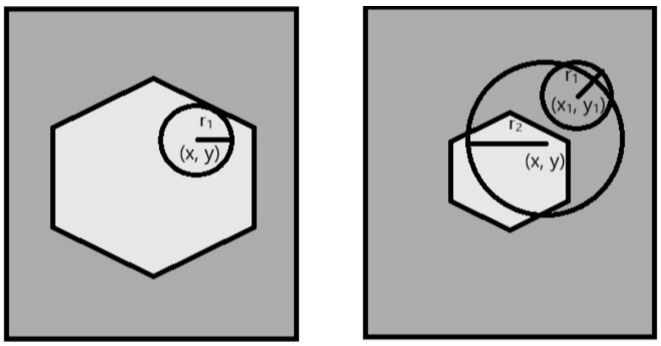
2.1 Assumptions

In this paper, we consider zooming is the only transformation of the image object. Translation and rotation of the image object are not considered in this paper. Moreover, we assume that the true image intensity function of the background of the image is constant. The resolutions of both the images i.e., reference and zoomed images are same, although the last assumption is not critical.

2.2 Description of the Problem

We have two images of the same object or scene. One is the actual image, we call the reference image, and the other one is the zoomed version of the reference image. Our goal is to register these two images i.e., we want to find an one-to-one mapping between these two images or in other words, we want to derive a mathematical function such that once we put the function on one image it will return the other image.

First, we briefly describe the zooming process. A specific region (e.g., the center of the image) of the image is being selected as target region for zooming. Then that region is being zoomed while keeping the resolution same as that of the reference image.



(a) (b)

Figure 2: (a) Zoomed image, (b) Reference image.

2.3 The Proposed Image Registration Procedure

We propose the following steps for performing the registration procedure for zooming:
Step 1: For a given pixel (x, y) in the zoomed image, we consider a circular neighborhood of radius r_1 around that pixel, denoted as $O(x, y, r_1)$.
Step 2: Then, in the reference image, around the same coordinates as that given pixel, we consider a circular neighborhood of radius r_2, denoted as $O(x, y, r_2)$, where, $r_2 \geqslant r_1$. For any pixel (x_1, y_1) in that neighborhood $O(x, y, r_2)$, we consider a circular neighborhood of radius r_1 around (x_1, y_1), denoted as $O(x_1, y_1, r_1)$.
Step 3: Next, we compare the two neighborhoods, i.e., $O(x, y, r_1)$ of zoomed image, and all possible choices of $O(x_1, y_1, r_1)$ corresponding to the reference image on the basis of three measures: mean square difference (MSD), mean absolute difference (MAD) and correlation (CC) and choose the neighborhood $O(x_1, y_1, r_1)$ of the reference image which is most similar to the neighborhood $O(x, y, r_1)$ of the zoomed image with respect to these three measures. Subsequently we map the pixel (x, y) of the zoomed image to that (x_1, y_1) of the reference image. These three measures are defined below:

$$\mathrm{MSD}((x_1, y_1); (x, y)) = \frac{1}{N}\sum (I_R(x_1 + s, y_1 + t) - I_Z(x + s, y + t))^2,$$

$$\mathrm{MAD}((x_1, y_1); (x, y)) = \frac{1}{N}\sum |I_R(x_1 + s, y_1 + t) - I_Z(x + s, y + t)|,$$

$$\mathrm{CC}((x_1, y_1); (x, y)) = \frac{\sum (I_R - \bar{I}_R)(I_Z - \bar{I}_Z)}{\sqrt{\sum (I_R - \bar{I}_R)^2}\sqrt{\sum (I_Z - \bar{I}_Z)^2}},$$

where all the above summations are taken over the set $\{(s, t) : (x_1 + s, y_1 + t) \in O(x_1, y_1, r_1)\}$, N is the number of pixels in $O(x_1, y_1, r_1)$, $I_{(.)}(a, b)$ denotes the intensity of the image pixel (a,b), R denotes the reference image, and Z denotes the zoomed image. Therefore, the transformed or mapped pixels are as follows:

$$\underset{(x_1, y_1) \in O(x, y, r_2)}{\arg\min} \mathrm{MSD}((x_1, y_1); (x, y)),$$

$$\underset{(x_1, y_1) \in O(x, y, r_2)}{\arg\min} \mathrm{MAD}((x_1, y_1); (x, y)),$$

$$\underset{(x_1,y_1)\in O(x,y,r_2)}{\arg\max}\ \mathrm{CC}((x_1, y_1); (x, y)),$$

corresponding to the measures MSD, MAD and CC, respectively. There are two tuning parameters of this procedure which are radii r_1 and r_2 of the neighborhoods. Performance of this procedure depends highly on the selection of these parameters. In the next section, we address the selection process of these parameters.

2.4 Selection of the Procedure Parameters

In the procedure described above, r_1 is the radius of the neighborhood $O(x, y, r_1)$ corresponding to some pixel (x, y) of the zoomed image and r_2 is the radius of the neighborhood $O(x, y, r_2)$ corresponding to the same pixel of the reference image. We use trial and error method to select r_1 and r_2. We find the mapping for some possible choices of r_1 and r_2 and create the registered version of the zoomed image. After that we compare the similarity between the reference image and the registered version of the zoomed image on the basis of MSD. Then we select those values of r_1 and r_2 for which $\mathrm{MSD}(Z, R, r_1, r_2)$ is minimized, i.e.,

$$(r_1, r_2) = \arg\min_{\widetilde{r}_1,\widetilde{r}_2} \sum_{(x,y)} (I_G(x, y, \widetilde{r}_1, \widetilde{r}_2) - I_R(x, y))^2,$$

where $I_G(x, y, \widetilde{r}_1, \widetilde{r}_2)$ is the intensity of the pixel (x, y) of registered version of the zoomed image, and $\widetilde{r}_1$ and $\widetilde{r}_2$ are the parameters used for the registration procedure.

