Feature Based Image Registration Using Non-Degenerate Pixels

Peihua Qiu and Chen Xing

School of Statistics

University of Minnesota

Minneapolis, MN 55455

(qiuxx008@umn.edu, xingx011@umn.edu)

September 17, 2012

Abstract

Image registration (IR) aims to geometrically match one image to another. It is extensively used in many imaging applications. Among many existing IR methods, one widely used group of methods are feature-based. By a feature-based method, a number of relevant image features are first extracted from the two images, respectively, and then a geometric matching transformation is found to best match the two sets of features. However, proper identification and extraction of image features turns out to be a challenging task. Generally speaking, a good image feature extraction method should have the following two properties: (i) the identified image features should provide us proper information to approximate the geometric matching transformation accurately, and (ii) they should be easy to identify by a computer algorithm so that the entire feature extraction procedure is computer automatic. In this paper, a new type of image features is studied, which has the two properties described above. Together with the widely used thin plate spline (TPS) geometric transformation model, it is shown that our feature-based IR method works effectively in various cases.

Key Words: Degeneration, feature extraction, image matching, local smoothing, thin plate splines, registration.

1 Introduction

Image registration (IR) is a process of geometrically matching up two images of a same scene. It has become an important tool for improving the quality of certain image-based technologies [19], and is widely used in various applications, including medical imaging [13], remote sensing [15], face or finger print recognition [16], image compression [8], image matching [3], and so forth. mathematically, the image registration problem can be described as follows. Assume that $Z_R(x, y)$ and $Z_M(x, y)$ are two images to register, $Z_R(x, y)$ is a reference image, and $Z_M(x, y)$ is a geometrically altered version of $Z_R(x, y)$ and is often called a moved image. Then, the major purpose of image registration is to find a geometric transformation $\mathbf{T}(x, y) = (T_1(x, y), T_2(x, y))$ such that $Z_M(x, y)$ and $Z_R(\mathbf{T}(x, y))$ are as close as possible, which is often accomplished by solving the following maximization problem:

$$\max_{\mathbf{T}\in\mathcal{T}} S(Z_M, Z_R(\mathbf{T})),$$

where S is a metric to measure the similarity between $Z_M(x, y)$ and $Z_R(\mathbf{T}(x, y))$, and $\mathcal{T}$ is a given transformation family. See [24] and [37] for a more detailed description of the image registration problem.

To find the geometric transformation $\mathbf{T}$, many existing methods focus on matching up some major features in the two images, instead of the two entire images. These methods are called feature based methods in this paper. By a feature based method, a number of relevant image features are first extracted from the two images, respectively, and then the geometric transformation $\mathbf{T}$ is found to best match the two sets of features [33], [37]. More specifically, a feature based method usually consists of the following three steps:

Step 1 (*Feature Extraction*) Different features are identified and extracted manually or automatically by a computer algorithm from the two images, respectively. The extracted features can be represented by some descriptor vectors, consisting of coordinates, intensities, gradients, and so forth.

Step 2 (*Feature Matching*) The two sets of extracted features are matched up by a matching algorithm based on a similarity measure.

Step 3 (*Transformation Estimation*) Based on the matching results obtained in Step 2, an estimator of $\mathbf{T}(x, y)$, denoted as $\widehat{\mathbf{T}}(x, y)$, is obtained by a function estimation/interpolation procedure, such as the thin plate spline (TPS) function interpolation procedure.

While all the above three steps are important and challenging, the first step (i.e., feature identification and extraction) is a critical one, because the final registration results depend directly on what features being extracted and how the features being extracted. In the literature, there are two types of features discussed: non-point features and point features. Non-point features include curves, edges, line segments, and regions, which are usually determined by thresholding and segmentation [30],[21], whereas point features are just sets of individual points. Examples of point features include landmark points, control points, centroid points, corner points, and so forth. Note that most features described above are not specifically defined for image registration. For instance, most edge features are defined for edge detection, image segmentation, and some other image processing purposes. For this reason, they may not be appropriate for image registration.

Besides feature based IR methods, there are some IR methods based directly on the observed image intensities of the two images $Z_R(x, y)$ and $Z_M(x, y)$. See, e.g., [6], [32], [2], [1], [36], and [26] for related discussion. There are also a number of IR methods that try to overcome certain limitations of both the feature-based and the intensity-based methods. See, e.g., [35], [20].

Recently, Xing and Qiu [36] find that the image registration problem is *ill-posed* in the sense that the transformation $\mathbf{T}$ cannot be properly defined in certain regions, including the regions where the image intensity function is straight or where the image intensity function has a step edge but it is flat on both sides of the edge curve. Such pixels are called degenerate pixels in that paper. Based on that concept, Xing and Qiu proposed an intensity based IR method, which tried to estimate the transformation $\mathbf{T}$ based directly on the observed image intensities around all detected non-degenerate pixels. From that paper, it seems that when we try to handle the image registration problem, we should focus on all non-degenerate pixels where $\mathbf{T}$ is well defined. Because such non-degenerate pixels contain all useful information about the transformation $\mathbf{T}$, they are ideal feature points for image registration. In this paper, we try to further study the properties of the non-degenerate pixels as feature points for image registration. To this end, we consider different feature-based image registration procedures, using the non-degenerate pixels and some other feature points proposed in the literature as their feature points, respectively. To make the comparison of their performance as fair as possible, in all these procedures except the one mentioned below, the feature matching (i.e., Step 2 of a feature-based method described above) is accomplished by searching all the detected feature points in the two images, using the mean squared difference (MSD) dissimilarity metric. For estimation of the transformation $\mathbf{T}$ (i.e., Step 3 of a feature-based method), we use the popular thin plate spline (TPS) model, which was first introduced by Duchon [7] for function estimation and was applied to medical image processing by many authors, including Bookstein [4]. The TPS model is commonly used in function interpolation [34], and it can be used for describing various elastic deformations well.

The rest part of the article is organized as follows. In section 2, our feature based image registration procedure is described in detail. Various numerical examples are presented in Section 3. Some concluding remarks are given in Section 4.

2 Proposed Method

In this section, our proposed feature based IR method is described in three parts. Non-degenerate pixels are introduced in Section 2.1. The TPS model for estimating the transformation $\mathbf{T}$ is discussed in Section 2.2. The whole IR procedure is described in Section 2.3.

2.1 Non-Degenerate Pixels

Let $R(x, y)$ and $M(x, y)$ be the true image intensity functions of the reference and moved images, respectively. It is assumed that there is a geometric transformation $\mathbf{T}(x, y) = (T_1(x, y), T_2(x, y))$ such that

$$R(T_1(x, y), T_2(x, y)) = M(x, y), \text{ for } (x, y) \in \Omega_M,$$

where Ω_M is the design space of the moved image. For simplicity, we further assume that $\Omega_M = [0, 1] \times [0, 1]$. The major goal of image registration is to estimate $\mathbf{T}(x, y)$ from observed image intensities of the two images defined by

$$\begin{aligned} Z_M(x_i, y_j) &= M(x_i, y_j) + \varepsilon_M(x_i, y_j), \\ Z_R(x_i, y_j) &= R(x_i, y_j) + \varepsilon_R(x_i, y_j), \\ &\quad \text{for } i, j = 1, 2, \ldots, n, \end{aligned}$$

where $\{(x_i, y_j)\}$ are pixel locations, and $\{\varepsilon_M(x_i, y_j)\}$ and $\{\varepsilon_R(x_i, y_j)\}$ are i.i.d. random errors in the two images with mean 0 and variances σ_M^2 and σ_R^2, respectively. To this end, we write

$$\mathbf{T}(x, y) = (x, y) + (b(x, y), c(x, y)),$$

where $b(x, y) = T_1(x, y) - x$ and $c(x, y) = T_2(x, y) - y$. Then, estimation of $\mathbf{T}(x, y)$ is equivalent to estimation of $(b(x, y), c(x, y))$.

In cases when both $b(x,y)$ and $c(x,y)$ are small and R has the first-order partial derivatives around (x,y), by the Taylor's expansion, we have

$$R(\mathbf{T}(x,y)) \approx R(x,y) + R'_x(x,y)b(x,y) + R'_y(x,y)c(x,y),$$

where "$\approx$" denotes equality after the high-order terms have been ignored on the right-hand side. Therefore, we can choose $(b(x,y), c(x,y))$ to minimize the residual

$$M(x,y) - \left[R(x,y) + R'_x(x,y)b(x,y) + R'_y(x,y)c(x,y)\right],$$

where we have used the relationship $R(\mathbf{T}(x,y)) = M(x,y)$. However, $M(x,y)$ and $R(x,y)$ are unobservable, and what we observed are the image intensities $\{Z_M(x_i,y_j)\}$ and $\{Z_R(x_i,y_j)\}$ which have random noise involved. To smooth out random noise while estimating $(b(x,y), c(x,y))$, we can replace $R'_x(x,y)$ and $R'_y(x,y)$ by their local linear kernel (LLK) estimators commonly used in statistical nonparametric regression [9], which are defined by

$$\begin{aligned}
\widehat{R}'_x(x,y) &= \frac{\sum_{i,j=1}^n (x_i - x) Z_R(x_i,y_j) K_{h_n}(x,y)}{\sum_{i,j=1}^n (x_i - x)^2 K_{h_n}(x,y)} \\
\widehat{R}'_y(x,y) &= \frac{\sum_{i,j=1}^n (y_j - y) Z_R(x_i,y_j) K_{h_n}(x,y)}{\sum_{i,j=1}^n (y_j - y)^2 K_{h_n}(x,y)},
\end{aligned}$$

where $K_{h_n}(x,y) = K(\frac{x_i - x}{h_n}, \frac{y_j - y}{h_n})$, K is a bivariate density kernel function with a unit circular support, and h_n is a bandwidth parameter. Then, consider the following weighted least square problem:

$$\min_{b(x,y),c(x,y)} \sum_{i,j=1}^n \{Z_M(x_i,y_j) - [Z_R(x_i,y_j) + \widehat{R}'_x(x_i,y_j)b(x,y) + \widehat{R}'_y(x_i,y_j)c(x,y)]\}^2 K_{h_n}(x,y). \quad (1)$$

