On Nonparametric Image Registration

Peihua Qiu and Chen Xing
School of Statistics
University of Minnesota

Abstract

Image registration aims to map one image to another of a same scene. It is a fundamental task in many imaging applications. Most existing image registration methods assume that the mapping transformation has a parametric form or satisfies certain regularity conditions (e.g., it is a smooth function with continuous first or higher order derivatives). They often estimate the mapping transformation globally by solving a global minimization/maximization problem. Such global smoothing methods usually cannot preserve singularities (e.g., discontinuities) and other features of the mapping transformation well. Further, the ill-posed nature of the image registration problem, namely, the mapping transformation is not well defined at certain places, including the place where the true image intensity surface is straight, is not handled properly by such methods. In this paper, we suggest solving the image registration problem locally, by first studying the local properties of a mapping transformation. To this end, some concepts for describing such local properties are suggested, and a local smoothing method for estimating the mapping transformation is proposed. Because of the flexibility of local smoothing, our method does not require any parametric form or other global regularity conditions on the mapping transformation. Both theoretical and numerical studies show that it is effective in various applications. Supplementary materials are available online.

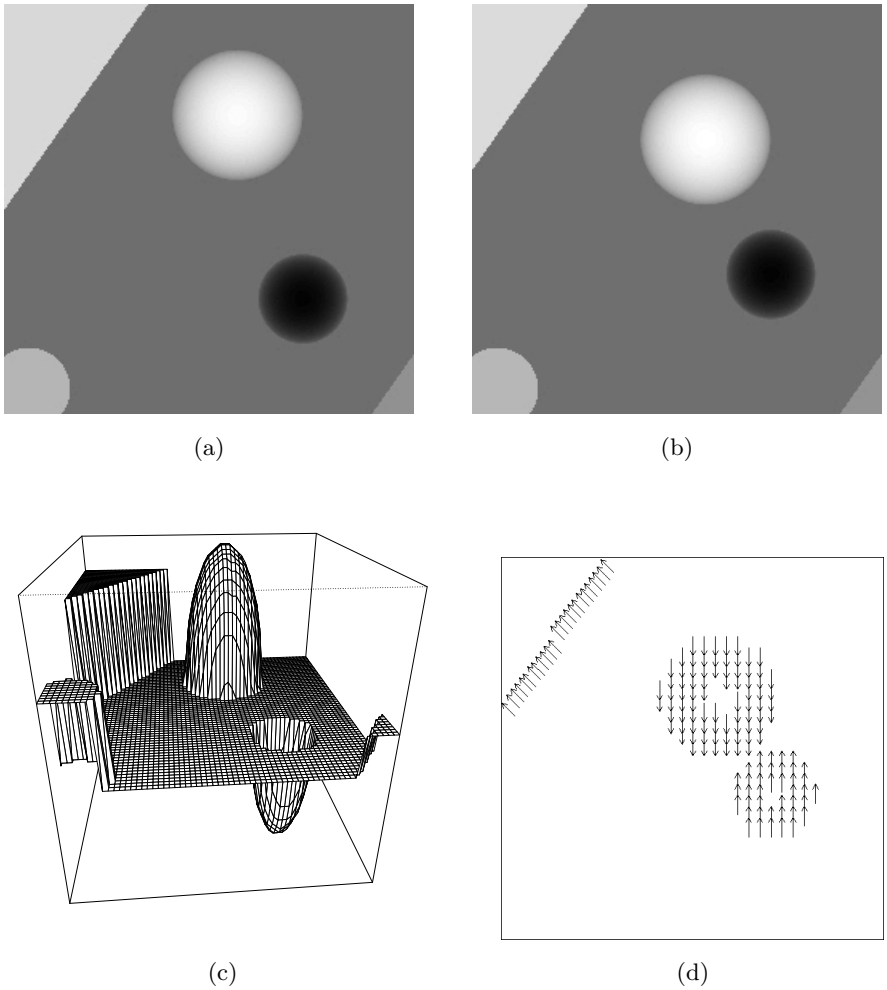
Key Words: Degenerate pixels; Discontinuity; Image mapping; Image transformation; Local approximation; Local smoothing; Nonparametric transformation; Statistical properties.

1 Introduction

Assume that two MRI images of a brain tumor patient were taken at a same spot of his/her head at two different times, to study the growth of the tumor. Although the two images are from the same spot of the head, pixels on one image may not geometrically match up the corresponding ones on the other image well, for various reasons, including the difference of the relative movement between the

head and the imaging device at the two different imaging times. In order to properly compare the two images for locating their differences accurately, the two images should be geometrically matched up beforehand. Image registration (IR) is specifically for this purpose (cf., Zitova and Flusser 2003, Modersitzki 2009). It is an indispensable step for many imaging applications, including medical imaging (Klein et al. 2009), remote sensing (Li et al. 1995), finger print or face recognition (Liu et al. 2006), image compression (Dufaux and Konrad 2000), video enhancement (Irani and Peleg 1993), and so forth. As a demonstration, Figure 1 presents two images to register. The image in plot (b) is obtained from the one in plot (a), after certain image objects (e.g., the two circular objects in the middle) are moved in their locations. In the literature, the image in plot (a) is often called a reference image, and the one in plot (b) is called a moved image. The image intensity surface of the reference image is shown in plot (c), and the true geometrical transformation, denoted as $\mathbf{T}$, that describes the geometrical movement from the reference image to the moved image is shown in plot (d). One major goal of IR is to estimate the geometrical transformation that describes the geometrical movement from the reference image to the moved image, and then match up the two images well.

IR has been discussed extensively in the computer science and medical imaging literatures. One group of existing methods first select two sets of features in the two images under consideration, and then find a geometrical transformation $\mathbf{T}$ to best match the two sets of features (Althof et al. 1997, Davis et al. 1997). Commonly used features include landmarks or control points that can be selected manually or automatically by a computer (Wu et al. 2006), edge lines or curves that are often detected by gradient-based methods (Hsieh et al. 1997), and regions, centroids or templates that are usually determined by ways of thresholding and segmentation (Saeed 1998). Because feature extraction is often a time-consuming and challenging task with much arbitrariness involved, recent IR research focuses more on the search of the transformation $\mathbf{T}$ based directly on the observed image intensities of the two images, instead of their features. Such methods are often referred to as the intensity-based image registration (IBIR) methods in the literature (Bhagalia et al. 2009). IBIR methods typically search for a transformation $\mathbf{T}$ in a certain transformation family, such that the similarity (or dissimilarity) of the transformed image to another image is maximized (or minimized). In the literature, both parametric and nonparametric transformation families have been considered. Commonly used parametric transformation families, such as the rigid-body and affine motion models, assume that the difference between the two images involves a global motion



(a) (b)

(c) (d)

Figure 1: (a) A reference image. (c) A moved image. (c) Image intensity surface of the reference image shown in plot (a). (d) The true geometrical transformation $\mathbf{T}$.

that can be modeled parametrically (e.g., Denton et al. 1999). For instance, the rigid-body model assumes that the distance between any two points in an image is unchanged after the geometrical transformation, and the transformation $\mathbf{T}$ in 2-D cases can be modeled parametrically with two translation parameters and one rotation parameter. In practice, however, the geometrical movement from the reference image to the moved image is often too complicated to be described properly by a parametric transformation family. In the MRI example discussed at the beginning of the paper, for instance, besides the relative movement between the head and the imaging device, many organs and tissues (e.g., the brain, the blood flow in blood vessels) inside the head would move or pulse over time. This type of geometrical movement is often local and difficult to describe by a parametric model. To handle such IR problems, nonparametric transformation families have been considered in the literature, which do not assume any specific parametric forms. Instead, it is often assumed

that $\mathbf{T}$ is a continuous function satisfying various regularity conditions. For instance, the free-form deformation (DDF) technique considers transformation families consisting of B-splines with knots properly chosen beforehand (Szeliski and Coughlan 1997, Tustison et al. 2009). Other IBIR methods using nonparametric transformation families include the diffeomorphic image registration method (e.g., Avants et al. 2008), image registration using Fourier transformations (e.g., Pan et al. 2009), information-theoretical image registration (e.g., Rajwade et al. 2009), image registration based on physical models (e.g., Wang and Staib 2000), among some others. Klein et al. (2009) made a comprehensive evaluation of 14 IBIR algorithms using various test images.

Most existing IBIR methods mentioned above estimate the geometrical transformation $\mathbf{T}$ using different global smoothing methods, although some smoothing methods (e.g., the B-spline methods) allow more general $\mathbf{T}$ than the others (e.g., the parametric transformation methods). Global smoothing methods are often convenient to implement, and they are efficient as well in cases when the assumed parametric models or other global regularity conditions are satisfied. However, validity of such model assumptions should be justified properly in practice, which is challenging and currently lacking. As discussed recently by Xing and Qiu (2011), the IR problem is *ill-posed* in the sense that the transformation $\mathbf{T}$ cannot be properly defined at certain places, including the place where the image intensity functions of the two related images are linear. For instance, in the example of Figure 1, it is impossible to define $\mathbf{T}$ in the gray background region of the reference image, because the observed image intensities in that region are all the same and thus it is impossible to know exactly how the pixels in that region are moved. However, the ill-posed nature of the IR problem is difficult to discover by a global smoothing method, because the related global maximization (or minimization) algorithm can always find a solution even when $\mathbf{T}$ is actually undefined at certain places. This explains the reason why the ill-posed nature of the IR problem is not adequately discussed in the literature yet.

It is our belief that, to handle the IR problem properly, we should first study the major properties of the geometrical transformation $\mathbf{T}$, including its ill-posed nature. Although the ill-posed nature of $\mathbf{T}$ was first discussed by Xing and Qiu (2011), that issue has not been studied systematically yet. For instance, Xing and Qiu (2011) did not provide formal definitions of the places where $\mathbf{T}$ is well defined and the places where $\mathbf{T}$ is partially defined. It did not discuss how to detect these places in a statistically consistent way either. Furthermore, there is no related discussion in that paper about the statistical properties of the estimated $\mathbf{T}$. The major goal of this

paper is to overcome these limitations and provide a more detailed discussion about the properties of $\mathbf{T}$. To this end, several concepts, including the 2-D degenerate pixels, 2-D partial degenerate pixels, 1-D degenerate pixels, and 1-D partial degenerate pixels, are proposed for describing the local properties of $\mathbf{T}$. The relationship among these concepts is also studied. Based on this study about $\mathbf{T}$, a local smoothing method is proposed for solving the IBIR problem. The local smoothing nature of the proposed method makes it possible not to impose restrictive assumptions on $\mathbf{T}$. As a matter of fact, our method does not require any parametric form or other global regularization conditions on $\mathbf{T}$. Both theoretical and numerical studies show that it is effective in various applications. It should be pointed out that, because of the local smoothing nature of the proposed IBIR method, it can only handle cases when the magnitude of $\mathbf{T}$ (i.e., the geometrical movement from the reference image to the moved image) is relatively small. In many real applications (cf., the two applications discussed in Section 5), the geometrical movement involved is small; thus, our proposed method can be applied to these applications directly. In cases when the magnitude of $\mathbf{T}$ is large, however, a global smoothing method based on a parametric transformation family (e.g., the affine motion transformation family) should be used first to account for part of $\mathbf{T}$ with a large magnitude, and then our proposed method can be used to fine tune the parametric transformation locally.

