# CHAPTER 8
# Image Registration

J. Michael Fitzpatrick
*Vanderbilt University*

Derek L. G. Hill
*King's College London*

Calvin R. Maurer, Jr.
*University of Rochester*

## Contents

## 8.1 Introduction

Registration is the determination of a geometrical transformation that aligns points in one view of an object with corresponding points in another view of that object or another object. We use the term "view" generically to include a three-dimensional image, a two-dimensional image, or the physical arrangement of an object in space. Three-dimensional images are acquired by tomographic modalities, such as computed tomography (CT), magnetic resonance (MR) imaging, single-photon emission computed tomography (SPECT), and positron emission tomography (PET). In each of these modalities, a contiguous set of two-dimensional slices provides a three-dimensional array of image intensity values. Typical two-dimensional images may be x-ray projections captured on film or as a digital radiograph or projections of visible light captured as a photograph or a video frame. In all cases, we are concerned primarily with digital images stored as discrete arrays of intensity values. In medical applications, which are our focus, the object in each view will be some anatomical region of the body. (See Volume I of this handbook for a discussion of medical imaging modalities.) The two views are typically acquired from the same patient, in which case the problem is that of *intrapatient registration*, but *interpatient registration* has application as well (see Chapter 17).

### *8.1.1 Operational goal of registration*

From an operational view, the inputs of registration are the two views to be registered; the output is a geometrical transformation, which is merely a mathematical mapping from points in one view to points in the second. To the extent that corresponding points are mapped together, the registration is successful. The determination of the correspondence is a problem specific to the domain of objects being imaged, which is, in our case, the human anatomy. To make the registration beneficial in medical diagnosis or treatment, the mapping that it produces must be applied in some clinically meaningful way by a system that will typically include registration as a subsystem. The larger system may combine the two registered images by producing a reoriented version of one view that can be "fused" with the other. This fusing of two views into one may be accomplished by simply summing intensity values in two images, by imposing outlines from one view over the gray levels of the other, or by encoding one image in hue and the other in brightness in a color image. Regardless of the method employed, image fusion should be distinguished from image *registration*, which is a necessary first step before fusion can be successful.

The larger system may alternatively use the registration simply to provide a pair of movable cursors on two electronically displayed views linked via the registering transformation so that the cursors are constrained to visit corresponding points. This latter method generalizes easily to the case in which one view is the physical patient and one of the movable "cursors" is a physical pointer held by the surgeon. The registration system may be part of a robotically controlled treatment

system whose guidance is based on registration between an image and the physical anatomy. Drills, for example, may be driven robotically through bone by following a path determined in CT and registered to the physical bone. Gamma rays produced by a linear accelerator or by radioactive isotopes may be aimed at tissue that is visible in MR but hidden from view during treatment with the aiming being accomplished via automatic calculations based on a registering transformation. Registration also serves as a first step in multimodal segmentation algorithms that incorporate information from two or more images in determining tissue types (see Chapter 2). Fusion, linked cursors, robotic controls, and multimodal segmentation algorithms exploit knowledge of a geometrical relationship between two registered views in order to assist in diagnosis or treatment. Registration is merely the determination of that relationship. The goal of registration is thus simply to produce as output a geometrical transformation that aligns corresponding points and can serve as input to a system further along in the chain from image acquisition to patient benefit.

### *8.1.2 Classification of registration methods*

There are many image registration methods, and they may be classified in many ways [1–3]. Maintz has suggested a nine-dimensional scheme that provides an excellent categorization [1]. We condense it slightly (see below) into the following eight categories: image dimensionality, registration basis, geometrical transformation, degree of interaction, optimization procedure, modalities, subject, and object [1].

"Image dimensionality" refers to the number of geometrical dimensions of the image spaces involved, which in medical applications are typically three-dimensional but sometimes two-dimensional. The "registration basis" is the aspect of the two views used to effect the registration. For example, the registration might be based on a given set of point pairs that are known to correspond or the basis might be a set of corresponding surface pairs. Other loci might be used as well, including lines or planes (a special case of surfaces). In some cases, these correspondences are derived from objects that have been attached to the anatomy expressly to facilitate registration. Such objects include, for example, the stereotactic frame and point-like markers, each of which have components designed to be clearly visible in specific imaging modalities. Registration methods that are based on such attachments are termed "prospective" or "extrinsic" methods and are in contrast with the so-called "retrospective" or "intrinsic" methods, which rely on anatomic features only. Alternatively, there may be no known correspondences as input. In that case, intensity patterns in the two views will be matched, a basis that we call "Intensity."

The category, "geometrical transformation", is a combination of two of Maintz's categories, the "nature of transformation" and the "domain of transformation". It refers to the mathematical form of the geometrical mapping used to align points in one space with those in the other. We take up these transformations in detail in

the next section. "Degree of interaction" refers to the control exerted by a human operator over the registration algorithm. The interaction may consist simply of the initialization of certain parameters, or it may involve adjustments throughout the registration process in response to visual assessment of the alignment or to other measures of intermediate registration success. The ideal situation is the fully automatic algorithm, which requires no interaction. "Optimization procedure" refers to the standard approach in algorithmic registration in which the quality of the registration is estimated continually during the registration procedure in terms of some function of the images and the mapping between them. The optimization procedure is the method, possibly including some degree of interaction, by which that function is maximized or minimized. The ideal situation here is a closed-form solution which is guaranteed to produce the global extremum. The more common situation is that in which a global extremum is sought among many local ones by means of iterative search.

"Modalities" refers to the means by which the images to be registered are acquired. Examples are listed at the beginning of this chapter. Registration methods designed for like modalities are typically distinct from those appropriate for differing modalities. Registration between like modalities, such as MR-MR, is called "intramodal" or "monomodal" registration; registration between differing modalities, such as MR-PET, is called "intermodal" or "multimodal" registration. "Subject" refers to patient involvement and comprises three subcategories: intrapatient, interpatient, and atlas, the latter category comprising registrations between patients and atlases, which are themselves typically derived from patient images (e.g., [4]). "Object" refers to the particular region of anatomy to be registered (e.g., head, liver, vertebra).

To build a registration hierarchy based on these eight categorizations, one categorization must be placed at the top level, which, in the organization of this chapter, is the registration basis. The three categories of registration basis mentioned above are examined in three major sections below: Section 8.3, Point-based methods, Section 8.4, Surface-based methods, and Section 8.5, Intensity-based methods. Before we begin these discussions of the basis for registration, however, we take up in the next section the category of geometrical transformation. We examine in detail the rigid transformations and other transformations that preserve the straightness of lines. These transformations are most appropriate for intrapatient registration of rigid anatomy. The remaining transformations, the "curved" transformations, are used for interpatient registration, for the registration of anatomy to an atlas, and for intrapatient registration of tissue that has been deformed between the acquisition of the two views. These situations are examined in Chapters 3 and 17.

## 8.2 Geometrical transformations

Each view that is involved in a registration will be referred to a coordinate system, which defines a space for that view. Our definition of registration is based

on geometrical transformations, which are mappings of points from the space $X$ of one view to the space $Y$ of a second view. The transformation $\mathcal{T}$ applied to a point in $X$ represented by the column vector $\mathbf{x}$ produces a transformed point $\mathbf{x}'$,

$$\mathbf{x}' = \mathcal{T}(\mathbf{x}) \; . \tag{8.1}$$

If the point $\mathbf{y}$ in $Y$ corresponds to $\mathbf{x}$, then a successful registration will make $\mathbf{x}'$ equal, or approximately equal, to $\mathbf{y}$. Any nonzero displacement $\mathcal{T}(\mathbf{x}) - \mathbf{y}$ is a registration error. The set of all possible $\mathcal{T}$ may be partitioned into rigid and nonrigid transformations with the latter transformations further divided into many subsets. This top-level division makes sense in general because of the ubiquity of rigid, or approximately rigid, objects in the world. It makes sense for medical applications in particular because of the rigid behavior of many parts of the body, notably the bones and the contents of the head (not during surgery). It is also a simple class with only six parameters completely specifying a rigid transformation in three dimensions. (We note here that, while one and two-dimensional motion is possible, such limited motion is sufficiently rare that we will ignore it in this chapter.)

### *8.2.1 Rigid transformations*

Rigid transformations, or rigid mappings, are defined as geometrical transformations that preserve all distances. These transformations also preserve the straightness of lines (and the planarity of surfaces) and all nonzero angles between straight lines. Registration problems that are limited to rigid transformations are called *rigid registration* problems. Rigid transformations are simple to specify, and there are several methods of doing so. In each method, there are two components to the specification, a translation and a rotation. The translation is a three-dimensional vector $\mathbf{t}$ that may be specified by giving its three coordinates $t_x, t_y, t_z$ relative to a set of $x, y, z$ Cartesian axes or by giving its length and two angles to specify its direction in polar spherical coordinates. There are many ways of specifying the rotational component, among them Euler angles, Cayley-Klein parameters, quaternions, axis and angle, and orthogonal matrices [5–9]. In our discussions we will utilize orthogonal matrices. With this approach, if $\mathcal{T}$ is rigid, then

$$\mathbf{x}' = R\mathbf{x} + \mathbf{t} \; , \tag{8.2}$$

where $R$ is a $3\times3$ orthogonal matrix, meaning that $R^tR = RR^t = I$ (the identity). Thus $R^{-1} = R^t$. This class of matrices includes both the *proper* rotations, which describe physical transformations of rigid objects, and *improper* rotations, which do not. These latter transformations both rotate and reflect rigid objects, so that, for example, a right-handed glove becomes a left-handed one. Improper rotations can be eliminated by requiring $\det(R) = +1$.

Proper rotations can be parameterized in terms of three angles of rotation, $\theta_x, \theta_y, \theta_z$, about the respective Cartesian axes, the so-called "Euler angles." The

rotation angle about a given axis is, with rare exception, considered positive if the rotation about the axis appears clockwise as viewed from the origin while looking in the positive direction along the axis. The rotation of an object (as opposed to the coordinate system to which it is referred) about the $x$, $y$, and $z$ axes, in that order leads to

$$\begin{aligned} R &= \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta_y\cos\theta_z & -\cos\theta_x\sin\theta_z + \sin\theta_x\sin\theta_y\cos\theta_z & \sin\theta_x\sin\theta_z + \cos\theta_x\sin\theta_y\cos\theta_z \\ \cos\theta_y\sin\theta_z & \cos\theta_x\cos\theta_z + \sin\theta_x\sin\theta_y\sin\theta_z & -\sin\theta_x\cos\theta_z + \cos\theta_x\sin\theta_y\sin\theta_z \\ -\sin\theta_y & \sin\theta_x\cos\theta_y & \cos\theta_x\cos\theta_y \end{bmatrix} \end{aligned}$$

with the three matrices in the first line representing the rotations $R_z(\theta_z)$, $R_y(\theta_y)$, and $R_x(\theta_x)$ about $z$, $y$, and $x$, respectively (in reverse order because they are applied from right to left). Other angular parameterizations are sometimes used, including all permutations of the order of $R_x$, $R_y$, and $R_z$. General rotations can also be produced by three rotations about only two of the Cartesian axes provided that successive rotations are about distinct axes. The most common of these is $R = R_z(\theta_{z2})R_x(\theta_x)R_z(\theta_{z1})$.

The *axis-and-angle* specification provides a parameterization of $R$ comprising a unit vector representing an axis of rotation $\hat{\omega}$ and a single angle of rotation $\theta$ about that axis. The rotational axis $\hat{\omega}$ is the (lone) real eigenvector of $R$. The rotational angle can be calculated from the relation $\cos\theta = (\mathsf{trace}\ R - 1)/2$. With this parameterization,

$$R = \begin{bmatrix} {\omega_x}^2 V + C & \omega_x\omega_y V - \omega_z S & \omega_x\omega_z V + \omega_y S \\ \omega_x\omega_y V + \omega_z S & {\omega_y}^2 V + C & \omega_y\omega_z V - \omega_x S \\ \omega_x\omega_z V - \omega_y S & \omega_y\omega_z V + \omega_x S & {\omega_z}^2 V + C \end{bmatrix}, \tag{8.3}$$

where the $\omega_x, \omega_y, \omega_z$ are the components of $\hat{\omega}$, $V = 1 - \cos\theta$, $C = \cos\theta$, and $S = \sin\theta$ [9].

The *quaternion* specification provides a parameterization that is closely related to the axis and angle. The quaternion is a vector $\mathbf{q}$ consisting of four elements, $q_0, q_x, q_y, q_z$, and obeying special rules for multiplication [7, 8]. A quaternion of unit length can be used to represent a rotation, where the elements are equal respectively to $\cos(\theta/2)$, $\omega_x \sin(\theta/2)$, $\omega_y \sin(\theta/2)$, and $\omega_z \sin(\theta/2)$ [7, 8]. The parameterization in Eq. (8.3) can be converted to a quaternion parameterization with the identities, $2\sin(\theta/2)\cos(\theta/2) = \sin\theta$ and $2\sin^2(\theta/2) = 1 - \cos\theta$,

$$R = \begin{bmatrix} q_0^2 + q_x^2 - q_y^2 - q_z^2 & 2q_xq_y - 2q_0q_z & 2q_xq_z + 2q_0q_y \\ 2q_xq_y + 2q_0q_z & q_0^2 - q_x^2 + q_y^2 - q_z^2 & 2q_yq_z - 2q_0q_x \\ 2q_xq_z - 2q_0q_y & 2q_yq_z + 2q_0q_x & q_0^2 - q_x^2 - q_y^2 + q_z^2 \end{bmatrix}. \tag{8.4}$$

### *8.2.2 Nonrigid transformations*

Nonrigid transformations are important not only for applications to nonrigid anatomy, but also for interpatient registration of rigid anatomy and intrapatient registration of rigid anatomy when there are nonrigid distortions in the image acquisition procedure. In all cases, it is preferable to choose transformations that have physical meaning, but in some cases, the choice is made on the basis of convenient mathematical properties.

#### *8.2.2.1 Scaling transformations*

The simplest nonrigid transformations are rigid except for scaling,

$$\mathbf{x}' = RS\mathbf{x} + \mathbf{t} \tag{8.5}$$

and

$$\mathbf{x}' = SR\mathbf{x} + \mathbf{t}\ , \tag{8.6}$$

where $S = \text{diag}(s_x, s_y, s_z)$ is a diagonal matrix whose elements represent scale factors along the three coordinate axes. Because $RS$ is not in general equal to $SR$, these equations represent two different classes of transformations. Such transformations may be needed to compensate for calibration errors in image acquisition systems. They are appropriate, for example, when gradient strengths are in error in MR. The diagonal elements of $S$ then become the respective correction factors for the $x$, $y$, and $z$ gradients (see Section 8.2.3, below). If the scaling is isotropic, the transformation is a *similarity transformation*,

$$\mathbf{x}' = sR\mathbf{x} + \mathbf{t}\ , \tag{8.7}$$

where $s$ is a positive scalar, sometimes known as a "dilation" (for values less than one as well). This transformation preserves the straightness of lines and the angles between them. Both Eq. (8.5) and Eq. (8.6) reduce to Eq. (8.7) when $s_x = s_y = s_z = s$. The coupling of scaling with the rigid transformation is effective when registrations must account for erroneous or unknown scales in the image acquisition process.

#### *8.2.2.2 Affine transformations*

The scaling transformations are special cases of the more general *affine* transformation,

$$\mathbf{x}' = A\mathbf{x} + \mathbf{t}\ , \tag{8.8}$$

in which there is no restriction on the elements $a_{ij}$ of the matrix $A$. The affine transformation preserves the straightness of lines, and hence, the planarity of surfaces, and it preserves parallelism, but it allows angles between lines to change. It is

an appropriate transformation class when the image may have been skewed during acquisition as, for example, when the CT gantry angle is incorrectly recorded.

The affine transformations and their associated special cases are sometimes represented by means of *homogeneous coordinates*. In this representation, both $A$ and $\mathbf{t}$ are folded into one 4×4 matrix $M$ whose elements are defined as follows: $m_{ij} = a_{ij}, i = 1,2,3, j = 1,2,3$; $m_{i4} = t_i, i = 1,2,3$; $m_{4j} = 0, j = 1,2,3$; and $m_{44} = 1$. To accomplish the transformation, augmented vectors $\mathbf{u}$ and $\mathbf{u}'$ are used for which $u_i = x_i$ and $x'_i = u'_i$ for $i = 1,2,3$ and $u_4 = u'_4 = 1$.

$$\mathbf{u}' = \begin{pmatrix} u'_1 \\ u'_2 \\ u'_3 \\ 1 \end{pmatrix} = M\mathbf{u} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & t_1 \\ a_{21} & a_{22} & a_{23} & t_2 \\ a_{31} & a_{32} & a_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ 1 \end{pmatrix} . \tag{8.9}$$

While the use of homogeneous coordinates does not produce any extra power or generality for rigid transformations, it does simplify notation, especially when rigid transformations must be combined with projective transformations, which we consider next.

#### *8.2.2.3 Projective transformations*

So far, the nonrigid transformations that we have considered, all of which are affine transformations, preserve parallelism. The more general nonrigid transformations include the *projective* transformations, which preserve the straightness of lines and planarity of surfaces, and the *curved* transformations, which do not. The projective transformations, which have the form,

$$\mathbf{x}' = (A\mathbf{x} + \mathbf{t})/(\mathbf{p} \cdot \mathbf{x} + \alpha) , \tag{8.10}$$

can be written simply in homogeneous coordinates,

$$\mathbf{u}' = \begin{pmatrix} u'_1 \\ u'_2 \\ u'_3 \\ u'_4 \end{pmatrix} = M\mathbf{u} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & t_1 \\ a_{21} & a_{22} & a_{23} & t_2 \\ a_{31} & a_{32} & a_{33} & t_3 \\ p_1 & p_2 & p_3 & \alpha \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ 1 \end{pmatrix} , \tag{8.11}$$

where, as for the affine transformation, $u_i = x_i$ for $i = 1,2,3$, $u_4 = 1$, and $m_{ij} = a_{ij}, i = 1,2,3, j = 1,2,3$, but $u'_4$ is no longer necessarily equal to 1, $m_{4j} = p_j, j = 1,2,3$, $m_{44} = \alpha$, and $x'_i = u'_i/u'_4, i = 1,2,3$. The linearity of Eq. (8.11) can provide considerable simplification for the projective transformations and the *perspective projections*. This latter class of transformations relate three-dimensional anatomy to two-dimensional planar images acquired by means of a camera and are taken up in the next section.

#### *8.2.2.4 Perspective transformations*

Images obtained by x-ray projection, endoscopy, laparoscopy, microscopy, and direct video acquisition are all two-dimensional views of three-dimensional objects rendered by means of projecting light rays or x-rays from a three-dimensional scene onto a two-dimensional plane. The geometrical transformation, which we call a *perspective projection*, produced by each of these modalities is equivalent to that of photography.

These perspective projections are a subset of the projective transformations of Eqs. (8.10) and (8.11). The projective transformations, unlike the perspective projections, do not, in general, transform $\mathbf{x}$ to a plane. Furthermore, the affine portion of the transformation is typically assumed to be the identity for perspective projections. Specializing now to perspective projections, we let $f = 1/|\mathbf{p}|$ in Eq. (8.10), and let $\hat{\mathbf{p}}$ be a unit vector in the direction of the projection axis, $\mathbf{p}$. These substitutions lead to

$$\mathbf{x}' = f\mathbf{x}/(\mathbf{x} \cdot \hat{\mathbf{p}} + \alpha f) \; . \tag{8.12}$$

If $\alpha$ is nonzero, then Eq. (8.12) does not, in fact, transform $\mathbf{x}$ to a plane and, hence, is not a perspective projection. Perspective projection can be produced, however, by zeroing the component of $\mathbf{x}'$ in the direction of $\mathbf{p}$:

$$\mathbf{x}' \longrightarrow \mathbf{x}' - (\mathbf{x}' \cdot \hat{\mathbf{p}})\hat{\mathbf{p}} \; . \tag{8.13}$$

Equation (8.12) and substitution (8.13) give the general form of the transformation produced when a photograph of a three-dimensional scene is acquired with a "pinhole camera," which is a camera in which a small hole substitutes for the lens system. Fig. 8.1 shows a schematic camera. A ray from a point $\mathbf{x}$ in the scene is projected through the pinhole onto a film screen, which is perpendicular to the axis of projection $\mathbf{p}$ and located at a distance $f$ from the pinhole. Fortunately, all the systems mentioned above can be approximated by the pinhole camera system by identifying the unique point relative to the lens system through which light rays travel undeflected, or the point from which the x rays emanate. That point, also known as the "center of perspectivity" [10], plays the role of the pinhole. Because the film image is inverted, it is convenient to treat instead an equivalent upright "image" located in front of the camera, which is also perpendicular to $\mathbf{p}$ and located the same distance $f$ from the pinhole, as shown in the figure. The transformed point, $\mathbf{x}'$, of Eq. (8.12), followed by substitution (8.13) lies in that plane.