3 Theoretical Justification

Proposition. *Let (x, y) be an edge pixel of an image of resolution $n \times n$ and there is a zoomed version of the same image. In the zoomed image, the pixel (x, y) has moved to (x_1, y_1), where $x_1 = x + c_1$ and $y_1 = y + c_2$, $(c_1, c_2) \in \mathbb{R}^2$. When we apply our feature mapping procedure with window widths r_1 and r_2 to register these two images, then (x_1, y_1) in the zoomed image is being mapped to (x_2, y_2) in the registered image. If $n \to \infty$ and $r_2 \to \infty$ such that $r_2/n \to 0$, then*

$$\frac{1}{n^2}[(x_2 - x)^2 + (y_2 - y)^2] \xrightarrow{P} 0,$$

where $\xrightarrow{P}$ denotes convergence in probability.

Proof. Fix any small $\varepsilon > 0$. Using Markov's inequality,

$$P\left[\frac{1}{n^2}\{(x_2 - x)^2 + (y_2 - y)^2\} > \varepsilon\right] \leqslant \frac{1}{\varepsilon} E\left[\frac{1}{n^2}\{(x_2 - x)^2 + (y_2 - y)^2\}\right].$$

Therefore,

$$\frac{E(x_2 - x)^2}{n^2\varepsilon} = \frac{E[x_2 - E(x_2) + E(x_2) - x]^2}{n^2\varepsilon} = \frac{V(x_2) + (E(x_2) - x)^2}{n^2\varepsilon}. \tag{1}$$

Next, we find upper bounds for $V(x_2)$ and $(E(x_2) - x)^2$. We first consider two scenarios where $V(x_2)$ and $(E(x_2) - x)^2$ are extreme. Let, in the neighborhood of width r_2, x_2 has realizations $(x_1 - r_2)$, $(x_1 - r_2 + 1)$, …, $(x_1 - 1)$, x_1, $(x_1 + 1)$, …, $(x_1 + r_2 - 1)$, $(x_1 + r_2)$. In the first scenario, $(2r_2 + 1)$ realizations are distributed as follows. Here, r_2 realizations are at $(x_1 - r_2)$, 1 realization

at x_1 and r_2 realizations at $(x_1 + r_2)$. In this case, $E(x_2) = x_1$ and $V(x_2) = 2{r_2}^3/(2r_2 + 1)$. In the second scenario, all $(2r_2 + 1)$ realizations are at $(x_1 + r_2)$. Therefore, in this case, $E(x_2) = (2r_2 + 1)(x_1 + r_2)/(2r_2 + 1) = (x_1 + r_2)$. Hence, from Eqn. (1), under the given conditions,

$$\frac{V(x_2) + (E(x_2) - x)^2}{n^2\varepsilon} \leqslant \frac{2{r_2}^3}{(2r_2 + 1)n^2\varepsilon} + \frac{(x_1 + r_2 - (x_1 - c_1))^2}{n^2\varepsilon} \to 0.$$

Similar result can be proved for $E(y_2 - y)^2/(n^2\varepsilon)$. Hence the proposition is proved. □

4 Numerical Studies

In order to apply this method we simulate an image of resolution 64×64 and the zoomed version of the same image where the resolutions of the image and zoomed version of the image are equal. In the simulated reference image, the central square region is of resolution 42×42 and intensities of its pixels are all 1, and intensities of all other image pixels are 0. Then, in order to complete the construction of the simulated image, we add a Gaussian noise with mean value 0 and standard deviation 0.01 to the intensities of each pixel. In the zoomed image, the resolution of the central region is 52×52. The resolutions of the pixels in the region covering the central region are all 0 as in the simulated reference image. Finally, we add a Gaussian noise of mean value 0 and standard deviation 0.01 to each pixel of the zoomed image.

Then for each combination of the values of r_1 and r_2, we run the proposed registration procedure for 10 times and find out the sample mean and sample standard deviation of the mean squared error (MSE) values corresponding to the similarity measures MSD, MAD and CC. In order to choose optimal values of r_1 and r_2, we run this procedure for 7 choices of r_2 viz. 3 to 9 while keeping r_1 at 3. Here we see that, for these choices of r_1 and r_2 minimum value of MSE appears at $(3, 5)$ for MSD, MAD and CC measure. So, we anticipate that the optimal choice of r_2 is 5. Then, in order to find the optimal choice of r_1, we run this procedure for 4 choices of r_1 viz. 2 to 5 when r_2 is at 5. Finally, we choose that combination of r_1 and r_2 for which the calculated MSE is the smallest. We present those MSE values in the Table 1.

We see from Table 1 that the optimal choices of r_1 and r_2 are 2 and 5 for registration procedure using MSD and MAD measures. And the optimal choices for r_1 and r_2 for CC method are 5 and 5, respectively. Therefore, using these optimal parameters for the specified methods, we run the algorithm and create the registered images. These registered images thus created contain some disjoint pixels (very small in number) where there are no mapping found according to this registration procedure and as a result of that intensity of these pixels remain unavailable. Then we take a neighborhood of width 2 around these pixels and perform local smoothing. Finally, we get the registered images corresponding to different measures. We follow the same process for the rest of the examples mentioned in this paper. These images are shown in the second row of Figure 3.

In order to check the usefulness of these methods, we compare registered images under all three methods with the reference image and create three matrices named 'anomaly' under respective methods. We compare each pixel of registered image with corresponding pixel of the original (reference) image and on the basis of that comparison we put some value to the respective location of anomaly matrices. Here inputs of the anomaly matrices are obtained as below. If absolute difference of image intensities of registered and the original image is less than 15 percent of range of intensities of zoomed image, then anomaly matrix will get value 0 at that location and will get value 1 elsewhere. Then from these anomaly matrices we create three

Table 1: Table for comparing the MSD values for the simulated image. The rationale behind presenting only this set of (r_1, r_2) values is provided in the second paragraph of Section 4.