The rest of the paper is organized as follows. In Section 2, we provide a statistical description of the IR problem, and define several concepts for describing local properties of the mapping transformation $\mathbf{T}$. In Section 3, our proposed IBIR method is introduced in detail. Its statistical properties are discussed in Section 4. Numerical performance of the proposed method is investigated in Section 5. Several concluding remarks are given in Section 6. Some technical details, including proofs of two theorems presented in Section 4, certain numerical results, computer source codes of our proposed method, and the data related to the two real-image examples considered in Section 5, are given online as supplementary materials.

2 The IR Problem and the Mapping Transformation

In the literature, the IR problem is often described as a problem to geometrically match up two images. However, to study this problem systematically, a statistically more precise description is necessary. Let $R(x, y)$ and $M(x, y)$ be two true images to register, and $M(x, y)$ be a geometrically moved version of $R(x, y)$. Then, there is a geometrical transformation $\mathbf{T}(x, y) = (T_1(x, y), T_2(x, y))$

such that

$$M(T_1(x,y),T_2(x,y)) = R(x,y), \qquad \text{for } (x,y) \in \Omega_R, \tag{1}$$

where Ω_R is the design space of the image R. For simplicity, we assume that $\Omega_R = [0,1]\times[0,1]$, and the design space of the image M is also assumed to be $\Omega_M = [0,1]\times[0,1]$. In practice, both $\mathbf{T}$ and the true images R and M are unobservable. The major goal of IR is to estimate the transformation $\mathbf{T}$ from the observed images of R and M, such that the two images can be well matched up. In the IR literature, most people consider pointwise noise only in the observed images. In such cases, the observed images of R and M can be described by

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

where $\{(x_i,y_j), i,j = 1,2,\ldots,n\}$ are pixel locations (or design points in the statistical literature), and $\{\varepsilon_R(x_i,y_j), i,j = 1,2,\ldots,n\}$ and $\{\varepsilon_M(x_i,y_j), i,j = 1,2,\ldots,n\}$ are i.i.d. random errors in the two images with mean 0 and unknown variances σ_R^2 and σ_M^2, respectively. In practice, the pixel locations are always regularly spaced in rows and columns, although the number of rows and the number of columns could be different. In model (2), the number of rows and the number of columns are assumed to be the same for simplicity of presentation, and our proposed method can also work well in cases when the two numbers are different. So, in model (2), the pixel locations are actually assumed to be $\{(x_i,y_j) = (i/n, j/n), i.j = 1,2,\ldots,n\}$. Besides pointwise noise, in practice, observed images may also contain spatial blur and other types of contamination. See Hall and Qiu (2007) for a related discussion.

The IR problem described in (1) and (2) has the following boundary problem: when a given pixel (x,y) is close to the border of Ω_R, it is possible that $\mathbf{T}(x,y)$ is outside Ω_M. In such cases, $M(\mathbf{T}(x,y))$ is not defined. Therefore, the transformation $\mathbf{T}(x,y)$ is well defined by (1) in $\Omega_{R,T}$ only, where $\Omega_{R,T} = \{(x,y) : (x,y) \in \Omega_R, \mathbf{T}(x,y) \in \Omega_M\}$. To define $\mathbf{T}(x,y)$ in the whole Ω_R, an appropriate extension of the design spaces is necessary.

Intuitively, in cases when the image intensity function $R(x,y)$ is a constant in the whole design space Ω_R, the transformation $\mathbf{T}(x,y)$ is not well defined, because the moved image intensity function $M(x,y)$ must be the same constant and $\mathbf{T}(x,y)$ can be any function defined in $\Omega_{R,T}$. More generally, the transformation $\mathbf{T}(x,y)$ is not well defined in cases when the 2-dimensional (2-D) function $R(x,y)$ is degenerate, and it is actually a 1-dimensional (1-D) function. For instance, in cases when

$R(x, y) = \phi(1 + 2x)$, where ϕ is a 1-D function, the geometric move along the y direction cannot be uniquely determined based on $R(x, y)$ and $M(x, y)$. Consequently, $\mathbf{T}(x, y)$ is not uniquely defined. In practice, it might be rare that the image intensity function $R(x, y)$ is degenerate in the entire design space Ω_R. However, it could be degenerate in small regions of Ω_R. To this end, we make the definitions described below.

Definition 1 A pixel (x, y) is called a *2-D degenerate pixel* of an image I if there is a neighborhood $O(x, y)$ of (x, y) such that the image intensity function of I is a constant in $O(x, y)$.

Definition 2 A pixel (x, y) is called a *2-D partial degenerate pixel* of an image I if there is a neighborhood $O(x, y)$ of (x, y) such that the image intensity function of I is a constant in a connected proper subregion of $O(x, y)$ whose closure contains (x, y), but there is no neighborhood of (x, y) in which the image intensity function of I is a constant.

Figure 2 demonstrates the concepts of 2-D degenerate pixels and 2-D partial degenerate pixels. In plot (a), the image intensity function of I is a constant in the entire neighborhood $O(x, y)$ of (x, y). So, (x, y) is a 2-D degenerate pixel of I. In plot (b), the image intensity function of I is a constant only in a proper subregion of $O(x, y)$. So, (x, y) is a 2-D partial degenerate pixel of I in this case.

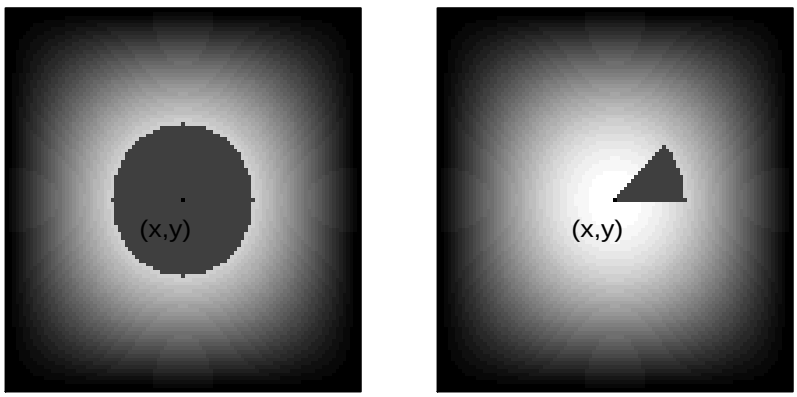


Figure 2: Examples of 2-D degenerate and 2-D partial degenerate pixels. (a) The pixel (x, y) is a 2-D degenerate pixel of the image. (b) The pixel (x, y) is a 2-D partial degenerate pixel of the image.

Let $D_{2,I}$ be the set of all 2-D degenerate pixels of I, and $PD_{2,I}$ be the set of all 2-D partial

degenerate pixels of I. A subregion of the design space Ω_I of the image I is called non-trivial if it has a positive Lebesgue area. Then, we have the following results.

Theorem 1 Assume that there are only a countable number of disjoint, connected, non-trivial subregions $\{D_{2,I}^{(j)}, j = 1, 2, \ldots\}$ of Ω_I, in each of which the image intensity function of I is a constant. Then, (i) $D_{2,I} = \bigcup_{j=1}^{\infty} \left(D_{2,I}^{(j)} \backslash \partial D_{2,I}^{(j)} \right)$, and (ii) $PD_{2,I} = \bigcup_{j=1}^{\infty} \partial D_{2,I}^{(j)}$, where $\partial D_{2,I}^{(j)}$ denotes the boundary set of $D_{2,I}^{(j)}$ for each j.

The proof of Theorem 1 is quite straight forward, and can be briefly described as follows. To prove (i), for a given pixel $(x, y) \in D_{2,I}$, by Definition 1, we can find a neighborhood $O(x, y)$ such that the image intensity function of I is a constant in $O(x, y)$. Because $O(x, y)$ has a positive Lebesgue area, (x, y) must be contained in one of $\{D_{2,I}^{(j)}, j = 1, 2, \ldots\}$ and it is obvious that (x, y) is an interior point of that region. Therefore, $D_{2,I} \subseteq \bigcup_{j=1}^{\infty} \left(D_{2,I}^{(j)} \backslash \partial D_{2,I}^{(j)} \right)$. On the other hand, for a pixel $(x, y) \in D_{2,I}^{(j)} \backslash \partial D_{2,I}^{(j)}$ with a given j, it must be a 2-D degenerate pixel because it is an interior point of $D_{2,I}^{(j)}$ and we can find a neighborhood within $D_{2,I}^{(j)}$ such that the image intensity function of I is a constant. Therefore, $D_{2,I} \supseteq \bigcup_{j=1}^{\infty} \left(D_{2,I}^{(j)} \backslash \partial D_{2,I}^{(j)} \right)$. To show the result (ii), let us consider any pixel $(x, y) \in \partial D_{2,I}^{(j)}$ with a given j. In such cases, obviously (x, y) is not a degenerate pixel. For an open circle $C(x, y; r)$ that is centered at (x, y) with radius r, let us consider $C(x, y; r) \bigcap D_{2,I}^{(j)}$. Obviously, it is a connected proper subregion of $D_{2,I}^{(j)}$ with its closure containing (x, y), and the image intensity of I is a constant in its interior region. Therefore, $(x, y) \in PD_{2,I}$. So, $\bigcup_{j=1}^{\infty} \partial D_{2,I}^{(j)} \subseteq PD_{2,I}$. On the other hand, if $(x, y) \in PD_{2,I}$, then by definition, there is a neighborhood in which a connected subregion belongs to one of $\{D_{2,I}^{(j)}, j = 1, 2, \ldots\}$, and (x, y) is contained by the closure of the subregion. Without loss of generality, let us assume that the subregion belongs to $D_{2,I}^{(j)}$ for a given j. Then, (x, y) is contained in the closure of $D_{2,I}^{(j)}$. Because (x, y) is not a 2-D degenerate point of I, it is not an interior point of $D_{2,I}^{(j)}$. Consequently, $(x, y) \in \partial D_{2,I}^{(j)}$. So, $PD_{2,I} \subseteq \bigcup_{j=1}^{\infty} \partial D_{2,I}^{(j)}$. Result (ii) then follows.

In practice, image intensity functions of most images are piecewise continuous, different continuity regions are separated by edge curves, and they are countable. In each continuity region, it should be reasonable to assume that there are only a countable number of disjoint, connected, non-trivial subregions in each of which the intensity function is a constant. Therefore, results in Theorem 1 should describe $D_{2,I}$ and $PD_{2,I}$ of a typical image well.

Intuitively, if (x, y) is a 2-D degenerate pixel of the reference image R, then the image intensity

function of R is a constant in a connected open set, denoted as $D^*(x,y)$, that contains (x,y). Then, the image intensity function of M should be the same constant in $\mathbf{T}(D^*(x,y))$. As long as $\mathbf{T}(D^*(x,y))$ contains more than one point, $\mathbf{T}(x,y)$ is not well defined, because any point in $\mathbf{T}(D^*(x,y))$ could be a reasonable value of $\mathbf{T}(x,y)$. This intuition can be generalized in the following way. If a pixel (x,y) satisfying the condition that there is a neighborhood $O(x,y)$ and a line $\mathbf{l}$ containing (x,y) such that the image intensity function of R is a constant along $\mathbf{l}$ in $O(x,y)$, then $\mathbf{T}(x,y)$ is not well defined either, as long as $\mathbf{T}(O_{\mathbf{l}}(x,y))$ contains at least two points, where $O_{\mathbf{l}}(x,y)$ denotes the set of all points in $O(x,y)$ on the line $\mathbf{l}$. Based on this observation, we have the following definitions.

Definition 3 A pixel (x,y) is called a *1-D degenerate pixel* of an image I if it is not a 2-D degenerate pixel of I, and there is a neighborhood $O(x,y)$ of (x,y) and a line $\mathbf{l}$ containing (x,y) such that the image intensity function of I is a constant on $\mathbf{l}$ within $O(x,y)$. The set of all 1-D degenerate pixels of I is denoted as $D_{1,I}$.