Figure 8.2 shows the relationship between the parameters in Eq. (8.12) and the imaging system both (a) for lens systems and (b) for x-ray systems. The geometric difference is that $\mathbf{x} \cdot \hat{\mathbf{p}} + \alpha f$ is greater than $f$ for lens systems and less than $f$ for x-ray systems. The parameter $f$ is called the *focal length*, or, alternatively, the "camera constant" or "principal distance." The name "focal length" is derived from lens systems. It is meant to imply that the lens is adjusted so that all light

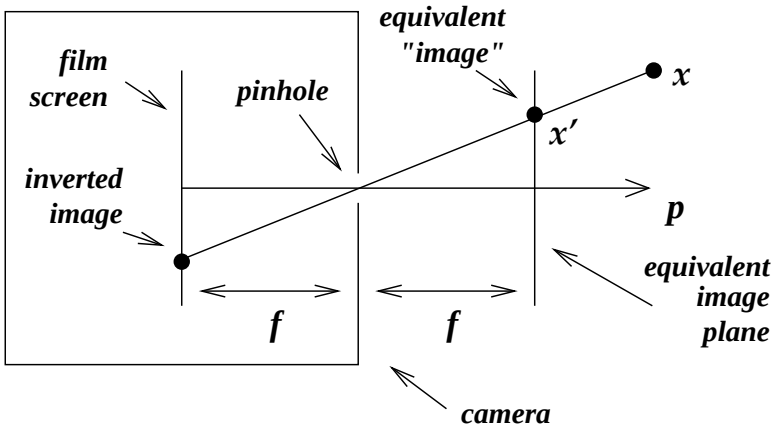



**Figure 8.1:** A schematic of a pinhole camera. A single point $\mathbf{x}$ in the scene is shown being imaged. The projection axis $\mathbf{p}$ points to the right. The image is produced within the camera on a film screen that is perpendicular to $\mathbf{p}$ and located a distance $f$ from the pinhole. Analysis is simplified by treating the transformation as if an equivalent, but upright, image were produced instead on a plane placed outside the camera at the same orientation and same distance from the pinhole. The point $\mathbf{x}$ is transformed to $\mathbf{x}'$ on this plane.

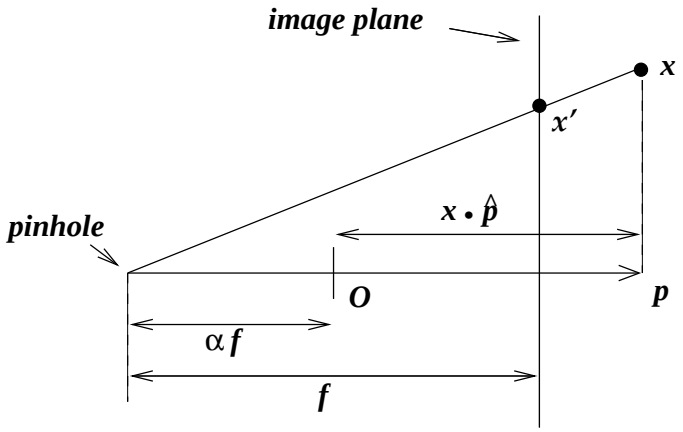
emanating from a given point in front of the camera and passing through its lens will be focused to a single point on a screen located at that distance from the effective pinhole. The focusing is only approximate and varies in quality with the distance of the anatomy from the lens. The value of $\alpha$ in Eq. (8.12) is determined by the placement of the origin. Typically, the origin is placed at the pinhole, for which $\alpha = 0$, or at the intersection of $\mathbf{p}$ and the image plane, for which $\alpha = 1$.

#### *8.2.2.5 Curved transformations*

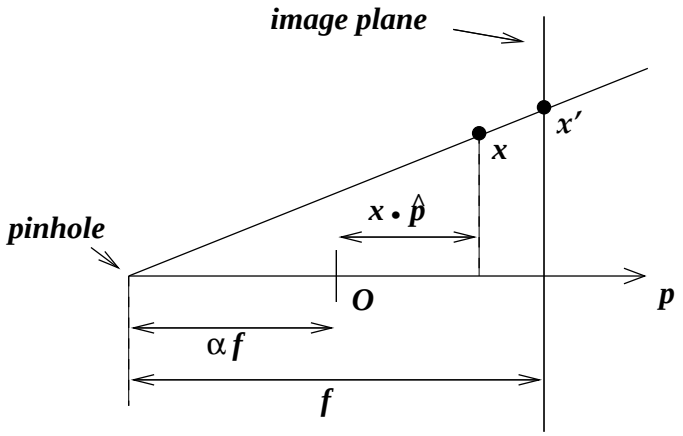
Curved transformations are those that do not preserve the straightness of lines. In curved transformations, the simplest functional form for $\mathcal{T}$ is a polynomial in the components of $\mathbf{x}$ [11, 12],

$$\mathbf{x}' = \sum_{ijk}^{IJK} \mathbf{c}_{ijk} x^i y^j z^k \,, \tag{8.14}$$

where $\mathbf{c}_{ijk}$ is the three-element vector of coefficients for the $i, j, k$ term in the polynomial expression for the three components $x', y', z'$ of $\mathbf{x}'$. Modifications may be employed that include all terms for which $i+j+k \leq M$. These transformations are rarely used with values of $I$, $J$, and $K$ greater than 2 or M greater than 5 because of spurious oscillations associated with high-order polynomials.




(a) Pinhole Camera System. $\mathbf{x} \cdot \hat{\mathbf{p}} + \alpha f > f$.




(b) X-ray Projection. $\mathbf{x} \cdot \hat{\mathbf{p}} + \alpha f < f$.

**Figure 8.2:** Perspective projection for the pinhole camera system and for x-ray projection. The point $\mathbf{x}$ is projected to $\mathbf{x}'$. The origin of the coordinate system is located at $O$. The projection axis $\mathbf{p}$ points to the right. Typical values for $\alpha$ are 0 or 1.

These oscillations can be reduced or eliminated by employing piecewise polynomials. The resulting transformations are defined by first partitioning the space into a set of three-dimensional rectangles by means of three sets of cut planes, each perpendicular to one of the Cartesian axes. Then, within each rectangle $i, j, k$ of dimensions, $u_i, v_j, w_k$, $\mathcal{T} = P_{ijk}^{(x)}(x)P_{ijk}^{(y)}(y)P_{ijk}^{(z)}(z)$, where each $P()$ is a univariate polynomial of degree $m$ defined only over that rectangle. (Polynomial forms that are more complex than products of univariate polynomials in $x$, $y$, and $z$ are available as well [13].) By judicious choice of the polynomials' coefficients, it is possible to ensure that the polynomials join smoothly across the rectangle interfaces in the sense that $\mathcal{T}$ is $m-1$ times continuously differentiable. Such transformations are called *splines*, and the corners $\mathbf{x}_{ijk}$ of the rectangles are called "knots." The properties of splines have been carefully studied in approximation theory [14]. The most common choice for $m$ is 3, producing the "cubic spline," for which $\mathcal{T}$ and its first and second derivatives are continuous. Splines of degree $m$ are often expressed in terms of a convenient basis set of polynomials of the same degree, called *B-splines* (see Chapter 6 for their definition). Noting that the coordinates of every $\mathbf{x}$ must be equal to $u_i \times (l+\alpha), v_j \times (m+\beta), w_k \times (n+\gamma)$, for some unique set $l, m, n, \alpha, \beta, \gamma$ where $l, m, n$ are integers and $0 \leq \alpha, \beta, \gamma \leq 1$, the expansion of the cubic spline in terms of the B-splines can be expressed as follows:

$$\mathbf{x}'(l, m, n, \alpha, \beta, \gamma) = \sum_{ijk} B_{i-l}(\alpha) B_{j-m}(\beta) B_{k-n}(\gamma) \mathbf{c}_{ijk}, \tag{8.15}$$

where each $B_q(t)$ is a segment of the cubic B-spline with $B_q(t) = 0$ unless $-1 \leq q \leq 2$. Simple algorithms are available for determining the form of the $B_q(t)$ (see for example [13]), but for the special case in which the knot spacing along a given direction is uniform, $u_i = u, v_j = v, w_k = w$, they have an especially simple form [13, 15, 16],

$$\begin{aligned} B_{-1}(t) &= (\ -t^3 + 3t^2 - 3t + 1)/6, \quad B_0(t) = (3t^3 - 6t^2 + 4)/6, \\ B_1(t) &= (-3t^3 + 3t^2 + 3t + 1)/6, \quad B_2(t) = \ \ t^3/6. \end{aligned} \tag{8.16}$$

A transformation that has been heavily used for two-dimensional problems is the *thin-plate spline*, which was originally called the "surface spline." This form was first proposed by Harder in 1972 [17] for designing aircraft wings. (Their mathematical properties are explored by Duchon [18], who coined the term "thin-plate spline," and Meinguet [19].) The thin-plate splines were first employed to describe deformations within the two-dimensional plane by Goshtasby in 1988 [20].[1] Goshtasby's formulation, which is now widely employed in the image processing literature, is as follows:

$$\mathbf{x}' = A\mathbf{x} + \mathbf{t} + \sum_{i}^{N} \mathbf{c}_i r_i^2 \ln r_i^2 \ , \tag{8.17}$$

[1] For suggestions as to their use in comparing two-dimensional shapes see [21].

where $r_i = |\mathbf{x} - \mathbf{x}_i|$ and $\mathbf{x}_i$ is a control point. Unlike the rectangular grid of knots required for the cubic-splines, the control points can be placed arbitrarily, a feature that is of considerable advantage in the registration of medical images. For three-dimensional transformations, the thin-plate spline has a simpler form in which $r_i^2 \ln r_i^2$ in Eq. (8.17) is replaced by $r_i$. For both the two-dimensional and the three-dimensional forms, the affine portion of Eq. (8.17) is a necessary part of the transformation. Without this component there may be no set of $\mathbf{c}_i$ that satisfies the equation at all $N$ points. With the affine part included, it is always possible, by means of the imposition of a set of side conditions on the $\mathbf{c}_i$ (see Section 8.3.4), to ensure that a solution exists for any arrangement of points.

The thin-plate splines are examples of the more general category of *radial basis functions*, which, for two-dimensional and three-dimensional spaces, have the same form as Eq. (8.17), but with $r_i^2 \ln r_i^2$ replaced by a general function $f(r_i)$. Other examples of $f(r_i)$ that have been employed to interpolate among points in two or three dimensions (not necessarily for registration) include the "volume spline" $r_i^3$, the "multiquadric" $(r_i^2 + b)^{1/2}$ $(b > 0)$, its inverse, and the Gaussian, $\exp(-r_i^2/a)$ [22]. Because none of these example functions has compact support, a change in any control point has an affect on the transformation at all points. Radial functions with compact support have been employed to limit this effect to points in a neighborhood surrounding a given control point [23, 24].

Other curved transformations have been employed, including solutions to the equations of continuum mechanics describing elastic and fluid properties attributed to the anatomy being registered [25–31]. These equations, which are derived from conservation of mass, momentum, and energy and from experimentally measured material properties, involve the displacement vector, $\mathbf{x}' - \mathbf{x}$, and the first and second spatial derivatives of its components. The nonrigid transformations that result from the numerical solution of these partial differential equations are appropriate for intrapatient registration when the anatomy is nonrigid, especially when surgical resection has changed its shape. In these cases, the major problem in registration is the determination of the material properties and forces that act on the tissue. With that information available, the solution of the differential equations may be carried out numerically by means of finite-difference or finite-element methods. These transformations have been also used for interpatient registration and for the closely related problem of mapping an atlas to a patient, and they are taken up in detail in Chapter 17.

### *8.2.3 Rectification*

Nonrigid transformations may be necessary in order to rectify images that are distorted in the acquisition process. Images are distorted when the transformation from physical object to image fails to have the ideal form assumed for the acquisition process. We define *rectification* as the process of transforming an acquired image into a "rectified" image, for which the resulting overall transformation from

physical object to rectified image has the ideal form. It is possible to perform registration in the face of such distortions, but if the distortion can be determined, a more efficient approach is to perform rectification to remove the distortion as a separate preprocessing step before registration. Efficiency is achieved by this division of the problem into two steps because additional information from scanner calibrations, phantom images, or other corrective measures can be used to reduce the size of the space of transformations to be searched during registration.

While for three-dimensional images the ideal transformation is normally rigid, as given by Eq. (8.2), in some cases, as for example when registration is to be effected between serial images acquired on the same system within a short time, Eq. (8.7) may be acceptable. Then, the scaling $s$ can be expected to remain constant between acquisitions and can thus be set equal to one for the purposes of registration. The simplest three-dimensional rectification problem is compensation for an isotropic scaling. In this case, rectification can be accomplished by means of the transformation $\mathbf{x}' = s\mathbf{x}$, which is a special case of Eq. (8.7). A slightly more complex example is provided by gantry tilt in CT imaging. The resultant skew produced in a volume CT image can be removed by transformations of the form $\mathbf{x}' = A\mathbf{x}$, which is a special case of Eq. (8.8) (i.e., no translation).

For projection by a camera system, the ideal transformation is Eq. (8.12), or the simpler form given below by Eq. (8.30), but, in many cases, the projection image captured by a camera system is distorted in the image plane. Such distortions occur not only for light focused by lens systems but for x-ray images produced in conjunction with image intensifiers. That distortion may be rectified by means of a two-dimensional transformation within the image plane before registration is attempted. The required transformation is typically a member of the class of curved transformations. The rectification step in that case may employ any of the curved transformations given by Eqs. (8.14), (8.15), or (8.17). The transformation may be determined by acquiring an image of a phantom consisting of a regularly spaced grid of fiducial points arranged in a plane perpendicular to the projection axis. That image is then compared with the expected image based on a model of the phantom and the ideal transformation for the camera. The difference between the pattern of imaged points and the pattern of model points is used in a point-based registration to determine the transformation that will restore the regular pattern to the image (see Section 8.3.4). An important special case is the rotationally symmetric "pin cushion" distortion exhibited near the image periphery by wide-angle lenses, such as those employed for endoscopy, and by image intensifiers in digital radiographic systems (see Chapter 3, Section 3.2.2 of Volume 1 of this handbook). In this case, the two-dimensional transformation reduces to a one-dimensional transformation in the radial distance from a central point, and the distortion correction reduces to a simple univariate interpolation [32]. Once the transformation is determined, it can be employed to rectify any subsequent anatomical image acquired with the same acquisition setup. Thus, by means of the separate rectification step, the task

of distortion correction is separated from the task of registering specific anatomical image pairs.

MR imaging is subject to geometric distortions whose rectification requires curved transformations. These distortions arise from nonuniform gradient fields due to imperfections in the scanner and from spatial variation in the static magnetic field that may be due either to imperfections in the scanner or to induced magnetization of the anatomy. Rectifying transformations can be developed for scanner imperfections by acquiring images of phantoms that include regularly spaced, three-dimensional grids of point-like objects in analogy to the two-dimensional rectification of camera images in projection imaging. For distortions arising from the susceptibility of the anatomy to magnetization, special protocols must be employed involving MR pulse sequences specifically designed for rectification [33,34]. These transformations have a special form in which all displacements are parallel to the so-called "readout" gradient. Because the magnitude of the displacements and their sense (forward or backward) varies from point to point, the resultant transformation is in general curved. Removal of these distortions by a rectification step can replace a complicated nonrigid registration problem with a rigid one.

For three-dimensional images, rectification may appear to be equivalent to registration between the image view and the physical view. In fact, the rectification problem is considerably simpler. The goal of registration from an image to physical space is to map points in the image view to the corresponding points in the physical view, while the goal of rectification is merely to bring all points in an image to within an ideal transformation of their corresponding physical points. The simplification is that the latter transformation may remain unknown. Thus, the goal of rectification is to find a transformation, which, *when composed with some ideal transformation*, will map corresponding points together. The determination of the unknown transformation is necessary to complete the registration process. A simple example is provided by errors in the magnitude of MR gradients. With such errors, the imaging process transforms physical points according to Eq. (8.6). The appropriate class of transformations for registering such distorted images to undistorted images is then Eq. (8.5) (see also page 473). If, however, the correcting scale factors, $s_x, s_y, s_z$, can be determined by means of calibration (see Algorithm 8.3 below), then rectification can be accomplished by applying Eq. (8.5) with $S = \text{diag}(s_x, s_y, s_z)$ and with any arbitrary rotation $R$ and translation $\mathbf{t}$. Typically $R = I$ and $\mathbf{t} = 0$ are chosen for simplicity, in which case, the rectifying transformation is $\mathbf{x}' = S\mathbf{x}$. The transformation between the resultant rectified image and the physical object now has the ideal form for the acquisition process, namely Eq. (8.2), as does the registering transformation between any two such rectified images.

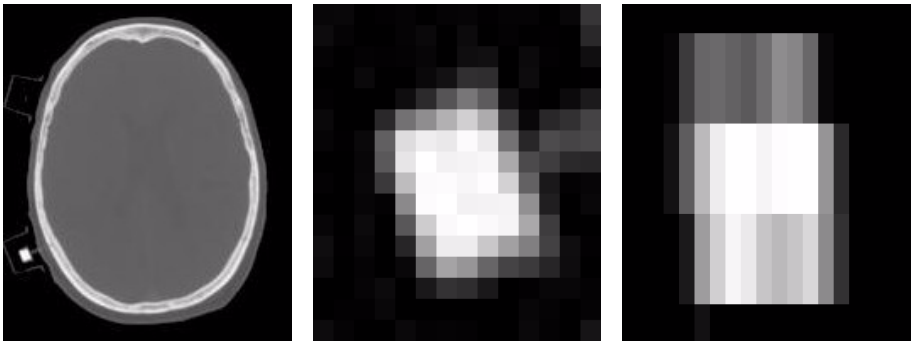
### 8.3 Point-based methods

If some set of corresponding point pairs can be identified *a priori* for a given pair of views, then registration can be effected by selecting a transformation that aligns the points. Because such points are taken as being reliable for the purposes of registration, they are called *fiducial points*, or *fiducials*. To be reliable, they must lie in clearly discernible features, which we will call *fiducial features*. The determination of a precise point within a feature is called *fiducial localization*. The transformation that aligns the corresponding fiducial points will then interpolate the mapping from these points to other points in the views.

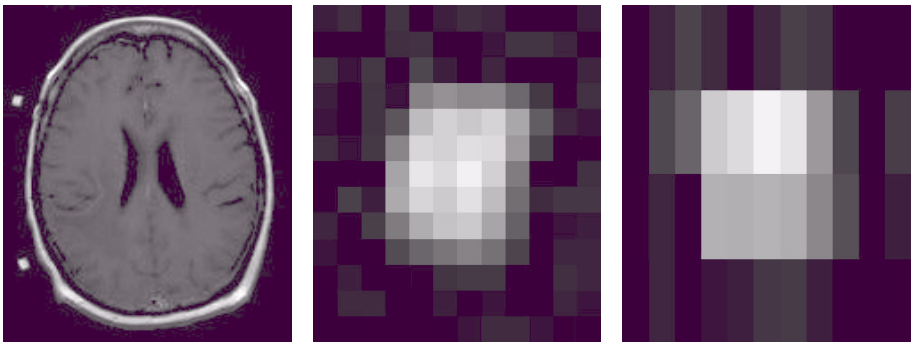
The fiducial localization process may be based on interactive visual identification of anatomical landmarks, such as the junction of two linear structures, e.g., the central sulcus with the midline of the brain or the intersection of a linear structure with a surface, e.g., the junction of septa in an air sinus, etc. [35]. Alternatively, the feature may be a marker attached to the anatomy and designed to be accurately localizable by means of automatic algorithms. In either case, the chosen point will inevitably be erroneously displaced somewhat from its correct location. This displacement in the determination of the fiducial point associated with a fiducial feature is commonly called the *fiducial localization error* (**FLE**). Such errors will occur in both image spaces. They cannot ordinarily be observed directly, but they can be observed indirectly through the registration errors that they cause.