	Mean (standard deviation)		
(r_1, r_2)	L_1 norm	L_2 norm	Correlation
(3,3)	0.3327	0.3362	0.4724
	(0.0013)	(0.0012)	(0.0035)
(3,4)	0.2337	0.2339	0.4026
	(0.0015)	(0.0011)	(0.0046)
(3,5)	0.1016	0.1004	0.314
	(0.0042)	(0.0029)	(0.0073)
(3,6)	0.1111	0.1126	0.3598
	(0.0046)	(0.0038)	(0.0031)
(3,7)	0.1192	0.1186	0.3673
	(0.002)	(0.0034)	(0.0043)
(3,8)	0.1198	0.122	0.3714
	(0.0036)	(0.0035)	(0.0121)
(3,9)	0.1255	0.1305	0.4081
	(0.0017)	(0.0019)	(0.0053)
(3,10)	0.1332	0.1387	0.4136
	(0.006)	(0.005)	(0.0046)
(1,5)	0.105	0.1087	0.4069
	(0.0035)	(0.0032)	(0.0046)
(2,5)	**0.0827**	**0.0809**	0.3691
	(0.0024)	**(0.0021)**	(0.0056)
(4,5)	0.1223	0.1201	0.2391
	(0.0013)	(0.0018)	(0.0035)
(5,5)	0.1408	0.1385	**0.1636**
	(0.0023)	(0.0012)	**(0.0029)**

anomaly images under these registration procedures. These images are presented in the third row of Figure 3.

Numerical studies on the 'girl' image of resolution 128×128: For our next example, we consider a 'girl' image of resolution 128×128 and select its central region of resolution 116×116. We then zoom it while keeping the resolution same as the original one. We show these two images in the first row of Figure 4.

Next, we add a Gaussian noise of mean value 0 and standard deviation 0.01 to each pixel of the two images before applying the method. We add this noise because while applying this method we found out that in some neighborhoods we get standard deviation values as 0 and as a result correlation value is becoming indefinite. Then for each combination of the values of r_1 and r_2, we run the proposed registration procedure for 10 times and find out the sample mean and sample standard deviation of the mean squared error (MSE) values corresponding to the similarity measures MSD, MAD and CC. In order to choose optimal values of r_1 and r_2 we run this procedure for 8 choices of r_2 viz. 3 to 10 while keeping r_1 at 3 and find out the sample mean

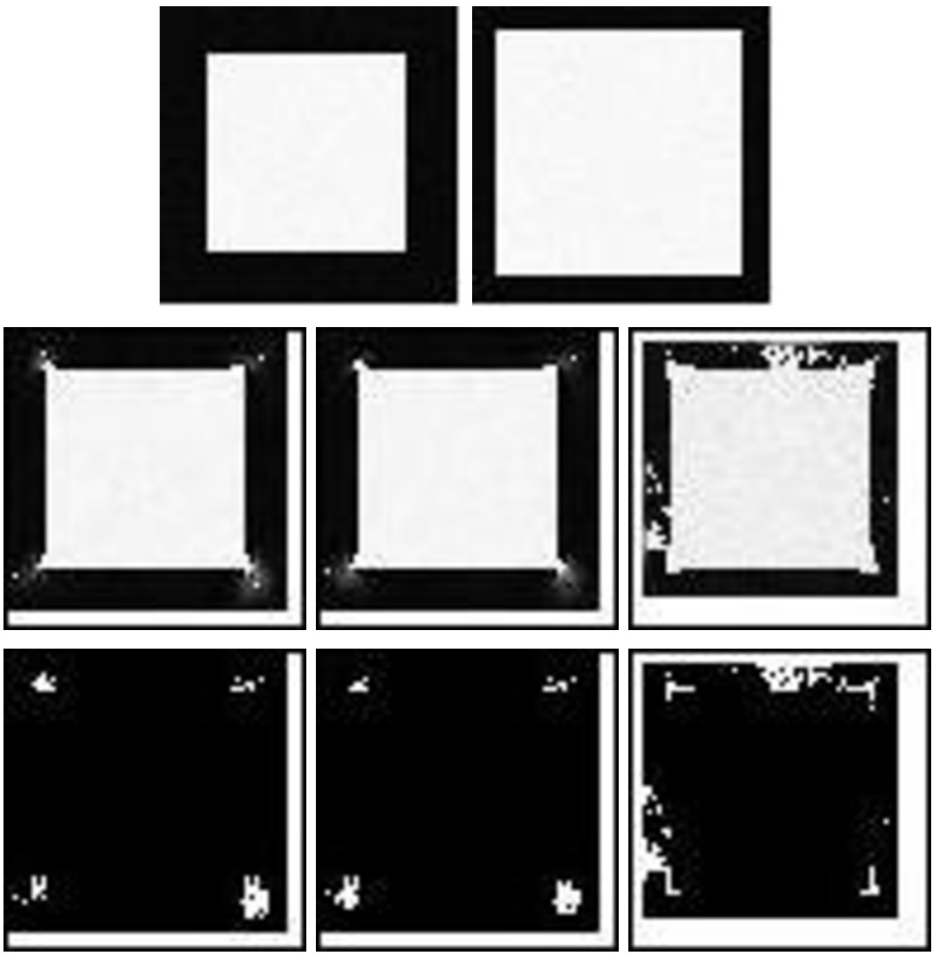


Figure 3: These figures are regarding the simulated image. Row 1 (from left to right): the reference image and the zoomed image; row 2 (from left to right): registered images under L_1-norm, L_2-norm and CC method, respectively; row 3 (from left to right): anomaly images under L_1-norm, L_2-norm and CC method, respectively.

and sample standard deviation of MSE values corresponding to the measures MSD, MAD and CC. Here we see that, for these choices of r_1 and r_2 minimum value of MSE appears at $(3, 6)$ for MSD and MAD measures and at $(3, 5)$ for CC measure. So, we anticipate that the optimal choice of r_2 is either 5 or 6. Then, in order to find the optimal choice of r_1, we run this procedure for 5 choices of r_1 viz. 2 to 6 when r_2 is 6 and for 4 choices of r_1 viz. 2 to 5 when r_2 is 5. Finally, we choose that value of r_1 and r_2 for which the calculated MSE value is the smallest. We present those MSE values in Table 2.