Definition 4 A pixel (x,y) is called a *1-D partial degenerate pixel* of an image I if it is not a 1-D or 2-D degenerate pixel of I, and there is a neighborhood $O(x,y)$ of (x,y) and a line $\mathbf{l}$ containing (x,y) such that the image intensity function of I is a constant on one side of $\mathbf{l}$ in $O(x,y)$ starting from (x,y). The set of all 1-D partial degenerate pixels of I is denoted as $PD_{1,I}$.

For a 1-D degenerate pixel (x,y) of the image I, by definition, there is a neighborhood $O(x,y)$ of (x,y) and a line $\mathbf{l}$ containing (x,y) such that the image intensity function of I is a constant on $\mathbf{l}$ in $O(x,y)$. Assume that $\mathbf{l}$ has a normal direction of $\mathbf{a}$ and the intensity function of I is continuous in $O(x,y)$. Then, the 2-D image intensity function of I is asymptotically a 1-D function in $O(x,y)$, and it has the expression

$$I(\mathbf{s}) = \psi(\mathbf{a}'\mathbf{s}) + o(\|\mathbf{s} - (x,y)\|), \qquad \text{for } \mathbf{s} \in O(x,y),$$

where ψ is a non-degenerate univariate function, and $\|\cdot\|$ is the Euclidean norm. In such cases, it is obvious that the geometrical transformation $\mathbf{T}(x,y)$ in the normal direction $\mathbf{a}$ is possible to define, but it is not uniquely defined in the tangent direction of $\mathbf{l}$ as long as $\mathbf{T}(O_{\mathbf{l}}(x,y))$ contains at least two points. Similar to Theorem 1, under some regularity conditions, we can establish the results that $D_{1,I}$ consists of a countable number of disjoint, open, line segments, and the set of endpoints of the line segments is just $PD_{1,I}$.

3 Proposed Intensity-Based Image Registration Method

From the discussion in Section 2, the geometrical transformation $\mathbf{T}$ is not well defined in $D_{2,R}$, the set of all 2-D degenerate pixels of R. At a 1-D degenerate pixel of R, it is only partially defined. In this section, we describe our proposed IBIR method in detail, which focuses mainly on all 2-D non-degenerate pixels of R, the set of which is denoted as $\overline{D}_{2,R}$. Our description is divided into three parts. In Section 3.1, a local smoothing procedure is suggested for detecting all 2-D non-degenerate pixels of R. Based on the detected 2-D non-degenerate pixels, an estimator of the geometrical transformation $\mathbf{T}$ is defined in Section 3.2. Then, in Section 3.3, some practical guidelines are given regarding selection of some procedure parameters used in our method.

As mentioned in Section 2, the IR problem has a boundary problem. Although the boundary problem here is different from the boundary problem that is common to most local smoothing procedures (cf., Qiu 2005, Section 2.3), some methods proposed in the local smoothing literature to extend the design space, including the periodic extension, symmetric extension, and so forth, can also be considered here. In all numerical examples in Section 4, we adopt the symmetric extension approach as follows. Suppose that we want to extend the design space of the image R from $\Omega_R = [0,1] \times [0,1]$ to $\Omega_{R,\epsilon} = [-\epsilon, 1+\epsilon] \times [-\epsilon, 1+\epsilon]$, with $\epsilon > 0$. Then, by the symmetric extension, we define $R(x,y) = R(-x,y)$ when $x \in [-\epsilon, 0)$ and $y \in [0,1]$. In other boundary regions, the extension is defined similarly.

3.1 Detection of the 2-D non-degenerate pixels

To detect 2-D non-degenerate pixels, there are different possible approaches. One natural approach is to make a judgment based on the sample variance of the observed image intensities in a neighborhood of a given pixel. If the sample variance is large, then it is an indication that the given pixel might be a 2-D non-degenerate pixel. Otherwise, it is unlikely a 2-D non-degenerate pixel. An alternative approach is to investigate the change of the underlying image intensity surface along its gradient direction at the given pixel. Because the image intensity surface increases the fastest along the gradient direction, if its change along that direction is small, then it is an indication that the given pixel might be a 2-D degenerate pixel. The latter approach focuses on the directional change which is about the variability of a univariate directional function, while the former approach checks the variability of a bivariate function. Since the two approaches use the same amount of

data, the latter approach would be statistically more efficient, compared to the former approach. Further, by investigating the change of the image intensity surface along the direction orthogonal to the gradient direction, which is called the normal direction hereafter, it is possible to detect 1-D degenerate pixels. Because of these two benefits, we adopt the latter approach in this paper, which is described below.

At a given pixel $(x, y) \in \Omega_R$, let us consider a circular neighborhood with radius h_n^*, denoted as $O(x, y; h_n^*)$. In $O(x, y; h_n^*)$, a local plane is fitted, using the local linear kernel (LLK) smoothing procedure accomplished by the following minimization problem:

$$\min_{a,b,c} \sum_{i,j=1}^{n} \{Z_R(x_i, y_j) - [a + b(x_i - x) + c(y_j - y)]\}^2 K^* \left(\frac{x_i - x}{h_n^*}, \frac{y_j - y}{h_n^*} \right), \tag{3}$$

where $K^*(u, v)$ is a 2-D density kernel function with the support $\{(u, v) : \sqrt{u^2 + v^2} \leq 1\}$. The solution of (3) is denoted as $(\widehat{a}(x, y), \widehat{b}(x, y), \widehat{c}(x, y))$, and it is the LLK estimate of $(R(x, y), R'_x(x, y), R'_y(x, y))$. Then, the gradient vector $\mathbf{G}(x, y) = (R'_x(x, y),\ R'_y(x, y))'$ of R at (x, y) can be estimated by $\widehat{\mathbf{G}}(x, y) = (\widehat{b}(x, y), \widehat{c}(x, y))'$, and the normal vector $\mathbf{n}(x, y) = (-R'_y(x, y), R'_x(x, y))'$ can be estimated by $\widehat{\mathbf{n}}(x, y) = (-\widehat{c}(x, y), \widehat{b}(x, y))$, as demonstrated in Figure 3(a).

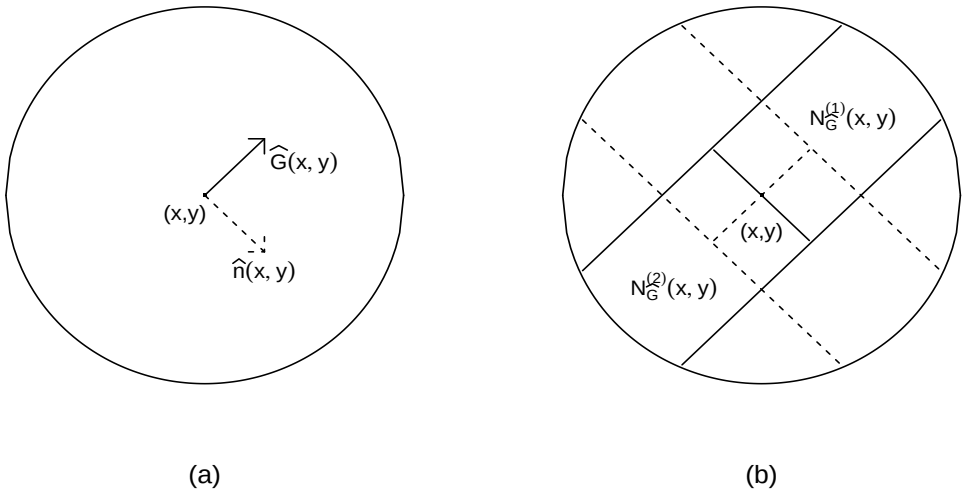


Figure 3: (a) The estimated gradient vector $\widehat{\mathbf{G}}(x, y)$ and the estimated normal vector $\widehat{\mathbf{n}}(x, y)$ are obtained by the local linear kernel smoothing procedure (3) in the neighborhood $O(x, y; h_n^*)$. (b) The two one-sided narrow bands $N_{\widehat{\mathbf{G}}}^{(1)}(x, y)$ and $N_{\widehat{\mathbf{G}}}^{(2)}(x, y)$ along the estimated gradient direction $\widehat{\mathbf{G}}(x, y)$ (the bands formed by solid lines), and the two one-sided narrow bands along the estimated normal direction $\widehat{\mathbf{n}}(x, y)$ (the bands formed by dashed lines).

To investigate the directional change of the underlying image intensity surface of R along the

estimated gradient direction $\widehat{\mathbf{G}}(x,y)$ at (x,y), we consider a narrow band in $O(x,y;h_n^*)$ formed by two lines along the direction of $\widehat{\mathbf{G}}(x,y)$, as demonstrated in Figure 3(b) by the two longer solid lines. The two lines are on two different sides of (x,y), and the Euclidean distance from (x,y) to each line is $\rho_n h_n^*$ where $\rho_n \in (0, 0.5]$ is a positive number. Then, the narrow band is divided into two parts by a line passing (x,y) in the normal direction (cf., the shorter solid line in Figure 3(b)). The two one-sided parts are denoted as $N^{(1)}_{\widehat{\mathbf{G}}}(x,y)$ and $N^{(2)}_{\widehat{\mathbf{G}}}(x,y)$, respectively. Then, we compute the local constant kernel estimator of $R(x,y)$ from the observed image intensities in each part (i.e., the solution to a of (3) when b and c are set to be 0 and when (x_i, y_j) are constrained to the region considered). The two one-sided estimators are

$$\begin{aligned} a^{(1)}_{\widehat{\mathbf{G}}}(x,y) &= \frac{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} Z_R(x_i,y_j) K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)} \\ a^{(2)}_{\widehat{\mathbf{G}}}(x,y) &= \frac{\sum_{(x_i,y_j)\in N^{(2)}_{\widehat{\mathbf{G}}}(x,y)} Z_R(x_i,y_j) K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(2)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}. \end{aligned} \tag{4}$$

Then, the difference between $a^{(1)}_{\widehat{\mathbf{G}}}(x,y)$ and $a^{(2)}_{\widehat{\mathbf{G}}}(x,y)$ can be used for measuring the directional change of the image intensity surface of R along the estimated gradient direction $\widehat{\mathbf{G}}(x,y)$. So, the pixel (x,y) is detected as a 2-D degenerate pixel of R if

$$U_n(x,y) = \left| a^{(1)}_{\widehat{\mathbf{G}}}(x,y) - a^{(2)}_{\widehat{\mathbf{G}}}(x,y) \right| \le u_n, \tag{5}$$

where $u_n > 0$ is a threshold value. Otherwise, it is detected as a 2-D non-degenerate pixel of R. The set of all detected 2-D degenerate pixels of R are denoted as $\widehat{D}_{2,R}$, and the set of all detected 2-D non-degenerate pixels of R are denoted as $\widehat{\overline{D}}_{2,R}$. It should be noted that the difference between two one-sided local constant kernel estimators is used in (5), instead of the difference between two one-sided local linear kernel estimators, because the latter could be small even when the image intensity surface of R changes fast along the direction of $\widehat{\mathbf{G}}$.