Intuitively, by (1), $(b(x,y), c(x,y))$ are estimated by minimizing the weighted residual mean squares in a circular neighborhood of (x,y) with radius h_n, denoted as $O(x,y;h_n)$, and the weights are

controlled by the kernel function. The solution of (1) is easy to derive to be

$$\begin{pmatrix} \widehat{b}(x,y) \\ \\ \widehat{c}(x,y) \end{pmatrix} = \frac{\begin{pmatrix} K_{22}(x,y), & -K_{12}(x,y) \\ \\ -K_{12}(x,y), & K_{11}(x,y) \end{pmatrix} \begin{pmatrix} K_1^*(x,y) \\ \\ K_2^*(x,y) \end{pmatrix}}{K_{11}(x,y)K_{22}(x,y) - K_{12}^2(x,y)}, \tag{2}$$

where, for $s, t = 1, 2$,

$$K_{st}(x,y) = \sum_{i,j=1}^{n} \left[\widehat{R}'_x(x_i,y_j)\right]^s \left[\widehat{R}'_y(x_i,y_j)\right]^t K_{h_n}(x,y)$$

and

$$K_1^*(x,y) = \sum_{i,j=1}^{n} \left[Z_M(x_i,y_j) - Z_R(x_i,y_j)\right] \widehat{R}'_x(x_i,y_j) K_{h_n}(x,y),$$
$$K_2^*(x,y) = \sum_{i,j=1}^{n} \left[Z_M(x_i,y_j) - Z_R(x_i,y_j)\right] \widehat{R}'_y(x_i,y_j) K_{h_n}(x,y).$$

It is obvious that the estimators in (2) are not well defined when

$$K_{11}(x,y)K_{22}(x,y) - K_{12}^2(x,y) = 0. \tag{3}$$

Qiu and Xing [25] have proven that, if there is a continuously differentiable univariate function ψ and a constant ρ such that

$$R(x',y') = \psi(\rho x' + y'), \text{ for any } (x',y') \in O(x,y;h_n), \tag{4}$$

then (3) is asymptotically true. Intuitively, if R satisfies the condition (4), then it is actually a degenerate univariate function, in the sense that its value is a constant along the line segment $\rho x' + y' = \rho_0$ within the neighborhood where ρ_0 is a given number. For that reason, Qiu and Xing call all pixels that satisfy the condition (3) the *degenerate pixels* of the image R, and all other pixels the *non-degenerate pixels* of R. The degenerate pixels and non-degenerate pixels of the image M can be defined similarly. From this discussion, it can be seen that the geometric transformation $\mathbf{T}(x,y)$ is not well defined around a degenerate pixel. In that sense, the image registration problem is *ill-posed*.

2.2 The Thin Plate Spline Model

The thin plate spline (TPS) model is a basic function interpolation approach that has been studied extensively in the numerical mathematics literature [7], [4]. It has been applied to the area of curve matching and transformation [22]. For image registration, assume that we have successfully matched two sets of feature points in the two images $R(x, y)$ and $M(x, y)$. More specifically, assume that we have m feature points in the reference image $M(x, y)$, denoted as $(u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m)$. For each feature point (u_j, v_j) in $M(x, y)$, it is matched to the feature point $\widehat{\mathbf{T}}(u_j, v_j)$ in $R(x, y)$, for $j = 1, 2, \ldots, m$. Namely, for $j = 1, 2, \ldots, m$, we have

$$\widehat{\mathbf{T}}(u_j, v_j) = (u_j, v_j) + (\widehat{b}(u_j, v_j), \widehat{c}(u_j, v_j)). \tag{5}$$

The above expression defines $\widehat{\mathbf{T}}(x, y)$, or equivalently $(\widehat{b}(x, y), \widehat{c}(x, y))$, at the m feature points only. The major purpose of the TPS modeling is to do function interpolation, so that $(\widehat{b}(x, y), \widehat{c}(x, y))$ is defined in the entire design space Ω_M after function interpolation. By this approach, we assume that $(\widehat{b}(x, y), \widehat{c}(x, y))$ go through the two sets of feature points in the two images (i.e., the expression (5) holds), and both $\widehat{b}(x, y)$ and $\widehat{c}(x, y)$ are smooth. The smoothness of $\widehat{b}(x, y)$ is measured by the so-called *bending energy*, defined by

$$I_{\widehat{b}} = \int\int_{\Omega_R} \left[\left(\frac{\partial^2 \widehat{b}}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 \widehat{b}}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 \widehat{b}}{\partial y^2} \right)^2 \right] dxdy.$$

The interpolated function $\widehat{b}(x, y)$ defined in the entire design space Ω_M is then obtained by minimizing the above bending energy under the constraint of (5). Similarly, we can obtain the interpolated function $\widehat{c}(x, y)$ defined in the whole space Ω_M by minimizing its bending energy $I_{\widehat{c}}$. This process to obtain the TPS model explains the reason why it has its current name. Namely, the surface in the TPS model acts like a thin metal plate compressed by the bending energy to go through the m given points.

By some algebraic manipulations, the interpolated geometric transformation $(\widehat{b}(x,y), \widehat{c}(x,y))$ by the TPS approach has the following expression:

$$\begin{aligned}
\widehat{b}(x,y) &= a_{11} + a_{12}x + a_{13}y + \\
&\quad \sum_{j=1}^{m} \omega_{1j} U\left(\|(u_j, v_j) - (x,y)\|\right), \\
\widehat{c}(x,y) &= a_{21} + a_{22}x + a_{23}y + \\
&\quad \sum_{j=1}^{m} \omega_{2j} U\left(\|(u_j, v_j) - (x,y)\|\right),
\end{aligned} \tag{6}$$

where $U(x) = x^2 \log x$, for $x \neq 0$, and $U(0) = 0$ is a basis function, $\|\cdot\|$ is the Euclidean norm, and $a_{11}, a_{12}, a_{13}, \omega_{1j}, a_{21}, a_{22}, a_{23}, \omega_{2j}$, for $j = 1, 2, \ldots, m$, are coefficients. Let

$$\mathbf{K} = \begin{pmatrix} U_{11} & U_{12} & \cdots & U_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ U_{m1} & U_{m2} & \cdots & U_{mm} \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} \omega_{11} & \omega_{21} \\ \vdots & \vdots \\ \omega_{1m} & \omega_{2m} \end{pmatrix}$$

$$\begin{aligned}
\mathbf{P} &= \begin{pmatrix} 1 & u_1 & v_1 \\ \vdots & \vdots & \vdots \\ 1 & u_m & v_m \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix}, \\
\mathbf{v} &= \begin{pmatrix} \widehat{b}(u_1, v_1) & \widehat{c}(u_1, v_1) \\ \vdots & \vdots \\ \widehat{b}(u_m, v_m) & \widehat{c}(u_m, v_m) \end{pmatrix},
\end{aligned}$$

where $U_{j_1 j_2} = U\left(\|(u_{j_1}, v_{j_1}) - (u_{j_2}, v_{j_2})\|\right)$, for $j_1, j_2 = 1, 2, \ldots, m$. Then, from (5) and (6), we have the following equation for the coefficient matrices $\mathbf{w}$ and $\mathbf{a}$:

$$\mathbf{K}\mathbf{w} + \mathbf{P}\mathbf{a} = \mathbf{v}.$$

To make sure that both $I_{\widehat{b}}$ and $I_{\widehat{c}}$ are finite, $\mathbf{w}$ should also satisfy the condition

$$\mathbf{P}^T \mathbf{w} = \mathbf{0}$$

Therefore, $\mathbf{w}$ and $\mathbf{a}$ can be determined from the equation system

$$\begin{pmatrix} \mathbf{K} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{w} \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} \mathbf{v} \\ \mathbf{0} \end{pmatrix}. \tag{7}$$

In the literature, there is also some research which uses a regularization parameter to control the smoothness of the estimated functions $(\widehat{b}(x,y), \widehat{c}(x,y))$ [27]. This can be achieved by considering the following minimization problems:

$$\min_{b(x,y)} \sum_{j=1}^{m} \left\{ \widehat{T}_1(u_j, v_j) - [u_j + b(u_j, v_j)] \right\}^2 + \lambda I_b,$$
$$\min_{c(x,y)} \sum_{j=1}^{m} \left\{ \widehat{T}_2(u_j, v_j) - [v_j + c(u_j, v_j)] \right\}^2 + \lambda I_c,$$

where $\widehat{\mathbf{T}}(u_j, v_j) = (\widehat{T}_1(u_j, v_j), \widehat{T}_2(u_j, v_j))$, and $\lambda > 0$ is a regularization parameter that controls the trade-off between the interpolation condition and the smoothness. When λ is 0, the above problem becomes the one with exact interpolation, whereas when λ approaches ∞, it reduces to a least square model for fitting straight planes. To accomplish the above regularized TPS, we can simply replace $\mathbf{K}$ by $\mathbf{K} + \lambda\mathbf{I}$ in (7), where $\mathbf{I}$ is an identity matrix [34], [10].

2.3 Proposed Feature Based Image Registration Procedure

Our proposed feature based IR procedure consists of three steps, described in detail below.

Step 1 (Non-Degenerate Feature Points Extraction) The non-degenerate feature points in the images $R(x,y)$ and $M(x,y)$ are extracted using the expression (3) as follows. Let $u_n > 0$ be a threshold value. Then, the point set

$$\begin{aligned} \overline{D}_R \;=\; & \{(x,y) : (x,y) \in \Omega_R, \\ & K_{11}(x,y)K_{22}(x,y) - K_{12}^2(x,y) \geq u_n\} \end{aligned}$$

denotes all extracted non-degenerate feature points in the image $R(x, y)$. Similarly, the point set

$$\begin{aligned}\overline{D}_M &= \{(x,y) : (x,y) \in \Omega_M, \\ &\quad K_{11}(x,y)K_{22}(x,y) - K_{12}^2(x,y) \geq u_n\}\end{aligned}$$

denotes all extracted non-degenerate feature points in the image $M(x, y)$. Note that, in the above expressions, $K_{11}(x, y), K_{22}(x, y)$, and $K_{12}(x, y)$ are computed from the observed image intensities of the two images, respectively, in the two different cases.