In this example, we see from Table 2 that the optimal choices of r_1 and r_2 are 2 and 6 for registration procedure using MSD and MAD measures and optimal choices for r_1 and r_2 for registration procedure using CC method are 5 and 6, respectively. therefore, using these optimal parameters for the specified methods, we run the algorithm and create the registered images. Those images are placed side by side in the second row of Figure 4.

Next, we compare the original image with each of the registered images and create three anomaly matrices under three registration procedures. Here inputs of the anomaly matrices are

Table 2: Table for comparing the MSD values for the 'girl' image. The rationale behind presenting only this set of (r_1, r_2) values is provided in the second paragraph of the numerical studies on the 'girl' image.

	Mean (standard deviation)		
(r_1, r_2)	L_1 norm	L_2 norm	Correlation
(3,3)	0.1172	0.1164	0.1256
	(0.0026)	(0.0023)	(0.0023)
(3,4)	0.1091	0.109	0.1244
	(0.0027)	(0.0031)	(0.0031)
(3,5)	0.0988	0.0995	0.1221
	(0.0022)	(0.002)	(0.002)
(3,6)	0.0972	0.0965	0.1238
	(0.0022)	(0.0021)	(0.0021)
(3,7)	0.1027	0.102	0.1249
	(0.002)	(0.002)	(0.002)
(3,8)	0.1064	0.1061	0.1311
	(0.0021)	(0.0017)	(0.0017)
(3,9)	0.1143	0.1133	0.1385
	(0.002)	(0.0017)	(0.0017)
(3,10)	0.1187	0.1175	0.1472
	(0.0017)	(0.0013)	(0.0013)
(2,6)	**0.0742**	**0.0753**	0.1189
	(0.0029)	**(0.0031)**	(0.0043)
(3,6)	0.0826	0.0835	0.1087
	(0.0029)	(0.0025)	(0.0023)
(4,6)	0.0921	0.0924	0.1032
	(0.002)	(0.0019)	(0.0029)
(5,6)	0.1009	0.1006	**0.1025**
	(0.0026)	(0.0025)	**(0.0026)**
(6,6)	0.1073	0.1067	0.1071
	(0.0021)	(0.0022)	(0.0022)
(2,5)	0.084	0.0863	0.1256
	(0.0029)	(0.0026)	(0.0034)
(4,5)	0.1004	0.1003	0.1094
	(0.002)	(0.0022)	(0.0019)
(5,5)	0.1074	0.107	0.1119
	(0.0025)	(0.0023)	(0.0028)

obtained as below. If absolute difference of image intensities of registered and reference images is less than 15 percent of range of intensities of zoomed image, then anomaly matrix will get value 0 at that location and will get value 1 elsewhere. From these anomaly matrices, we create three anomaly images under these registration procedures. The images are presented in the third row of Figure 4.

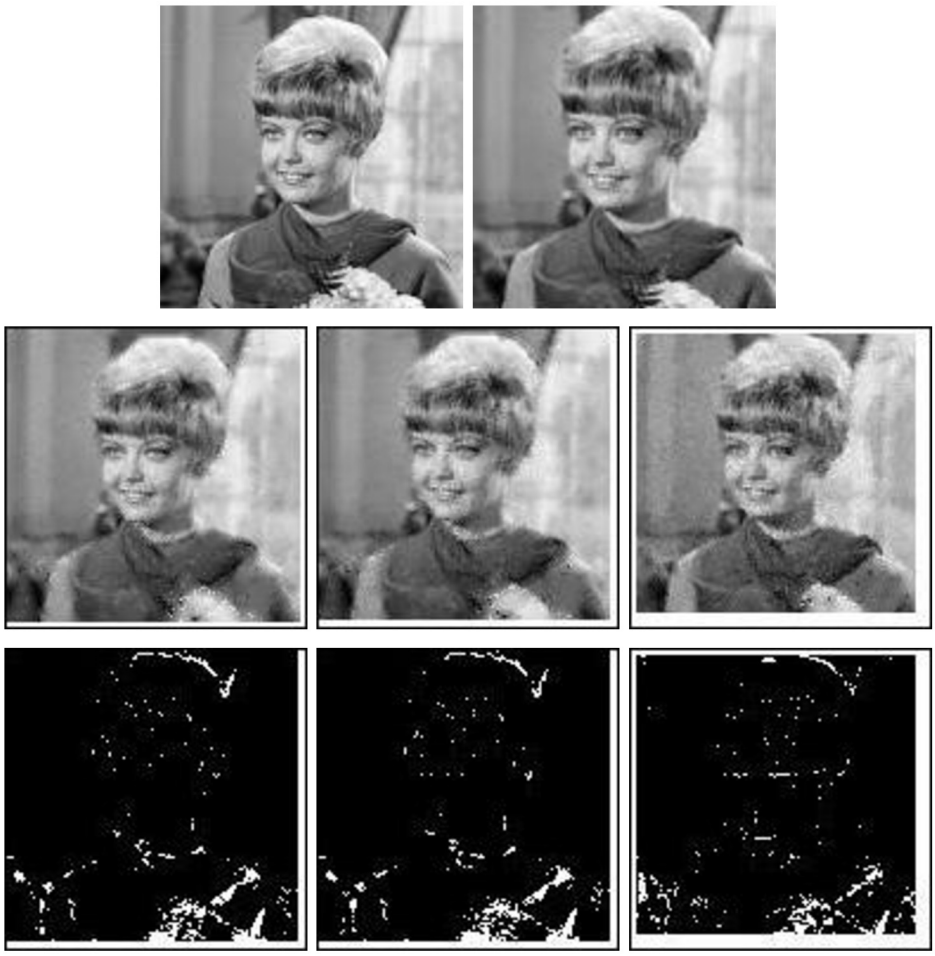


Figure 4: These figures are regarding the 'girl' image. Row 1 (from left to right): the reference image and the zoomed image; row 2 (from left to right): registered images under L_1-norm, L_2-norm and CC method, respectively; row 3 (from left to right): anomaly images under L_1-norm, L_2-norm and CC method, respectively.