Among all detected 2-D non-degenerate pixels of R, we can further detect 1-D degenerate pixels of R as follows. Let $V_n(x,y)$ be defined in the same way as $U_n(x,y)$, except that $N^{(1)}_{\widehat{\mathbf{G}}}(x,y)$ and $N^{(2)}_{\widehat{\mathbf{G}}}(x,y)$ in (4) are replaced by $N^{(1)}_{\widehat{\mathbf{n}}}(x,y)$ and $N^{(2)}_{\widehat{\mathbf{n}}}(x,y)$, respectively, where $N^{(1)}_{\widehat{\mathbf{n}}}(x,y)$ and $N^{(2)}_{\widehat{\mathbf{n}}}(x,y)$ are two one-sided narrow bands in $O(x,y;h_n^*)$ along the normal direction $\mathbf{n}(x,y)$ (cf., the two one-sided bands shown in Figure 3(b) by the dashed lines). Then, for a given pixel $(x,y) \in \widehat{\overline{D}}_{2,R}$, it is detected as a 1-D degenerate pixel of R if

$$V_n(x,y) \le v_n, \tag{6}$$

where $v_n > 0$ is a threshold value. Otherwise, it is detected as a 1-D non-degenerate pixel of R. The set of all detected 1-D degenerate pixels of R are denoted as $\widehat{D}_{1,R}$, and the set of all detected 1-D non-degenerate pixels of R are denoted as $\widehat{\overline{D}}_{1,R}$. Obviously, $\widehat{D}_{1,R} \bigcup \widehat{\overline{D}}_{1,R} = \widehat{\overline{D}}_{2,R}$.

The procedures (5) and (6) for detecting 1-D and 2-D degenerate pixels are based on certain ideas used in the research area of jump regression analysis (JRA) (cf., Qiu 2005). For some recent developments in JRA, see Joo and Qiu (2009), Mukherjee and Qiu (2011), Qiu (2009), Sun and Qiu (2007), and the references cited therein.

3.2 Estimation of the mapping transformation T

As discussed in Section 1, the geometrical mapping transformation $\mathbf{T}$ is not well defined at a 2-D degenerate pixel of R. So, we first define $\mathbf{T}$ at the detected 2-D non-degenerate pixels of R. Let $\widehat{\overline{D}}_{2,M}$ be the set of all detected 2-D non-degenerate pixels of M. Then, for a given detected 2-D non-degenerate pixel $(x, y) \in \widehat{\overline{D}}_{2,R}$, define

$$\widehat{\mathbf{T}}(x, y) = \arg \min_{(x_i, y_j) \in O(x,y;r_n) \bigcap \widehat{\overline{D}}_{2,M}} WSSD\left((x, y), (x_i, y_j); h_n\right), \tag{7}$$

where

$$WMSD\left((x, y), (x_i, y_j); h_n\right) = \frac{\sum_{\sqrt{s^2+t^2} \le h_n} \left[Z_R(x+s, y+t) - Z_M(x_i+s, y_j+t)\right]^2 K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2} \le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \tag{8}$$

is the weighted mean squared difference (WMSD) between the observed intensities of the reference image R in the neighborhood $O(x, y; h_n)$ and the observed intensities of the moved image M in the neighborhood $O(x_i, y_j; h_n)$. In (7) and (8), r_n and h_n are two positive radiuses, K is a 2-D density kernel function, and the two summations on the right-hand-side of (8) are constrained to discrete pixel locations, although it is not explicit in notation. Namely, the summations are over all (s, t) satisfying the conditions that $\sqrt{s^2 + t^2} \le h_n$ and $\{(x + s, y + t)\}$ (or equivalently, $\{(x_i + s, y_j + t)\}$) are pixel locations.

In (7) and (8), for a given detected 2-D non-degenerate pixel (x, y) of R, we search all the detected 2-D non-degenerate pixels (x_i, y_j) of M in the circular neighborhood $O(x, y; r_n)$, such that the WMSD value defined in two corresponding neighborhoods $O(x, y; h_n)$ and $O(x_i, y_j; h_n)$ of the two images reaches the minimum. Therefore, by this procedure, our major goal is to best

match the two images around the two sets of detected 2-D non-degenerate pixels using the local smoothing procedure (8), which is consistent with the ultimate goal of most IR applications to best match the two images in question. Because of the local smoothing nature of our procedure, a term regularizing the smoothness of $\mathbf{T}$, which is commonly seen in global smoothing procedures (e.g., the smoothing spline procedure), is not included explicitly in (7). As a matter of fact, the procedure (7)–(8) implicitly assumes that $\mathbf{T}$ can be approximated by a constant transformation in the local neighborhood $O(x, y; h_n)$. There are several benefits to use a local smoothing procedure here. One major benefit is that it does not require global smoothing assumptions on $\mathbf{T}$; thus, it is quite flexible and can handle certain cases when $\mathbf{T}$ is not uniformly smooth in the entire design space. Another major benefit is that its computation is relatively simple. See Chapter 2 of Qiu (2005) for an introduction about local smoothing and global smoothing procedures and their major properties.

Although the transformation $\mathbf{T}$ is not well defined at a 2-D degenerate pixel of R, for the convenience of comparing R with M, in practice it is helpful to define $\mathbf{T}$ in the entire design space Ω_R. To this end, let $(\widetilde{x}, \widetilde{y})$ be the pixel in $\widehat{\widehat{D}}_{2,R}$ that is closest to a given pixel $(x, y) \in \widehat{D}_{2,R}$, $\mathbf{q} = \widehat{\mathbf{T}}(\widetilde{x}, \widetilde{y}) - (\widetilde{x}, \widetilde{y})$, and $\widetilde{\mathbf{T}}(x, y) = (x, y) + \mathbf{q}$. Then, we define

$$\widehat{\mathbf{T}}(x, y) = \begin{cases} \widetilde{\mathbf{T}}(x, y), & \text{if } WMSD((x, y), \widetilde{\mathbf{T}}(x, y); \widetilde{h}_n) \leq WMSD((x, y), (x, y); \widetilde{h}_n) \\ (x, y), & \text{otherwise,} \end{cases} \tag{9}$$

where $\widetilde{h}_n \geq 0$ is a bandwidth that could be different from h_n. In (9), $\widetilde{\mathbf{T}}(x, y)$ is determined by the estimated transformation at $(\widetilde{x}, \widetilde{y})$. So, by (9), the estimator $\widehat{\mathbf{T}}(x, y)$ at a 2-D degenerate pixel (x, y) of R is defined to be the identity transformation (i.e., $\widehat{\mathbf{T}}(x, y) = (x, y)$) or the estimated transformation $\widetilde{\mathbf{T}}(x, y)$ defined at the nearest 2-D non-degenerate pixel $(\widetilde{x}, \widetilde{y})$, whichever gives the smaller WMSD value.

3.3 Practical guidelines for choosing procedure parameters

In our proposed IBIR procedure, there are some procedure parameters. Proper selection of these parameters is important because they could affect the performance of the proposed procedure in a substantial way. For the IR problem, procedure parameters can always be chosen such that a registration performance measure (e.g., the RRMS measure defined in Section 5 below) is minimized. However, an exhaustive search could take much computing time. For one set of parameter values,

it takes about 10 seconds for the proposed procedure to register two 256×256 images on a PC of Intel(R) 2.40GHz Core(TM)2 Duo CPU with 2982 MByte memory. We have tried many different images of the resolutions 128×128, 256×256, and 512×512. Next, we provide some empirical guidelines for choosing the parameters, based on our numerical experience. When reading these guidelines, readers are reminded that Ω_R is assumed to be $[0,1] \times [0,1]$. Also, selection of procedure parameters always depends the specific images under study. So, in practice, when choosing the parameters, we can try the recommended values first, and then make some adjustments around the recommended values when necessary. In that way, much computation could be saved in many cases.

Selection of h_n^* and ρ_n: In detection of 2-D and 1-D degenerate pixels, there are two parameters h_n^* and ρ_n involved (cf., (3)–(5)). h_n^* is a bandwidth of the local smoothing procedure (3). If it is chosen too small, then the procedure (3)–(5) would be sensitive to noise. On the other hand, if h_n^* is chosen too large, then some local image structures would be smoothed out by (3), and the procedure (3)–(5) would be insensitive to these local structures. We recommend choosing h_n^* from the values $\{0.015, 0.02, 0.025\}$. The parameter ρ_n controls the widths of the one-sided narrow bands shown in Figure 3(b). The IBIR results would be reasonably good if we choose ρ_n to be one of the two values $\{0.2, 0.5\}$.

Selection of u_n: The threshold value u_n used in (5) controls the amount of detected 2-D non-degenerate pixels. If it is chosen larger, then there are fewer detected 2-D non-degenerate pixels. Our numerical experience shows that this parameter can be chosen such that $|\widehat{\overline{D}}_{2,R}|/|\Omega_R| \in [0.1, 0.15]$, where $|\widehat{\overline{D}}_{2,R}|$ denotes the number of detected 2-D non-degenerate pixels and $|\Omega_R|$ denotes the total number of pixels in Ω_R.

Selection of r_n, h_n and $\widetilde{h}_n$: The size of r_n is related to the expected magnitude of the geometrical transformation (i.e., $d_E(\mathbf{T}(x,y),(x,y))$). In most IR problems, we find that it is good enough to choose $r_n \in [0.05, 0.25]$. If the expected magnitude of the geometrical transformation is larger than 0.25, then we can consider to combine the proposed IBIR procedure with a global smoothing method, as discussed in the second last paragraph of Section 1. For h_n, we can choose it to be $\min\{r_n, s_n\}$, where s_n is a number between 0.01 and 0.1. The parameter $\widetilde{h}_n$ should be chosen small. Usually, it is good enough to choose $\widetilde{h}_n$ such that $O(x,y;\widetilde{h}_n)$ only contains t_n pixels with $t_n \in \{1,5,9\}$. That is, the pixel itself (1), or that pixel and those directly adjacent (5), or the pixel and those directly or diagonally adjacent (9).

4 Statistical Properties of the Proposed Method

In this section, we discuss some statistical properties of the proposed IBIR procedure. Theorem 2 below establishes the almost sure consistency of the detected 2-D and 1-D degenerate pixels, while Theorem 3 builds the pointwise consistency of the estimated geometrical transformation $\widehat{\mathbf{T}}$ at 2-D non-degenerate pixels of R. First, we discuss the properties of the detected 2-D degenerate pixels and the detected 1-D degenerate pixels of the reference image R. To this end, we introduce the following notations where all pointsets are actually subsets of the grid $\{(x_i, y_j), i, j = 1, 2, \ldots, n\}$, although it is not made explicit in notation:

- For a connected 2-D point set $A \subseteq \Omega_R$, we use $A(h_n^*)$ to denote the set of its interior points $\{(x, y) : (x, y) \in A, d_E((x, y), \partial A) \leq h_n^*\}$ that are within h_n^* of its boundary ∂A, and $A_{h_n^*}$ to denote the set of its boundary points $\{(x, y) : (x, y) \in \Omega_R, d_E((x, y), A) \leq h_n^*\}$ that are within h_n^* of A, where d_E denotes the Euclidean distance.

- $\mathbf{G}_{2,\epsilon_n} = \{(x, y) : (x, y) \in \overline{D}_{2,R}, |R'_{\mathbf{G}}(x, y)| < \epsilon_n$, where $R'_{\mathbf{G}}(x, y)$ is the directional derivative of R at (x, y) in the gradient direction $\mathbf{G}(x, y)$, and $\epsilon_n > 0$ is a number that converges to 0 as $n \to \infty$.

- $\mathbf{n}_{1,\epsilon_n} = \{(x, y) : (x, y) \in \overline{D}_{2,R} \backslash D_{1,R}, |R'_{\mathbf{n}}(x, y)| < \tau_n\}$, where $R'_{\mathbf{n}}(x, y)$ is the directional derivative of R at (x, y) in the normal direction $\mathbf{n}(x, y)$, and $\tau_n > 0$ is a number that converges to 0 as $n \to \infty$.