Marker-based registration has the considerable advantage over landmark-based registration that the fiducial feature is independent of anatomy. Automatic algorithms for locating fiducial markers can take advantage of knowledge of the marker's size and shape in order to produce a consistent fiducial point within it [36]. Typically, the fiducial point chosen by a localization algorithm will lie near its center. Hence, the point is often referred to as the fiducial "centroid". However, registration accuracy depends only on the degree to which the chosen points correspond in the two views. Random errors in the localized position will be caused by noise in the image and by the varying positioning of the marker relative to the voxel grid, but, for reasonable localization algorithms, the mean of the fiducial points chosen should be the same in the two views relative to a coordinate system fixed in the marker. Because of this consistency, the effective mean displacement, $\langle \mathbf{FLE} \rangle$, in each view is zero. (We use $\langle x \rangle$ to indicate the expected value of $x$.) The variance $\langle \mathrm{FLE}^2 \rangle$ may be appreciable, however. (Note that we use a bold font to indicate the vector displacement **FLE** and a normal font to indicate the magnitude FLE.)

The goal of fiducial design and of the design of the accompanying fiducial localization algorithm is to produce a small variance. In general, as the marker volume becomes larger, and as the signal per volume produced in the scanner by its contents becomes larger, FLE will become smaller. Figure 8.3 shows sample images of fiducial markers. Because of the discrete nature of the digital images, it is important to use a marker that is larger than an image voxel. Figure 8.3 shows a hollow marker that is of cylindrical shape. Its inside dimensions are height, 5 mm



(a) Transaxial CT (b) Magnified CT (c) Magnified CT, coronal reformatting



(d) Transaxial MR (e) Magnified MR (f) Magnified MR, coronal reformatting

**Figure 8.3**: Sample CT and MR images of a fiducial marker. The marker is a cylinder of height 5 mm and diameter 7 mm. (a) Transverse CT of a human head with an attached marker at the lower left. (Part of a second marker is just visible at the upper left. The structures surrounding the markers are protective housings that are visible in CT but not MR.) Voxel dimensions, 0.65 by 0.65 by 3.0 mm. (b) Magnification of the marker in (a). (c) Magnification of the marker in (a) after reformatting into the coronal orientation. (d) Transverse MR of head with markers at upper and lower left. Voxel dimensions, 1.1 by 1.1 by 4.0 mm. (e) Magnification of the upper marker in (d). (f) Magnification of the upper marker in (d) after reformatting into the coronal orientation.

and diameter, 7 mm. The marker is filled with a liquid that is imageable in CT and MR [37]. Both CT and an MR (Spin-Echo, T1-weighted) transaxial images are shown in Fig. 8.3. The CT voxel dimensions are 0.65 by 0.65 by 3.0 mm; the MR voxel dimensions are 1.1 by 1.1 by 4.0 mm. The marker is larger than a CT voxel, but only slightly so in the slice direction. Thus, the rendition is good within a CT slice, as seen in (b), but poor in the slice direction, as seen in (c). The MR voxels are larger than the CT voxels. Thus, the renditions are relatively poor, both within an MR slice, as seen in (e) and in the slice direction, as seen in (f).

The rendition of the marker in an image is related to the potential accuracy with which it can be localized. Smaller markers, as measured in voxels, will be more poorly localized than larger ones. The shape of a marker that is smaller than a voxel cannot be represented at all. More importantly, with regard to registration accuracy, such a marker can be situated entirely within a voxel with the result that the brightness "pattern" is independent of the marker's position within the voxel. Markers that span two or more voxels will ordinarily (absent reconstruction artifacts) produce a signal for each occupied voxel that is a monotonically increasing function of the volume of intersection of the voxel and the signal-producing contents of the marker. This effect is illustrated in Fig. 8.4. To first order, ignoring noise and reconstruction artifacts, the function will be linear,

$$I = aV + I_0 \,, \tag{8.18}$$

where $I$ is the voxel intensity and $V$ is the volume of intersection. $I_0$ is the intensity for an empty voxel, sometimes called a "background" voxel. While $I_0 = 0$ for most modalities, it is sometimes nonzero (CT, for example) and must be accounted for. If $I_i$ is measured for one voxel $i$ of a set of voxels that includes all those occupied by a given marker, an approximate centroid can be calculated by a weighted sum,

$$\mathbf{x} = \sum_i^n (I_i - I_0)\mathbf{x}_i / \sum_i^n (I_i - I_0) \,, \tag{8.19}$$

where $\mathbf{x}_i$ is the centroid of voxel $i$. (Note that the value of $a$ is not needed for this calculation.)

Equation (8.19) provides a good approximation, but it is flawed by the implicit assumption that the centroid of the marker-filled portion of each voxel is coincident with the centroid $\mathbf{x}_i$ of the voxel itself. The error can be appreciated in Fig. 8.4 by comparing the marker-filled portions of voxels $a$ and $b$, at the right and left, respectively. Voxel $a$ is almost completely filled, while only the right third of voxel $b$ is filled. The centroid of the marker-filled portion of $a$ lies approximately at the voxel centroid, $\mathbf{x}_a$, while the centroid of the marker-filled portion of $b$ lies somewhere between the white dotted border of the marker and the right edge of the voxel, which is considerably to the right of centroid $\mathbf{x}_b$. Thus, in Eq. (8.19), the contribution of voxel $b$ contributes an erroneous left shift to the marker centroid.

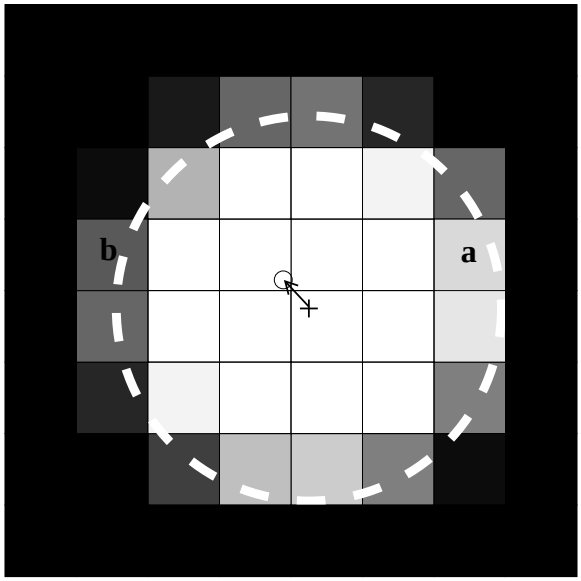



**Figure 8.4:** An illustration in two dimensions of intensity-weighted localization applied to a circular fiducial marker. The white, gray, and black squares represent voxels. The shadings of the voxels indicate their intensities, lighter shading indicating higher intensity. The marker is isolated from the anatomy such that none of the voxels contains both marker and anatomical structure. The black voxels around the periphery are empty and thus have the background intensity $I_0$. The dotted white outline is the edge of the marker. The small "plus" sign is the physical centroid of the marker. The small circle is the estimate of the centroid chosen by intensity-weighted localization. The small arrow is the fiducial localization error **FLE**. Standard intensity-weighted localization implicitly assumes that the centroid of the marker-filled portion of each voxel is coincident with the centroid of the voxel itself. This assumption is almost correct for voxel **a** but not for **b**.

This error can be appreciable and can be corrected for one-dimensional images, but only heuristic algorithms are available for two and three dimensions [36].

As indicated in Fig. 8.4, the localization error can be expected to be considerably smaller than the dimension of a voxel if the visible portion of the marker occupies many voxels. Large marker size (relative to voxel size) provides three benefits. First, the fraction of the voxels within the marker that are only partially filled tends to be smaller for larger markers, so their errors make a relatively smaller contribution to Eq. (8.19). Second, the number of partially filled voxels (as opposed to the fraction) tends to increase with marker size, resulting in a greater likelihood of cancellation among the erroneous shifts. Third, a larger marker results in an averaging of image noise over a larger number of voxels, both partially and completely filled. If the noise is additive and uncorrelated among the voxels, its effect on Eq. (8.19) can be expected to be smaller. Thus, larger markers tend to exhibit smaller FLEs. Brighter markers also have smaller FLEs because of the smaller contribution of image noise relative to marker intensity in Eq. (8.18).

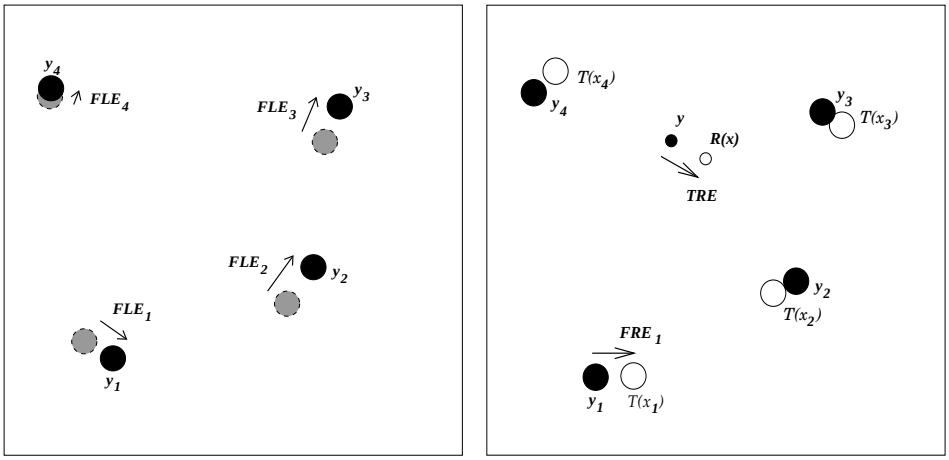
As mentioned at the beginning of Section 8.2, any nonzero displacement $\mathcal{T}(\mathbf{x})-\mathbf{y}$ between a transformed point $\mathcal{T}(\mathbf{x})$ and its corresponding point $\mathbf{y}$ is a registration error. To the extent that FLE is small and that the form of the transformation correctly describes the motion of the object, the alignment of the fiducial points in the two views will lead to small registration errors for all points. If the transformation is selected from some constrained set (as, for example, the rigid transformations), then, because of FLE, it will ordinarily not be possible to achieve a perfect alignment of fiducials. The resultant misalignment may, in some cases, be used as feedback to assess whether or not the registration is successful (but see the warnings on page 471). A common measure of overall fiducial misalignment, is the root-mean-square (RMS) error. This error, which we will call the *fiducial registration error*, or FRE, is defined as follows. First, we define an individual fiducial registration error,

$$\mathbf{FRE}_i = \mathcal{T}(\mathbf{x}_i) - \mathbf{y}_i \,, \tag{8.20}$$

where $\mathbf{x}_i$ and $\mathbf{y}_i$ are the corresponding fiducial points in views $X$ and $Y$, respectively, belonging to feature $i$, as depicted in Fig. 8.5b. We define FRE in terms of the magnitudes of the $\mathbf{FRE}_i$.

$$\mathrm{FRE}^2 = (1/N)\sum_i^N w_i^2 \mathrm{FRE}_i^2 \,, \tag{8.21}$$

where $N$ is the number of fiducial features used in the registration and $w_i^2$ is a non-negative weighting factor, which may be used to decrease the influence of less reliable fiducials. For example, if $\langle \mathrm{FLE}_i^2 \rangle$ is the expected squared fiducial localization error for fiducial $i$, then we may choose to set $w_i^2 = 1/\langle \mathrm{FLE}_i^2 \rangle$, where $\mathrm{FLE}_i$ is the fiducial localization error for fiducial $i$.




(a) Schematic of point-based registration illustrating *fiducial localization error* (FLE). Black circles represent positions $y$ at which points are determined by the localization process in one of two spaces involved in the registration process. The light circles represent the actual positions.

(b) Schematic of point-based registration illustrating two measures of registration error. Black circles represent positions $y$ in one space. The unfilled circles represent positions $x$ in the other space after they have been mapped by the registering transformation $\mathcal{T}$. The larger, numbered circles are the points used to effect the registration. *Fiducial registration error* (FRE) is the alignment error between these. *Target registration error* (TRE) is the registration error at a point (smaller circles) not used in the registration.

**Figure 8.5:** Errors in Point-Based Rigid Registration

Figure 8.5b also depicts *target registration error*, or **TRE**, which is, simply, registration error calculated at some point of interest,

$$\mathbf{TRE}(\mathbf{x}) = \mathcal{T}(\mathbf{x}) - \mathbf{y} \,. \tag{8.22}$$

The term "target" is meant to suggest that the point is the subject of some diagnosis or treatment. (As with FLE we use a bold font for both FRE and TRE to indicate the displacement and a normal font to indicate magnitude.)

### 8.3.1 Points in rigid transformations

If the transformation to be determined is constrained to be rigid, then Eq. (8.21) can be written as

$$\mathrm{FRE}^2 = (1/N) \sum_i^N w_i^2 |R\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i|^2 \,. \tag{8.23}$$

If the $\mathbf{FLE}_i$ are random errors with zero means and isotropic distributions for all fiducials, then an optimum registration can be achieved by minimizing $\mathrm{FRE}^2$ with $w_i = 1/\langle \mathrm{FLE}_i^2 \rangle$. The minimization of Eq. (8.23) is known as the "Orthogonal Procrustes" problem in the statistics literature.[2] Closed-form solutions for this problem have been available from that discipline since the first one was published by Green in 1952 [39]. The problem is also important in the theory of shape [40–42]. The solution is unique unless the fiducial configuration is linear (in which case $\mathrm{FRE}^2$ is independent of rotations about the fiducial line). Thus, we assume henceforth that the configuration is nonlinear, and, hence, that $N \geq 3$. Algorithm 8.1 provides a simple, reliable method of solution.

**Algorithm 8.1: Point-based, rigid registration**

Find $R$ and $\mathbf{t}$ to minimize $\sum_i^N w_i^2 |R\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i|^2$ :

1. Compute the weighted centroid of the fiducial configuration in each space:

$$\begin{aligned} \bar{\mathbf{x}} &= \sum_i^N w_i^2 \mathbf{x}_i / \sum_i^N w_i^2 \\ \bar{\mathbf{y}} &= \sum_i^N w_i^2 \mathbf{y}_i / \sum_i^N w_i^2 \,. \end{aligned}$$

[2] The term "Procrustes" was originally pejorative. It was first used by Hurley and Cattell [38] in 1962 to express disapproval of a perceived tendency of some to distort one set of observations to support the claim that they fit another set. Hurley and Cattell were drawing an analogy to the habits of the character of the same name from Greek mythology, who stretched, squeezed, or otherwise altered visitors to fit his guest bed. The term is now used in the statistical theory of shape with no negative connotation attached. The "orthogonal" version involves only rigid motion.

2. Compute the displacement from the centroid to each fiducial point in each space:

$$\begin{aligned} \tilde{\mathbf{x}}_i &= \mathbf{x}_i - \bar{\mathbf{x}} \\ \tilde{\mathbf{y}}_i &= \mathbf{y}_i - \bar{\mathbf{y}} \,. \end{aligned}$$

3. Compute the weighted fiducial covariance matrix:

$$H = \sum_i^N w_i^2 \tilde{\mathbf{x}}_i \tilde{\mathbf{y}}_i^t \,,$$

4. Perform singular value decomposition of $H$:

$$H = U \Lambda V^t \,,$$

where $U^t U = V^t V = I$, $\Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$, and $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$.

5. $R = V \text{diag}(1, 1, \det(VU)) U^t$.

6. $\mathbf{t} = \bar{\mathbf{y}} - R\bar{\mathbf{x}}$.

---

The diagonal matrix interposed between $V$ and $U^t$ in step 5 is there to insure that $R$ is a proper rotation. It is necessary only when the fiducial points are approximately planar or when the expected size of FLE is on the order of the separation of the fiducials.

(Historical Note: The use of singular value decomposition for this problem was introduced in 1966 by Schönemann along with the first proof of solution for general fiducial configurations [43]. That method was later rediscovered independently by others [44–46]. These methods all permitted improper rotations. The use of $\det(VU)$ to restrict the solution to proper rotations was presented without proof later in 1966 by Farrell [44]. It was rediscovered and proved in 1991 by Umeyama [47]. The original method by Green, which is based on the matrix square root, was later independently derived as well [48, 49]. Green's method was first proved to be a general solution in 1978 by Sibson [50]. A third solution method based on the unit quaternion was discovered in 1986 simultaneously by Faugeras [51] and Froimowitz [52] and then independently by Horn [8]. A recent comparison of these methods plus a method using dual quaternions, whose elements are dual numbers [53], shows no substantial differences among them [54].)

The application of Algorithm 8.1 minimizes $\text{FRE}^2$, but, as indicated in Fig. 8.5, finite fiducial localization error FLE can be expected to make both FRE and TRE nonzero. The relationships among the expected values of $\text{FLE}^2$, $\text{FRE}^2$, and $\text{TRE}^2$ are known to an excellent approximation for the most common case, in which the

$\mathbf{FLE}_i$ are random, their probability distributions are isotropic, independent, and identical for all $N$ fiducials, and a uniform weighting is used in the registration ($w_i^2 = 1$). The simplest relationship is that between the expected values of $\mathrm{FLE}^2$ and $\mathrm{FRE}^2$,

$$\langle \mathrm{FRE}^2 \rangle \approx (1 - 2/N) \langle \mathrm{FLE}^2 \rangle \,, \tag{8.24}$$

where the approximation ignores terms involving $\mathrm{FLE}^4$ and higher [55]. This equation and Eqs. (8.25) through (8.29), which follow below, apply both to the situation in which $\langle \mathrm{FLE}^2 \rangle$ arises from the distribution of fiducial localization error in one only space, with the fiducial localization error in the other space being negligible, and to the situation in which there is error in both spaces, in which case $\langle \mathrm{FLE}^2 \rangle$ equals $\langle \mathrm{FLE}_a^2 \rangle + \langle \mathrm{FLE}_b^2 \rangle$, where the subscripts $a$ and $b$ refer to the two spaces [56].

It should be noted that the relationship given by Eq. (8.24) is *completely independent of the fiducial configuration*. Furthermore, the same holds for the probability distribution of $\mathrm{FRE}^2$ [55]. Thus, for a given number of fiducials and given distribution for **FLE**, $\mathrm{FRE}^2$ will be distributed in the same way regardless of whether the fiducials are clustered within a few millimeters of each other or are miles apart, nor do, planar, regular polyhedral, or approximately linear arrangements exhibit any differences with regard to FRE. An advantage of this independence of configuration is that an estimate of $\langle \mathrm{FLE}^2 \rangle$ can be obtained for a given fiducial design and image acquisition protocol by performing a set of registrations involving pairs of images of possibly differing configurations and numbers of fiducials. An estimate for the expected value of $\mathrm{FLE}^2$ can be obtained from the weighted average,

$$\langle \mathrm{FLE}^2 \rangle \approx (1/M) \sum_{m}^{M} N_m / (N_m - 2) \times \mathrm{FRE}^2(m) \,, \tag{8.25}$$

where $M$ is the number of registrations performed and $N_m$ is the number of fiducials involved in registration $m$, and $\mathrm{FRE}^2(m)$ is the FRE for registration $m$.

Unlike $\langle \mathrm{FRE}^2 \rangle$, the value of $\langle \mathrm{TRE}^2 \rangle$ does depend on the fiducial configuration, and it depends on the target position as well. It is most easily stated in terms of quantities measured relative to the principal axes of the fiducial configuration:

$$\langle \mathrm{TRE}^2 \rangle \approx \frac{1}{N} \left( 1 + \frac{1}{3} \sum_{k=1}^{3} \frac{d_k^2}{f_k^2} \right) \langle \mathrm{FLE}^2 \rangle \,, \tag{8.26}$$

where $d_k$ is the distance of the target from principal axis $k$, and $f_k$ is the RMS distance of the fiducials from the same axis [57]. This approximation, like that of Eq. (8.24), ignores terms involving $\mathrm{FLE}^4$ and higher. An equivalent approximation was first suggested by Hawkes [58, 59].