Numerical studies on the 'peppers' image of resolution 256×256: For our next example, we consider the 'peppers' image of resolution 256×256 and select its central region of resolution 246×246 and then zoom it while keeping the resolution same as the original one. We show these two images in first row of Figure 5.

Next, for each combination of the values of r_1 and r_2, we run the proposed registration procedure for 10 times and find out the sample mean and sample standard deviation of the mean squared error (MSE) values corresponding to the similarity measures MSD, MAD and CC. In order to choose optimal values of r_1 and r_2, we run this procedure for 7 choices of r_2 viz. 3 to 9 while keeping r_1 at 3 and find out the sample mean and sample standard deviation of MSE values corresponding to the measures MSD, MAD and CC. Here we see that, for these choices of r_1 and r_2 minimum value of MSE appears at $(3, 6)$ for MSD, MAD and CC measure. So, we anticipate that the optimal choice of r_2 is 6. Then, in order to find the optimal choice of r_1, we run this procedure for 5 choices of r_1 viz. 2 to 6 when r_2 is 6. Finally, we choose that

Table 3: Table for comparing the MSD values for the 'peppers' image. The rationale behind presenting only this set of (r_1, r_2) values is provided in the second paragraph of the numerical studies on the 'peppers' image.

	Mean (standard deviation)		
(r_1, r_2)	L_1 norm	L_2 norm	Correlation
(3,3)	0.104	0.1047	0.1211
	(0.0023)	(0.0023)	(0.0019)
(3,4)	0.0927	0.0932	0.1165
	(0.0021)	(0.0021)	(0.0022)
(3,5)	0.122	0.121	0.1453
	(0.0017)	(0.0018)	(0.0015)
(3,6)	0.0691	0.0685	0.0952
	(0.003)	(0.0029)	(0.0031)
(3,7)	0.0907	0.0901	0.1202
	(0.002)	(0.0021)	(0.0016)
(3,8)	0.0966	0.0958	0.1296
	(0.002)	(0.002)	(0.002)
(3,9)	0.1025	0.1013	0.1378
	(0.002)	(0.0015)	(0.0017)
(2,6)	**0.0628**	**0.0626**	0.1032
	(0.0029)	**(0.0024)**	(0.0059)
(4,6)	0.0774	0.0773	**0.0939**
	(0.0029)	(0.0028)	**(0.0026)**
(5,6)	0.0873	0.0872	0.0965
	(0.0025)	(0.0025)	(0.0025)
(6,6)	0.0965	0.0964	0.1018
	(0.0025)	(0.0025)	(0.0024)

value of r_1 and r_2 for which the calculated MSE value is the smallest. We present those MSE values in Table 3.

In this example, we see from the Table 3 that the optimal choices of r_1 and r_2 are 2 and 5 for registration procedure using MSD and MAD measures and optimal choices for r_1 and r_2 for CC method are 4 and 5, respectively. Therefore, using these optimal parameters for the specified methods, we run the algorithm and create the registered images. Those images are placed side by side in the second row of Figure 5.

Next, we compare the original image with each of the registered images and create three anomaly matrices under three registration procedures. Here inputs of the anomaly matrices are obtained as below. If absolute difference of image intensities of registered and reference image is less than 15 percent of range of intensities of zoomed image, then anomaly matrix will get value 0 at that location and will get value 1 elsewhere. Then from these anomaly matrices we create three anomaly images under these registration procedures. These images are presented in the third row of Figure 5.