- J denotes the set of step edge curves of R (i.e., the image intensity function of R has jumps at points in J), and $J1$ denotes the set of roof edge curves of R (i.e., the first-order derivatives of R have jumps at points in $J1$).

- $\Omega_2(R, T, n) = \Omega_{R,T} \setminus \left(\mathbf{G}_{2,\epsilon_n} \bigcup J_{h_n^*} \bigcup J1_{h_n^*}\right)$,

- $\Omega_1(R, T, n) = \overline{D}_{2,R} \setminus \left(\mathbf{n}_{1,\epsilon_n} \bigcup J_{h_n^*} \bigcup J1_{h_n^*}\right)$,

- $J_n = \{(x, y) : (x, y) \in J$, (x, y) is not a crossing point of several step edge curves, the edge curve containing (x, y) has a tangent line at (x, y), and the jump size of R at (x, y) is at least $\epsilon_n h_n^*\}$

- $J1_n = \{(x, y) : (x, y) \in J1$, (x, y) is not a crossing point of several roof edge curves, the roof

edge curve containing (x, y) has a tangent line at (x, y), and the jump size in the slope of R at (x, y) is at least ϵ_n}.

- The Hausdorff distance between two point sets A and B are defined to be $d_H(A, B) = \max\left\{\sup_{(x,y)\in A}\inf_{(x',y')\in B} d_E((x, y), (x', y')), \sup_{(x,y)\in B}\inf_{(x',y')\in A} d_E((x, y), (x', y'))\right\}$.

Theorem 2 *Assume that there are only a countable number of disjoint, connected, non-trivial subregions* $\{D_{2,R}^{(j)}, j = 1, 2, \ldots\}$ *of* Ω_R*, in each of which the image intensity function of* R *is a constant. The image intensity function* R *has continuous second order derivatives in* $\Omega\backslash(J\bigcup J1)$*. The noise satisfies the condition* $E|\varepsilon_R(x_1, y_1)|^3 < \infty$*. The kernel function* K^* *is Lipschitz-1 continuous in its support. The bandwidth* h_n^**, the quantities* ρ_n *and* ϵ_n*, and the threshold value* u_n *satisfy the conditions that (i)* $h_n^* = o(1)$*, (ii)* $\rho_n = o(1)$*, (iii)* $1/(nh_n^*\sqrt{\rho_n}) = o(1)$*, (iv)* $\log^2 n/(nh_n^{*3}\rho_n^{3/2}) = O(1)$*, (v)* $u_n/(\epsilon_n h_n^*) = o(1)$*, (vi)* $h_n^{*2}/u_n = o(1)$*, and (vii)* $\epsilon_n = o(1)$*. Then, we have*

$$d_H\left(\widehat{D}_{2,R}\bigcap\Omega_2(R, T, n), D_{2,R}\bigcap\Omega_2(R, T, n)\right) = o(1), \qquad a.s., \tag{10}$$

and when n *is large enough,*

$$\left(J_n\bigcup J1_n\right) \subseteq \widehat{\overline{D}}_{2,R}, \qquad a.s. \tag{11}$$

If we further assume that $v_n = o(1), v_n/(\tau_n h_n^*) = o(1)$*, and* $h_n^{*2}/v_n = o(1)$*, then*

$$d_H\left(\widehat{D}_{1,R}\bigcap\Omega_1(R, T, n), D_{1,R}\bigcap\Omega_1(R, T, n)\right) = o(1), \qquad a.s. \tag{12}$$

Intuitively, when n gets larger, the point set $\Omega_2(R, T, n)$ converges to the point set $\Omega_{R,T}\backslash(P\bigcup J\bigcup J1)$, where P denotes the point set of all isolated points in Ω_R at which the first order derivatives of R are 0. Note that the 2-D Lebesgue measure of $P\bigcup J\bigcup J1$ is 0. Therefore, the equation (10) says that $\widehat{D}_{2,R}$ converges to $D_{2,R}$ almost surely in Hausdorff distance, after a point set with Lebesgue measure 0 is excluded. Expression (11) says that most points in J and $J1$ are actually detected as 2-D non-degenerate pixels. Equation (12) builds the almost sure consistency of the detected 1-D degenerate pixels in Hausdorff distance.

For the estimated geometrical transformation $\widehat{\mathbf{T}}(x, y)$, because of the ill-posed nature of the IR problem, we only discuss the properties of $\widehat{\mathbf{T}}(x, y)$ when $(x, y) \in \Omega_{R,\mathbf{T}}\bigcap\overline{D}_{2,R}$. Namely, we confine ourselves in cases when (x, y) is a 2-D non-degenerate pixel and when it is not in the boundary region of the design space of R.

Theorem 3 *Besides the conditions in Theorem 2, assume that* (x, y) *is a given pixel in* $(\Omega_{R,\mathbf{T}} \bigcap \overline{D}_{2,R}) \backslash (D_{1,R} \bigcup P \bigcup J \bigcup J1)$, $r_n > d_E((x, y), \mathbf{T}(x, y))$, $h_n = o(1)$, $1/(nh_n) = o(1)$, *the kernel function* K *is a Lipschitz-1 continuous and circularly symmetric function in its unit circular support,* $E|\varepsilon_R(x_1, y_1)|^6 < \infty$, $E|\varepsilon_M(x_1, y_1)|^6 < \infty$, *the two components of* $\mathbf{T}$ *have continuous first order derivatives in* $O(x, y, h_n)$, *for any point* $(x', y') \in (O(x, y, r_n) \backslash O(\mathbf{T}(x, y), h_n))$ *the image intensity function of* M *satisfies the condition that there is a subregion of area* $q_n > 0$ *in* $O(\mathbf{T}(x, y), h_n)$ *such that the absolute differences between the values of* M *in that subregion of* $O(\mathbf{T}(x, y), h_n)$ *and the values of* M *in the corresponding subregion of* $O(x', y', h_n)$ *are at least* τ_n, q_n *and* τ_n *satisfy the conditions that* $h_n^4/(q_n\tau_n^2) = o(1)$ *and* $h_n \log(n)/(nq_n\tau_n^2) = o(1)$, *then we have*

$$d_E(\widehat{\mathbf{T}}(x, y), \mathbf{T}(x, y)) = O(h_n), \qquad a.s. \tag{13}$$

By (13) in Theorem 3, if (x, y) is a 2-D non-degenerate pixel and it is a 1-D non-degenerate pixel as well, then $\widehat{\mathbf{T}}(x, y)$ is an almost surely consistent estimator of $\mathbf{T}(x, y)$ when r_n is chosen large enough so that $O(x, y, r_n)$ contains $\mathbf{T}(x, y)$ in the image of M, in cases when there is no pixel in $O(x, y, r_n) \backslash O(\mathbf{T}(x, y), h_n)$ such that M is almost identical in $O(\mathbf{T}(x, y), h_n)$ and in the corresponding neighborhood of that pixel (i.e., $\mathbf{T}(x, y)$ is not identifiable in $O(x, y, r_n)$) and when other regularity conditions hold. In our proposed IBIR procedure, r_n should be chosen beforehand. By Theorem 3, in order to estimate $\mathbf{T}$ properly in the entire design space, r_n should be chosen large. However, if r_n is chosen too large, there will be more pixels in the design space that are not identifiable. So, there is a trade-off between the two considerations. In practice, for a given IR problem, if we believe that the geometrical transformation $\mathbf{T}$ is large at certain places in the design space, then we can first apply a global smoothing approach to account for part of $\mathbf{T}$ with large magnitude and then our proposed IBIR procedure is used afterwards to make local adjustments, as described in the second last paragraph of Section 1.

5 Numerical Study

In this section, we present some numerical results concerning the performance of the proposed IBIR procedure, in comparison with several existing IBIR procedures. The existing procedures considered are the directly manipulated free-form deformation (DMFFD) method (Tustison et

al. 2009) and the symmetric diffeomorphic image normalization (SyN) method (Avants et al. 2008). Both methods are based on global smoothing; they provide estimators of the geometrical transformation $\mathbf{T}$ in the entire design space Ω_R, even when $\mathbf{T}$ is degenerate in certain regions of Ω_R. Besides certain global smoothness conditions on $\mathbf{T}$, they do not require $\mathbf{T}$ to follow a parametric model, such as the rigid body and affine motion models. Results of both methods depend on the similarity metric that they use. As suggested in the related papers, three similarity metrics, including mean squared difference (MSD), cross-correlation (CC), and mutual information (MI), are used respectively in the DMFFD method, and two similarity metrics, including CC and the pure cross correlation (PCC), are used respectively in the SyN method. Therefore, a total of five existing procedures are considered here, which are denoted as DMFFD-MSD, DMFFD-CC, DMFFD-MI, SyN-CC, and SyN-PCC. These existing procedures are implemented using the release 1.9 of the software package ANTS, available at http://www.picsl.upenn.edu/ANTS/. Our proposed nonparametric IBIR procedure is denoted NEW. In all numerical examples presented in this section, no global smoothing methods based on parametric transformation families (e.g., the rigid-body and affine motion transformation families) are used before executing NEW, because the magnitude of $\mathbf{T}$ is quite small in these examples and the global smoothing methods are unnecessary. To evaluate the performance of all six procedures, we use two popular measures in the IR literature, including the root residual mean squares (RRMS) and the cross correlation (CC). When evaluating an estimator $\widehat{\mathbf{T}}$ of the geometrical transformation $\mathbf{T}$, RRMS is defined to be

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

Basically, RRMS is the Euclidean distance between the observed reference image $\{Z_R(x_i, y_j)\}$ and the restored reference image $\{Z_M(\widehat{\mathbf{T}}(x_i, y_j))\}$. Therefore, if its value is smaller, then the registration is regarded better. The CC measure is defined to be the Pearson's sample correlation coefficient of the bivariate data $\{(Z_R(x_i, y_j), Z_M(\widehat{\mathbf{T}}(x_i, y_j)))\}$. Intuitively, if the estimator $\widehat{\mathbf{T}}$ is good, then $Z_M(\widehat{\mathbf{T}}(x, y))$ would be close to $Z_R(x, y)$. Consequently, the CC measure would be close to its maximum value 1. So, by this measure, the registration is better if the CC value is larger, although this measure cannot reflect possible scale difference between $Z_R(x, y)$ and $Z_M(\widehat{\mathbf{T}}(x, y))$. In procedure NEW, the kernel functions $K^*(u, v)$ and $K(u, v)$ are both chosen to be the truncated bivariate Gaussian density function with support $\{(u, v) : u^2 + v^2 \leq 1\}$. The design spaces of the two images to register are extended as described at the beginning of Section 3 with $\epsilon = 0.3$. Its other procedure parameters are chosen as suggested in Section 3.3 for the smallest RRMS values.

For all five competitors, we try all possible values of their procedure parameters and use the ones resulting in the smallest RRMS values.

We first consider an artificial reference image shown in Figure 1(c) as a 3-D surface, and in Figure 1(a) as an image. We can see that it has several step edge curves at three corners and two circular roof edge curves in the middle. The two circular regions in the middle constitute the set of the true 2-D non-degenerate pixels, the several step and roof edge curves constitute the set of the true 1-D non-degenerate pixels, and the flat region is the set of the true 2-D degenerate pixels. Figure 1(b) presents the moved image in which the white circular region in the middle has moved downwards by 0.06, the dark circular region has moved upwards by 0.06, the white triangular region at the upper-left corner has moved upwards by 0.04 and to the left by 0.04 as well, and the remaining parts are unchanged. The true geometrical transformation $\mathbf{T}$ is shown in Figure 1(d).