FRE is sometimes used as an indication of accuracy in a point-based registration system. Unfortunately, it may in fact be a very poor indicator of accuracy. First

of all, as we have seen from Eqs. (8.26) and (8.24), $\langle \mathrm{TRE}^2 \rangle$ is a function of the fiducial configuration, while $\langle \mathrm{FRE}^2 \rangle$ is not. Thus, for example, an approximately linear fiducial configuration will result in a very small value for one of the $f_k$ in the denominator of Eq. (8.26). If the target marker is at an appreciable distance from the line of fiducials, then the corresponding $d_k$ in the numerator will not be small. The result will be a large $\langle \mathrm{TRE}^2 \rangle$, even if $\langle \mathrm{FLE}^2 \rangle$, is small. For this same configuration, $\langle \mathrm{FRE}^2 \rangle$, on the other hand, will be small. By contrast, if the target is located near the centroid of a non-linear fiducial configuration, then all the $d_k$ will be small, none of the $f_k$ will be small, and $\langle \mathrm{TRE}^2 \rangle$ can be smaller than $\langle \mathrm{FRE}^2 \rangle$. Second, if the fiducial configuration and target position are kept constant from registration to registration, variation in FRE is no indicator of variation in TRE for a properly working system. No correlation between these two quantities for a given system has been observed. Thus, for a given registration with a properly working point-based registration system, a value of FRE that is lower or higher than expected for that system does not indicate that the value of TRE for that registration is lower or higher than expected. Third, the registration error $\mathrm{FRE}_i$ for a specific marker $i$ is a perverse indicator of TRE in the vicinity of that marker. The problem can be seen from the following relationship, which, like Eqs. (8.24) and (8.26), is accurate to third order in FLE [57]:

$$\langle \mathrm{TRE}^2(\mathbf{x}_i) \rangle \approx \langle \mathrm{FLE}^2 \rangle - \langle \mathrm{FRE}_i^2 \rangle \,. \tag{8.27}$$

A consequence of this equation is that, for a given fiducial configuration, TRE tends to be worse in the vicinity of those fiducials whose alignments tend to be better! (See [57] for an explanation of this anti-intuitive behavior.)

If interpreted correctly, FRE can, however, provide useful feedback. It can be seen from Eqs. (8.24) and (8.26) that systems with a smaller $\langle \mathrm{FRE}^2 \rangle$ will have a smaller $\langle \mathrm{TRE}^2 \rangle$. These expected values are in fact proportional and do show that a system that tends to exhibit a lower FRE will also tend to produce a lower TRE. A very large FRE on a given registration with a point-based system that usually exhibits a small FRE may indicate that at least one part of the registration system, which includes image acquisition, fiducial localization, and fiducial registration, may not be working properly. A formal statistical decision about the likelihood that a given system with a known $\langle \mathrm{FRE}^2 \rangle$ is performing correctly can be based on the probability that the observed value of FRE will occur at random. For that calculation, it is necessary to use the probability distribution of FRE, which is chi-square with 3N-6 degrees of freedom [55, 57].

### 8.3.2 Points in scaling transformations

If the transformation to be determined is a similarity transformation, i.e., is of the form of Eq. (8.7), then Eq. (8.21) becomes

$$\text{FRE}^2 = (1/N)\sum_i^N w_i^2 |sR\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i|^2 \,. \tag{8.28}$$

A simple extension of Algorithm 8.1 determines the scaling factor $s$, rotation $R$, and translation $\mathbf{t}$ that minimize $\text{FRE}^2$. The extension is given by Algorithm 8.2 [60].

**Algorithm 8.2: Point-based registration: Similarity Transformations**

Find $R$, $\mathbf{t}$, and $s$ to minimize $\sum_i^N w_i^2 |sR\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i|^2$ :

1. Set $s = 1$ and determine $R$ using steps 1 through 5 of Algorithm 8.1.
2. Calculate $s$: $\quad s = \frac{\sum_i^N w_i^2 R\tilde{\mathbf{x}}_i \cdot \tilde{\mathbf{y}}_i}{\sum_i^N w_i^2 \tilde{\mathbf{x}}_i \cdot \tilde{\mathbf{x}}_i}$ .
3. $\mathbf{t} = \bar{\mathbf{y}} - sR\bar{\mathbf{x}}$.

A weakness of this solution is that the scaling factor $s$ does not have a reciprocal relationship with the factor in the reverse problem. Thus, if the point sets, $\{\mathbf{x}_i\}$ and $\{\mathbf{y}_i\}$ are exchanged, the new scale factor will not necessarily equal $1/s$. The analog to Eq. (8.24) for similarity transformations is

$$\langle \text{FRE}^2 \rangle \approx (1 - 7/(3N))\langle \text{FLE}^2 \rangle \,, \tag{8.29}$$

and the probability distribution of FRE is chi-square with $3N - 7$ degrees of freedom [56].

There are no known closed-form solutions for the anisotropic scaling problems, in which the scaling matrix $S$, the rotation matrix $R$, and the translation vector $\mathbf{t}$ are to be determined for transformations of the form of Eqs. (8.5) or (8.6). Instead, an iterative search is required. For scaling of the form of Eq. (8.5) the problem is especially simple. This problem, which is encountered when the $x$, $y$, and $z$ scales of an imaging device are incorrect to differing degrees, as, for example, with inaccurate MR gradients (see page 462 in Section 8.2.3), reduces to a search of the three-dimensional space of diagonal elements of $S$, as described by Algorithm 8.3.

**Algorithm 8.3: Point-based registration: Nonisotropic scaling**

Find $R$ , $\mathbf{t}$, and $S$ to minimize $\sum_i^N w_i^2 |RS\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i|^2$ :

1. Perform steps 1 and 2 of Algorithm 8.1 to determine the centroids, $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$, in each space and the displacement of each fiducial point $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{y}}_i$ from the centroid in its respective space.
2. Set iterative count $n = 0$.
3. Choose an initial scaling matrix $S^{(0)}$.
4. Repeat the following steps:
   (a) Set $\tilde{\mathbf{x}}_i^{(n)} = S^{(n)}\tilde{\mathbf{x}}_i$.
   (b) Perform steps 3 through 5 of Algorithm 8.1 to find $R$.
   (c) Add one to $n$.
   (d) Determine a new value of $S^{(n)}$.
5. Stop when FRE $<$ threshold or $n >$ maximum iteration count.
6. $\mathbf{t} = \bar{\mathbf{y}} - RS\bar{\mathbf{x}}$.

While the problem is reduced to a three-dimensional search for the optimal $S$, no method is specified for conducting that search. In particular, no methods are given for steps 3 or 4d of Algorithm 8.3. A review of methods for multidimensional optimization can be found in [61].

For scaling of the form of Eq. (8.6), the problem is complicated by the fact that $S$ is applied to the rotated $\tilde{\mathbf{x}}_i$. As a result, the substitution in $\tilde{\mathbf{x}}_i^{(n)} = S^{(n)}\tilde{\mathbf{x}}_i$ in step 4a is no longer correct. Instead, a six-dimensional parameter space must be searched. The space consists of the diagonal elements of $S$ and three independent parameters to specify $R$, e.g., three Euler angles (see Section 8.2.1). Once these parameters are determined, the translation is $\mathbf{t} = \bar{\mathbf{y}} - SR\bar{\mathbf{x}}$.

### *8.3.3 Points in perspective projections*

When tomographic images are used in conjunction with images acquired by means of x-ray projection, endoscopy, laparoscopy, microscopy, direct video, or photography, or when any of these latter modalities are used to guide surgery or other therapy to a target location, a 2-D-to-3-D registration problem must be solved. When several corresponding point pairs can be identified in two-dimensional and three-dimensional views of a single rigid object, the problem can be solved by

means of point-based, rigid registration. This problem is common in robot and computer vision applications and has been thoroughly studied in that domain, where it is known as the "object-pose estimation problem" [6, 10, 62, 63].

The form of the transformation associated with these projections from three dimensions to two is the perspective transformation. That transformation is described in Section 8.2.2.4 in terms of the pinhole camera and is given by Eq. (8.12) and substitution (8.13). Without loss of generality, the problem can be treated by orienting the coordinate system so that $\hat{\mathbf{p}}$ is aligned with the $z$ axis and by placing the origin at the pinhole. With these choices, we have $\mathbf{x} \cdot \hat{\mathbf{p}} = z$ and $\alpha = 0$. Eq. (8.12) then simplifies to

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \frac{f}{z} \begin{pmatrix} x \\ y \\ z \end{pmatrix} , \tag{8.30}$$

which transforms all points to the $z' = f$ plane. Solving Eq. (8.30) for $x$, $y$, and $z$ leads to $\mathbf{x} = \lambda \mathbf{x}'$, where $\lambda$ is an arbitrary scalar. Thus, a given two-dimensional image point, $(x', y')^t$ lying in the $z' = f$ image plane is projected to a line in three-dimensional space that passes through the three-dimensional origin (the pinhole) and through the three-dimensional image point, $(x', y', f)^t$.

Thus, it is not possible to determine a unique three-dimensional position for a single point from its perspective projection alone. Some additional information is required to determine where on the projected line the point lies. At least one other constraint must be provided to resolve the ambiguity. A common constraint in computer vision that is uncommon in medical applications is $ax+by+cz+d = 0$, which means that the point is known to lie on a given plane (e.g., the floor, a wall, etc.). Another constraint restricts the line to intersect a "silhouette curve," which is the apparent outline of a three-dimensional surface when viewed in perspective from a given point. The use of this latter constraint for registration of tomographic images of the anatomy to radiographic images is described in Chapter 6. In the most common point-based object-pose estimation problem, the projections $(x'_i, y'_i, f)^t$ of $N$ three-dimensional points $\mathbf{x}_i, i = 1, \ldots, N$ are used to determine all $N$ $\mathbf{x}_i$ simultaneously. The additional information needed to solve the problem is provided by (a) stipulating that the $N$ $\mathbf{x}_i$ all lie on the same rigid object and (b) giving their three-dimensional positions for some orientation of that object. Thus, (a) it is known that, for some $R$ and $\mathbf{t}$, $R\mathbf{x}_i + \mathbf{t} = \mathbf{y}_i$, and (b) the points $\mathbf{y}_i, i = 1, \ldots, N$ are given. In medical applications, these points are typically localized from a tomogram of rigid anatomy. The pose estimation is accomplished by searching for the $R$ and $\mathbf{t}$ that solve the $2N$ equations represented by the $x$ and $y$ components of Eq. (8.30) for the $N$ points. The equations are typically solved in the least-square sense, and because of the nonlinear constraints on the form of $R$, they must be solved iteratively. A common approach is carried out in terms of an $R$ that is parameterized by rotation angles, such as the Euler angles (see Section 8.2.1): start with an approximate

solution $\mathbf{x}_i, i = 1, \ldots, N$, calculate the perspective projections of the $\mathbf{x}_i$, determine the sum of squared errors between the calculated projections and the actual projections, linearize the resulting expression (assumes small error), solve for the first-order corrections to the angles and $\mathbf{t}$, and iterate [10]. If the camera constant $f$ is unknown, then a similar approach is required with $f$ included as an additional unknown. This latter problem is known as the "interior orientation" problem in photogrammetry [10].

### *8.3.4 Points in curved transformations*

When points are used for registration with curved transformations, it is usually possible to align all fiducial points exactly, but the mapping of the non-fiducial points can vary strongly with the form of transformation used. Knowledge of the genesis of the morphological changes from one view to the other should be used to determine the class of curved transformations to be employed whenever possible, as, for example, with rectification (see Section 8.2.3) or with interpatient registration (see Chapters 3 and 17). Otherwise, $\mathcal{T}$ is typically chosen from some set of continuous, differentiable functions, such as the polynomials, the polynomial splines, or the thin-plate splines. For example, an interpolating polynomial (see Section 8.2.2) of the form of Eq. (8.14) can reduce FRE to zero at $(I+1)(J+1)(K+1)$ points. The thin-plate spline of Eq. (8.17) reduces FRE to zero at $N$ points. The cubic spline of Eq. (8.15), which is typically implemented in terms of B-splines, can reduce FRE to zero at a given set of points, but only for very uniform point distributions or for very closely spaced knots. These restrictions have limited the use of B-splines as interpolants in medical applications of point-based registration primarily to rectification problems based on phantoms with uniformly spaced fiducials (but see Section 8.5.3.5 for an application to intensity-based registration). In these situations, the knot spacing is typically forced to be uniform by the machining of the phantom. For this case, the cubic B-spline expansion can be given in terms of the simple segments of splines given by Eqs. (8.16).

With the polynomial, cubic-spline, and thin-plate spline transformations [and with the affine transformation of Eq. (8.8)], the point-based registration problem reduces to the determination of a set of coefficients. There is in fact a separate problem for each of the three components of $\mathbf{x}'$. For each class of transformation, the problem can be solved individually for each component by finding the solution to a set of simultaneous linear algebraic equations for coefficients that can, in each case, be organized into a one-dimensional vector $\mathbf{a}$. The meaning of the elements of $\mathbf{a}$ varies with each class of transformation. The vector for the $x$ coordinates, for example, is uniquely determined by using the $x, y, z$ coordinates of the $N$ fiducial localizations $\mathbf{x}_i$ on the right side of Eqs. (8.14), (8.15), or (8.17) and substituting the corresponding $x$ coordinates of the $N$ fiducial localizations $\mathbf{x'}_i$ into the left side. The form of the resultant equation is in every case,

$$M\mathbf{a} = \mathbf{b}. \tag{8.31}$$

Thus, the point-based registration problem with any of these transformations is reduced to the solution of three sets of linear algebraic equations—one set each for $x$, $y$, and $z$. While there are more coefficients involved, this situation is simpler than that of the rigid-body transformation problem (see Algorithm 8.1) because it lacks the nonlinear side-conditions that insure that $R$ is orthogonal. For the polynomial transformation of Eq. (8.14), with points $\mathbf{x}_n$, $n = 1, \ldots, N$ where $N$ equals $(I + 1)(J + 1)(K + 1)$, the matrix $M$ is $N \times N$, and its elements are the polynomials $x_n^i y_n^j z_n^k$. $M$ is nonsingular but tends to be ill-conditioned [64]. (Techniques for increasing numerical accuracy with polynomial interpolation can be found in [61].) For the cubic-spline it is necessary to impose some additional conditions along with Eq. (8.15) in order to produce as many equations as unknowns. A common condition is to restrict the second derivatives to be zero at the borders, producing the so-called "natural splines." $M$ is again nonsingular. For the thin-plate splines, both two-dimensional and three dimensional, side conditions are necessary for each component in order to insure that Eq. (8.31) has a solution. With the $x$ component of $\mathbf{x}'$ again serving as the example, the conditions are for the two-dimensional case [17]

$$\sum_i^N c_i = \sum_i^N x_i c_i = \sum_i^N y_i c_i = 0 \tag{8.32}$$

with a similar condition for $z$ in the three-dimensional case. When these conditions are incorporated into Eq. (8.31) along with Eq. (8.17), $M$ is $(N + 3) \times (N + 3)$ for two dimensions, or $(N + 4) \times (N + 4)$ for three dimensions, and nonsingular.

Transformations derived from the equations of continuum mechanics can be used with point-based registration as well. In this case the disparity between corresponding points may be used to determine the forces that drive the displacements. The correct form of the force as a function of these disparities is not clear, nor is the means to interpolate that force to other points. Recently the three-dimensional thin-plate spline has been employed for this purpose for the registration of breast images [30].

It should be pointed out that the use of curved transformations to align all fiducial points exactly, i.e., to interpolate them, is sometimes not a sensible approach. The problem is that the localization of the points will inevitably include some localization error. Aligning these erroneous points exactly will result in the propagation of the error throughout the space. To accommodate localization error, the polynomial and cubic-spline interpolants should be employed with more fiducial points than coefficients to allow for a least-squares fit in analogy with the rigid-body case. In that case, Eq. (8.31) has more unknowns than equations. The least-squares approximation may then be found with standard techniques [61, 64].

## 8.4 Surface-based methods

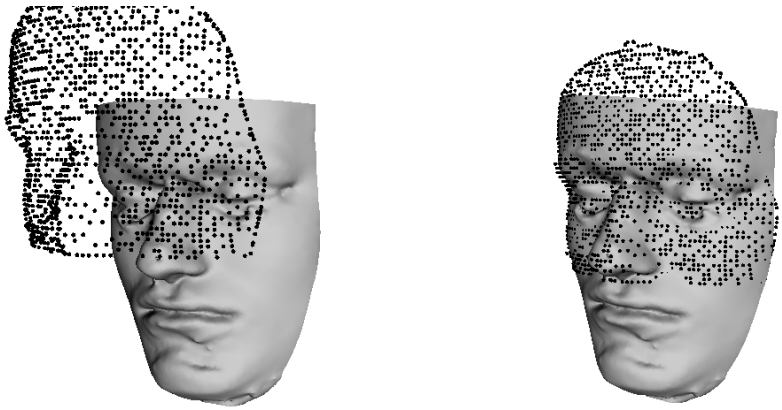
The 3-D boundary surface of an anatomic object or structure is an intuitive and easily characterized geometrical feature that can be used for medical image registration. Surface-based image registration methods involve determining corresponding surfaces in different images (and/or physical space) and computing the transformation that best aligns these surfaces.

The skin boundary surface (air-skin interface) and the outer cranial surface are obvious choices that have frequently been used for both image-to-image (e.g., CT-MR, serial MR) and image-to-physical registration of head images (see Fig. 8.6). The surface representation can be simply a point set (i.e., a collection of points on the surface), a faceted surface (e.g., triangle set), an implicit surface, or a parametric surface (e.g., B-spline surface). Extraction of a surface such as the skin or bone is relatively easy and fairly automatic for head CT and MR images. Extraction of many soft tissue boundary surfaces is generally more difficult and less automatic. Image segmentation algorithms can generate 2-D contours in contiguous image slices that are linked together to form a 3-D surface, or they can generate 3-D surfaces directly from the image volume. Surface extraction from images is covered in detail in Chapter 6. The related problem of two-dimensional contour extraction is treated in Chapter 3. In physical space, skin surface points can be easily determined using laser range finders; stereo video systems; and articulated mechanical, magnetic, active and passive optical, and ultrasonic 3-D localizers. Bone surface points can be found using tracked A-mode [65] and B-mode [66] ultrasound probes. The computer vision sensors, 3-D localizers, and tracked A-mode ultrasound probes produce surface point sets. Tracked B-mode probes produce a set of 2-D images (or a single compounded 3-D image) from which bone surface points need to be segmented.

Surfaces can provide basic features for both rigid-body and nonrigid registration. A central and difficult question that must be addressed by any nonrigid surface-based registration algorithm is how deformation of the contents of an object is related to deformation of the surface of the object. Most of the surface-based registration algorithms that have been reported are concerned with rigid-body transformation, occasionally with isotropic or nonisotropic scaling. Thus, in this section, we are concerned primarily with rigid-body registration. The registration of 2-D projection images and 3-D tomographic images is considered in Chapter 6.

### *8.4.1 Disparity functions*

There is a large body of literature in computer vision concerned with the surface-based registration problem. Some of this work has addressed limited classes of shapes, e.g., polyhedral models [69] and piecewise-superquadric models [70, 71]. Other authors rely on the existence of specific simple features or make assumptions about the global shape (see [72] for an extensive review). For example, the work in [51] assumes the existence of reasonably large planar regions within a free-form



(a) (b)

**Figure 8.6:** Illustration of surface-based registration of the head using the facial skin surface. The dots represent skin surface points acquired with a four-camera photogrammetry system. The surface rendering represents a triangle set model of the skin surface extracted from an MR image volume. The left panel (a) shows the initial position of the data sets. The right panel (b) shows the data sets after registration. The registration was performed using an independent implementation [67] of the iterative closest point algorithm [68]. The surfaces overlap only partially, which is a common situation in surface-based registration. The problem of partial overlap was dealt with by setting the weights of outliers to zero after the first search converged before running a second search.

shape. Some approaches represent the surface with generalized cylinders [73] or extended Gaussian images [74]. The latter method allows matching of convex and restricted sets of nonconvex shapes based on surface normal histograms.