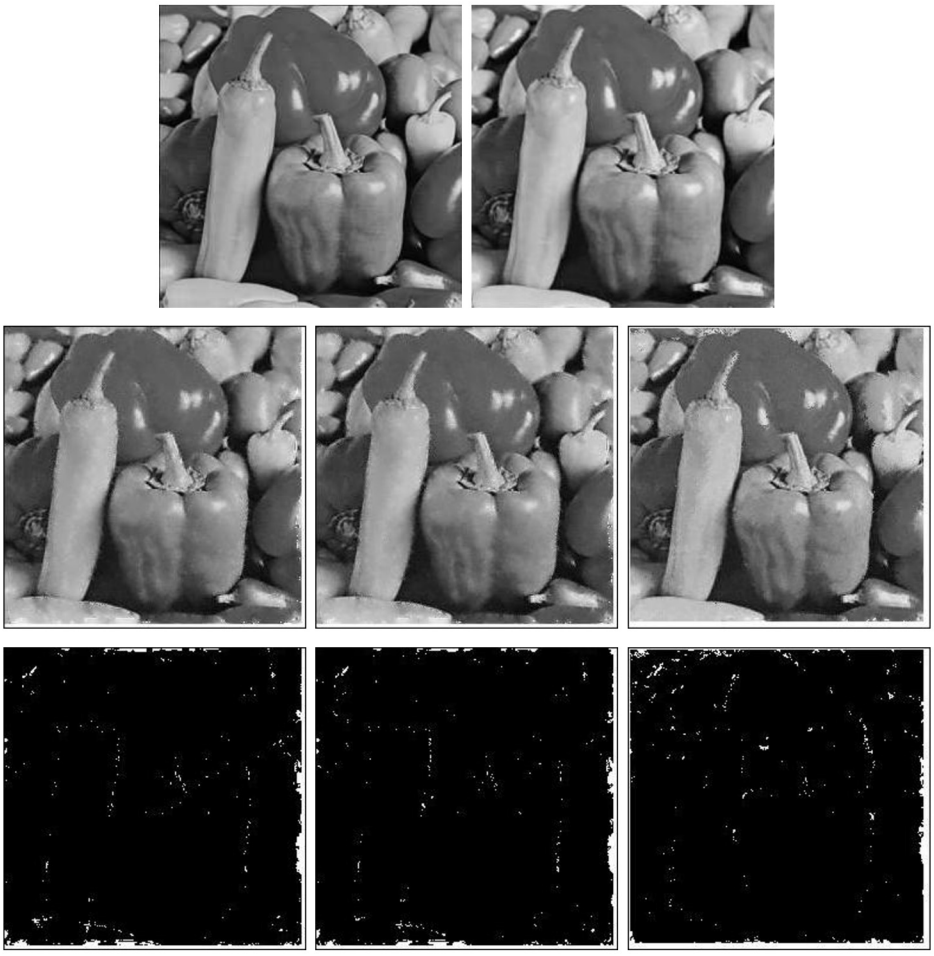


Figure 5: These figures are regarding the 'peppers' image. Row 1 (from left to right): the reference image and the zoomed image; row 2 (from left to right): registered images under L_1-norm, L_2-norm and CC method, respectively; row 3 (from left to right): anomaly images under L_1-norm, L_2-norm and CC method, respectively.

This method might not work satisfactorily in cases where images have too much uniformity in a neighborhood. Then this method might fail to distinguish between two similar pixels and might end up overlapping one with other. This tile image and its zoomed version, presented in the first row of Figure 6, are one such example where this method might not work well.

In this example, we use the above method and find out that the optimal choices of r_1 and c are 2 and 2, respectively for the registration procedure using MAD, MSD and CC measure. Therefore, using these optimal parameters for these methods, we run the algorithm and create the registered images. Those images are placed side by side in the second row of Figure 6.

In this example also, we compare the original image with each of the registered images and create three anomaly matrices under three registration procedures. Here inputs of the anomaly matrices are obtained as below. If the absolute difference of image intensities of registered and reference image is less than 15 percent of range of intensities of zoomed image, then anomaly matrix will get value 0 at that location and will get value 1 elsewhere. Then from these anomaly

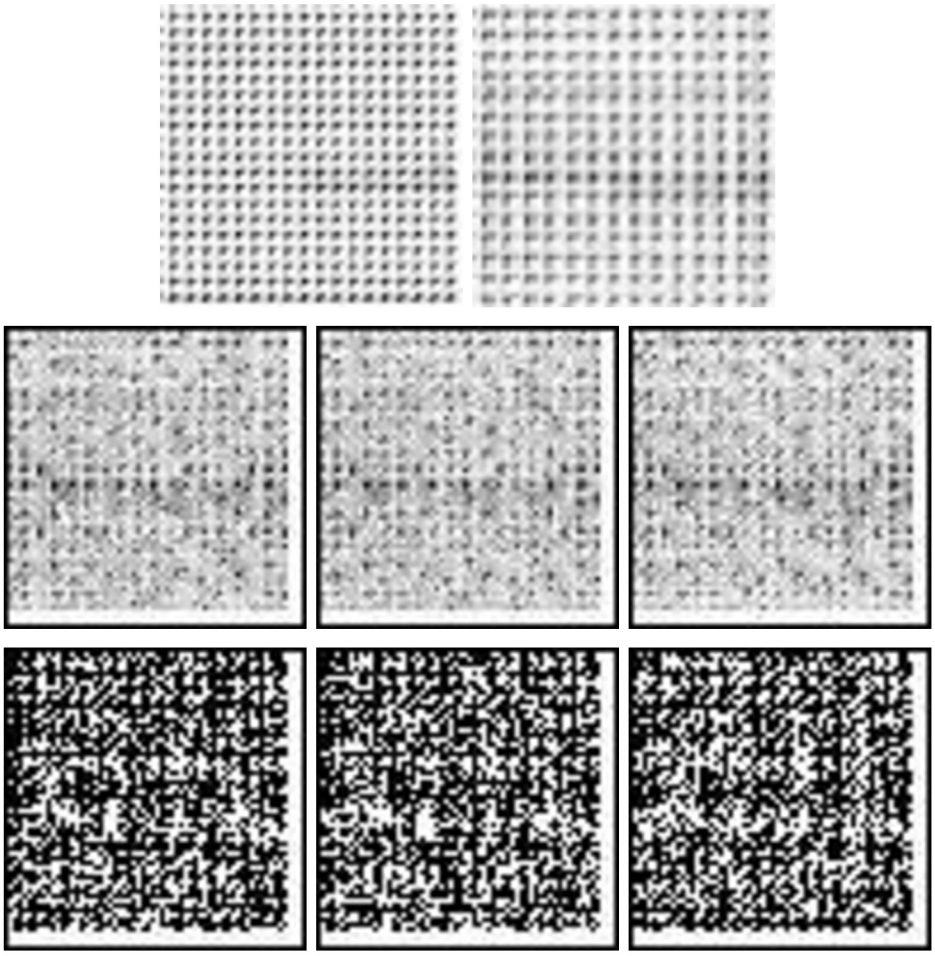


Figure 6: These figures are for the 'tile' image. Row 1 (from left to right): the reference image and the zoomed image; row 2 (from left to right): registered images under L_1-norm, L_2-norm and CC method, respectively; row 3 (from left to right): anomaly images under L_1-norm, L_2-norm and CC method, respectively.

matrices we create three anomaly images under these registration procedures. These images are presented in the third row of Figure 6.

5 Comparison with State-of-the-Art Methods

In this section, we compare our method with a number of state-of-the-art image registration techniques. These methods are based on certain feature matching algorithms. There are several popular feature matching algorithms in the literature, viz., Scale Invariant Feature Transform (SIFT) (Lowe, 2004), Speeded Up Robust Features (SURF) (Bay et al., 2008), Oriented Fast and Rotated Brief (ORB) (Rublee et al., 2011), Binary Robust Invariant Scalable points (BRISK) (Leutenegger et al., 2011), KAZE (Alcantarilla et al., 2012), Accelerated-KAZE (AKAZE) (Alcantarilla and Solutions, 2011) to name a few. These methods work on the principle of detection and description of the features in the reference and transformed images. Then the relationship

Table 4: Mean squared error (MSE) for different methods.