In this example, we consider two image resolutions $n = 128$ and 256, and four noise levels $\sigma_R = \sigma_M$ =0, 1, 5, and 10. For each combination of n and σ_R, we compute the RRMS and the corresponding CC values of all six procedures, based on 10 replications. These values are shown in Figure 4. From the plots of the figure, it can be seen that our proposed procedure NEW outperforms the five competitors in terms of both RRMS and CC, and the outperformance is in a quite large margin in all cases. In the case when $n = 256$ and $\sigma_R = 5$, one set of restored reference images, defined to be $Z_M(\widehat{\mathbf{T}}(x, y))$, and the corresponding difference images, defined to be $Z_R(x, y) - Z_M(\widehat{\mathbf{T}}(x, y))$, of the six procedures are shown in Figure 5. From the images in this figure, it can be seen that (i) procedure NEW restores the reference image well, (ii) procedures DMFFD-MSD, DMFFD-CC, and SyN-CC can restore the two circular regions in the middle to their original positions well, while procedures DMFFD-MI and SyN-PCC cannot satisfactorily restore their original positions, and (iii) all five competing procedures have difficulty to handle boundary regions. The detected 2-D non-degenerate pixels of the reference and moved images from one simulation run in cases when $\sigma_R = \sigma_M$ =0 and 5 are shown in Figure 6, from which it can be seen that the true 2-D non-degenerate pixels are almost perfectly detected when $\sigma_R = \sigma_M$ =0, and there are some false detected 2-D non-degenerate pixels when $\sigma_R = \sigma_M$ =5. The estimated geometrical transformation $\mathbf{T}$ is shown in the two plots of Figure 7 in cases when $\sigma_R = \sigma_M = 0$ and 5. It can be seen that it is quite close to the true $\mathbf{T}$ (cf., Figure 1(d)), especially in the case when $\sigma_R = \sigma_M = 0$. Note that certain pixels around the centers of the two central circular regions are not detected as 2-D non-degenerate pixels because the true image intensity surfaces there are

quite flat.

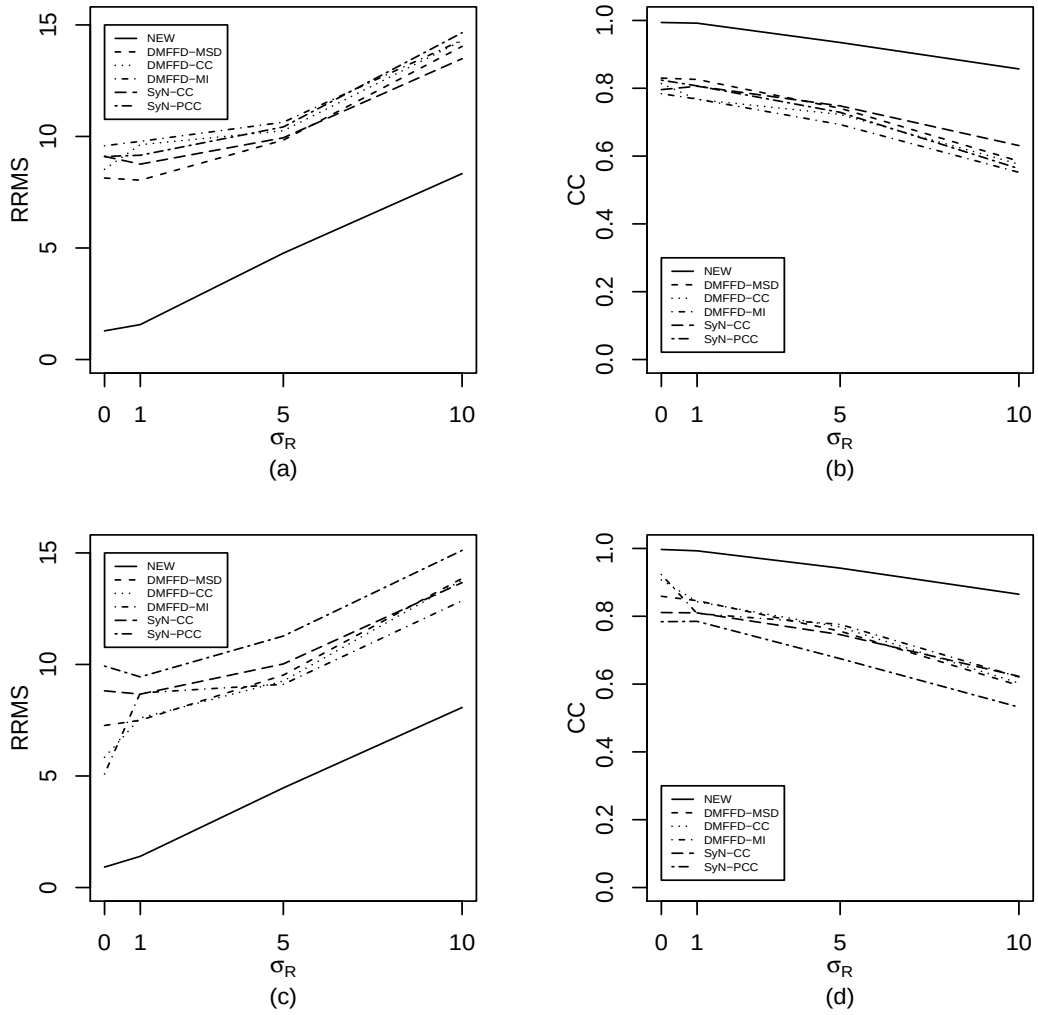


Figure 4: RRMS and CC values of the six IBIR methods NEW, DMFFD-MSD, DMFFD-CC, DMFFD-MI, SyN-CC, and SyN-PCC. (a) RRMS values when $n = 128$, (b) CC values when $n = 128$, (c) RRMS values when $n = 256$, and (d) CC values when $n = 256$.

In the above example, the pointwise noise is assumed independent at different pixels. In practice, the noise could be spatially correlated. In the supplementary file, another example is considered, in which spatially correlated noise is added to the two images shown in Figure 1(a)-(b) and all other aspects of the experiment are the same as those in the above example. The computed RRMS and CC values of all six procedures show that NEW outperforms all other five procedures in a quite large margin in this example as well in terms of both RRMS and CC.

Next, we consider two real IR problems. The reference and moved images of the first problem are shown in Figure 8, which are the satellite images taken in 1990 and 1999, respectively, in the

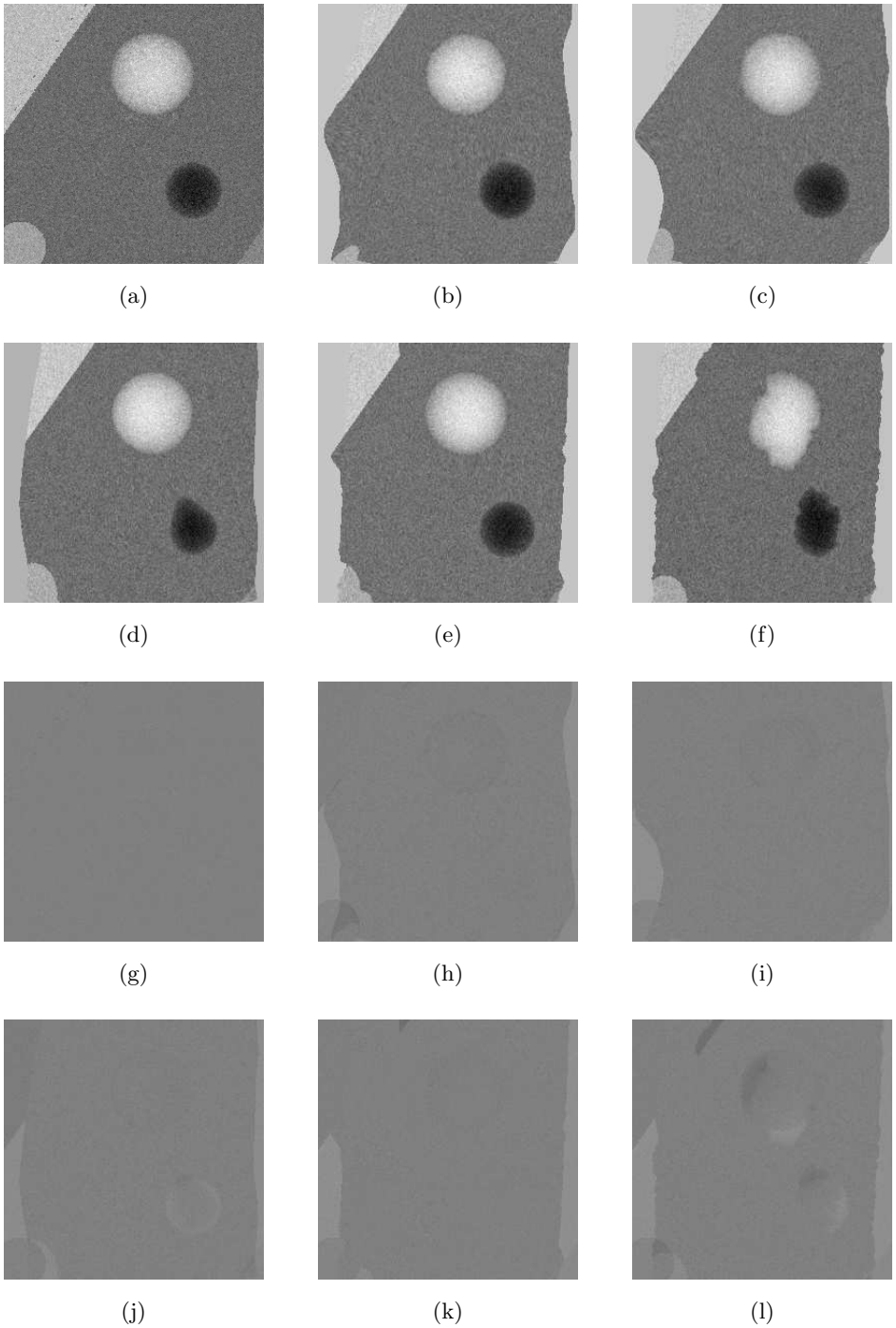


Figure 5: Restored reference images (plots (a)–(f)), defined to be $Z_M(\widehat{\mathbf{T}}(x,y))$, and the corresponding difference images (plots (g)–(l)), defined to be $Z_R(x,y) - Z_M(\widehat{\mathbf{T}}(x,y))$, of the procedures NEW, DMFFD-MSD, DMFFD-CC, DMFFD-MI, SyN-CC, and SyN-PCC, respectively, in the case when $n = 256$, $\sigma_R = 5$, and the original reference and moved images are shown in Figure 1(a)-(b). The gray scale of all images in this figure is kept the same.