Step 2 (Feature Points Matching) In this paper, we match the feature points using the mean squared difference (MSD) metric that is commonly used in the IR literature. Let $|\overline{D}_R|$ denote the number of detected feature points in R, and $|\overline{D}_M|$ denote the number of detected feature points in M. The following discussion is under the condition that $|\overline{D}_R| \geq |\overline{D}_M|$. In cases when $|\overline{D}_R| < |\overline{D}_M|$, the roles of $R(x, y)$ and $M(x, y)$ should be switched. For any extracted feature point $(x, y) \in \overline{D}_M$, its matched feature point in $\overline{D}_R$, denoted as $\widehat{\mathbf{T}}(x, y)$, is defined by the minimizer of

$$\min_{(x',y')\in\overline{D}_R, \|(x',y')-(x,y)\|\leq r_0} MSD\left((x,y);(x',y');d_n\right), \tag{8}$$

where

$$\begin{aligned}&MSD\left((x,y),(x',y');d_n\right) = \\ &\frac{1}{\widetilde{N}} \sum_{\sqrt{s^2+t^2}\leq d_n} \left[Z_M(x+s,y+t) - Z_R(x'+s,y'+t)\right]^2,\end{aligned}$$

$d_n > 0$ and $r_0 > 0$ are two radius parameters, and $\widetilde{N}$ is the total number of pixels in the circular neighborhood $O(x, y; d_n)$. From the above expression, $MSD((x, y); (x', y'); d_n)$ is just the MSD measure of the observed image intensities of the two images in the neighborhoods $O(x, y; d_n)$ and $O(x', y'; d_n)$, respectively. In (8), for the given feature point (x, y) in $\overline{D}_M$, we search all the feature points in $\overline{D}_R$ whose Euclidean distances from (x, y) are smaller than or equal to r_0 such that the MSD measure is minimized. Statistically, the probability for (8) to have more than one minimizer

is 0 under some regularity conditions (e.g., the noise involved in the related images has a continuous distribution such as the normal distribution). But, for completeness of our algorithm, in cases when there are several minimizers in (8), we choose the one closest to (x, y) as its final matched feature point in $\overline{D}_R$. Careful selection of the final matched feature point, however, is unimportant in such cases because the event for (8) to have multiple minimizers is rare and this issue would hardly affect the overall performance of the proposed procedure. If there are no points in $\overline{D}_R$ that are within r_0 from (x, y), then (8) has no solution. In such cases, remove (x, y) from $\overline{D}_M$. After $\widehat{\mathbf{T}}(x, y)$ is defined, we further define $\widehat{b}(x, y) = \widehat{T}_1(x, y) - x$ and $\widehat{c}(x, y) = \widehat{T}_2(x, y) - y$.

Step 3 (Interpolation by the TPS Model) From step 2, we obtain a sequence of matched feature points $\{(u_j, v_j), j = 1, 2, \ldots, m\}$ in image M and $\{\widehat{\mathbf{T}}(u_j, v_j), j = 1, 2, \ldots, m\}$ in image R. Then, by the TPS interpolation procedure (6) and (7), we can obtain the geometric transformation $\widehat{\mathbf{T}}(x, y)$ defined in the entire design space Ω_M.

In the proposed feature based IR procedure described above, there are some parameters to choose. Data-driven selection of these parameters is an on-going research project. Here, we provide some practical guidelines based on our numerical experience. In Step 1 of the procedure, there are two parameters h_n and u_n involved. The parameter h_n is the bandwidth used in the weighted least square problem (1). We recommend choosing it from the values $\{0.015, 0.02, 0.025\}$. The parameter u_n is the threshold value used when defining the extracted non-degenerate feature points. If u_n is chosen larger, then less non-degenerate feature points will be extracted from the two related images. Proper determination of the number of feature points for best image registration is still an open research problem [11]. Based on our experience, the image registration would be reasonably good if u_n is chosen such that $|\overline{D}_R|/|\Omega_R| \in [0.1, 0.15]$, where $|A|$ denotes the number of pixels in the pointset A. In Step 2 of the procedure, there are two radius parameters r_0 and d_n involved. Based on our experience, r_0 can be chosen in the interval $[0.05, 0.25]$, and d_n can be chosen $\min\{r_0, s_n\}$,

where s_n is a number in $[0.01, 0.1]$.

3 Numerical Study

We have tried many different numerical examples to evaluate the performance of our proposed feature based image registration procedure, denoted as NEW hereafter. In this section, we present three of them. The results from other examples are similar to those of the examples presented here. In all the examples, for comparison purposes, besides the feature points described in Section 2.1, we also consider four other types of features that are commonly used in various applications. The first is the edge pixels detected by the Canny's edge detector [5], the second is the corner points detected by the FAST (Features from Accelerated Segment Test) algorithm [28], [29], the third type of feature points is those identified by the Harris' corner detector [12], [31], and the last type is those extracted by the SIFT (Scale-Invariant Feature Transform) algorithm [17], [18]. The Canny's edge detector has been commonly used in various imaging applications since 1986. The FAST algorithm is a corner detector that examines the so-called Bresenham circle of a given radius around a candidate pixel. If at least k contiguous pixels have their intensities all above or below the intensity at the candidate pixel by a given amount specified by a threshold value, then the candidate pixel is classified as a corner pixel. Various numerical experiments show that a reasonable choice of k is 9 [29]. The software package of the FAST algorithm can be downloaded from http://svr-www.eng.cam.ac.uk/ er258/work/fast.html. The Harris' corner detector is another popular algorithm for identifying corners, which uses eigenvalues of the structure tensor matrix for determining the candidate corner pixels, described briefly below. Let (x_i, y_j) be a given pixel on an image I, and W_{ij} be a window of (x_i, y_j). Then, the "auto-correlation function" of the image I

at (x_i, y_j) is defined by

$$\begin{aligned} \mathrm{AC}(x_i, y_j) &= \sum_{(x_i,y_j)\in W_{ij}} [I(x_i, y_j) - I(x_i + \Delta x, y_j + \Delta y)] \\ &\approx \sum_{(x_i,y_j)\in W_{ij}} \left\{ I(x_i, y_j) - \left[I(x_i, y_j) + I'_x(x_i, y_j)\Delta x + I'_y(x_i, y_j)\Delta y \right] \right\} \\ &= (\Delta x, \Delta y) H(x_i, y_j) (\Delta x, \Delta y)' \end{aligned}$$

where $(\Delta x, \Delta y)$ is a shift, "$\approx$" in the second line implies that some higher-order terms have been ignored when $I(x_i + \Delta x, y_j + \Delta y)$ is replaced by its first-order Taylor's expansion, the image I is usually filtered by a Gaussian filter beforehand, and

$$H(x_i, y_j) = \begin{pmatrix} \sum_{(x_i,y_j)\in W_{ij}} \left(I'_x(x_i, y_j)\right)^2, & \sum_{(x_i,y_j)\in W_{ij}} I'_x(x_i, y_j) I'_y(x_i, y_j) \\ \sum_{(x_i,y_j)\in W_{ij}} I'_x(x_i, y_j) I'_y(x_i, y_j), & \sum_{(x_i,y_j)\in W_{ij}} \left(I'_y(x_i, y_j)\right)^2 \end{pmatrix}$$

is a 2×2 matrix describing the local structure of I around the pixel (x_i, y_j). The pixel (x_i, y_j) is identified as a corner if both eigenvalues of $H(x_i, y_j)$ are large. From this brief description, it can be seen that the "auto-correlation function" defined above and the weighted least square problem (1) are completely different, although both of them have the Taylor's expansion involved. The former is for measuring the local structure of a single image around a given pixel while the latter is for measuring the difference between two images around a given pixel after one image is geometrically adjusted. A matlab code of the Harris' corner detector can be downloaded from http://www.csse.uwa.edu.au/ pk/research/matlabfns/. The feature points extracted by the SIFT algorithm have become popular in recent years for various purposes, including object recognition. They have a nice property that they are invariant to scale and affine transformations. The software of the SIFT algorithm can be downloaded from http://www.cs.ubc.ca/ lowe/keypoints/.

With different types of feature points, the feature based image registration is proceeded as described in Section 2.3, except that, for the extracted SIFT features, because each of them has a set of descriptors associated, the Euclidean distance (instead of the MSD) of the associated

descriptors is used in Step 2 of the registration procedure for feature points matching. The image registration procedure with the five different types of features is labeled as NEW, Canny, FAST, Harris, and SIFT, respectively. To numerically evaluate a given set of image registration results, we use the root residual mean square (RRMS) and the cross correlation (CC) measures, both of which are popular in the IR literature. The RRMS measure is defined by

$$\text{RRMS} = \left\{ \frac{1}{n^2} \sum_{i,j=1}^{n} \left[Z_M(x_i, y_j) - Z_R(\widehat{\mathbf{T}}(x_i, y_j)) \right]^2 \right\}^{1/2},$$

where $\widehat{\mathbf{T}}$ denotes the estimated geometric transformation by a given image registration procedure. It is obvious that RRMS is proportional to the Euclidean distance between $\{Z_M(x_i, y_j)\}$ and $\{Z_R(\widehat{\mathbf{T}}(x_i, y_j))\}$. Thus, the smaller its value, the better. The CC measure is defined to be the Pearsons sample correlation coefficient of the bivariate data $\{(Z_M(x_i, y_j), Z_R(T(x_i, y_j)))\}$. Intuitively, if $\widehat{\mathbf{T}}(x, y)$ is a good estimator of $\mathbf{T}(x, y)$, then $Z_R(\widehat{\mathbf{T}}(x, y))$ would be close to $Z_M(x, y)$. Consequently, the CC value should be large. Therefore, by this measure, the larger its value, the better.