	Mean Squared Error (MSE)			
	L_1-norm	ORB	AKAZE	BRISK
simulated	**0.0827**	NA	NA	NA
girl	0.0742	0.0087	0.0318	**0.0049**
peppers	0.0628	0.0053	0.0238	**0.0037**
mountain	**0.0479**	0.0925	NA	NA

between these features from reference and transformed images are found using some algorithm e.g., RANSAC (Derpanis, 2010) and registered images are generated. In this paper, we take ORB, AKAZE and BRISK among these methods to compare with our best performing algorithm i.e., registration using L_1-norm. We first apply these methods on some images and create the registered images. Then we compute the MSE between reference image and the registered images obtained with respect to these methods and provide a table for comparison. We also provide the reference image, zoomed image and the registered images side by side to facilitate comparison between these methods.

Here we consider four images viz., simulated image, girl image, peppers image and mountain image for the comparison purpose. The reason for taking the mountain image is that, this image has very less features. As these methods (ORB, AKAZE and BRISK) are all feature based, we want to compare their performance with our intensity based approach on the basis of this image as well. Next, we present the table containing MSE values computed for different methods.

From Table 4 we see that, in case of simulated image, the feature-based methods are failing miserably as the number of features in the simulated image is sparse. It is to be noted that the number of features detected by these competing methods are small in number. The feature based methods (especially BRISK) outperform our proposed method in case of the 'girl' and 'peppers' images because there are numerous features in these images. Again because of sparsity of the features according to their algorithms, the feature-based methods perform poorly in case of the 'mountain' image. In this case, our proposed method is the clear winner. Note that the competing feature-based methods detect a small number of features according to their algorithms. From Figure 7 we see that, in case of simulated image the feature-based methods are unable to generate the registered images, whereas our proposed method is performing reasonably. In case of the 'girl' and 'peppers' images, the feature-based methods are clear winners. Finally, in case of the 'mountain' image, AKAZE and BRISK are not being able to generate the registered image due to scarcity of features. On the other hand, the ORB method generates a distorted registered image which is not comparable with the reference image. In this case also, our proposed method performs reasonably well and clearly outperforms some of the well known feature-based methods.

In addition to Table 4, we also provide the registered images generated by these methods in the Figure 7. In each row of the following figure we place the images in the order: reference image, zoomed image, registered image for registration using L_1-norm, registered image for ORB, registered image for AKAZE and registered image for BRISK. In the first row we provide images corresponding to simulated image and in the following rows we place the images for girl, peppers and mountain image respectively.

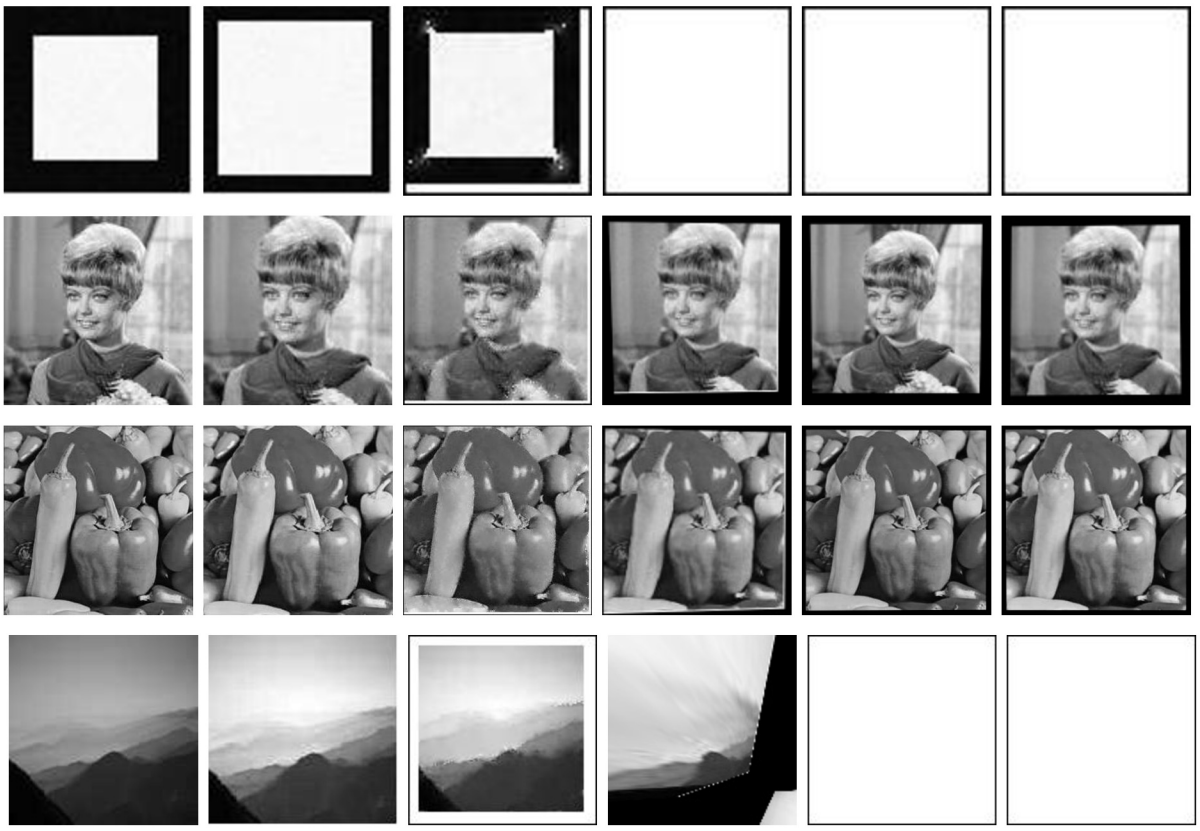


Figure 7: The reference image, the zoomed image, and the registered images under the proposed method, ORB, AKAZE and BRISK.