The approach for solving the surface-based registration problem that is frequently used in more recent computer vision literature (where it is often called the "free-form" surface matching problem), and that is normally used in the medical image processing community, is to search for the transformation that minimizes some disparity function or metric between the two surfaces $X$ and $Y$. The disparity function is generally a distance. In mathematics, the distance between two feature sets $A$ and $B$ is normally defined as the minimum distance between a point in $A$ and a point in $B$, namely, $d(A,B) = \min_a \min_b d(\mathbf{a},\mathbf{b})$, where $\mathbf{a} \in A$, $\mathbf{b} \in B$, and $d(\mathbf{a},\mathbf{b})$ is the Euclidean distance between $\mathbf{a}$ and $\mathbf{b}$. Thus $d(A,B)$ is small if one pair of points in these two sets are close. Another common measure is the Hausdorff distance. The Hausdorff distance from $A$ to $B$ is $d(A,B) = \max_a \min_b d(\mathbf{a},\mathbf{b})$, and the Hausdorff distance between two sets $A$ and $B$ is $d_H(A,B) = \max(d(A,B), d(B,A))$. Thus $d_H(A,B)$ is small if and only if every point of $A$ is close to a point in $B$ and vice versa. Neither of these distance measures is very useful as a disparity function. The Hausdorff distance is not meaningful in practice because the corresponding surfaces generally overlap only partially, and because it is sensitive to statistical outliers. The disparity function normally used for surface-based image registration is an average, and optionally weighted, distance between points on one surface and corresponding points on the other surface. Let $\{\mathbf{x}_j\}$ for $j = 1, \ldots, N_x$ be a set of $N_x$ points on the surface $X$. The general approach is to search for the transformation that minimizes the disparity function

$$d(\mathcal{T}(X), Y) = \sqrt{\sum_{j=1}^{N_x} w_j^2 d^2(\mathcal{T}(\mathbf{x}_j), Y)} = \sqrt{\sum_{j=1}^{N_x} w_j^2 ||\mathcal{T}(\mathbf{x}_j) - \mathbf{y}_j||^2}\,, \tag{8.33}$$

where

$$\mathbf{y}_j = \mathcal{C}(\mathcal{T}(\mathbf{x}_j), Y) \tag{8.34}$$

is a point on the surface $Y$ "corresponding" to the point $\mathbf{x}_j$, $\mathcal{C}$ is a "correspondence" function (e.g., closest point operator), and $\{w_j\}$ is a set of weights associated with $\{\mathbf{x}_j\}$. We note that Eq. 8.33 is similar to Eq. 8.21 of point-based registration. The principal difference between point-based registration and surface-based registration is in the availability of point correspondence information. It is the lack of exact point correspondence information that causes surface-based registration algorithms to be based on iterative search. Eq. 8.34 merely provides approximate point correspondence information for a particular $\mathcal{T}$ during an iterative search.

The point set $\{\mathbf{x}_j\}$ and the surface $Y$ have been called, respectively, the "hat" and "head" [75], the "dynamic" and "static" feature sets [76], and the "data" point

set and "model" surface shape [68]. Typically one surface contains more information than the other. The surface from the image that covers the larger volume of the patient and/or has the highest resolution is generally picked as the model shape.[3] For example, when using the skin surface to perform CT-to-physical space registration, the triangle set representation of the CT skin surface typically contains $\sim 10^4$–$10^5$ vertices, whereas the number of physical space skin surface points typically is $\sim 10^2$–$10^3$. In this case, the CT triangle set is chosen as the model surface shape and the physical space point set is chosen as the data point set. We note that Eq. 8.33 is generally not symmetric in $X$ and $Y$.

In point-based registration (Eq. 8.21), if the fiducial localization error $\mathbf{FLE}_i$ for each point is independently random and normally distributed around the true position, then the maximum likelihood estimate of the transformation parameters is obtained by weighting the $i$th point by $1/\langle \mathrm{FLE}_i^2 \rangle$ in Eq. 8.21. In surface-based registration (Eq. 8.33), statistical independence of errors is unlikely. For example, if an isointensity surface is extracted from an image, and the isointensity value used is too high or too low, the extracted surface relative to the actual surface will be either slightly dilated or slightly eroded. The skin is a movable and deformable structure, and local deformations tend to be highly correlated. Physical space surface points acquired with a sensor can have biased error due to miscalibration. Nonetheless, weights can be useful to reduce the influence of less reliable surface points. For example, the gradient magnitude of an edge in an image can provide information about surface segmentation error at that point. Often a triangle set representation of a surface is extracted and simplified to reduce storage requirements and execution time. Some surface simplification algorithms provide an estimate of error arising from the simplification process. Many sensors and tracking devices have less accuracy at the edges of the working volume. Weights could potentially be used to account for the sensitivity of the registration to the perturbation of individual surface points (e.g., see the geometrical constraint analysis in [77]). For example, a few points in areas of high curvature might be more important to the registration than many points in relatively planar regions. Weights can be used to account for nonuniform sampling density [67]. Finally, weights can also be used to deal with outliers that can arise from nonoverlapping sections of surfaces, as seen at the top of Fig. 8.6(b), poor segmentation, and erroneous sensor data [67, 78]. For example, outliers can be defined as points whose distance from the closest corresponding surface point is more than two or three standard deviations above the mean distance. The weights of outliers can be reduced or set to zero, either after each iteration, or after one search converges before running another search.

Many variations of Eq. 8.33 are possible. For example, mean distance can be

[3] The terms "data" and "model" arise from an industrial application: registration of digitized data from unfixtured rigid objects obtained using high-accuracy noncontact devices with an idealized geometrical (e.g., computer-aided design) model prior to shape inspection (Besl & McKay wrote [68] while working at General Motors Research Laboratories).

used rather than root-mean-square distance (i.e., L1 norm versus L2 norm) [79]. Outliers can be handled using a thresholded distance, or by using a sigmoidal distance function (which is essentially a gradually tapered thresholded distance). If surface points are collected in physical space using a probe with a spherical ball at the tip, the recorded surface points $\mathbf{x}_j$ are displaced from the actual surface $Y$ by the radius $r$ of the ball-point tip. This is true for any orientation of the localization probe relative to the surface normal at the point of contact.[4] In this case, a more appropriate disparity function is [78]

$$d(\mathcal{T}(X), Y) = \sqrt{\sum_{j=1}^{N_x} w_j^2 (||\mathcal{T}(\mathbf{x}_j) - \mathbf{y}_j|| - r)^2} \,. \tag{8.35}$$

### 8.4.2 *Head and hat algorithm*

The first investigators to apply surface-based registration to a medical problem were Pelizzari, Chen, and colleagues [75]. They used their "head and hat" algorithm to register CT, MR, and PET images of the head. The "hat" is a skin surface point set $\{\mathbf{x}_j\}$. The "head" is a polygon set model of the skin surface $Y$ created by segmenting contours in contiguous transverse image slices. They define $\mathbf{y}_j$ as the intersection with the "head" $Y$ of a line joining the transformed "hat" point $\mathcal{T}(\mathbf{x}_j)$ and the centroid of the "head" $Y$. The intersection is efficiently calculated by reducing the 3-D line-polyhedron intersection problem to a 2-D line-polygon intersection problem. The transformation $\mathcal{T}$ that minimizes Eq. 8.33 is found using a standard gradient descent technique. The major limitations of this technique are due to the particular distance used, i.e., the distance from the surface point to the surface intersection along a line passing through the surface centroid. This definition of distance requires that the surface be approximately spherical. It also requires that a good initial transformation be supplied as input to the transformation parameter search. Finally, it is probably related to the observation by the authors and others that the search frequently terminates in local minima and thus requires substantial user interaction.

### 8.4.3 *Distance definitions*

A more general definition of distance between a point and a surface is the distance between the point and the closest point on the surface. That is, the correspondence function $\mathcal{C}$ in Eq. 8.34 is the closest point operator, and $\mathbf{y}_j$ is the point on the surface $Y$ closest to the transformed point $\mathcal{T}(\mathbf{x}_j)$. The closest point and distance

[4] A localization probe is often calibrated by placing the ball-point tip in a hemispherical divot and pivoting the probe about the center of the divot. The position of the probe tip relative to the coordinate system of the probe is determined by finding the most invariant point (in a least squares sense) in these pivot motions. This invariant point is the center of the ball-point tip. That is, the ball-point tip pivots about the center of the ball rather than a point on the surface of the ball.

calculation depends on the surface representation. For example, a common representation is a triangle set. Let $t$ be the triangle defined by the three vertices $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$. The distance between the point $\mathbf{x}_j$ and the triangle $t$ is

$$d(\mathbf{x}_j, t) = \min_{u+v+w=1} ||u\mathbf{r}_1 + v\mathbf{r}_2 + w\mathbf{r}_3 - \mathbf{x}_j|| \,, \tag{8.36}$$

where $u \in [0,1]$, $v \in [0,1]$, and $w \in [0,1]$. The required closed-form computations are straightforward. Let $T = \{t_i\}$ for $i = 1, \ldots, N_t$ be a set of $N_t$ triangles. The distance between the point $\mathbf{x}_j$ and the triangle set $T$ is given by

$$d(\mathbf{x}_j, T) = \min_{i \in \{1,\ldots,N_t\}} d(\mathbf{x}_j, t_i) \,. \tag{8.37}$$

The closest point $\mathbf{y}_j$ on the the triangle set $T$ satisfies the equality $d(\mathbf{x}_j, \mathbf{y}_j) = d(\mathbf{x}_j, T)$. Another relatively common representation is a parametric surface $S = \{\mathbf{r}(u,v)\}$, where, for example, the components of $\mathbf{r}$ are spline functions. The point-to-parametric-entity distance $d(\mathbf{x}_j, S)$ is easily computed with a 2-D minimization in $u$ and $v$ using derivatives of $\mathbf{r}$ (e.g., Newton's method). A good initial value $(u_0, v_0)$ is necessary. Initial values can be obtained by approximating the parametric surface with a triangle set, tagging each triangle vertex with the corresponding $(u,v)$ value of the parametric surface, and then, for each $\mathbf{x}_j$, finding the closest triangle vertex and using as the initial value the value stored with that vertex.

The most computationally expensive step in the registration process is finding the closest points. Given a data shape point set with $N_x$ points and a model shape with $N_y$ geometrical primitives (point, line segments, and triangles), the computational complexity of finding the closest points using an exhaustive search is $O(N_x N_y)$. One way to speed up the search process is to use a special data structure. One such data structure is a multidimensional binary search tree ($k$-D tree, where $k$ is the dimensionality of the search space) [80,81]. For 3-D image registration, $k = 3$. The $k$-D tree is a binary tree in which each node represents a subset of data records (encoding in our case, geometrical primitives) and a partitioning of that subset. Each nonterminal node has two children that represent the two subsets defined by the partitioning. The terminal nodes represent mutually exclusive small subsets of the records called buckets. A 3-D tree divides space into a collection of rectangular parallelepipeds that correspond to the terminal nodes. This data structure provides an efficient method for examining only those points closest to a given point. A $k$-D tree can be constructed in $O(N_y \log N_y)$ time. Each closest point search can be performed in $O(\log N_y)$ time. Thus the computational complexity of finding $N_x$ closest points at each step of an iterative transformation parameter search using a $k$-D tree is $O(N_x \log N_y)$.

### *8.4.4 Distance transform approach*

The calculation of point-to-surface distance is computationally intensive, even when using special data structures and other optimizations. A computationally efficient alternative is to use a distance transform (DT). A DT of a binary image $I$

is an assignment to each voxel $\mathbf{v}$ of the distance between $\mathbf{v}$ and the closest feature voxel in $I$. A DT of a binary image where the feature voxels are surface voxels is a gray-level image in which each voxel $\mathbf{v}$ has a value that is the distance from the center of $\mathbf{v}$ to the center of the nearest surface voxel. Thus a DT provides a method for precomputing and storing point-to-surface distance. Normally squared distance is stored. Then, at each step of an iterative transformation parameter search, the value of the disparity function in Eq. 8.33 is computed simply by summing the values of the voxels in the squared distance image that contain the transformed points $\{\mathcal{T}(\mathbf{x}_j)\}$. One limitation of this approach is that a DT is spatially quantized, i.e., a DT image contains exact point-to-surface distance only at regularly spaced lattice points (centers of voxels). A slight improvement over using the distance at the nearest lattice point can be achieved by using a trilinear interpolation of the distances at the nearest eight lattice points. Nonetheless, the surface is fundamentally represented by the point set consisting of the centers of all feature (surface) voxels, and thus subvoxel surface position information is lost. Spatial quantization might be the reason that registrations produced by surface-based methods using DTs have been reported to be considerably less accurate than registrations produced by surface-based methods not using DTs (e.g., see [82]). Integer approximations of Euclidean distance (e.g., the chamfer 3-4-5 algorithm [83]) are frequently used to reduce the DT computation time, but several reasonably efficient algorithms for computing an exact Euclidean DT now exist (e.g., [84]). Surface-based registration using a DT was first applied to medical images by [85].

### *8.4.5 Iterative closest point algorithm*

All surface-based registration algorithms must search for the transformation $\mathcal{T}$ that minimizes the disparity function in Eq. 8.33 or a variation thereof. This is a general nonlinear minimization problem that is typically solved using one of the common gradient descent techniques (e.g., see [61]). The search will typically converge to, or very close to, the correct minimum of the disparity function minimum if the initial transformation is within about 20–30 degrees and 20–30 mm of the correct solution. To help minimize the possibility of the search getting stuck in a local minimum, many investigators perform the search in a hierarchical coarse-to-fine manner.

Besl & McKay [68] presented an algorithm which reduces the general nonlinear minimization problem to an iterative point-based registration problem. Their iterative closest point (ICP) algorithm is a general-purpose, representation-independent, shape-based registration algorithm that can be used with a variety of geometrical primitives including point sets, line segment sets, triangle sets (faceted surfaces), and implicit and parametric curves and surfaces. One shape is assigned to be the "data" shape and the other shape to be the "model" shape. For surface-based registration, the shapes are surfaces. The data shape is decomposed into a point set (if it is not already in point set form). Then the data shape is registered to the model

shape by iteratively finding model points closest to the data points, registering the two point sets (e.g., using Algorithm 8.1), and applying the resulting transformation to the data points. Details of this approach are provided in Algorithm 8.4.[5] This algorithm converges to a local minimum of the disparity function in Eq. 8.33. The proof of convergence depends on the correspondence function $\mathcal{C}$ being the closest point operator.

### 8.4.6 *Weighted geometrical feature algorithm*

Most registration methods that align 3-D image volumes by matching geometrical features such as points or surfaces use a single type of feature. It is possible and potentially advantageous to perform registration using multiple geometrical features simultaneously. One approach that was suggested by Collignon et al. [86] is to define a new disparity function that is a weighted sum of the point-based (Eq. 8.21) and surface-based (Eq. 8.33) disparity functions. Let $\{X_i\}$ for $i = 1, \ldots, N_s$ be a set of $N_s$ data shapes to be registered with another set of model shapes $\{Y_i\}$. For example, $i = 1$ might be a point set and $i = 2$ a surface. Then we can define a new disparity function

$$d(\mathcal{T}) = \sum_{i=1}^{N_s} \sqrt{\sum_{j=1}^{N_{X_i}} w_{ij}^2 ||\mathcal{T}(\mathbf{x}_{ij}) - \mathbf{y}_{ij}||^2} \,, \tag{8.38}$$

where

$$\mathbf{y}_{ij} = \mathcal{C}_i(\mathcal{T}(\mathbf{x}_{ij}), Y_i) \,. \tag{8.39}$$

The term under the square root symbol is analogous to Eqs. 8.21 and 8.33. The weights $\{w_{ij}\}$ can incorporate both *intra*-shape weighting and *inter*-shape weighting. A similar approach was suggested by Maurer *et al.* [67], but with the outer summation inside the square root symbol, that is,

$$d(\mathcal{T}) = \sqrt{\sum_{i=1}^{N_s} \sum_{j=1}^{N_{X_i}} w_{ij}^2 ||\mathcal{T}(\mathbf{x}_{ij}) - \mathbf{y}_{ij}||^2} \,. \tag{8.40}$$

The transformation $\mathcal{T}$ that minimizes either of these hybrid disparity functions can be found using any of the common gradient descent techniques. A consequence of putting the outer summation inside the square root symbol is that Eq. 8.40 can be minimized using an extension of the ICP algorithm called the weighted geometrical feature (WGF) algorithm [67]. Details of this approach are provided in Algorithm 8.4.

[5] Algorithm 8.4 details the more general weighted geometrical feature (WGF) algorithm. The ICP algorithm is a special case of the WGF algorithm where the number of shapes $N_s = 1$.

The WGF algorithm is potentially useful for image-to-physical registration. For example, the position of a bone-implanted marker can be determined much more accurately than that of a skin-affixed marker or an anatomic landmark. A disadvantage of using bone-implanted markers is that an invasive pre-operative procedure is required to implant each marker. By combining surface information, the WGF algorithm allows registration to be performed using only one or two such markers (point-based registration requires a minimum of three noncollinear points). For example, very accurate (TRE $\sim$ 1–2 mm) point-and-surface-based CT-to-physical registration of the head was achieved with one bone-implanted marker and 25 surface points [78].

**Algorithm 8.4: Weighted geometrical feature (WGF) rigid-body registration**

Find the rigid-body transformation $\mathcal{T}$ that minimizes the disparity function in Eq. 8.40:

1. Initialization: $k = 1, \mathbf{x}_{ij}^{(0)} = \mathbf{x}_{ij}, \mathbf{x}_{ij}^{(1)} = \mathcal{T}^{(0)}(\mathbf{x}_{ij}^{(0)})$,

   where $\mathcal{T}^{(0)}$ is some initial transformation. The variable $k$ and the superscript in parentheses are iteration indices. The algorithm can be repeated using multiple initial transformations to solve the local minimum problem.

2. Iteratively apply the following steps, incrementing $k$ after each loop, until convergence within a tolerance $\epsilon$ is achieved:

   (a) For each shape $X_i$, compute the closest points $\mathbf{y}_{ij}^{(k)} = \mathcal{C}_i(\mathbf{x}_{ij}^{(k)}, Y_i)$ for $j = 1, \ldots, N_{X_i}$.

   (b) Compute the transformation $\mathcal{T}^{(k)}$ between the initial point set, $\{\mathbf{x}_{ij}^{(0)}\}$, and the current set, $\{\mathbf{y}_{ij}^{(k)}\}$, using the weights $\{w_{ij}\}$. This step is effected by means of Algorithm 8.1 with the points for all shapes collected in each of the two point sets to produce two corresponding point sets.

   (c) Apply the transformation to produce registered points $\mathbf{x}_{ij}^{(k+1)} = \mathcal{T}^{(k)}(\mathbf{x}_{ij}^{(0)})$.

   (d) Terminate the iterative loop when $d(\mathcal{T}^{(k)}) - d(\mathcal{T}^{(k+1)}) < \epsilon$, where $d(\mathcal{T})$ is given by Eq. 8.40.

## 8.5 Intensity-based methods

According to the classifications introduced earlier in this chapter, image intensity is an alternative "registration basis" to points or surface features. It has recently become the most widely used registration basis for several important applications. In this context, the term intensity is invariably used to refer to the scalar values in image pixels or voxels. The physical meaning of the pixel or voxel value depends on the modalities being registered and is very often not a direct measure of optical power (the strict definition of intensity).

Intensity-based registration involves calculating a transformation between two images using the pixel or voxel values alone. In its purest form, the registration transformation is determined by iteratively optimizing some "similarity measure" calculated from all pixel or voxel values. Because of the predominance of three-dimensional images in medical imaging, we refer to these measures as *voxel similarity* measures. In practice, many intensity-based registration algorithms use only a subset of voxels and require some sort of preprocessing. For example, the algorithm may run faster if only a subset of voxels are used. This subset can be chosen on a regular grid, or be randomly chosen. It is normal in these circumstances to blur the images before sampling to avoid aliasing in the subsampled images, and the amount of blurring used may be application dependent. Alternatively, an algorithm may work reliably only if the similarity measure is calculated from the voxels in a defined region of interest in the image, rather than all voxels. In this case, some sort of pre-segmentation of the images is required, and this is likely to depend both on the modalities being registered and the part of the body being studied. In some other intensity-based algorithms, the similarity measures work on derived image parameters such as image gradients, rather than the original voxel values.

For retrospective registration, a major attraction of intensity-based algorithms is that the amount of preprocessing or user-interaction required is much less than for point-based or surface-based methods. As a consequence, these methods are relatively easy to automate. The need for preprocessing does, however, mean that many intensity-based algorithms are restricted to a quite limited range of images. One of the aims of recent research in this area has been to devise general algorithms that will work on a wide variety of image types, without application-specific preprocessing.

Intensity-based registration algorithms can be used for a wide variety of applications: registering images with the same dimensionality, or different dimensionality; both rigid transformations and registration incorporating deformation; and both intermodality and intramodality images. Most algorithms are applicable to only a subset of these applications, but some are quite generally applicable.

In this section, we review the main similarity measures used for intensity-based image registration and describe the applications in which they are used. All these algorithms are iterative, so each similarity measure needs to be optimized. Issues specifically related to optimization are discussed in Section 8.5.2.