In all the examples described below, each image has 128×128 pixels with the image intensities in the range $[0, 255]$. In the proposed procedure NEW, a 2-D density kernel function K is used in (1). In all examples, K is chosen to be the truncated bivariate Gaussian density function with support $\{(u, v) : u^2 + v^2 \leq 1\}$. This kernel function is commonly used in the computer science literature (cf., Qiu 2005, Section 7.4). In the statistical literature, people traditionally use the Epanecknikov kernel function (cf., Qiu 2005, Section 2.3). Our experience is that our proposed procedure NEW performs similarly using either of them. In Step 2 of the image registration procedure described in Section 2.3, there are two parameters d_n and r_0 (cf., expression (8)). They are chosen to be 0.05 and 0.15, respectively, in all examples for the four methods NEW, Canny, FAST, and Harris, because they give reasonably good results in all cases considered. In the TPS interpolation procedure used in Step 3 of the image registration procedure described in Section 2.3, there is a smoothing parameter $\lambda > 0$ involved. We have tried many different values of λ in our numerical study, and find that

$\lambda = 8.0$ would give reasonably good results in all cases considered. So, this value is used in all examples for all five methods.

In the first example, the reference image is a texture image that is downloaded from

http://www.freeseamlesstextures.com/texture_gallery/index.htm

and shown in Fig 1(a). This image is then moved to the right by 0.05 and moved upward by 0.06, where its design space is normalized to be $\Omega_R = [0, 1] \times [0, 1]$. The resulting moved image is shown in Fig 1(b).

[Figure 1 about here.]

Then, we apply the five methods NEW, Canny, FAST, Harris, and SIFT to the observed images. Each method has a number of parameters involved. More specifically, the method NEW has two parameters h_n and u_n used in detecting feature points. The method Canny has three parameters involved: a scale parameter used in the Gaussian density function for image preprocessing and two threshold values. The method FAST has two parameters: a threshold value and a boolean parameter to control the implementation of a nonmaximal suppression. The Harris' corner detector has three parameters: a scale parameter in the Gaussian filter, a radius parameter in the non-maximal suppression algorithm, and a threshold value for determining the corner candidates. The SIFT algorithm is accomplished by a binary executable file provided by its author, with no parameters to adjust. In order to make the comparison of the five types of features meaningful, we try to choose the parameters of different registration methods properly such that the numbers of matched pairs of the detected feature points in the two images are about the same among the five methods. In our numerical study, we try to set the number of matched pairs of the detected feature points at a low, a medium, and a high level. However, the method SIFT can only give a single number of matched pairs of the detected feature points in each example because it has no parameters to adjust, and that number is usually small. Among the remaining four methods, the method FAST has the least

flexibility in generating different numbers of matched pairs of the detected feature points. So, we set the specific number of matched pairs determined by SIFT as the first level, and determine the other two levels by FAST. Then, the numbers of matched pairs of the other three methods are adjusted to be as close to these three levels as possible. In the texture image example, the three levels of the number of matched pairs of the detected feature points are determined by SIFT to be 106, and by FAST to be 38 and 295. The closest numbers of the matched pairs by NEW are 38, 106, and 296, those by Canny are 38, 106, and 296, those by Harris are 38, 106, and 296, and those by FAST are 38, 107, and 296.

The extracted feature points by the five methods are shown in the five rows of Fig 2, respectively. In each of the first four rows, the extracted feature points in the reference and moved images are shown in three pairs, with the numbers of matched pairs of the detected feature points increase from the left to the right. For instance, in the first row, plots (a) and (b) show the 38 matched pairs of the detected feature points in the reference and moved images, respectively, obtained by the method New, plots (c) and (d) show the 106 matched pairs of the detected feature points by the same method, and plots (e) and (f) show the 296 matched pairs of the detected feature points by this method. Because the texture reference image does not contain any objects with clear outlines, the detected feature points by all five methods do not display clear patterns either. However, if we look at the plots in this figure carefully, we would find that the detected feature points by the methods NEW and Canny can still reveal certain structures of the reference and moved images, while the detected feature points by the other three methods are mostly random.

[Figure 2 about here.]

The restored moved images $Z_R(\widehat{\mathbf{T}}(x,y))$, obtained after applying the estimated transformation $\widehat{\mathbf{T}}(x,y)$ onto the reference image Z_R, are presented in Fig 3, along with the residual images defined as $Z_R(\widehat{\mathbf{T}}(x,y)) - Z_M(x,y)$. From the images shown in this figure, it can be seen that the method NEW

indeed performs better than the other four methods at all three levels of the number of matched pairs of detected feature points, because it has less non-zero residuals. Numerical performance measures RRMS and CC of the five methods are presented in Table 1, along with a summary of their detected feature points in the reference and moved images. From the table, it can be seen that the proposed method NEW has the smallest RRMS values and the largest CC values in all cases, consistent with the residual images in Fig 3.

[Figure 3 about here.]

[Table 1 about here.]

In the second example, the reference image is a real fingerprint image that is downloaded from http://www.foxnews.com/story/0,2933,290215,00.html and shown in Fig 4 (a). This image is then twisted arbitrarily using the photoshop software package, and the resulting moved image is shown in Fig 4 (b). This example tries to simulate a real application in which a fingerprint image is altered by a nonparametric geometrical transformation arbitrarily.

[Figure 4 about here.]

[Figure 5 about here.]

In this example, all procedure parameters are chosen in the same way as that in the previous example. The three levels of the number of matched pairs of detected feature points is determined first by the SIFT method to be 41, and then by the FAST method to be 185 and 394. The closest numbers of the matched pairs by the method NEW are 41, 185, and 395, the closest numbers of the matched pairs by the method Canny are 41, 187, and 396, the closest numbers by the method Harris are 41, 186, and 395, and the closest numbers by the method FAST are 42, 185, and 394. The detected feature points by the five methods are shown in Fig 5, and the individual images in

the figure are organized in the same way as those in Fig 2. In this example, the main structure of the fingerprint cannot be described well by step edges; thus, feature points extracted by the method Canny (cf., the images in the 2nd row of Fig 5) cannot provide much useful information for image registration either. From the images, it seems that the detected features by the methods FAST, Harris, and SIFT are mostly random and do not provide much useful information for image registration. These results are consistent with the results of the numerical measures RRMS and CC that are presented in Table 2. From the table, it can be seen that the method NEW performs uniformly better than its peers in all cases considered. The matched feature points are presented in Fig 6, where the feature points in the reference image are indicated by the red color and those in the moved image are indicated by the green color. The restored moved images and the corresponding residual images are shown in Fig 7, from which it can be seen that the residual images of the proposed method NEW contain the least pixels with non-zero intensities in all cases considered.

[Figure 6 about here.]

[Figure 7 about here.]

[Table 2 about here.]

In the last example, both the reference and moved images are real rock images obtained from a series of motion pictures that can be downloaded from

http://vasc.ri.cmu.edu//idb/html/motion/marked_rocks/index.html.

The two images are shown in Fig 8. If we observe the two images carefully, then we can notice that the moved image is taken after the camera is moved to the the right a little bit.

[Figure 8 about here.]

As in the previous two examples, the three levels of the number of matched pair of detected feature points are first determined by the method SIFT to be 43, and then by the method FAST

to be 107 and 229. The closest numbers of the matched pairs by the method NEW are 43, 106, and 221, the closest numbers of the matched pairs by the method Canny are 43, 106, and 221, the closest numbers by the method Harris are 43, 106, and 221, and the closest numbers by the method FAST are 42, 107, and 229. The detected feature points by the five methods are shown in Fig 9, which is organized in the same way as that in Fig 2. From the images in the figure, it can be seen that, when the number of matched pairs increases, the detected feature points by the methods New, Canny, and FAST concentrate at some specific positions in the two images, while the detected feature points by the method Harris are scattered in the entire images. The matched feature points are presented in Fig 10, organized in the same way as that of Fig 6. The restored moved images and the corresponding residual images are shown in Fig 11, from which it can be seen that our method NEW performs the best in all cases considered. The results of the numerical measures RRMS and CC are presented in Table 3 which confirms that the method New indeed performs the best in this example.

[Figure 9 about here.]

[Figure 10 about here.]

[Figure 11 about here.]

[Table 3 about here.]

4 Concluding Remarks

In this paper, we have proposed an image registration procedure based on the detected non-degenerate feature points. From their definition, the non-degenerate feature points are relevant for characterizing the geometrical transformation $\mathbf{T}(x, y)$ that is involved in the image registration

problem. Therefore, they should be good feature points for image registration. Our numerical examples presented in the previous section have confirmed this intuition.

There are still some issues in our proposed method that should be addressed in the future research. For instance, in each of the presented numerical examples, we set three levels for the number of matched pairs of the extracted feature points. In real applications, however, it is often a challenging task to determine how many feature points to use. Intuitively, the more feature points extracted, the more information for approximating the geometrical transformation $\mathbf{T}(x, y)$. Consequently, the feature-based image registration can be better achieved, although more computation is involved in such cases. However, our numerical examples demonstrate that large numbers of extracted feature points cannot always guarantee a better image registration (cf., Tables 1-3). Based on our numerical experience, the best numbers of extracted feature points in the reference and moved images depend on the complexity of the transformation $\mathbf{T}(x, y)$ and on the image structure as well. If $\mathbf{T}(x, y)$ is simple (e.g., a linear motion transformation), then a small number of extracted feature points in each image might be good enough; but, that would not be sufficient if $\mathbf{T}(x, y)$ has a more complex structure. Some existing research in the literature is dedicated to selection of a subset of extracted feature points for image registration without sacrificing the quality [11], [14]. In our opinion, proper determination of the numbers of extracted feature points from the observed image intensities is still an open problem, and requires much future research. In our proposed feature-based image registration procedure, there are a number of parameters involved (e.g., d_n and r_0 in (8)). Data-driven parameter selection is not discussed in this paper, although some empirical guidelines were given at the end of section 2. It might be possible to choose these parameters by using the cross-validation (CV), generalized cross-validation (GCV), and other parameter selection procedures that are discussed extensively in the statistical literature [23]. Adaptation of such parameter selection procedures to the image registration problem requires much future research as

well.

Acknowledgments: We thank the handling editor, Dr. Hengyong Yu, and three referees for many constructive comments and suggestions, which greatly improved the quality of the paper. This research is supported in part by an NSF grant.