6 Concluding Remarks

In this paper, we propose a method to register an image to its zoomed counterpart. This method works reasonably in many cases described in this paper. As the resolution of the image increases, the performance of this method also improves, as justified in Section 3. In Section 4, the numerical performance of the proposed method on three different images of various resolutions also supports the statement. There are, however, certain drawbacks of this method which are discussed in the latter part of Section 4. In images where the number of features are small, our intensity-based method outperforms many state-of-the-art feature-based registration methods. A more practical image registration problem to work on is where translation and rotation are also involved along with zooming. Much future research is needed in this direction.

Supplementary Material

The supplementary materials contain the codes and relevant data.

References

Alcantarilla PF, Bartoli A, Davison AJ (2012). Kaze features. In: *Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VI 12*, 214–227. Springer.

Alcantarilla PF, Solutions T (2011). Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Transanctions on Pattern Analysis and Machine Intelligence*, 34(7): 1281–1298.

Alpert N, Bradshaw J, Kennedy D, Correia J (1990). The principal axes transformation—a method for image registration. *Journal of Nuclear Medicine*, 31(10): 1717–1722.

Ashburner J, Friston KJ (2007). Rigid body registration. In: *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, 49–62.

Bay H, Ess A, Tuytelaars T, Van Gool L (2008). Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3): 346–359. https://doi.org/10.1016/j.cviu.2007.09.014

Brown LG (1992). A survey of image registration techniques. *ACM Computing Surveys (CSUR)*, 24(4): 325–376. https://doi.org/10.1145/146370.146374

Crum WR, Hartkens T, Hill D (2004). Non-rigid image registration: Theory and practice. *The British Journal of Radiology*, 77(suppl_2): S140–S153. https://doi.org/10.1259/bjr/25329214

D'agostino E, Maes F, Vandermeulen D, Suetens P (2003). A viscous fluid model for multimodal non-rigid image registration using mutual information. *Medical Image Analysis*, 7(4): 565–575. https://doi.org/10.1016/S1361-8415(03)00039-2

Das S, Roy A, Mukherjee PS (2024). Image registration for zooming: A statistically consistent local feature mapping approach. *Stat*, 13(1): e664. https://doi.org/10.1002/sta4.664

De Castro E, Morandi C (1987). Registration of translated and rotated images using finite Fourier transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI–9(5): 700–703.

Derpanis KG (2010). Overview of the ransac algorithm. *Image Rochester NY*, 4(1): 2–3.

Eddyy WF, Fitzgerald M, Noll DC (1996). Improved image registration by using Fourier interpolation. *Magnetic Resonance in Medicine*, 36(6): 923–931. https://doi.org/10.1002/mrm.1910360615

Gaens T, Maes F, Vandermeulen D, Suetens P (1998). Non-rigid multimodal image registration using mutual information. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 1099–1106. Springer.

Hill DL, Batchelor PG, Holden M, Hawkes DJ (2001). Medical image registration. *Physics in Medicine and Biology*, 46(3): R1. https://doi.org/10.1088/0031-9155/46/3/201

Leutenegger S, Chli M, Siegwart RY (2011). Brisk: Binary robust invariant scalable keypoints. In: *2011 International Conference on Computer Vision*, 2548–2555. Ieee.

Lowe DG (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60: 91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Matabosch C, Salvi J, Fofi D, Meriaudeau F (2005). Range image registration for industrial inspection. In: *Machine Vision Applications in Industrial Inspection XIII*, volume 5679, 216–227. SPIE.

Oliveira FP, Tavares JMR (2014). Medical image registration: A review. *Computer Methods in Biomechanics and Biomedical Engineering*, 17(2): 73–93. https://doi.org/10.1080/10255842.2012.670855

Pataky TC, Goulermas JY, Crompton RH (2008). A comparison of seven methods of within-subjects rigid-body pedobarographic image registration. *Journal of Biomechanics*, 41(14):

3085–3089. https://doi.org/10.1016/j.jbiomech.2008.08.001

Pluim JP, Maintz JA, Viergever MA (2003). Mutual-information-based registration of medical images: A survey. *IEEE Transactions on Medical Imaging*, 22(8): 986–1004. https://doi.org/10.1109/TMI.2003.815867

Qiu P, Xing C (2013). Feature based image registration using non-degenerate pixels. *Signal Processing*, 93(4): 706–720. https://doi.org/10.1016/j.sigpro.2012.09.013

Reddy BS, Chatterji BN (1996). An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 5(8): 1266–1271. https://doi.org/10.1109/83.506761

Rublee E, Rabaud V, Konolige K, Bradski G (2011). Orb: An efficient alternative to sift or surf. In: *2011 International Conference on Computer Vision*, 2564–2571. Ieee.

Sotiras A, Davatzikos C, Paragios N (2013). Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging*, 32(7): 1153–1190. https://doi.org/10.1109/TMI.2013.2265603

Špiclin Ž, Bukovec M, Pernuš F, Likar B (2011). Image registration for visual inspection of imprinted pharmaceutical tablets. *Machine Vision and Applications*, 22: 197–206. https://doi.org/10.1007/s00138-007-0104-0

Wolberg G, Zokai S (2000). Robust image registration using log-polar transform. In: *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*, volume 1, 493–496. IEEE.

Wyawahare MV, Patil PM, Abhyankar HK, et al. (2009). Image registration techniques: An overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2(3): 11–28.

Xing C, Qiu P (2011). Intensity-based image registration by nonparametric local smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(10): 2081–2092. https://doi.org/10.1109/TPAMI.2011.26

Zitova B, Flusser J (2003). Image registration methods: A survey. *Image and Vision Computing*, 21(11): 977–1000. https://doi.org/10.1016/S0262-8856(03)00137-9