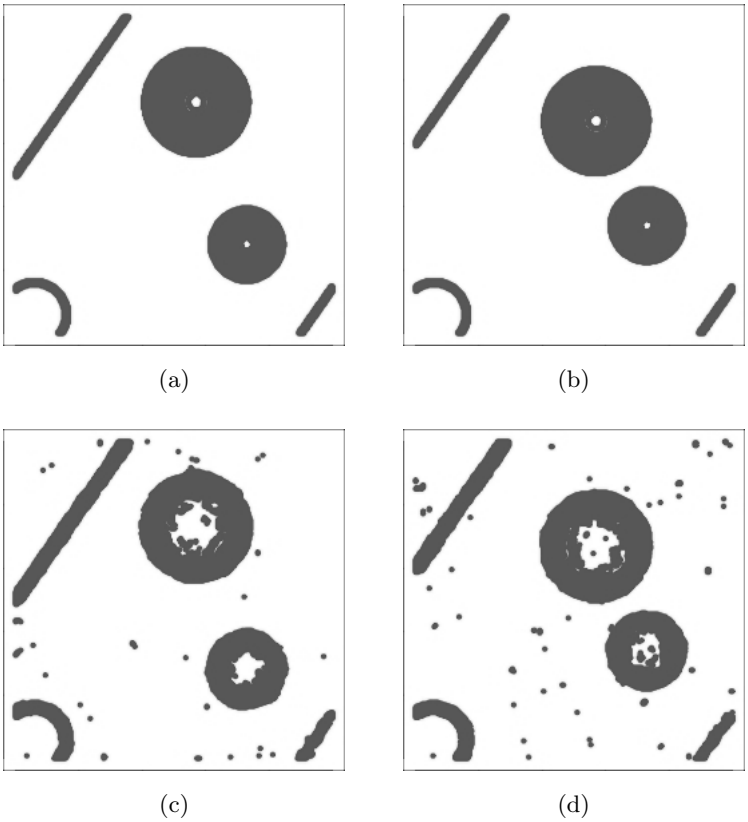
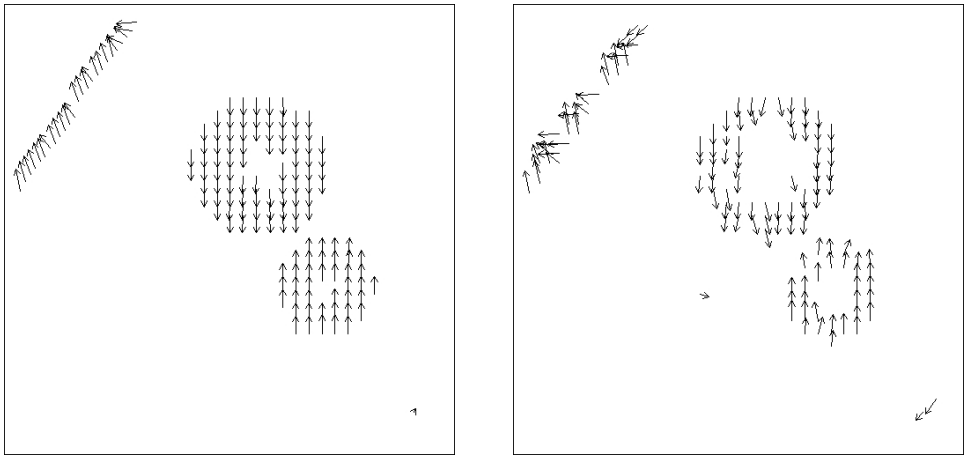


Figure 6: Detected 2-D non-degenerate pixels of the proposed method NEW in the example of Figure 1. (a) Reference image when $\sigma_R = 0$, (b) moved image when $\sigma_M = 0$, (c) reference image when $\sigma_R = 5$, (d) moved image when $\sigma_M = 5$.

Chicago area. From the figure, it can be seen that the two images have many geometrical differences. For instance, it seems that the left-side border of the reference image has moved to the right in the moved image (cf., the dark spots in the middle and lower parts around the left-side border). The left-side edge of the Lake Michigan seems smoother in the moved image, compared to left-side edge of the lake in the reference image. Also, after a 9-year period, there seems to be more dark spots in the moved image, which may reflect environmental worsening or changes of buildings and other constructions. Some dark spots in the reference image remain in the moved image; but, their sizes have changed (cf., some dark spots in the lower parts of the two images). These geometrical differences are difficult to describe by a parametric model, and some differences are local in the sense that they only affect small portions of the two images. Furthermore, all the geometrical differences described above are small in magnitude. Therefore, procedure NEW should be appropriate to use in this example. After applying the six IBIR procedures, their restored reference images and the corresponding difference images are shown in Figure 9. It can be seen that (i) procedure NEW



(a) (b)

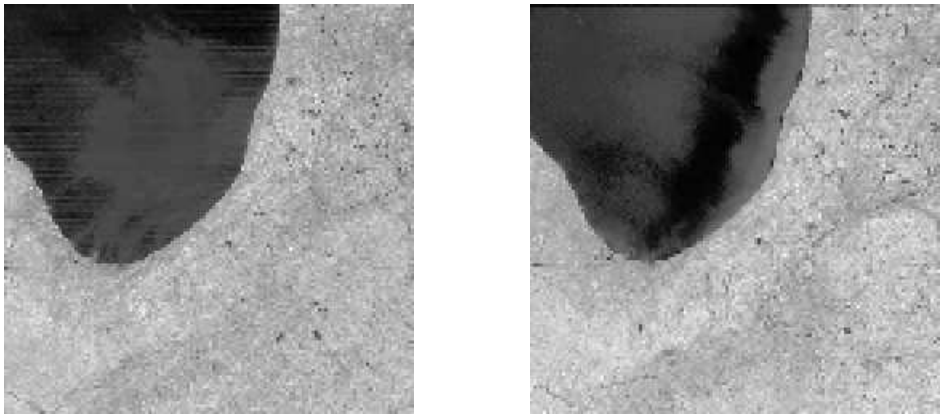
Figure 7: The estimated $\mathbf{T}$ of the proposed method NEW in the example of Figure 1 in cases when (a) $\sigma_R = \sigma_M = 0$, and (b) $\sigma_R = \sigma_M = 5$.

restores the reference image well, (ii) the five competing procedures have difficulty to restore the reference image at certain boundary regions, and (iii) some of them cannot restore the boundary of the Lake Michigan well (e.g., the bottom part of the Lake Michigan boundary is rough in plot (d)). The RRMS and CC measures of the six procedures are presented in the upper part of Table 1. We can see that procedure NEW is indeed better in terms of both RRMS and CC in this example.

Table 1: Performance measures RRMS (first line) and CC (second line) of the six IBIR methods in the Chicago and Teddy Bear examples.

	NEW	DMFFD-MSD	DMFFD-CC	DMFFD-MI	SyN-CC	SyN-PCC
Chicago	17.522	28.294	20.923	23.759	21.566	21.580
	0.953	0.894	0.937	0.917	0.935	0.933
Teddy Bear	5.473	9.689	9.443	7.260	7.540	8.587
	0.987	0.960	0.963	0.978	0.977	0.969

The second real IR problem concerns the two Teddy bear images shown in Figure 10. It can be seen from the two images that the position of the Teddy bear has been changed in the moved image. If we check the two images carefully, then it can be noticed that the background is moved only a little bit in a different direction, making the overall geometrical movement from the reference image to the moved image challenging to handle. The restored reference images and the corresponding



(a) (b)

Figure 8: (a) A reference satellite image taken in 1990 in the Chicago area. (b) A corresponding satellite image taken in 1999 in the same area.

difference images of the six IBIR procedures are shown in Figure 11. From the images in this figure, it seems that all six methods move the Teddy bear to its original position reasonably well. But, the five competing methods alter the background texture to different degrees (e.g., in plot (b)), which explains why the background texture is partially visible in their difference images. These five methods cannot handle the right boundary area well either. The RRMS and CC measures of the six procedures are presented in the lower part of Table 1. We can see that procedure NEW performs better in terms of both RRMS and CC in this example.

6 Concluding Remarks

The ill-posed nature of the IR problem that the related geometrical transformation $\mathbf{T}$ is not well defined at certain places of the design space is described in this paper. Several concepts, including the 2-D degenerate pixels, 2-D partial degenerate pixels, 1-D degenerate pixels, and 1-D partial degenerate pixels, have been defined and used for describing the local properties of the transformation $\mathbf{T}$. Based on these concepts, methods for detecting 2-D and 1-D non-degenerate pixels and for estimating $\mathbf{T}$ have been proposed. A major feature of our proposed methods is that they adopt the local smoothing approach, instead of the conventional regularization-based global smoothing approach, so that the ill-posed nature of the IR problem can be properly accommodated. Both theoretical arguments and numerical studies show that our proposed methods work well in various

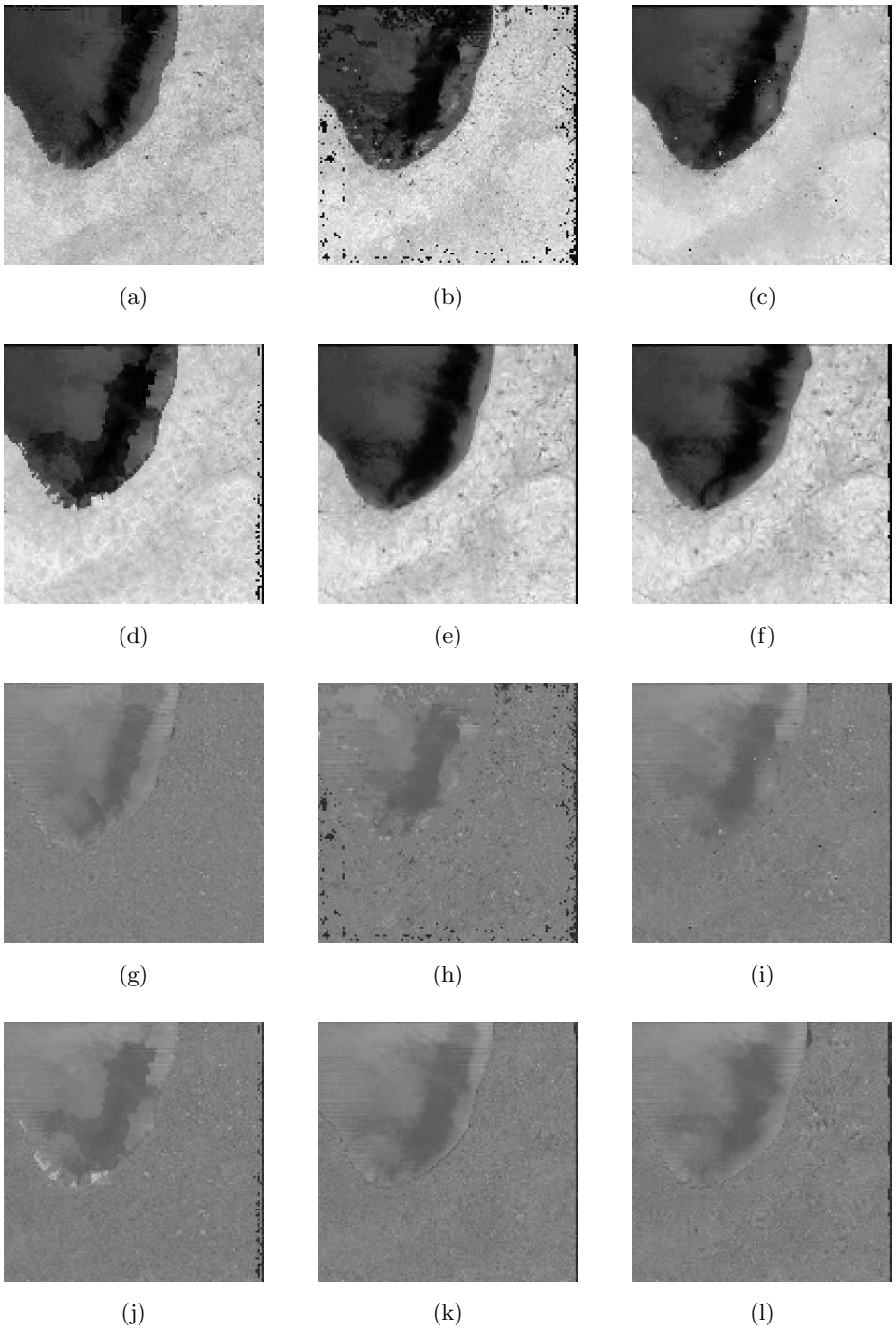
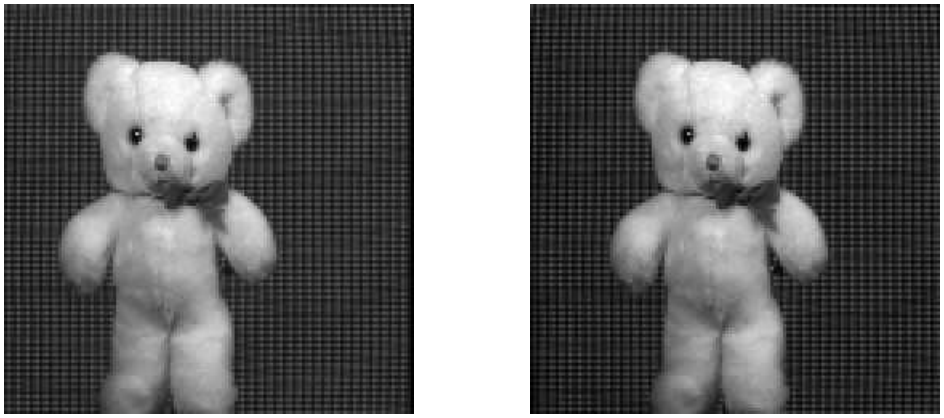