References

[1] B. Avants, C. Epstein, M. Grossman, and J. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12:26–41, 2008.

[2] F. Beg, M. Miller, A. Trouv'e, and L. Younes. Computing deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 6:139–157, 2005.

[3] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(24):509–522, Apr. 2002.

[4] F. Bookstein. Principal warps: Thin-plate splines and decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, June 1989.

[5] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[6] E. Denton, L. Sonoda, D. Rueckert, S. Rankin, C. Hayes, M. Leach, D. Hill, and D. Hawkes. Comparison and evaluation of rigid, affine, and nonrigid registration of breast mr images. *Journal of Computer Assisted Tomography*, 23:800–805, 1999.

[7] J. Duchon. Splines minimizing rotation-invariant semi-norms in sobolev spaces. In W. Schempp and K. Zeller, editors, *Constructive Theory of Functions of Several Variables*, pages 85–100. Springer-Verlag, Berlin, 1977.

[8] F. Dufaux and J. Konrad. Efficient, robust, and fast global motion estimation for video coding. *IEEE Transactions on Image Processing*, 9(3):497–501, Mar. 2000.

[9] J. Fan and I. Gijbels. *Local Polynomial Modeling and Its Applications*. Chapman and Hall, London, 1996.

[10] F. Girosi, M. Jones, and T. Poggio. Regularization theory and neural networks architectures. *Neural Computation*, 7(2):219–269, 1995.

[11] A. Goshtasby and G. C. Stockman. Point pattern matching using convex hull edges. *IEEE Transactions on Systems, Man and Cybernetics*, 15:631–637, 1985.

[12] C. Harris and M. Stephens. A combined corner and edge detector. *Conference, Proceedings of the 4th Alvey Vision*, pages 147–151, 1988.

[13] A. Klein, J. Andersson, B. Ardekani, and et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *NeuroImage*, 46:786–802, 2009.

[14] D. Lavine, B. Lambird, and L. Kanal. Recognition of spatial point patterns. *Pattern Recognition*, 16:289–295, 1983.

[15] H. Li, B. S. . Manjunath, and S. K. Mitra. A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, 4(3):320–334, Mar. 1995.

[16] L. Liu, T. Jiang, J. Yang, and C. Zhu. Fingerprint registration by maximization of mutual information. *IEEE Transactions on Image Processing*, 15(5):1100–1110, May 2006.

[17] D. G. Lowe. Object recognition from local scale-invariant features. *Proceedings of the International Conference on Computer Vision*, 2:1150–1157, 1999.

[18] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[19] J. Modersitzki. *Fair: Flexible Algorithms for Image Registration*. SIAM, Philadelphia, 2009.

[20] X. Papademetris, A. Jackowski, R. Schultz, L. Staib, and J. Duncan. Integrated intensity and point-feature nonrigid registration. In *In Proceedings of the 7th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 763–770, Saint-Malo, France, Sept. 2004.

[21] E. Persoon and K. Fu. Shape discrimination using fourier descriptors. *IEEE Transactions on Systems, Man and Cybernetics*, 7(3):170–179, Mar. 1977.

[22] M. Powell. A thin plate spline method for mapping curves into curves in two dimensions. *Computational Techniques and Applications (CTAC '95)*, 1995.

[23] P. Qiu. *Image Processing and Jump Regression Analysis*. John Wiley & Sons, New York, 2005.

[24] P. Qiu and T. Nguyen. On image registration in magnetic resonance imaging. In *2008 International Conference on BioMedical Engineering and Informatics*, pages 753–757. IEEE, May 2008.

[25] P. Qiu and C. Xing. Intensity based nonparametric image registration. Proceedings of the 2010 ACM SIGMM International Conference on Multimedia Information Retrieval, pages 221–225, Philadelphia, PA, 2010.

[26] P. Qiu and C. Xing. On nonparametric image registration. *Technometrics*, 55:xxx–xxx, 2013.

[27] K. Rohr, M. Fornefett, and H. Stiehl. Approximating thin-plate splines for elastic registration: integration of landmark errors and orientation attributes. In *Proceedings of the 16th International Conference on Information Processing in Medical Imaging*, volume 1613 of *Lecture Notes in Computer Science*, pages 252–265, Hungary, 1999.

[28] E. Rosten and T. Drummond. Fusing points and lines for high performance tracking. *IEEE International Conference on Computer Vision*, 2:1508–1511, Oct. 2005.

[29] E. Rosten and T. Drummond. Machine learning for high-speed corner detection. *European Conference on Computer Vision*, pages 430–443, May 2006.

[30] N. Saeed. Magnetic resonance image segmentation using pattern recognition, and applied to image registration and quantitation. *NMR in Biomedicine*, 11:157–167, 1998.

[31] J. Shi and C. Tomasi. Good features to track. *9th IEEE Conference on Computer Vision and Pattern Recognition*, 1994.

[32] N. Tustison, B. Avants, and J. Gee. Directly manipulated free-form deformation image registration. *IEEE Transactions on Image Processing*, 18:624–635, 2009.

[33] P. van den Elsen, E. Pol, and M. Viergever. Medical image matching a review with classification. *IEEE Engineering in Medicine and Biology*, 12:384–396, 1993.

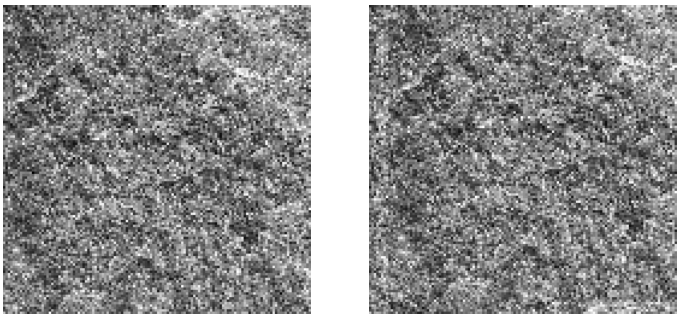
[34] G. Wahba. *Spline Models for Observational Data.* Soc. Industrial and Applied Math., 1990.

[35] Y. Wang and L. Staib. Physical model based non-rigid registration incorporating statistical shape information. *Medical Image Analysis*, 4(1):7–20, 2000.

[36] C. Xing and P. Qiu. Intensity based image registration by nonparametric local smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33:2081–2092, 2011.

[37] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000, 2003.

List of Figures



(a) (b)

Figure 1: Reference image (plot (a)) and moved image (plot (b)) in the texture image example.

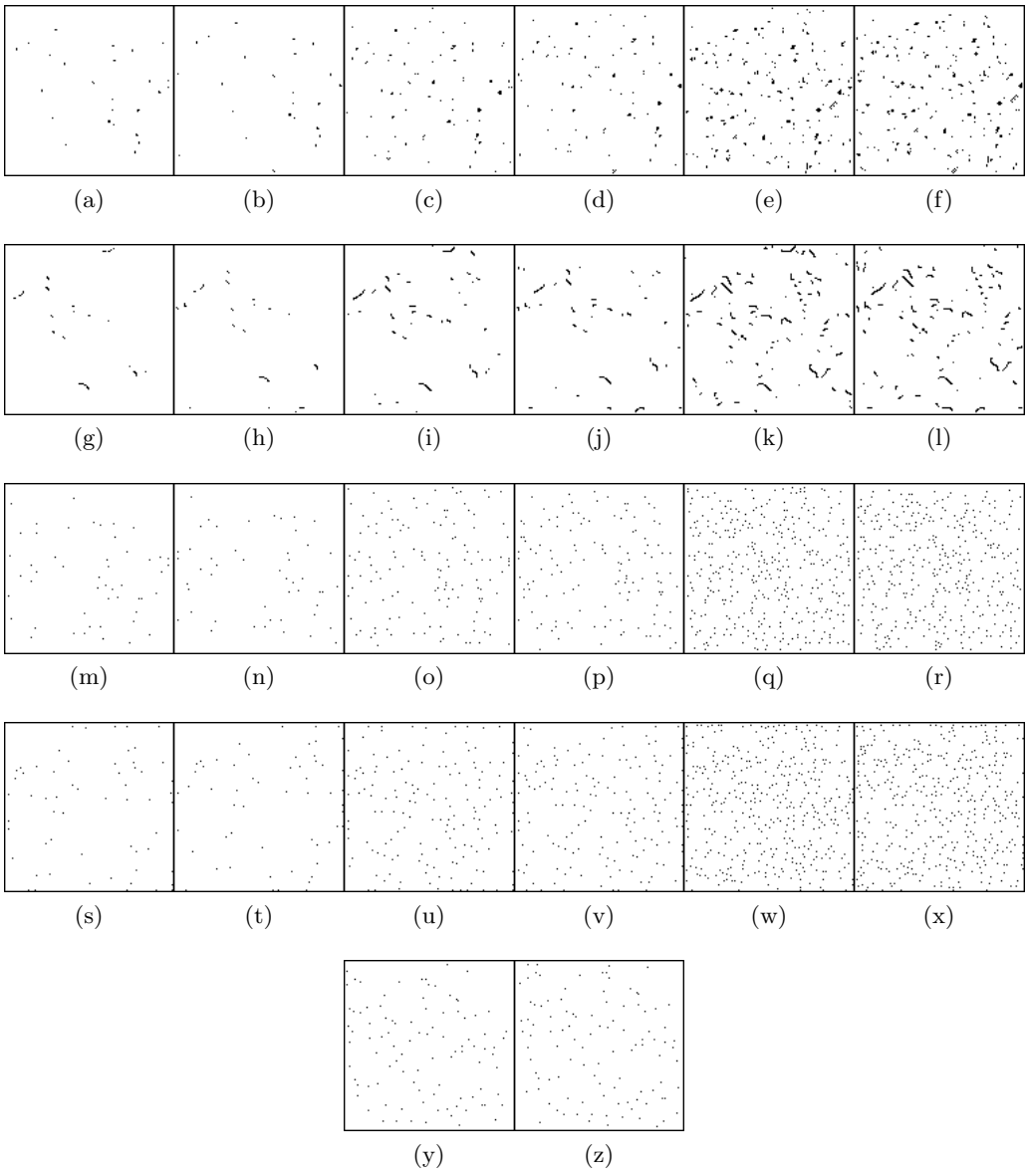


Figure 2: Detected feature points in the reference image (1st, 3rd, and 5th columns) and the moved image (2nd, 4th, and 6th columns) by the five methods NEW (1st row), Canny (2nd row), FAST (3rd row), Harris (4th row), and SIFT (5th row) in the texture image example. In each of the first four rows, results are presented in three pairs of plots, in which the numbers of matched pairs of detected feature points increase from the left to the right.