In this section we introduce some new notation. The images to be registered are $A$ and $B$. The sets of voxels in these images are $\{A(i)\}$ and $\{B(i)\}$, respectively. We will treat image $A$ as a reference image, and $B$ as an image that is iteratively transformed to $B' = \mathcal{T}(B)$ by successive estimates of the registration transformation $\mathcal{T}$. The transformation estimates will change the overlap between the images being registered. Voxel-similarity measures are invariably calculated for the set of voxels in the overlapping region of $A$ and $B'$, i.e., within $A \cap B'$, which is a function of $\mathcal{T}$ and so changes as the algorithm iterates. For some voxel-similarity measures, information from the intensity histogram is used, so we need to refer directly to intensity values in the image, rather than index voxels. Medical images may have 10 bits (1024 values), 12 bits (4096 values), or even 16 bits (65536 values) worth of intensity information per voxel. Many algorithms that use intensity information group voxel values into a smaller number of partitions, for example 64, 128, or 256 partitions. We refer to the sets of intensity partitions in images $A$ and $B'$ as $\{a\}$ and $\{b\}$, respectively, and the number of intensity partitions used as $N_a$ and $N_b$. Because the range of voxel intensities in an image is dependent on $\mathcal{T}$, $\{b\}$ may also be a function of $\mathcal{T}$.

### *8.5.1 Similarity measures*

#### *8.5.1.1 Image subtraction*

If the assumption is made that the images $A$ and $B$ being registered are identical, except for the misalignment, then an intuitively obvious similarity measure to use is the sum of squares of intensity differences (SSD). In this case, SSD will be zero when the images are correctly aligned and will increase with misregistration-registration error. In the slightly more realistic scenario in which the $A$ and $B$ differ only by Gaussian noise, then it can be shown that SSD is the optimum measure [87].

Certain image registration problems are reasonably close to this ideal case. For example, in serial registration of MR images, it is expected that the images being aligned will be identical except for small changes, which might result from disease progression or response to treatment. Similarly, in functional MR experiments, only a small number of the voxels are expected to change during the study, so all the images that need to be registered to correct for patient motion during the study are very similar to each other. If only a small fraction of the voxels being aligned are likely to have changed between image acquisitions, SSD is likely to work well. This approach has been used by Hajnal et al. [88] and is used in the SPM software by Friston et al. [89]. This approach can fail if the data diverges too much from the ideal case. For example, if a small number of voxels change intensity by a large amount, they can have a large effect on the change in squared intensity difference. For this reason, it is sometimes desirable to pre-segment parts of the image prior to registration. This preprocessing is commonly done for the scalp when carrying out serial MR brain registration, where the scalp can deform.

It is worth remembering that the assumption of Gaussian noise is frequently

broken in medical images. In MR magnitude images, the noise is Rician [90], which is approximately Gaussian for high intensity parts of the image but is a long way from Gaussian in low intensity regions. Also, artifacts caused by reconstruction or subject motion introduce noise that is non-Gaussian.

**Algorithm 8.5: Sum of Squares of Differences**

For images $A$ and $B$ with voxels $i$, find the transformation $\mathcal{T}$ to minimize:

$$\text{SSD} = \frac{1}{N}\sum_{i}^{N} |A(i) - B'(i)|^2 \qquad \forall i \in A \cap B' \,.$$

### 8.5.1.2 *Correlation Coefficient*

If the intensities in images $A$ and $B$ are linearly related, then the correlation coefficient CC can be shown to be the ideal similarity measure [87]. Once again, few registration applications will precisely conform to this requirement, but many intramodality applications come sufficiently close for this to be a useful measure.

**Algorithm 8.6: Correlation Coefficient**

For images $A$ and $B$ with voxels $i$, find the transformation $\mathcal{T}$ to maximize:

$$\text{CC} = \frac{\sum_i (A(i) - \bar{A})(B'(i) - \bar{B}')}{\{\sum_i (A(i) - \bar{A})^2 \sum_i (B'(i) - \bar{B}')^2\}^{1/2}} \qquad \forall i \in A \cap B' \,,$$

where $\bar{A}$ and $\bar{B}'$ are the mean values of voxels in image $A$ and the transformed image $B$, respectively.

#### 8.5.1.3 *Ratio-Image Uniformity*

An alternative intramodality registration measure was proposed by Woods in 1992 [91]. This algorithm was initially devised for registration of multiple PET images of the same subject and has subsequently been widely used for serial MR registration of the brain [92]. The name Ratio-Image Uniformity (RIU) has become used quite recently for this algorithm, and it is also known as Variation of Intensity Ratios (VIR). The RIU name does, however, explain well what the algorithm does.

For each estimate of the registration transformation, a ratio image $R$ is calculated by dividing each voxel in $A$ by each voxel in $B'$. The uniformity of $R$ is then determined by calculating the normalized standard deviation of $R$. The algorithm iteratively determines the transformation $\mathcal{T}$ that minimizes the normalized standard deviation, i.e., maximizes uniformity.

In some cases, in order to get good results, it may be necessary to preprocess the images to remove some anatomy. In the original PET-PET application, for example, it was necessary to segment the brain from the images, removing all extradural material so that both $A$ and $B$ include only brain voxels. This step is not necessary, however, for serial MR registration (see Section 8.5.3.2). Furthermore, an alternative registration will be obtained in almost all cases by reversing the roles of images $A$ and $B$ (i.e., the result will not be the inverse of the forward transformation). Woods proposed that possible bias of the registration solution produced by this asymmetry can be reduced by averaging the forward transformation with the inverse of the reverse transformation [91].

**Algorithm 8.7: Ratio-Image Uniformity**

For two images $A$ and $B'$, find $\mathcal{T}$ to minimize RIU, where RIU is calculated over the region $A \cap B'$ as follows:

1. Smooth and interpolate the images to have cubic voxels (optional), and determine the number of voxels $N$ in $B'$ that lie within $A \cap B'$.
2. Calculate the ratio image $R(i) = B'(i)/A(i)$.
3. Calculate the mean of $R$: $\mu_R = \frac{1}{N}\sum_i R(i)$.
4. Calculate the standard deviation of $R$: $\sigma_R = \frac{1}{N}\sum_i (R(i) - \mu_R)^2$.
5. RIU = the normalized standard deviation: $\sigma_R/\mu_R$.

#### 8.5.1.4 *Partitioned Intensity Uniformity*

Shortly after proposing the RIU method for intramodality registration, Woods proposed a modified version of the algorithm for MR-PET registration [93]. This second algorithm, which we call Partitioned Intensity Uniformity (PIU), makes an idealized assumption that "all pixels with a particular MR pixel value represent the same tissue type so that values of corresponding PET pixels should also be similar to each other." The algorithm therefore minimizes the normalized standard deviation of PET voxel values for each MR intensity value (or partitioned intensity value). The PIU algorithm can usefully be thought of in terms of intensity histograms. For each intensity partition $a$ in the MR image, there will be $n_A(a)$ voxels distributed throughout the MR image volume that have an intensity within this partition. For a given estimate $\mathcal{T}$ of the registration transformation, there will be $n_A(a)$ PET voxels that are co-located with these MR voxels. The intensities of these PET voxels could be plotted as a histogram. The PIU algorithm aims to find the transformation $\mathcal{T}$ that minimizes the spread of that histogram for all partitions. The algorithm can fail unless the histogram for each partition is unimodal.

For MR-PET registration of head images it is normal to remove extra-dural tissue before running the algorithm. This preprocessing step helps to avoid bimodal or trimodal histograms. Successful registrations are possible even with a relatively crude segmentation process [82]. As with the RIU method, the registration obtained when $A$ and $B$ are interchanged will rarely be consistent with the original registration. The choice of PIU measure will depend on the application.

**Algorithm 8.8: Partitioned Intensity Uniformity**

For two images $A$ and $B'$, find $\mathcal{T}$ to minimize:

$$\text{PIU} = \sum_{a \in \{a\}} \frac{n_A(a)}{N} \frac{\sigma_{B'}(a)}{\mu_{B'}(a)} ,$$

where $B'$ has been interpolated to have voxels of the same size and shape as $A$, the sum is carried out over the $N$ voxels in $A \cap B'$, $n_A(a)$ is the number of voxels in image $A$ with intensity $a$, and $\mu_{B'}(a)$ and $\sigma_{B'}(a)$ are the mean and standard deviation of the voxels in image $B'$ that co-occur with voxels whose intensities lie in partition $a$ in image $A$.

For MR-PET registration, 256 intensity partitions are typical.

#### 8.5.1.5 *Joint histograms and joint probability distributions*

Statistical classifiers have been widely used in MR image analysis for segmentation of multi-spectral data for many years. In these approaches, a joint histogram is constructed from two images that are correctly aligned (e.g.: the first and second echo images from a spin-echo acquisition). A joint histogram is $n$-dimensional where $n$ is the number of images used to generate it. The axes of the histogram are the intensities (or intensity partitions) in each image, and the value at each point in the histogram is the number of voxels with a particular combination of intensities in the different spectral components.

If the joint histogram is normalized, it becomes an estimate of the joint probability distribution function (PDF) of intensities in the $n$ images.

**Algorithm 8.9: Calculating a joint PDF**

For two images $A$ and $B'$ related by a transformation $\mathcal{T}$, calculate the two-dimensional PDF for intensity partitions $\{a\}$ and $\{b\}$:

1. Allocate an $N_a$ by $N_b$ array HIST$[j,k]$.
2. Initialize the histogram: HIST$[j,k] = 0$ for all $j,k$.
3. For each voxel $i \in A \cap B'$, calculate intensity values $A(i)$ and $B'(i)$, calculate the intensity partition numbers $a$ and $b$ corresponding to $A(i)$ and $B'(i)$, and increment HIST$[a,b]$.
4. Calculate $\sum_{j,k}$ HIST$[j,k]$.
5. Normalize the histogram to calculate the PDF:

$$\text{PDF}[j,k] = \frac{\text{HIST}[j,k]}{\sum_{j,k} \text{HIST}[j,k]} .$$

Note that elements in the array PDF are floating point, not integer.

For many image modality combinations, the PDF changes with $\mathcal{T}$. Example histograms as a function of $\mathcal{T}$ are shown in Figure 8.7. This observation lead several research groups to investigate similarity measures calculated from the PDF, which could be optimized to register the images. The attraction of this approach is that the observed change in PDF with $\mathcal{T}$ is qualitatively similar for many modality combinations; the signal in the PDF is "clustered" at registration but diffuses with misregistration.

The most successful similarity measures calculated from the PDF are based on information theory.

#### *8.5.1.6 Joint entropy*

The Shannon entropy $H$ is widely used as a measure of information in many branches of engineering. It was originally developed as part of information theory in the 1940s and describes the average information supplied by a set of symbols $\{s\}$ whose probabilities are given by $\{p(s)\}$.

$$H = -\sum_{s} p(s) \log p(s) \ .$$

If all symbols $s$ have equal probability, then entropy will be at maximum. If one symbol has a probability of 1 and all others have a probability of zero, then entropy will have a minimum value.

At first sight, image registration has little to do with measuring the amount of information being transmitted down a communication channel. The use of entropy and other information-theoretic measures for image registration came about, however, after inspection of joint histograms and PDFs. When the images are correctly aligned, the joint histograms have tight clusters, surrounded by large dark regions. These clusters disperse as the images become less well registered. The tight clusters in the histograms at registration represent a small number of symbols $s$ having high probabilities $p(s)$. Surrounding dark regions in the joint histogram represent large numbers of symbols with probability zero. As the clusters disperse, the high intensity regions of the joint histogram become less intense (their probabilities decrease), and previously dark regions in the histograms become brighter (there are fewer histogram entries with zero or very low probabilities). Misregistration, therefore, results in an increase in histogram entropy.

As a consequence of this observation, it was proposed that the entropy of the PDF calculated from images $A$ and $B'$ should be iteratively minimized to register these images.

Minimizing joint histogram entropy to register images can be thought of as an extension of PIU minimization, Algorithm 8.8. PIU minimization tends to minimize the spread of the histogram of voxels in image $B'$ for each intensity partition in image $A$. Provided this histogram is unimodal, then minimizing the spread will also minimize the entropy. The joint entropy measure has two advantages over PIU. Firstly, it minimizes the spread of clusters in two dimensions rather than just one. Secondly, minimizing entropy does not require that the histograms are unimodal in the way that minimizing variance does. For this reason, joint entropy would seem more generally applicable to multimodality registration than PIU, and also obviates the need to segment some parts of the image in order to avoid bi-modal histograms.

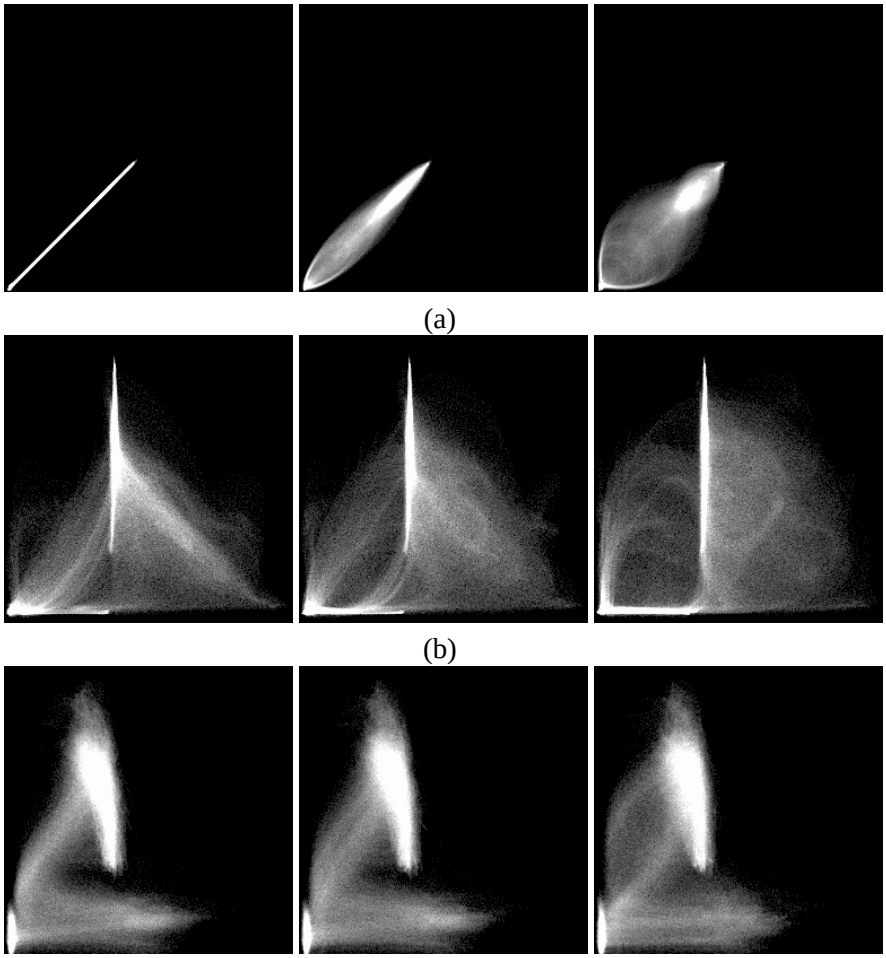



**Figure 8.7:** Example joint histograms, or joint probability distribution functions, from Hill et al. [94]. These are produced from images $A$ and $B'$ using Algorithm 8.9. Intensities from image $A$ are plotted along the horizontal axis, and those from image $B'$ on the vertical axis. In row (a), $A$ and $B$ are identical MR images of the head. In row (b), $A$ is a CT image of the head and $B$ an MR scan of the head, and in row (c), $A$ is an MR image of the head, and $B$ a PET image of the head. For all modality combinations, the histograms are calculated for three different transformations $\mathcal{T}$: the identity (left), a lateral translation of 2mm (center) and a lateral translation of 5 mm (right).

**Algorithm 8.10: Minimization of Joint Entropy**

For the PDF$[j,k]$ calculated from images $A$ and $B$ using algorithm 8.9, find the transformation $\mathcal{T}$ to minimize:

$$H = -\sum_{j,k} \text{PDF}[j,k] \log \text{PDF}[j,k] \ .$$

Note: For most registration applications, a slight modification to this measure is needed, as described in Sections 8.5.1.7 and 8.5.1.8.

### *8.5.1.7 Mutual information*

Entropy minimization as described in Section 8.5.1.6 is not a robust voxel-similarity measure for all types of image registration. The problem is that the PDF from which the joint entropy is calculated is defined only for the region of overlap between the two images, i.e., within $A \cap B'$. The range and distribution of intensity values in the portion of either image that overlaps with the other is a function of $\mathcal{T}$. The change in overlap with $\mathcal{T}$ can lead to histogram changes that mask the clustering effects described above.

The solution to this difficulty, proposed independently by researchers at Leuven, Belgium [95] and MIT in the US [96], is to use the information-theoretic measure mutual information (MI) instead of entropy $H$. MI normalizes the joint entropy with respect to the partial entropies of the contributing signals. In terms of image registration, this measure takes account of the change in the intensity histogram of images $A$ and $B'$ with $\mathcal{T}$.

**Algorithm 8.11: Maximizing Mutual Information**

Find the transformation $\mathcal{T}$ to maximize MI$(A, B')$ as follows:

1. Calculate the PDF$[j,k]$ from images $A$ and $B'$ using Algorithm 8.9.
2. Calculate the joint entropy $H(A, B')$ using the formula in Algorithm 8.10.
3. Calculate the marginal entropies $H(A)$ and $H(B')$:

$$H(A) = \sum_j \left( \sum_k \text{PDF}[j,k] \log \sum_l \text{PDF}[j,l] \right) \ ,$$

$$H(B') = \sum_k \left( \sum_i \text{PDF}[i,k] \log \sum_j \text{PDF}[j,k] \right) .$$

4. Calculate the mutual information $\text{MI}(A, B')$:

$$\text{MI}(A, B') = H(A) + H(B') - H(A, B') .$$

#### 8.5.1.8 *Normalization of mutual information*

Mutual information overcomes many of the shortcomings of joint entropy but can still fail for some types of clinical image, particularly those which contain large amounts of air (noise) around the outside of the subject. Improved performance of mutual information can be obtained by various normalization schemes. These algorithms are not taken from the information theory literature but have been arrived at through experiments on medical images. Despite its heuristic origins, the variant given below (from [97]) works extremely well in practice. Current validation studies have shown that it works at least as well as mutual information and in some cases performs better.

**Algorithm 8.12: Maximizing Normalized Mutual Information**

Perform Algorithm 8.11 but replace $\text{MI}(A, B')$ with:

$$\text{NMI}(A, B') = \frac{H(A) + H(B')}{H(A, B')} .$$

### 8.5.2 *Capture ranges and optimization*

#### 8.5.2.1 *Optimization*

In order to register two images using Algorithms 8.5 to 8.12, it is necessary to find the optimal value of the similarity measure over a parameter space with dimensionality defined by the number of degrees of freedom of $\mathcal{T}$. Each registration algorithm that makes use of a voxel-similarity measure tends to use a different optimization algorithm. There are two classes of optimization algorithms that can be used: those that use derivative information and those that do not. Where derivatives

can be calculated efficiently, they should be used to speed up optimization. A detailed discussion of optimization techniques is outside the scope of this chapter. A review of optimization techniques is given in Numerical Recipes in C [61]. Some techniques are also discussed in Chapter 1.

#### *8.5.2.2 Capture ranges*

In many optimization problems it is desirable to determine the globally optimal solution. For image registration, however, the desired optimum is frequently a local rather than global extremum of the voxel-similarity measure. This has several important consequences for medical image registration. The reason that the global optimum is not desirable can be understood by careful consideration of the behavior of the similarity measures for large transformations. For the sum of squares of difference (SSD) measure, we have previously assumed that for two images that are very similar the smallest value of SSD will arise when the images are correctly aligned and that SSD will increase with misregistration. However, if we consider two head images surrounded by low intensity noise, then we may be able to obtain a lower value of SSD if we misalign the images so completely that only the noise in the two images overlaps. This transformation produces a registration that is clearly incorrect, since no features of interest in the images overlap, but it is nevertheless a more nearly optimal registration according to the similarity measure. Similarly, with regard to the information-theoretic measures, we have been assuming that the clustering of features in the joint histogram decreases with increasing misregistration. Once again, if we transform the images sufficiently so that only background noise overlaps, we observe a very tight cluster in that area of the histogram corresponding to background intensities in both images. This transformation will produce a lower joint entropy, and a higher mutual information, than the correct alignment.

The fact that the desired optimum is local rather than global does not invalidate the use of voxel-similarity measures for medical image registration. It does, though, have implications for robust implementations.