Figure 9: Restored reference images (plots (a)–(f)) and the corresponding difference images (plots (g)–(l)) of the procedures NEW, DMFFD-MSD, DMFFD-CC, DMFFD-MI, SyN-CC, and SyN-PCC in the Chicago example. The gray scale of all images in this figure is kept the same.



(a) (b)

Figure 10: (a) A reference Teddy bear image. (b) A moved Teddy bear image.

cases.

In the proposed methods, there are still some issues that need to be addressed in our future research. For instance, in the proposed procedure for detecting 2-D degenerate pixels, our detection criterion focuses solely on the variation of the underlying image intensity function in the gradient direction (cf., Figure 3(b) and expressions (4) and (5)). It is unknown to us yet whether the detection can be improved if the variation of the image intensity function in the entire neighborhood $O(x, y; h_n^*)$ is taken into account. When estimating $\mathbf{T}$, in the proposed procedure (7)–(9), the related bandwidths r_n, h_n and $\widetilde{h}_n$ are chosen to be the same in the entire design space. If variable bandwidths are used, intuitively, estimation of $\mathbf{T}$ could be improved, which needs to be confirmed in the future research. In the theoretical front, consistency of the detected 2-D and 1-D degenerate pixels is established after the point set $\Omega(R, T, n)$ is excluded (cf., (10) and (12) in Theorem 2). Behaviors of the detected 2-D and 1-D degenerate pixels in $\Omega(R, T, n)$ is unknown yet. Further, Theorem 3 establishes the consistency of $\widehat{\mathbf{T}}(x, y)$ under certain conditions, and this consistency result is for individual 2-D non-degenerate pixels. It is unknown to us at this moment whether the conditions can be further weakened and whether the consistency result is true uniformly in certain subsets of the 2-D non-degenerate pixels.

Supplementary Materials

supplemental.pdf: This pdf file provides proofs of Theorems 2 and 3 presented in Section 4 of the paper, and some numerical results related to the IBIR procedures considered in Section 5.

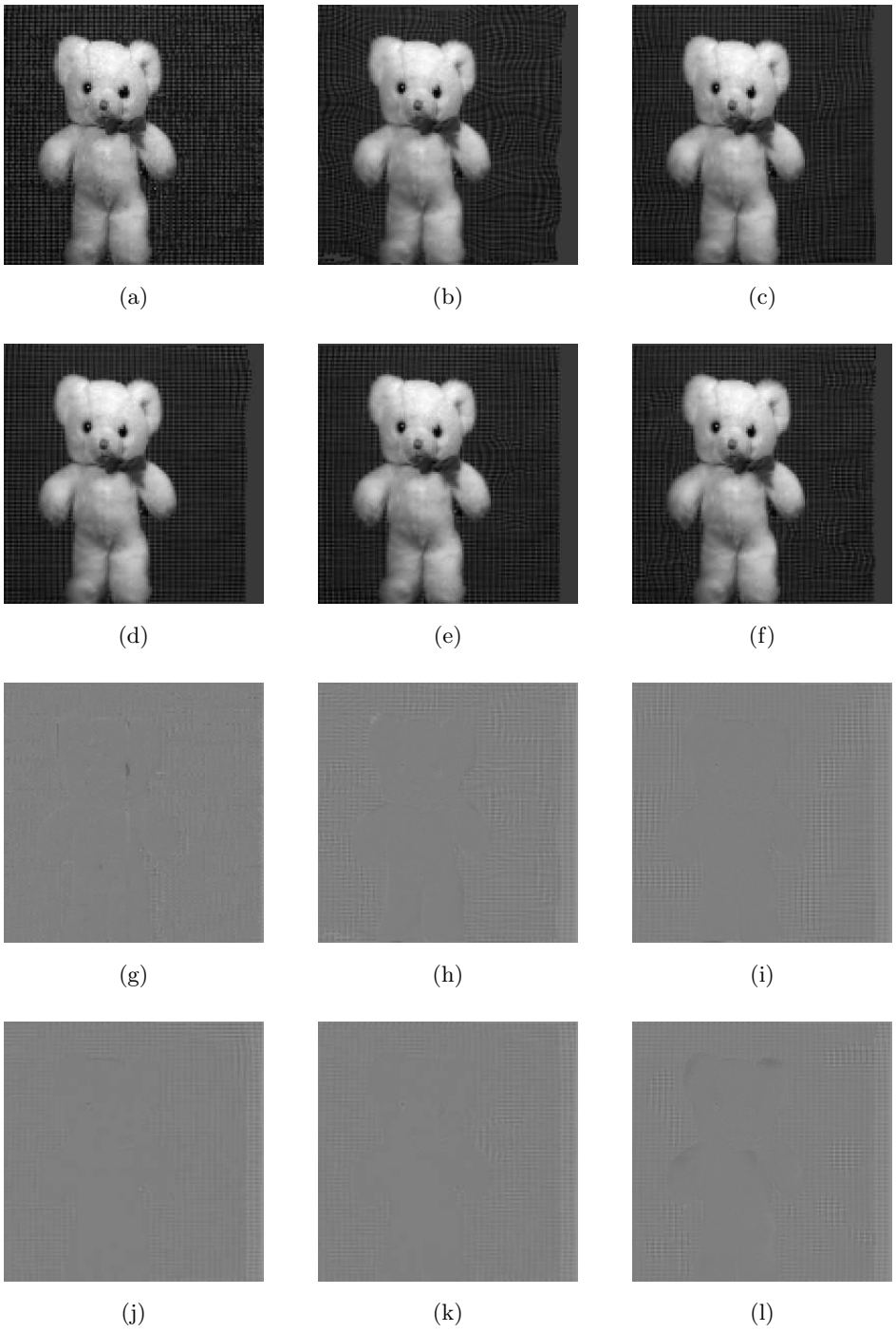


Figure 11: Restored reference images (plots (a)–(f)) and the corresponding difference images (plots (g)–(l)) of the procedures NEW, DMFFD-MSD, DMFFD-CC, DMFFD-MI, SyN-CC, and SyN-PCC in the Teddy bear example. The gray scale of all images in this figure is kept the same.

ComputerCodesAndData.zip: This zip file contains Fortran source codes of our proposed IBIR method and the image data of the Chicago satellite image and the Teddy bear examples.

Acknowledgments: The authors thank the editor, the associate editor and two 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

Althof, R.J., Wind, M.G.J., and Dobbins, J.T. (1997), "A rapid and automatic image registration algorithm with subpixel accuracy," *IEEE Transactions on Medical Imaging*, **16**, 308–316.

Avants, B.B., Epstein, C.L., Grossman, M., and Gee, J.C. (2008), "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain," *Medical Image Analysis*, **12**, 26–41.

Bhagalia, R., Fessler, J.A., and Kim, B. (2009), "Accelerated nonrigid intensity-based image registration using importance sampling," *IEEE Transactions on Medical Imaging*, **28**, 1208–1216.

Davis, M.H., Khotanzad, A., Flamig, D.P., and Harms, S.E. (1997), "Physics-based coordinate transformation for 3-D image matching," *IEEE Transactions on Medical Imaging,* **16**, 317–328.

Denton, E.R., Sonoda, L.I., Rueckert, D., Rankin, S.C., Hayes, C., Leach, M.O., Hill, D.L., and Hawkes, D.J. (1999), "Comparison and evaluation of rigid, affine, and nonrigid registration of breast MR images," *Journal of Computer Assisted Tomography,* **23**, 800–805.

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

Hall, P., and Qiu, P. (2007), "Nonparametric estimation of a point spread function in multivariate problems," *Annals of Statistics*, **35**, 1512–1534.

Hsieh, J.W., Liao, H.Y.M., Fan, K.C., Ko, M.T., and Hung, Y.P. (1997), "Image registration using a new edge-based approach," *Computer Vision and Image Understanding*, **67**, 112–130.

Irani, M., and Peleg, S. (1993), "Motion analysis for image enhancement: resolution, occlusion and transparency," *Journal of Visual Communication and Image Representation*, **4**, 324–335.

Joo, J., and Qiu, P. (2009), "Jump detection in a regression curve and its derivative," *Technometrics*, **51**, 289-305.

Klein, A., Andersson, J., Ardekani, B.A. et al. (2009), "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *NeuroImage,* **46**, 786–802.

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

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

Modersitzki, J. (2009), *Fair: Flexible Algorithms for Image Registration,* SIAM: Philadelphia.

Mukherjee, P.S., and Qiu, P. (2011), "3-D Image Denoising By Local Smoothing And Nonparametric Regression," *Technometrics*, **53**, 196–208.

Pan, W., and Qin, K., and Chen, Y. (2009), "An adaptable-multilayer fractional Fourier transform approach for image registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**, 400–412.

Qiu, P. (2005), *Image Processing and Jump Regression Analysis*, New York: John Wiley & Sons.

Qiu, P. (2009), "Jump-preserving surface reconstruction from noisy data," *Annals of the Institute of Statistical Mathematics*, **61**, 715–751.

Rajwade, A., Banerjee, A. and Rangarajan, A. (2009), "Probability density estimation using isocontours and isosurfaces: application to information-theoretic image registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**, 475–491.

Saeed, N. (1998), "Magnetic resonance image segmentation using pattern recognition, and applied to image registration and quantitation," *NMR in Biomedicine*, **11**, 157–167.

Sun, J., and Qiu, P. (2007), "Jump detection in regression surfaces using both first-order and second-order derivatives," *Journal of Computational and Graphical Statistics*, **16**, 289–311.

Szeliski, R., and Coughlan, J. (1997), "Spline-based image registration," *International Journal of Computer Vision*, **22**, 199–218.

Tustison, N.J., Avants, B.B., and Gee, J.C. (2009), "Directly manipulated free-form deformation image registration," *IEEE Transactions on Image Processing,* **18**, 624–635.

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

Wu, G., Qi, F., and Shen, D. (2006), "Learning-based deformable registration of MR brain images," *IEEE Transactions on Medical Imaging*, **25**, 1145–1157.

Xing, C., and Qiu, P. (2011), "Intensity Based Image Registration By Nonparametric Local Smoothing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**, 2081–2092.

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