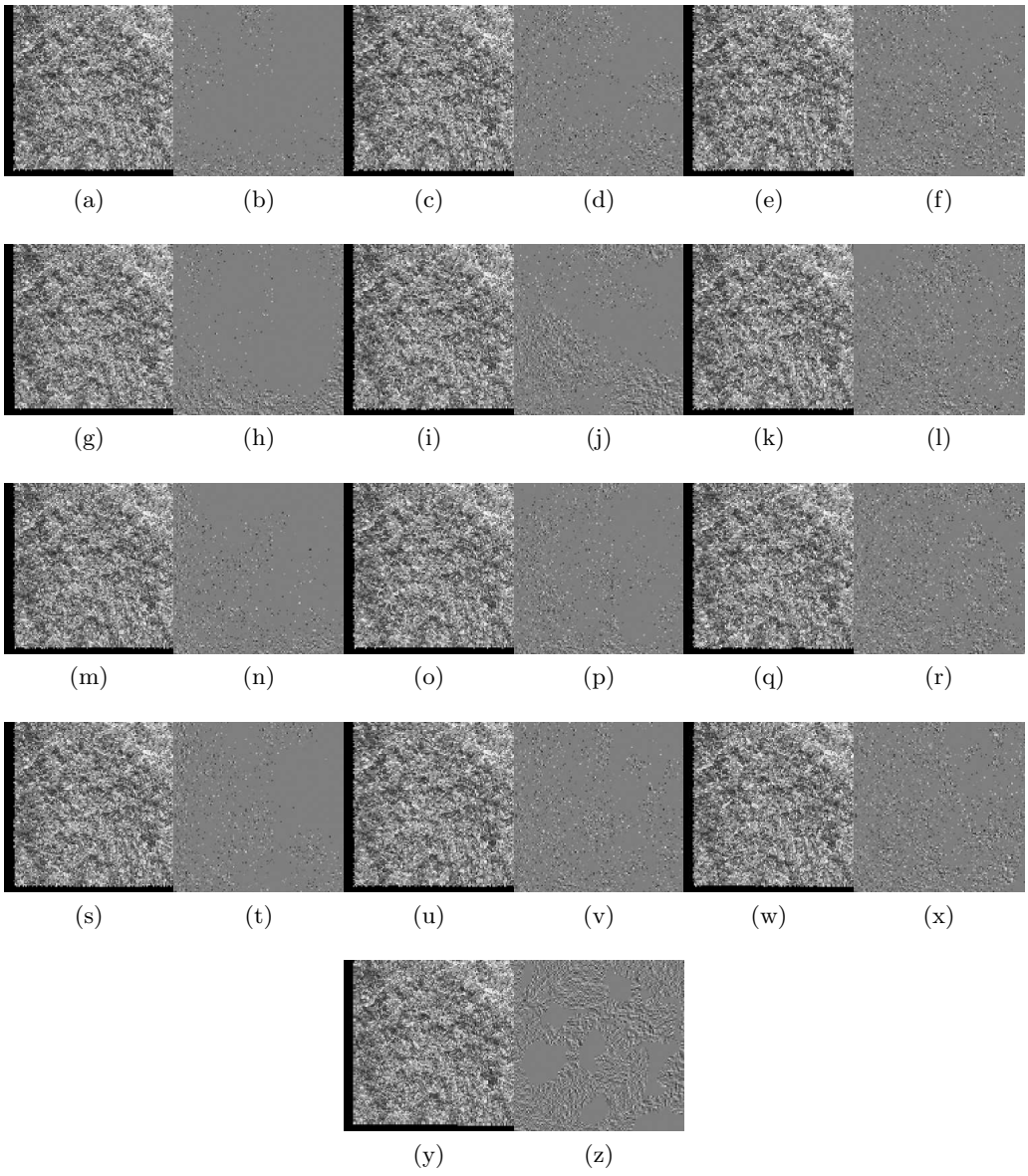


Figure 3: Restored moved images $Z_R(\widehat{\mathbf{T}}(x,y))$ and the corresponding residual images $Z_R(\widehat{\mathbf{T}}(x,y)) - Z_M(x,y)$ in the texture image example. The five rows are for the methods NEW, Canny, FAST, Harris, and SIFT, respectively. The pair of plots (a) and (b) corresponds to the pair of Fig 2(a)-(b), and the other pairs correspond to the other plots in Fig 2 in a similar way.

(a) (b)

Figure 4: (a) A reference fingerprint image. (b) A moved image.

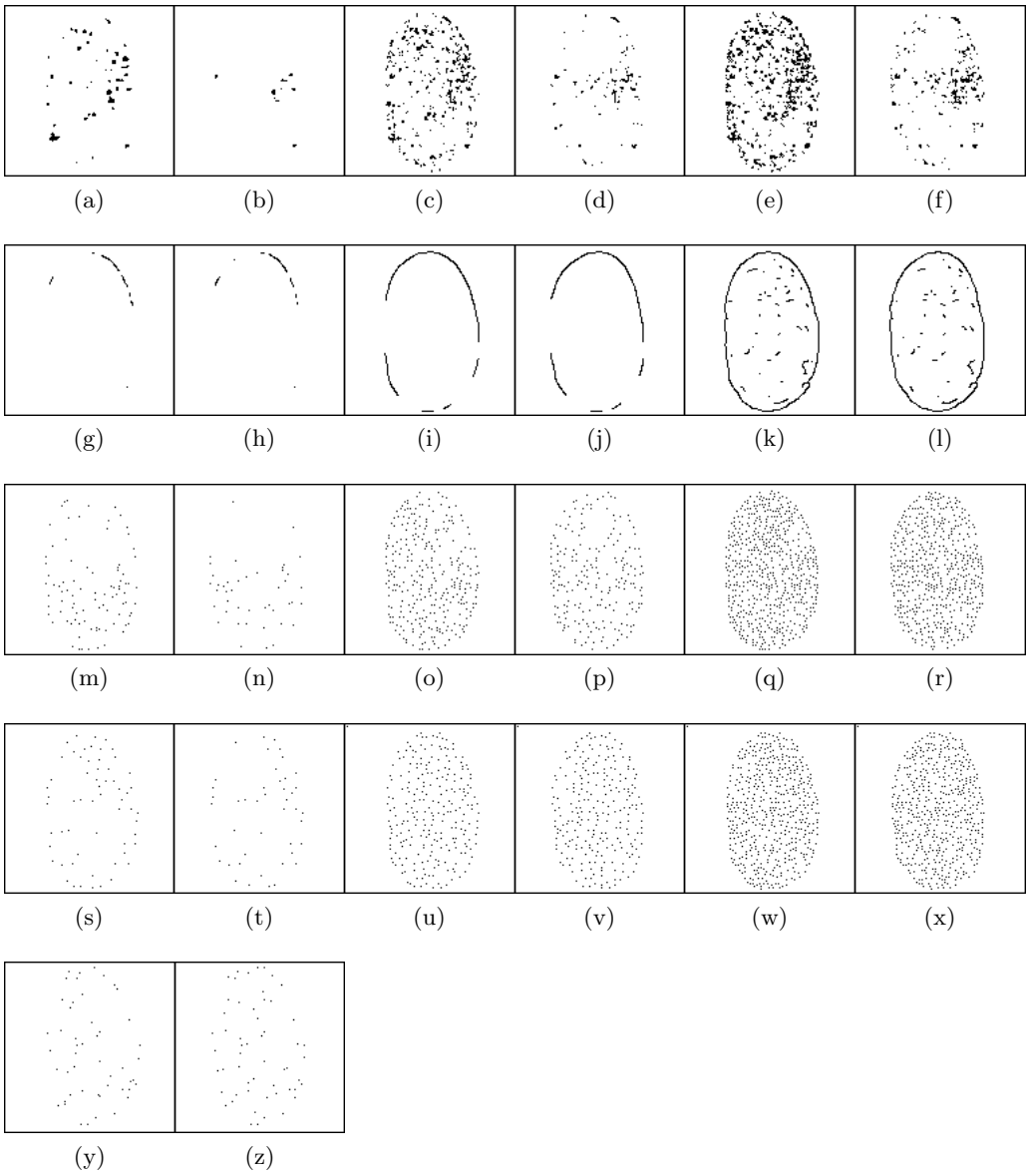


Figure 5: Detected feature points in the reference image (1st, 3rd, and 5th columns) and the moved image (2nd, 4th, and 6th columns) by the five methods NEW (1st row), Canny (2nd row), FAST (3rd row), Harris (4th row), and SIFT (5th row) in the fingerprint example. In each of the first four rows, results are presented in three pairs of plots, in which the numbers of matched pairs of detected feature points increase from the left to the right.