The correct local optimum value of the similarity measure will be the optimum value within a portion of parameter space termed the *capture range*. Provided the starting estimate of $\mathcal{T}$ is within the capture range, the registration algorithm can be expected to converge to the correct solution. The capture range will depend on the field of view and intensity information in the images, so cannot be known a priori. Intuitively, it might seem that the starting estimate should be closer to the correct solution, than to the solution in which only background voxels overlap, and it has been shown that for MR-CT and MR-PET registration, similarity measures will converge to the correct solution within about 20–30 mm and 20–30 degrees of this solution [98, 99]. These ranges should be treated as a "rule of thumb." It is essential when registering images using these algorithms to inspect the solution visually. A solution that lies outside the capture range is immediately recognizable

as incorrect, and a better solution can be found by re-starting the algorithm with a user-provided starting estimate that is nearer to the correct solution.

A further implication of the existence of a capture range is that the direct application of stochastic optimization algorithms such as simulated annealing and genetic algorithms are inappropriate for medical image registration using voxel-similarity measures.

#### *8.5.2.3 Optimizing similarity measures for sub-voxel registration*

The calculation of similarity measures for a given value of $\mathcal{T}$ in general requires interpolation of the images. For an iterative registration algorithm, fast interpolation (most commonly linear) is required. Because the interpolation is not perfect, it alters the intensity histogram. Linear interpolation, for example, causes a spatially varying, low-pass filtering of the images [100], whose interaction with edges will tend to smooth the histogram, and in spatially smooth areas will reduce noise and hence contribute to a sharpening of the histogram. The consequence of this effect can be an interpolation-dependent change in the value of the similarity measure, which can cause fluctuations of the similarity measure with the period of the voxel separation. When using an entropy-based measure, blurring of the image results in increased entropy. The blurring is zero at positions coinciding with voxel locations and is maximum at displacements of half a voxel from these positions. In multimodality registration, the required registration accuracy is often no better than the dimensions of a single voxel, so these interpolation errors are unimportant. In some multimodality applications, however, and also in intramodality applications such as serial MR registration, an accuracy considerably better than the dimension of a voxel is required. In these cases, the artifacts caused by interpolation can introduce local extrema in parameter space, leading to reduced registration accuracy. The interpolation artifacts are greatest for high spatial frequency features in the images. One solution to the problem is, therefore, to blur the images prior to carrying our registration, thereby reducing the artifacts [101]. The effect is also reduced if the starting estimate of $\mathcal{T}$ includes a rotational component, or if the voxel dimensions in the images being registered are different.

### *8.5.3 Applications of intensity-based methods*

#### *8.5.3.1 Types of transformation*

In the discussion of Algorithms 8.5 to 8.12, we have not stated the form of geometrical transformation $\mathcal{T}$ that we wish to find. Many of these measures have been applied to a wide variety of registration problems, including:

- rigid registration of 3-D images of the same subject taken at different times using the same modality
- rigid registration of 3-D images of the same subject taken at different times using different modalities

- rigid registration of 2-D projection images and 3-D images of the same subject taken at different times
- nonrigid registration of images of the same subject taken at different times
- nonrigid inter-subject registration of the 3-D images of the head.

Mutual information and normalized mutual information have been demonstrated to provide good quality image registration for all these types of applications. The validation of the registration quality is a particularly difficult problem when calculating nonrigid transformations and is a rather poorly defined problem when aligning images from different subjects. See Chapter 10 for a discussion of validation methods. In the sections below, we describe the application of voxel-similarity measures to a selection of medical image registration tasks.

#### *8.5.3.2 Serial MR*

It is increasingly common for a patient to have multiple magnetic resonance images acquired at different times, in order to monitor disease progression or response to treatment. Traditional radiological viewing of these images involves printing the images onto film and viewing them side-by-side on light boxes. This approach can make it difficult to identify small changes in the images between scans. A more sensitive technique is to subtract the images, generating difference images, and to view the difference images to identify regions of the images that have changed. The generation of difference images requires that the images acquired at the different times have identical contrast characteristics and that any drift in the scanner scaling parameters is corrected using suitable quality-assurance procedures, such as phantom scanning [102, 103]. Serial MR registration also requires accurate alignment, because registration errors of less than a voxel can lead to artifacts that can mask the changes of interest in the patient. The required registration accuracy will be a function of the spatial frequency content of the images and of the contrast-to-noise ratio. The required accuracy is often stated as of the order of tens or hundreds of microns.

Serial MR registration is almost invariably applied to the brain and makes use of a rigid transformation. Indeed, for visual inspection of difference images, it is undesirable to use nonrigid registration transformations because they may produce a transformation that removes the changes of interest.

With serial MR registration the images are likely to be very similar, and so the measures SSD and CC are widely used. RIU, MI, and NMI are also frequently used. The advantage of the information-theoretic measures in this case is that they are less sensitive to small numbers of voxels that change by large amounts, than are the SSD and CC measures. When using SSD or CC, data preparation by segmentation of extra-dural tissue, or small lesions that have changed substantially in intensity, is often essential. The information-theoretic approaches should be less sensitive

to these small numbers of outlying voxels, but where the changes are large, data pre-processing may still be required.

For many applications, the choice of an optimization algorithm that will iteratively find the optimum transformation to subvoxel accuracy is as important as the choice of similarity measure.

In serial MR registration, the choice of interpolation algorithm is also important. The use of tri-linear interpolation is equivalent to convolving the data with a triangular kernel of variable width, resulting in a spatially varying blurring of the data. When subtracting image $A$ from the blurred image $B'$, edge artifacts result. Higher order interpolation methods, such as those that incorporate truncated sinc kernels [88] or skew transformations [104], are more effective.

#### *8.5.3.3 MR and CT*

Registration of MR and CT images is most often applied to images of the head, and is used for surgery and radiotherapy planning. The registration transformation determined is usually a rigid one. On some occasions, however, it may be beneficial to search for an affine transformation because it can generate more accurate registration in cases when there are scaling or skew errors in the data. When the images being registered are subsequently going to be used for guiding treatment (image-guided surgery or radiation therapy), then it is not advisable to use an affine transformation unless one modality is known to have satisfactory geometrical properties and is treated as the reference image $A$. Otherwise, the images may be well registered but measurements made in the registered images may be inaccurate. It is preferable in these cases to calibrate the imaging devices as part of a routine quality-assurance procedure and to rectify the images prior to registration to remove any residual geometric distortion (see Section 8.2.3). The most problematic aspects of imaging geometry are geometric distortion in MR and errors in CT bed speed and gantry tilt angle. For 3-D gradient echo MR images, object dependent distortion is predominantly in the readout direction and can be minimized by selecting a high bandwidth-per-pixel (readout-gradient strength) in that direction (at a slight cost in signal to noise ratio). The residual scaling errors in MR due to gradient strength errors, and the CT errors, can be corrected using phantom experiments. For some CT systems, it is advisable to carry out phantom measurements with the CT couch loaded with weight similar to that of a patient.

Since MR and CT images have very different intensity distributions, image subtraction is never used for viewing the aligned images. As a result, the very high accuracy requirements of serial MR registration do not apply here. The accuracy requirement for MR-CT registration is usually governed by the accuracy of the treatment system that will subsequently be used. Since neurosurgery and radiosurgery systems are accurate to a small number of millimeters, registration accuracy of about 1 mm is normally sufficient.

A study by West et al. established that, if the images are free from motion

artifacts and have been rectified to remove scaling errors and distortion, registration accuracy of better than 1 mm is possible when information-theoretic measures are used [82]. It is important to note that the images used in that validation study had similar fields of view in both modalities, covering most of the head, and had very similar slice orientations. It is possible to register MR and CT images that have very different slice orientations and fields of views. For example, it is possible to register coronal CT images with axial MR images or to register CT images with a field of view of about 20 mm around the temporal bone to MR images of the entire head. In both these cases, the rigid transformation solution may include translations of several tens of millimeters and rotations of many tens of degrees. In such cases, it is advisable to provide the algorithm with an initial estimate of $\mathcal{T}$ to increase the likelihood that it is starting within its capture range (see Section 8.5.2). Also, since the overlap between images at the correct registration transformation includes only a small portion of one or both modalities, the registration accuracy in these cases may be lower than in the validation study by West et al. [82].

#### *8.5.3.4 MR or CT and PET*

Registration of MR or CT images with PET images of the head was one of the first applications of medical image registration [105]. Because of the relatively low resolution of PET images, it is frequently desirable to make use of anatomical detail from MR or CT images to assist in their interpretation. The resolution of PET images is of the order of 6–10 mm, and a registration accuracy of 2–4 mm is considered desirable. Although the low resolution of PET images reduces the required registration accuracy compared to MR-CT registration, there are several difficulties in this registration application. Firstly, some PET tracers are very specific to certain parts of the brain. Although several algorithms, including Partitioned Intensity Uniformity (Algorithm 8.8) and mutual information (Algorithm 8.11), have been shown to work well when registering MR images to PET fluoro-deoxyglucose images [82], the accuracy of registration of the more specific tracers will be tracer-dependent and may be much lower. Secondly, for patients with large lesions, the PET images can have very large intensity abnormalities. A lesion that makes a small difference to image intensities in an MR scan may make a very large difference to intensities in the PET image. Many of the images used in the validation study by West et al. [82] had large lesions of this sort, and the median registration accuracy of both PIU and MI was better than 4 mm, suggesting that these measures are reasonably robust to these intensity abnormalities.

#### *8.5.3.5 Nonrigid 3-D registration*

This chapter focuses on rigid registration algorithms because these are the most mature image registration methodologies. There is, however, considerable research going on in extending the use of intensity-based registration algorithms to nonrigid transformations. While we do not attempt in this section to a give a comprehensive

review of this field, we do explain how voxel-similarity measures can be extended beyond rigid or affine transformations.

It is possible to apply voxel-similarity measures to nonrigid registration by placing a regular grid of $N$ nodes across images $A$ and $B$ being registered. The position of these nodes in $B$ is then iteratively altered while optimizing a voxel-similarity measure. One approach is to calculate a rigid or affine transformation for a sub-image around each node, and then to smooth the resulting set of discrete transformations [106]. Alternatively, each node can be treated as a spline knot, and the nodes iteratively adjusted, with the the entire image being transformed using the spline transformation at each iteration. The approximating B-spline, which is described by Eqs. (8.15) and (8.16), has the advantage that change in position at a node has only a local effect on the deformation (extending about two node separations), and this approach is used in an implementation by Rueckert et al. [107].

The first approach involves multiple registrations of subimages. The accuracy of registration at each sub-image will be dependent on the number of voxels in the sub-images and the intensity information contained. Each subimage registration involves 6 or 12 degrees of freedom (for rigid and affine transformations respectively), and this is carried out $N$ times. Because each subimage contains only a small number of voxels, the information-theoretic approaches may suffer from a poor estimate of the PDF, and hence noisy values of the voxel-similarity measures. The SSD and CC measures will be less sensitive to this problem. The second approach involves three translational degrees of freedom per node, producing $3N$ degrees of freedom in total, with the similarity measure being calculated for the entire images, rather than just for subimages. This approach has the advantage that large numbers of voxels are used in each iteration, making the information-theoretic measures practicable.

These algorithms have been applied to both intrapatient registration in which there is tissue deformation and to interpatient registration. Example mean and variance images from a cohort of PET FDG patients are shown in Fig. 8.8. In this figure, an affine transformation and a spline interpolation are compared for use in interpatient registration. The similarity measure is, in both cases, normalized mutual information. The mean image is sharper, and the variance image has lower values when the transformation with larger numbers of degrees of freedom is used. Appropriate techniques for validation of these algorithms remain a matter of debate. By changing the spacing of the nodes, the algorithms are able to find a transformation with more degrees of freedom. If the number of nodes is too low, then the transformation may be too smooth to align the images. If the number of nodes is too high, then the transformations may inappropriately "correct" for features of interest, such as lesions.

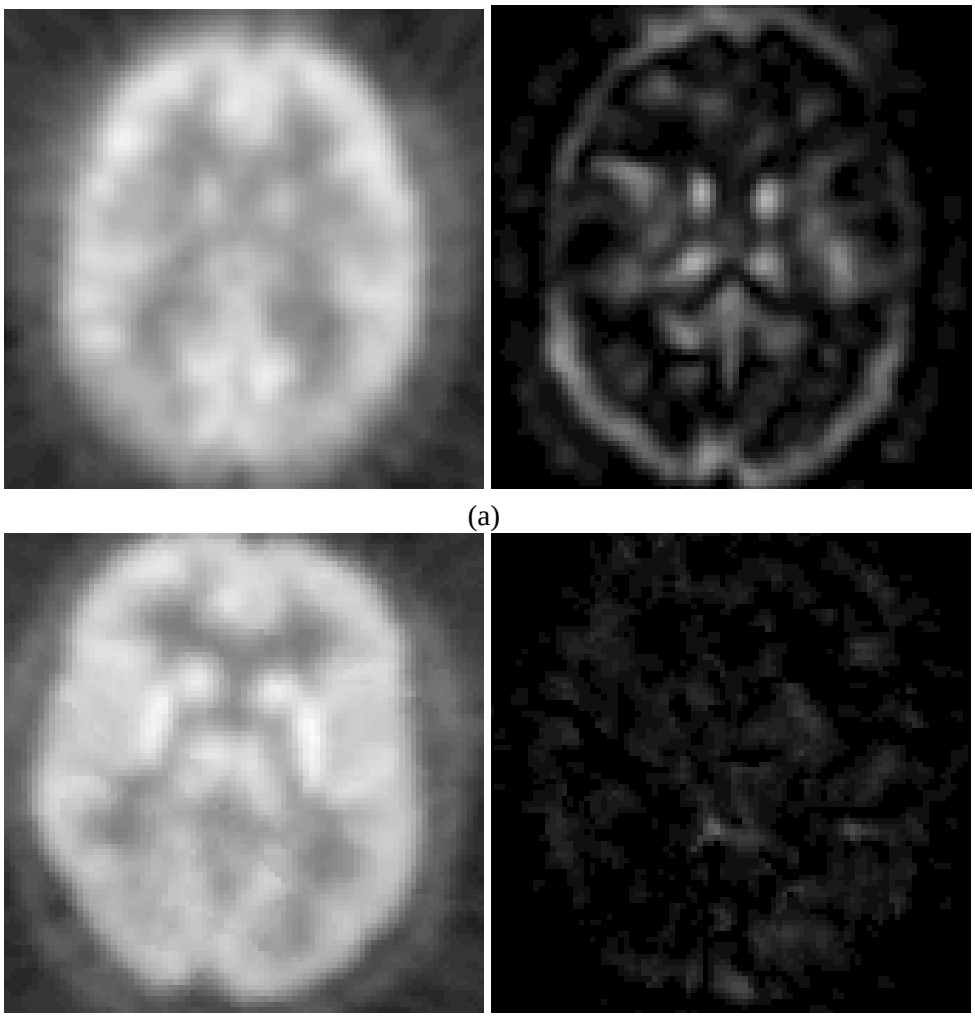



**Figure 8.8:** Mean and variance images generated from a cohort of five PET images of different subjects. Row (a) shows an example axial slice from the mean (left) and variance (right) images after registration using an affine transformation with 12 degrees of freedom. Row (b) shows the corresponding slices after registration using a transformation with 2160 degrees of freedom. The transformation in (b) is determined by iteratively adjusting the positions of B-spline control points on a 20 mm grid. Normalized mutual information was the voxel similarity measure used by both algorithms. The variance images are displayed with the same intensity scale in both cases.

#### *8.5.3.6 2-D-to-3-D registration*

Many medical imaging modalities generate images by perspective projection. (See Sections 8.2.2 and 8.2.2.4 above.) The most common of these is conventional radiography or fluoroscopy. However, with video endoscopy and microscopy becoming more widely used, video images are increasingly common. Aligning projection images with 3-D tomographic modalities is a type of 2-D-to-3-D registration and requires an estimate of both the 3-D orientation of the objects being registered and the projection transformation. For rigid structures, the 3-D orientation information is a rigid transformation, so $\mathcal{T}$ is the composition of a rigid transformation and the projection transformation.

In many medical imaging applications, the projection transformation can be measured as part of a calibration procedure, so the registration problem has just six degrees of freedom.

The approach used for registration of either video images or radiographs to tomographic images is to optimize a similarity measure while iteratively transforming the tomographic image. For each iteration, a simulated projection image is generated from the tomographic image, and the similarity measure is calculated from these. In this case, image $A$ is the projection image, and $B$ the tomographic image. For each estimation of the transformation $\mathcal{T}$, image $B'$ is a simulated projection generated from the tomographic image $B$. For example, when registering a radiograph $A$ to a CT scan $B$, a digitally reconstructed radiograph $B'$ is produced using $\mathcal{T}$ and a model of the x-ray image formation process to simulate an x-ray projection by integrating the CT voxel intensities. This approach has been used by several authors [108, 109]. When registering a video image $A$ to a tomographic image $B$, $B'$ can be a perspective surface rendering of the image structures in $B$ that are visible in $A$ [110, 111].

## 8.6 Conclusion

This chapter represents the current state of the field of medical image registration. Registration is defined as the determination of a geometrical transformation that aligns points in one view of an object with corresponding points in another view of that object or another object. Using this definition, the chapter treats registration as it is applied to images of anatomy that have been acquired by medical imaging modalities, especially the tomographic modalities, such as CT, MR, and PET. Because the body moves in three dimensions, the emphasis is on three-dimensional registration problems. Commonly used classes of geometrical transformations are presented with an emphasis on rigid and scaled transformations. Methods of registration are divided into point-based, surface-based, and intensity-based; well-established algorithms are described for all three categories, and both theoretical and practical discussions of their applications are provided.

The emphasis in this chapter is on rigid registration because, as of this writing, most of the work and most of the progress in registration has been made in this area.

Rigid registration is important because of its simplicity and because of its many applications to the relatively rigid head and to orthopedics, especially in image-guided surgery. Thanks to the researchers listed in the bibliography of this chapter, much is now known about this area, and many effective algorithms have been developed, some of the best of which are described above. These rigid registration problems include both prospective registration, in which fiducial markers are employed, and retrospective registration, in which they are not. Recent research has demonstrated that the most effective registration algorithms for many types of retrospective application are those based on intensities. Point-based and surface-based methods can also be used for these applications, but they require a greater degree of user interaction and have typically exhibited lower accuracy than the intensity-based methods. Techniques based on points and surfaces do, however, play an important role in the registration of images to physical space because the internal information necessary for intensity-based registration is typically unavailable in physical space. Research in point-based and surface-based approaches to this latter problem are central to the advancement of image-guided surgery and radiosurgery.

While the head and bones are rigid, most of the rest of the body is nonrigid, and, therefore, nonrigid transformations must be employed to register images of most of the anatomy. Furthermore, interpatient registration requires nonrigid transformation for all parts of the anatomy. These transformations are far more complex and varied than rigid ones. The appropriate form of nonrigid transformation is for intrapatient registration strongly affected by the mechanics of the tissue involved and for interpatient registration is determined by the natural variation in anatomic development. While there are still many important problems to be solved in the area of rigid registration, it seems likely that much of the future work in the field of medical image registration will focus on nonrigid registration.

## 8.7 Acknowledgments

Much of the work on the theory of errors in point-based registration reported in Section 8.3 is the result of a collaboration with Jay West of the Department of Electrical Engineering and Computer Science, Vanderbilt University. JMF is grateful to Professors David Hawkes and Derek Hill of the Radiological Sciences Department, King's College London for their collaboration with him on point-based registration and for their support of his visit to their laboratory during June–July, 1999. He also benefited from many discussions with Dr. Philippe Batchelor of that laboratory during his visit. The images in Fig. 8.3 were provided as part of the "Evaluation of Retrospective Registration" project, sponsored by the National Institutes of Health, grant number R01 NS33926. JMF gratefully acknowledges support for work on point-based registration provided by the National Science Foundation through grant number BES-9802982, by the UK Engineering and Physical Sciences Research Council Grant GR/M85753, entitled "Improving and Predicting the Accuracy of Image Guided Surgery," and by the Vanderbilt University Research Council. CRM

gratefully acknowledges support for work on surface-based registration provided by the Ronald L. Bittner Brain Tumor Research Fund. DLGH is grateful to his colleagues and students in Radiological Sciences at King's College London for their suggestions about Section 8.5. In particular, he is grateful to David Atkinson and Daniel Rueckert for useful discussions (both funded by the UK Engineering and Physical Sciences Research Council). The authors are grateful to Matthew Clarkson of the Radiological Sciences Department, King's College London for helping produce Fig. 8.6.