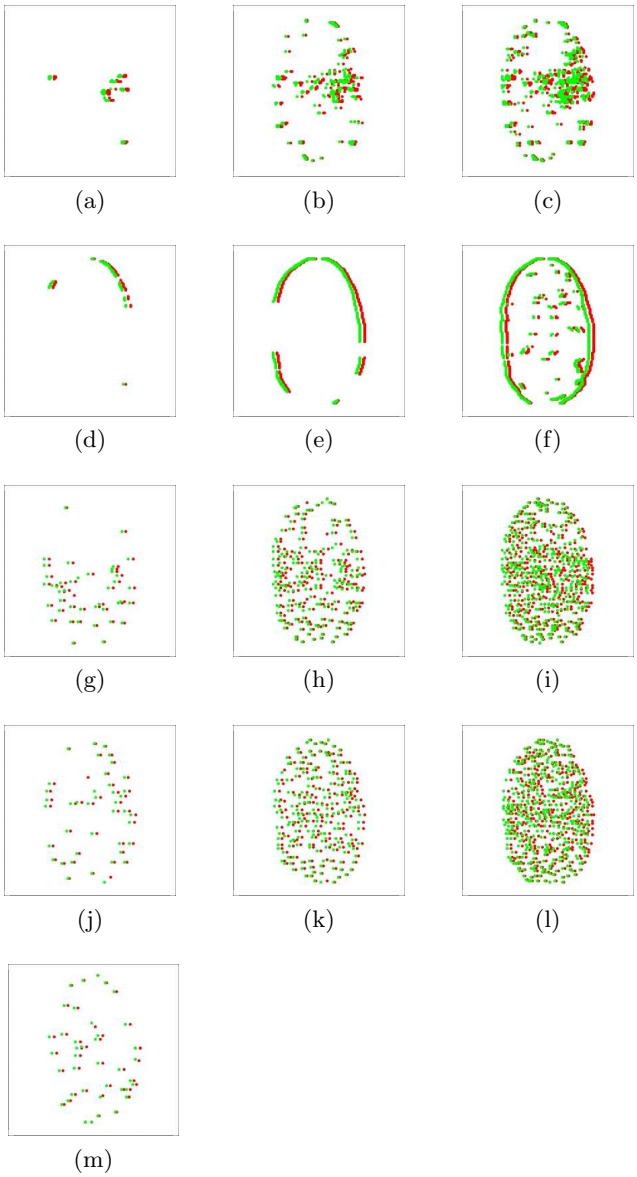


Figure 6: Matched feature points in the fingerprint example obtained from the results shown in Fig 5. Results in the five rows are for methods NEW, Canny, FAST, Harris and SIFT, respectively. Results in plot (a) are from the feature points shown in Fig 5(a)-(b), in which red points denote the feature points in the reference image and green points denote their matched feature points in the moved image. Results in other plots are organized in the same way.

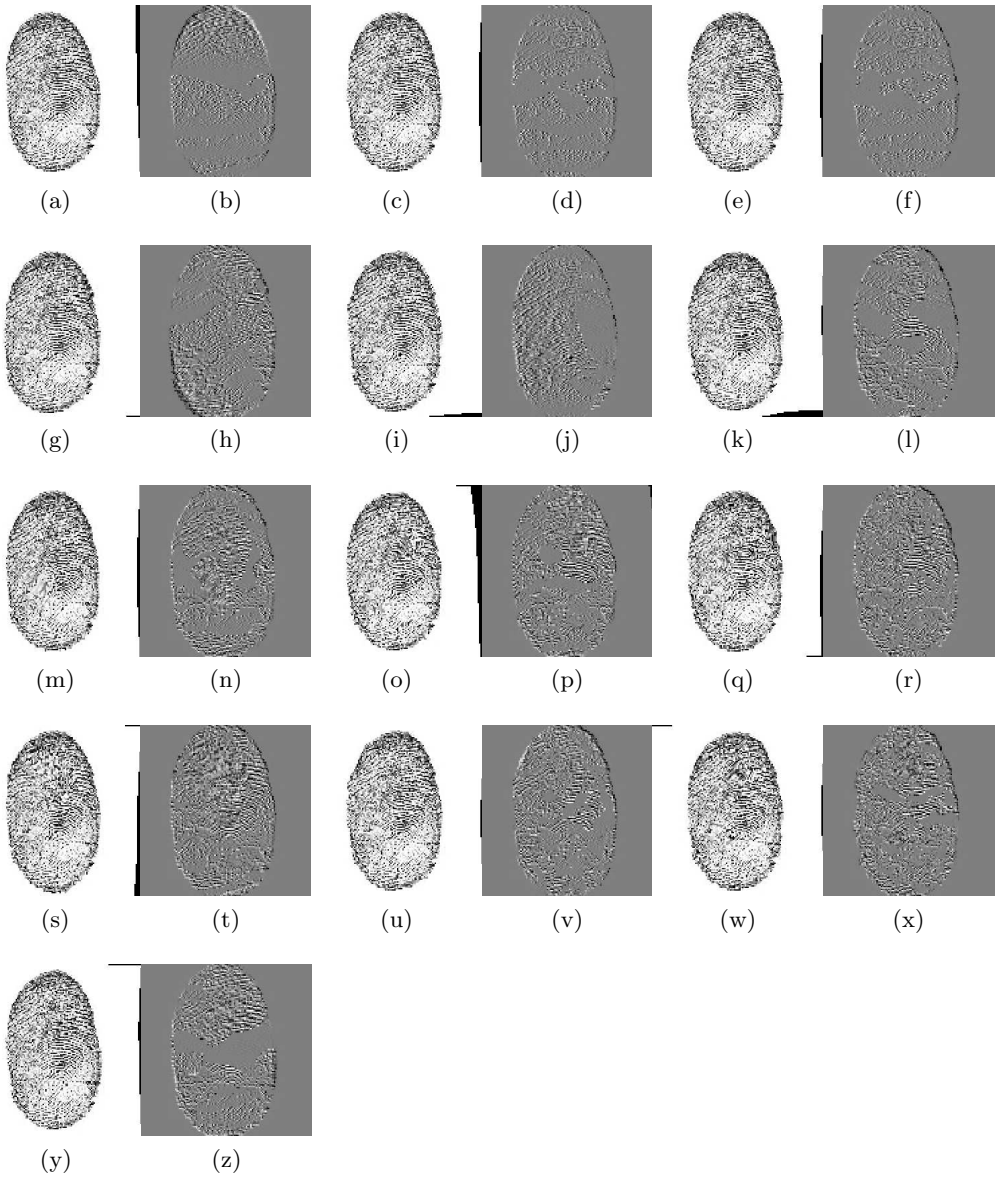


Figure 7: Restored moved images $Z_R(\widehat{\mathbf{T}}(x,y))$ and the corresponding residual images $Z_R(\widehat{\mathbf{T}}(x,y)) - Z_M(x,y)$ in the fingerprint example. The five rows are for the methods NEW, Canny, FAST, Harris, and SIFT, respectively. The pair of plots (a) and (b) corresponds to the case of Fig 5(a)-(b), and the other pairs correspond to the related results in Fig 5 similarly.

(a) (b)

Figure 8: (a) A reference rock image. (b) A moved image.

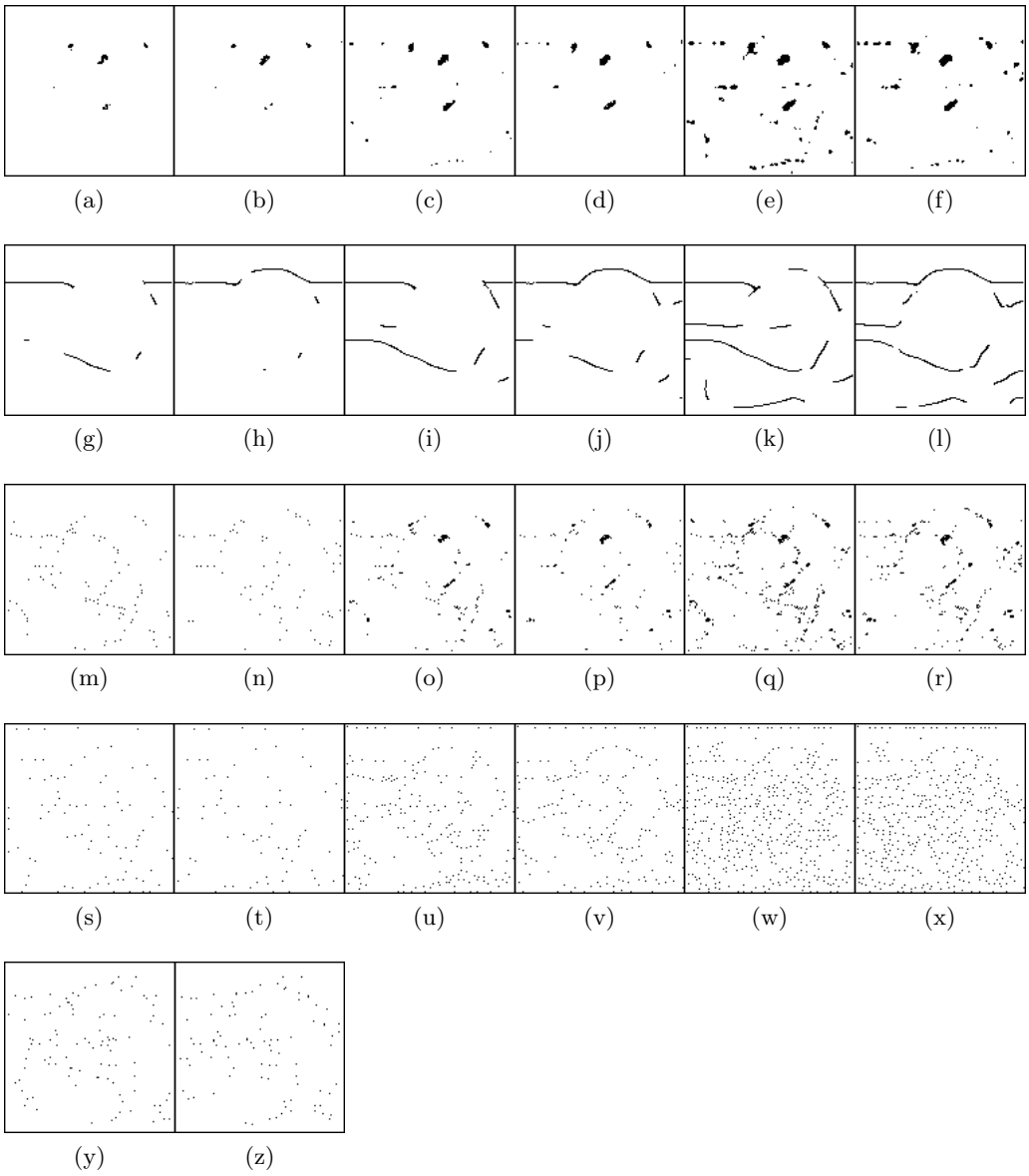


Figure 9: Detected feature points in reference image (1st, 3rd, and 5th columns) and the moved image (2nd, 4th, and 6th columns) by the five methods NEW (1st row), Canny (2nd row), FAST (3rd row), Harris (4th row), and SIFT (5th row) in the rock image example. In each of the first four rows, results are presented in three pairs of plots, in which the numbers of matched pairs of detected feature points increase from the left to the right.

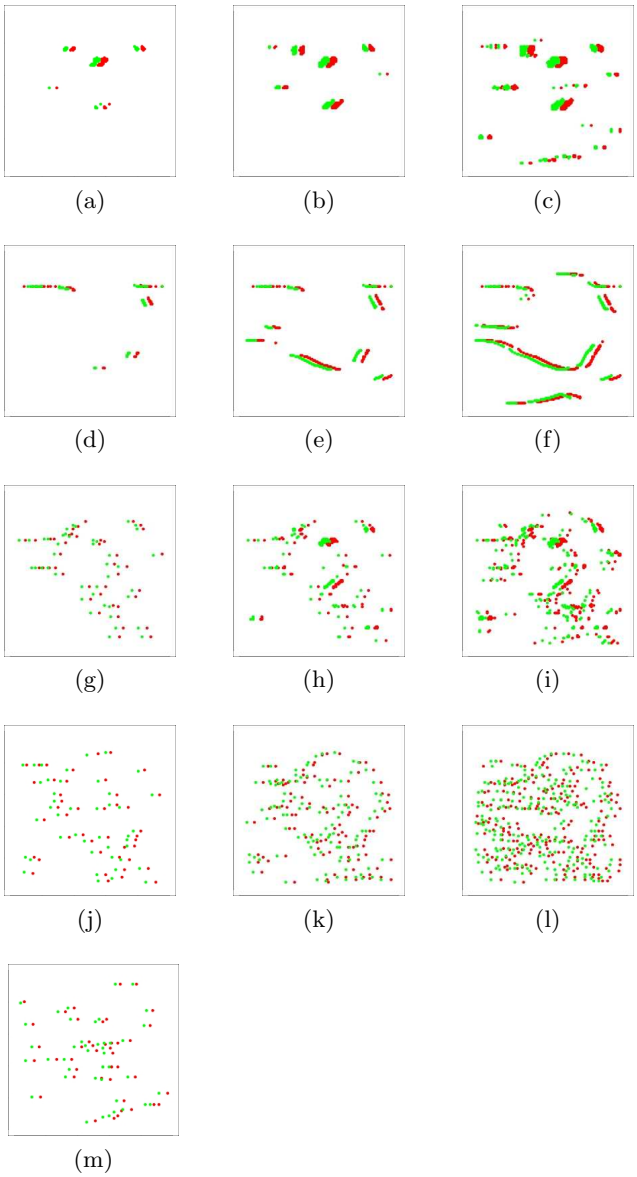


Figure 10: Matched feature points in the rock image example obtained from the results shown in Fig 9. Results in the five rows are for methods NEW, Canny, FAST, Harris and SIFT, respectively. Results in plot (a) are from the feature points shown in Fig 9(a)-(b), in which red points denote the feature points in the reference image and green points denote their matched feature points in the moved image. Results in other plots are organized in the same way.

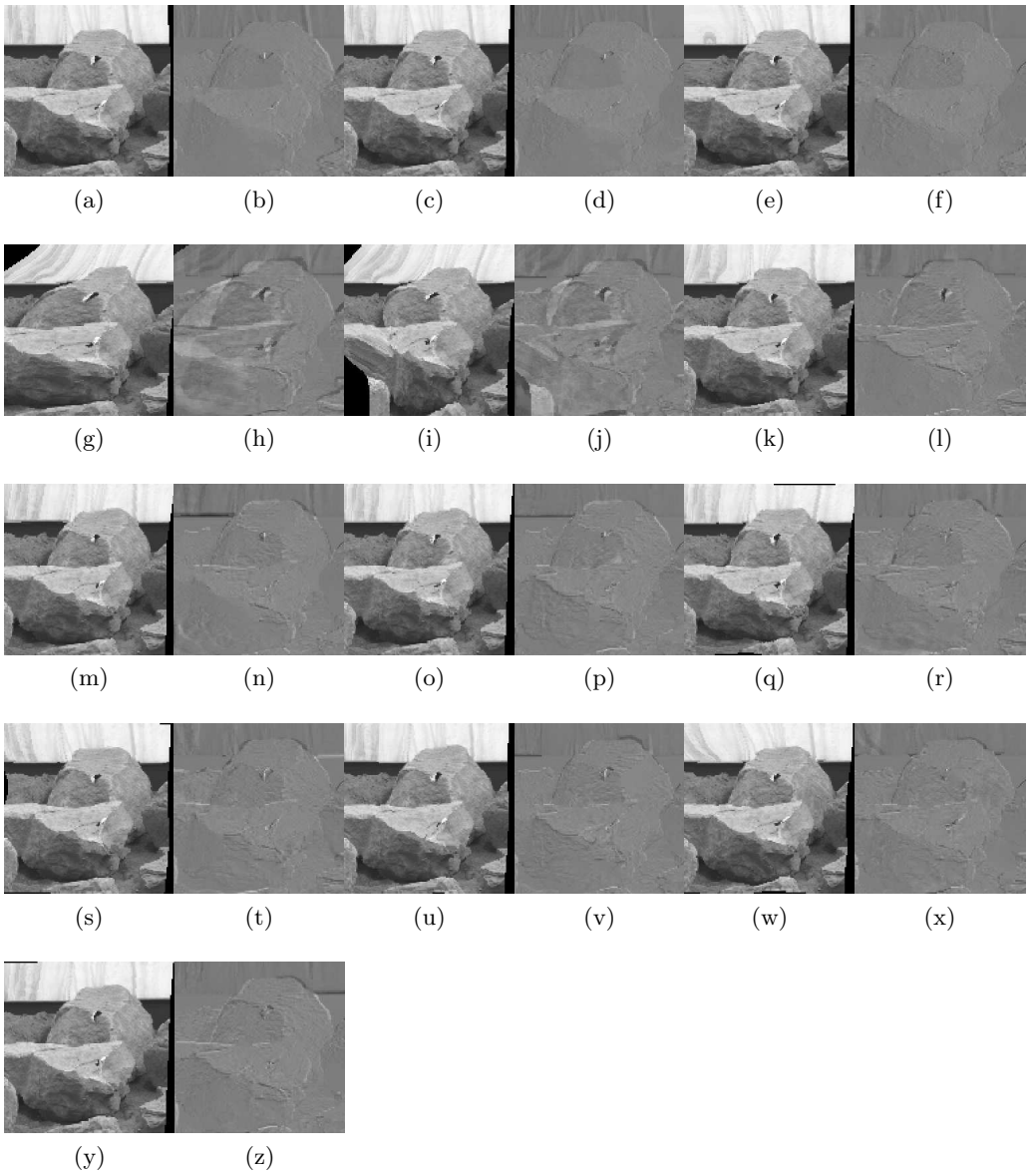


Figure 11: Restored moved images $Z_R(\widehat{\mathbf{T}}(x,y))$ and the corresponding residual images $Z_R(\widehat{\mathbf{T}}(x,y)) - Z_M(x,y)$ in the rock image example. The five rows are for the methods NEW, Canny, FAST, Harris, and SIFT, respectively. The pair of plots (a) and (b) corresponds to the case of Fig 9(a)-(b), and the other pairs correspond to the related results in Fig 9 similarly.

List of Tables

Table 1: Numerical performance measures RRMS and CC of the five methods NEW, Canny, FAST, Harris, and SIFT, along with a summary of their detected feature points, in the texture image example.

Method	# of Feature Points in Z_R	# of Feature Points in Z_M	# of Matched Feature Points	RRMS Value	CC Value
New	45	45	38	17.436060	0.921180
	128	127	106	26.966148	0.811951
	347	360	296	32.472932	0.727368
Canny	54	51	38	25.077406	0.836841
	144	140	106	34.823847	0.684244
	383	383	296	33.553160	0.709173
FAST	50	49	38	23.621764	0.855085
	132	129	107	27.044096	0.810149
	362	365	295	34.758800	0.685656
Harris	59	57	38	23.281763	0.859266
	150	155	106	28.940011	0.782706
	412	429	296	34.179141	0.696844
SIFT	133	141	106	48.563976	0.386922

Table 2: Numerical performance measures RRMS and CC of the five methods NEW, Canny, FAST, Harris, and SIFT, along with a summary of their detected feature points, in the fingerprint example.

Method	# of Feature Points in Z_R	# of Feature Points in Z_M	# of Matched Feature Points	RRMS Value	CC Value
New	224	41	41	50.981442	0.709939
	664	185	185	43.128353	0.793319
	1168	397	395	42.554701	0.798680
Canny	48	50	41	61.978028	0.572733
	227	234	187	53.123332	0.685827
	490	482	396	55.892717	0.653374
FAST	90	44	42	61.337823	0.577478
	299	211	185	63.961907	0.542112
	576	517	394	65.375967	0.521666
Harris	60	44	41	68.415590	0.475635
	243	243	186	60.888339	0.583962
	498	502	395	61.957272	0.576090
SIFT	59	61	41	59.477569	0.601160

Table 3: Numerical performance measures RRMS and CC of the five methods NEW, Canny, FAST, Harris, and SIFT, along with a summary of their detected feature points, in the rock image example.

Method	# of Feature Points in Z_R	# of Feature Points in Z_M	# of Matched Feature Points	RRMS Value	CC Value
New	59	56	43	14.429331	0.957693
	158	132	106	12.795238	0.966670
	409	314	221	12.788663	0.967155
Canny	140	144	43	27.821090	0.838532
	238	249	106	28.215268	0.837185
	427	435	221	17.845393	0.933753
FAST	98	72	42	18.594821	0.928695
	206	152	107	17.186702	0.939592
	419	334	229	17.178284	0.939301
Harris	89	72	43	17.097757	0.940096
	229	199	106	16.443445	0.944312
	478	447	221	16.089799	0.946788
SIFT	110	103	43	17.144078	0.939319