### 8.8 References

[1] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, pp. 1–36, 1998.

[2] C. R. Maurer, Jr. and J. M. Fitzpatrick, "A review of medical image registration," in *Interactive Image-Guided Neurosurgery* (R. J. Maciunas, ed.), pp. 17–44, Park Ridge, IL: American Association of Neurological Surgeons, 1993.

[3] P. A. van den Elsen, *Multimodality matching of brain images*. Ph.D. thesis, Utrecht University, Utrecht, The Netherlands, 1993.

[4] D. L. Collins, A. P. Zijdenbos, V. Killokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imaging*, vol. 17, pp. 463–468, 1998.

[5] H. Goldstein, *Classical Mechanics*. Reading, MA: Addison-Wesley, 1950.

[6] G. H. Rosenfield, "The problem of exterior orientation in photogrammetry," *Photogrammetric Engineering*, vol. 25, pp. 536–553, 1959.

[7] B. K. P. Horn, *Robot Vision*. Cambridge, MA: MIT Press, 1986.

[8] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Amer. A*, vol. 4, pp. 629–642, 1987.

[9] K. C. Fu, R. C. Gonzalez, and C. S. G. Lee, *Robotics: Control, Sensing, Vision, and Intelligence*. New York: McGraw-Hill, 1987.

[10] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, vol. II. Reading, MA: Addison-Wesley, 1993.

[11] A. Goshtasby, "Piecewise linear mapping functions for image registration," *Pattern Recognit.*, vol. 19, pp. 459–466, 1986.

[12] G. Q. Maguire, Jr., M. E. Noz, H. Rusinek, J. Jaeger, E. L. Kramer, J. J. Sanger, and G. Smith, "Graphics applied to medical image registration," *IEEE Comput. Graph. Appl.*, vol. 11, pp. 20–29, March 1991.

[13] E. V. Shikin and A. I. Plis, *Handbook on Splines for the User*. New York: CRC Press, 1995.

[14] L. L. Schumaker, *Spline Functions: Basic Theory*. New York: John Wiley & Sons, 1981.

[15] T. Pavlidis, *Algorithms for Graphics and Image Processing*. Rockville, Maryland: Computer Science Press, 1982.

[16] D. H. Ballard and C. M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1982.

[17] R. L. Harder and R. N. Desmarais, "Interpolation using surface splines," *J. Aircraft*, vol. 9, pp. 189–191, 1972.

[18] J. Duchon, "Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces," *R.A.I.R.O. Analyse Numérique*, vol. 10, pp. 5–12, 1977.

[19] J. Meinguet, "An intrinsic approach to multivariate spline interpolation at arbitrary points," in *Polynomial and Spline Approximation*, pp. 163–190, Boston: D. Reidel Publishing Company, 1978.

[20] A. Goshtasby, "Registration of images with geometric distortions," *IEEE Trans. Geosci. Remote Sensing*, vol. 26, pp. 60–64, 1988.

[21] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, pp. 567–585, 1989.

[22] R. Franke, "Scattered data interpolation: tests of some methods," *Mathematics of Computation*, vol. 38, pp. 181–200, 1982.

[23] G. M. Nielson, "Scattered data modeling," *IEEE Comput. Graph. Appl.*, vol. 13, pp. 60–70, 1993.

[24] M. Fornefett, K. Rohr, and H. S. Stiehl, "Radial basis functions with compact support for elastic registration of medical images," in *Biomedical Image Registration* (F. F. Pernuš, S. Kovačič, H. s. Stiehl, and M. A. Viergever, eds.), Ljubljana, Slovenia: Slovenian Pattern Recognition Society, 1999.

[25] C. Broit, *Optimal registration of deformed images*. Ph.D. thesis, University of Pennsylvania, Pennsylvania, USA, 1981.

[26] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Comput. Vision Graph. Image Processing*, vol. 46, pp. 1–21, 1989.

[27] J. C. Gee, M. Reivich, and R. Bajcsy, "Elastically deforming a three-dimensional atlas to match anatomical brain images," *J. Comput. Assist. Tomogr.*, vol. 17, pp. 225–236, 1993.

[28] P. Thompson and A. W. Toga, "A surface-based technique for warping three-dimensional images of the brain," *IEEE Trans. Med. Imaging*, vol. 15, pp. 402–417, 1996.

[29] G. E. Chirstensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Trans. Med. Imaging*, vol. 5, pp. 1435–1447, 1996.

[30] M. H. Davis, A. Khotanzad, D. P. Flamig, and S. E. Harms, "A physics-based coordinate transformation for 3-D image matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, pp. 317–328, 1997.

[31] M. I. Miga, K. D. Paulsen, J. M. Lemery, S. D. Eisner, A. Hartov, F. E. Kennedy, and D. W. Roberts, "Model-updated image guidance: initial clinical experiences with gravity-induced brain deformation," *IEEE Trans. Med. Imaging*, vol. 18, pp. 866–874, 1999.

[32] W. E. Smith, N. Vakil, and S. A. Maislin, "Correction of distortion in endoscope images," *IEEE Trans. Med. Imaging*, vol. 11, pp. 117–122, 1992.

[33] H. Chang and J. M. Fitzpatrick, "A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities," *IEEE Trans. Med. Imaging*, vol. 11, pp. 319–329, 1992.

[34] T. S. Sumanaweera, G. H. Glover, T. O. Binford, and J. R. Adler, "MR susceptibility misregistration correction," *IEEE Trans. Med. Imaging*, vol. 12, pp. 251–259, 1993.

[35] D. L. G. Hill, D. J. Hawkes, M. J. Gleeson, T. C. S. Cox, A. J. Strong, W.-L. Wong, C. F. Ruff, N. D. Kitchen, D. G. T. Thomas, J. E. Crossman, C. Studholme, A. J. Gandhe, S. E. M. Green, and G. P. Robinson, "Accurate frameless registration of MR and CT images of the head: Applications in surgery and radiotherapy planning," *Radiology*, vol. 191, pp. 447–454, 1994.

[36] M. Y. Wang, J. M. Fitzpatrick, and C. R. Maurer, Jr., "Design of fiducials for accurate registration of CT and MR volume images," *Medical Imaging 1995: Image Processing*, vol. Proc. SPIE 2434, pp. 96–108, 1995.

[37] C. R. Maurer, Jr., J. M. Fitzpatrick, M. Y. Wang, R. L. Galloway, Jr., R. J. Maciunas, and G. S. Allen, "Registration of head volume images using implantable fiducial markers," *IEEE Trans. Med. Imaging*, vol. 16, pp. 447–462, 1997.

[38] J. R. Hurley and R. B. Cattell, "The Procrustes program: Producing direct rotation to test a hypothesized factor structure," *Behav. Sci.*, vol. 7, pp. 258–262, 1962.

[39] B. F. Green, "The orthogonal approximation of an oblique structure in factor analysis," *Psychometrika*, vol. 17, pp. 429–440, 1952.

[40] D. G. Kendall, "A survey of the statistical theory of shape," *Statist. Sci.*, vol. 4, pp. 87–120, 1989.

[41] C. Goodall, "Procrustes methods in the statistical analysis of shape," *J. R. Statist. Soc. B*, vol. 53, pp. 285–339, 1991.

[42] F. L. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge: Cambridge University Press, 1991.

[43] P. H. Schönemann, "A generalized solution of the orthogonal Procrustes problem," *Psychometrika*, vol. 31, pp. 1–10, 1966.

[44] J. L. Farrell and J. C. Stuelpnagel, "Problem 65–1: A least squares estimate of satellite attitude," *SIAM Rev.*, vol. 8, pp. 384–386, 1966.

[45] G. H. Golub and C. F. van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins University Press, 1983.

[46] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, pp. 698–700, 1987.

[47] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, pp. 376–380, 1991.

[48] R. H. Wessner, "Problem 65–1: A least squares estimate of satellite attitude," *SIAM Rev.*, vol. 8, p. 386, 1966.

[49] B. K. P. Horn, H. M. Hilden, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *J. Opt. Soc. Amer. A*, vol. 5, pp. 1127–1135, 1988.

[50] R. Sibson, "Studies in the robustness of multidimensional scaling: Procrustes statistics," *J. R. Statist. Soc. B*, vol. 40, pp. 234–238, 1978.

[51] O. D. Faugeras and M. Hebert, "The representation, recognition, and locating of 3-D objects," *Int. J. Robotics Res.*, vol. 5, pp. 27–52, 1986.

[52] M. Froimowitz and S. Matthysse, "Geometrical correspondence between phenazocine and the enkephalins," *J. Med. Chem.*, vol. 29, pp. 573–578, 1986.

[53] M. W. Walker and L. Shao, "Estimating 3-D location parameters using dualnumber quaternions," *CVGIP: Image Understanding*, vol. 54, pp. 358–367, 1991.

[54] D. W. Eggert, A. Lorusso, and R. B. Fisher, "Estimating 3-D rigid body transformations: a comparison for four major algorithms," *Mach. Vision Appl.*, vol. 9, pp. 272–290, 1997.

[55] R. Sibson, "Studies in the robustness of multidimensional scaling: Perturbational analysis of classical scaling," *J. R. Statist. Soc. B*, vol. 41, pp. 217–229, 1979.

[56] S. P. Langron and A. J. Collins, "Perturbation theory for generalized Procrustes analysis," *J. R. Statist. Soc. B*, vol. 47, pp. 277–284, 1985.

[57] J. M. Fitzpatrick, J. West, and C. R. Maurer, Jr., "Predicting error in rigid-body, point-based registration," *IEEE Trans. Med. Imaging*, vol. 17, pp. 694–702, 1998.

[58] D. J. Hawkes, "The accuracy of 3D image registration using point landmarks," Tech. Rep. RS94/1, UMDS, London, 1995.

[59] P. J. Edwards, D. J. Hawkes, D. G. Hill, D. Jewell, R. Spink, A. Strong, and M. Gleeson, "Augmented reality in the stereo operating microscope for otolaryngology and neurosurgical guidance," *Computer Assisted Surgery*, vol. 1, pp. 172–178, 1995.

[60] P. H. Schönemann and R. M. Carroll, "Fitting one matrix to another under choice of a central dilation and a rigid motion," *Psychometrika*, vol. 35, pp. 245–255, 1970.

[61] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, England: Cambridge University Press, 2nd ed., 1992.

[62] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: John Wiley & Sons, 1973.

[63] W. M. Newman and R. F. Sproull, *Principles of Interactive Computer Graphics, 2nd ed.* New York: McGraw-Hill, 1979.

[64] D. Kincaid and W. Cheney, *Numerical Analysis*. Pacific Grove, CA: Brooks/Cole, 1991.

[65] C. R. Maurer, Jr., R. P. Gaston, D. L. G. Hill, M. J. Gleeson, M. G. Taylor, M. R. Felon, P. J. Edwards, and D. J. Hawkes, "AcouStick: A tracked A-mode ultrasonography system for registration in image-guided surgery," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI) 1999* (C. J. Taylor and A. C. F. Colchester, eds.), pp. 953–962, Berlin: Springer-Verlag, 1999.

[66] S. Lavallee, J. Troccaz, P. Sautot, B. Mazier, P. Cinquin, P. Merloz, and J.-P. Chirossel, "Computer-assisted spinal surgery using anatomy-based registration," in *Computer-Integrated Surgery: Technology and Clinical Applications* (R. H. Taylor, S. Lavallee, G. Burdea, and R. Mösges, eds.), Cambridge, MA: MIT Press, 1996.

[67] C. R. Maurer, Jr., G. B. Aboutanos, B. M. Dawant, R. J. Maciunas, and J. M. Fitzpatrick, "Registration of 3-D images using weighted geometrical features," *IEEE Trans. Med. Imaging*, vol. 15, pp. 836–849, 1996.

[68] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, pp. 239–256, 1992.

[69] W. E. L. Grimson and T. Lozano-Perez, "Model-based recognition and localization from sparse range or tactile data," *Int. J. Robotics Res.*, vol. 3, pp. 3–35, 1984.

[70] R. Bajcsy and F. Solina, "Three-dimensional object representation revisited," *Proc. 1st Int. Conf. Comput. Vision*, pp. 231–240, 1987.

[71] E. Hall, J. Tio, C. McPherson, and F. Sadjadi, "Measuring curved surfaces for robot vision," *Comput.*, vol. 15, pp. 42–54, 1982.

[72] P. J. Besl and R. C. Jain, "Three dimensional object recognition," *ACM Comput. Surveys*, vol. 17, pp. 75–145, 1985.

[73] R. Nevatia and T. O. Binford, "Description and recognition of curved objects," *Artificial Intell.*, vol. 8, pp. 77–98, 1977.

[74] P. Brou, "Using the gaussian image to find the orientation of objects," *Int. J. Robotics Res.*, vol. 3, pp. 89–125, 1984.

[75] C. A. Pelizzari, G. T. Y. Chen, D. R. Spelbring, R. R. Weichselbaum, and C.-T. Chen, "Accurate three-dimensional registration of CT, PET, and/or MR images of the brain," *J. Comput. Assist. Tomogr.*, vol. 13, pp. 20–26, 1989.

[76] T. Zuk, S. Atkins, and K. Booth, "Approaches to registration using 3D surfaces," *Medical Imaging 1994: Image Processing*, vol. Proc. SPIE 2167, pp. 176–187, 1994.

[77] D. A. Simon and T. Kanade, "Geometric constraint analysis and synthesis: Methods for improving shape-based registration accuracy," in *CVRMed-MRCAS '97* (J. Troccaz, E. Grimson, and R. Mösges, eds.), pp. 181–190, Berlin: Springer-Verlag, 1997.

[78] C. R. Maurer, Jr., R. J. Maciunas, and J. M. Fitzpatrick, "Registration of head CT images to physical space using a weighted combination of points and surfaces," *IEEE Trans. Med. Imaging*, vol. 17, pp. 753–761, 1998.

[79] C. R. Maurer, Jr., G. B. Aboutanos, B. M. Dawant, S. Gadamsetty, R. A. Margolin, R. J. Maciunas, and J. M. Fitzpatrick, "Effect of geometrical distortion correction in MR on image registration accuracy," *J. Comput. Assist. Tomogr.*, vol. 20, pp. 666–679, 1996.

[80] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Comm. ACM*, vol. 18, pp. 509–517, 1975.

[81] F. P. Preparata and M. I. Shamos, *Computational Geometry: An Introduction*. Berlin: Springer-Verlag, 1985.

[82] J. B. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Malandain, X. Pennec, M. E. Noz, G. Q. Maguire, Jr., M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods, "Comparison and evaluation of retrospective intermodality brain image registration techniques," *J. Comput. Assist. Tomogr.*, vol. 21, pp. 554–566, 1997.

[83] G. Borgefors, "Distance transformations in arbitrary dimensions," *Comput. Vision Graph. Image Processing*, vol. 27, pp. 321–345, 1984.

[84] C. T. Huang and O. R. Mitchell, "A Euclidean distance transform using grayscale morphology decomposition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, pp. 443–448, 1994.

[85] H. Jiang, R. A. Robb, and K. S. Holton, "A new approach to 3-D registration of multimodality medical images by surface matching," *Visualization in Biomedical Computing 1992*, vol. Proc. SPIE 1808, pp. 196–213, 1992.

[86] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal, "Registration of 3D multi-modality medical images using surfaces and point landmarks," *Pattern Recognit. Lett.*, vol. 15, pp. 461–467, 1994.

[87] P. A. Viola, *Alignment by maximization of mutual information*. Ph.D. thesis, Massachusetts Institute of Technology, 1995.

[88] J. V. Hajnal, N. Saeed, E. J. Soar, A. Oatridge, I. R. Young, and G. M. Bydder, "A registration and interpolation procedure for subvoxel matching of serially acquired MR images," *J. Comput. Assist. Tomogr.*, vol. 19, pp. 289–296, 1995.

[89] K. J. Friston, S. Williams, R. Howard, R. S. Frackowiak, and R. Turner, "Movement-related effects in fMRI time-series," *Magn. Reson. Med.*, vol. 35, pp. 346–355, 1996.

[90] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magn. Reson. Med.*, vol. 36, pp. 910–914, 1995.

[91] R. P. Woods, S. R. Cherry, and J. C. Mazziotta, "Rapid automated algorithm for aligning and reslicing PET images," *J. Comput. Assist. Tomogr.*, vol. 16, pp. 620–633, 1992.

[92] P. A. Freeborough and N. C. Fox, "The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI," *IEEE Trans. Med. Imaging*, vol. 16, pp. 623–629, 1997.

[93] R. P. Woods, J. C. Mazziotta, and S. R. Cherry, "MRI-PET registration with automated algorithm," *J. Comput. Assist. Tomogr.*, vol. 17, pp. 536–546, 1993.

[94] D. L. G. Hill and D. J. Hawkes, "Voxel similarity measures for automated image registration," *Visualization in Biomedical Computing 1994*, vol. Proc. SPIE 2359, pp. 205–216, 1994.

[95] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imaging*, vol. 16, pp. 187–198, 1997.

[96] W. M. Wells, III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, pp. 35–51, 1996.

[97] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, pp. 71–86, 1999.

[98] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated 3D registration of MR and CT images of the head," *Med. Image Anal.*, vol. 1, pp. 163–175, 1996.

[99] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated 3D registration of MR and PET brain images by multi-resolution optimisation of voxel similarity measures," *Med. Phys.*, vol. 24, pp. 25–35, 1997.

[100] J. Parker, R. V. Kenyon, and D. Troxel, "Comparision of interpolating methods for image resampling," *IEEE Trans. Med. Imaging*, vol. 2, pp. 31–39, 1983.

[101] J. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual information matching and interpolation artifacts," *Medical Imaging 1999: Image Processing*, vol. Proc. SPIE 3661-01, pp. 56–65, 1999.

[102] D. L. G. Hill, C. R. Maurer, Jr., C. Studholme, J. M. Fitzpatrick, and D. J. Hawkes, "Correcting scaling errors in tomographic images using a nine degree of freedom registration algorithm," *J. Comput. Assist. Tomogr.*, vol. 22, pp. 317–323, 1998.

[103] L. Lemieux and G. J. Barker, "Measurement of small inter-scan fluctuations in voxel dimensions in magnetic resonance images using registration," *Med. Phys.*, vol. 25, pp. 1049–1054, 1998.

[104] W. F. Eddy, M. Fitzgerald, and D. C. Noll, "Improved image registration by using Fourier interpolation," *Magn. Reson. Med.*, vol. 36, pp. 923–931, 1996.

[105] G. Q. Maguire, Jr., M. E. Noz, E. M. Lee, and J. H. Schimpf, "Correlation methods for tomographic images using two and three dimensional techniques," in *Information Processing in Medical Imaging 1985* (S. L. Bacharach, ed.), pp. 266–279, Dordrecht: Martinus Nijhoff Publishers, 1986.

[106] D. L. Collins, C. J. Holmes, T. M. Peters, and A. C. Evans, "Automatic 3D model-based neuroanatomical segmentation," *Human Brain Mapping*, vol. 3, pp. 190–208, 1995.

[107] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D. J. Hawkes, "Non-rigid registration of breast MR images using mutual information," in *Med. Image Comput. and Computer-Assisted Intervention, 1998* (A. C. F. Colchester, S. L. Delp, and W. M. Wells, III, eds.), pp. 1144–1152, Berlin: Springer-Verlag, 1998.

[108] L. Lemieux, N. D. Kitchen, S. W. Hughes, and D. G. T. Thomas, "Voxel-based localization in frame-based and frameless stereotaxy and its accuracy," *Med. Phys.*, vol. 21, pp. 1301–1310, 1994.

[109] G. Penny, J. Weese, L. J. A., P. Desmedt, D. L. G. Hill, and D. J. Hawkes, "A comparison of similarity measures for use in 2D-3D medical image registration," *IEEE Trans. Med. Imaging*, vol. 17, pp. 586–595, 1998.

[110] P. Viola and W. Wells, "Alignment by maximization of mutual information," in *Proc. 5th Int. Conf. Computer Vision*, pp. 16–23, 1995.

[111] M. J. Clarkson, D. Rueckert, D. L. G. Hill, and D. J. Hawkes, "Registration of multiple video images to pre-operative CT for image guided surgery," *Medical Imaging 1999: Image Processing*, vol. Proc. SPIE 3661-01, pp. 14–23, 1999.