Mutual-Information-Based Registration of Medical Images: A Survey

Josien P. W. Pluim*, *Member, IEEE*, J. B. Antoine Maintz, and Max A. Viergever, *Member, IEEE*

***Abstract*—An overview is presented of the medical image processing literature on mutual-information-based registration. The aim of the survey is threefold: an introduction for those new to the field, an overview for those working in the field, and a reference for those searching for literature on a specific application. Methods are classified according to the different aspects of mutual-information-based registration. The main division is in aspects of the methodology and of the application. The part on methodology describes choices made on facets such as preprocessing of images, gray value interpolation, optimization, adaptations to the mutual information measure, and different types of geometrical transformations. The part on applications is a reference of the literature available on different modalities, on interpatient registration and on different anatomical objects. Comparison studies including mutual information are also considered. The paper starts with a description of entropy and mutual information and it closes with a discussion on past achievements and some future challenges.**

***Index Terms*—Image registration, literature survey, matching, mutual information.**

I. Introduction

THERE are two things Collignon and colleagues and Viola and Wells probably did not foresee when they were working on a new idea in approximately 1994. First of all, that someone else had the same idea and, second, that this new idea would lead to a list of publications as long as the one in this paper, in only seven years. This survey covers the literature until spring 2002. Actually, the "true" list is longer—in the first place, because we have left out redundant publications and because we are bound to have missed some publications. Second, mutual-information-based registration has become common place in many clinical applications. There is a wealth of papers mentioning the use of the method as a step in a larger method or in an application. These papers were generally not included, except when the modality involved or the application was unusual.

In the following, we aim to introduce and explain mutual information and to give an overview of the literature on mutual-information-based registration for medical applications. We start at the basics, with the definition of entropy and its interpretation. We then turn to mutual information, presenting its history in image registration, its multiple forms of definition and its properties. For a recent general introduction to and review of medical image registration, including many references to mutual-information-based methods, we refer to [1] and [2].

The survey classifies methods into two main categories: methodological aspects and matters of application. The aspects of the method are subdivided into preprocessing, measure, transformation, and implementation, most of which have a further subclassification. Aspects of the application entail the image modalities, the subject of registration (a single person or different persons) and the object of registration (the imaged anatomy). We also classify according to the image dimensionality and the number of images involved in registration.

Finally, having considered a number of comparison studies, we discuss the results of seven years of research and also some challenges that still lie ahead.

Manuscript received December 5, 2002; revised February 28, 2003. This work was supported by The Netherlands Organization for Scientific Research (NWO). The Associate Editor responsible for coordinating the review of this paper and recommending its publication was C. Meyer. *Asterisk indicates corresponding author*.

*J. P. W. Pluim is with the University Medical Center Utrecht, Image Sciences Institute, Room E01.335, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands (e-mail: josien@isi.uu.nl).

J. B. A. Maintz is with the Institute of Information and Computing Sciences, Utrecht University, 3508 TB Utrecht, The Netherlands (e-mail: twan@cs.uu.nl).

M. A. Viergever is with the University Medical Center Utrecht, Image Sciences Institute, 3584 CX Utrecht, The Netherlands (e-mail: max@isi.uu.nl).

Digital Object Identifier 10.1109/TMI.2003.815867

II. Entropy

The desire for a *measure of information* (commonly termed *entropy*) of a message stems from communication theory. This field concerns the broadcast of a message from a sender to a receiver. The first attempts to arrive at an information measure of a message focused on telegraph and radio communication, sending Morse code or words. However, picture transmission (television) was already considered in the important paper by Hartley [3]. In 1928, he defined a measure of information of a message that forms the basis of many present-day measures. He considered a message a string of symbols, with s different possibilities for each symbol. If the message consists of n symbols, there are s^n different messages possible (assuming there are no syntactic rules). He sought to define an information measure that increases with message length. The measure s^n complies, but the amount of information would increase exponentially with the length of the message and that is not realistic. Hartley wanted a measure H that increases linearly with n, i.e., $H = Kn$, where K is a constant depending on the number of symbols s. He further assumed that, given messages of length n_1 and n_2 from s_1 and s_2 numbers of symbols, respectively, if $s_1^{n_1} = s_2^{n_2}$, i.e., the number of possible messages is equal, then the amount of information per message is also equal. These two restrictions led him to define the following measure of information:

$$H = n \log s = \log s^n \tag{1}$$

as is shown in the Appendix. Hartley's information measure depends on the number of possible outcomes: the larger the number of possible messages, the larger the amount of information you get from a certain message. If there is only a single message possible, you gain no information $(\log 1 = 0)$ from it, because you already knew you would receive that message. In this respect, the measure can also be viewed as a measure of *uncertainty*. When there are more different messages you could possibly receive, you are more uncertain which one you will actually receive. And, again, if there is only one, there is no uncertainty.

A drawback of Hartley's measure is that it assumes all symbols (and, hence, all messages of a given length) are equally likely to occur. Clearly, this will often not be the case. In the previous paragraph, for example, the letter "e" has occurred 229 times and the letter "q" only twice. Shannon introduced an adapted measure in 1948 [4], which weights the information per outcome by the probability of that outcome occurring. Given events $e_1, \ldots, e_m$ occurring with probabilities $p_1, \ldots, p_m$, the *Shannon entropy* is defined as

$$H = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i. \tag{2}$$

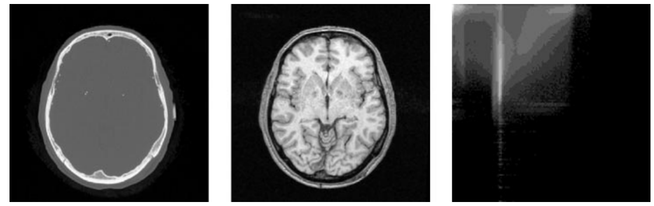
If we apply to Shannon's entropy the assumption that all outcomes are equally likely to occur, we get

$$H = -\sum \frac{1}{s^n} \log \frac{1}{s^n} = \sum \frac{1}{s^n} \log s^n = \log s^n \tag{3}$$

which is exactly Hartley's entropy.

Although the second definition of the Shannon entropy in (2) is more commonly used, the first one more clearly explains the meaning. The term $\log(1/p_i)$ signifies that the amount of information gained from an event with probability p_i is inversely related to the probability that the event takes place. The more rare an event, the more meaning is assigned to occurrence of the event. The information per event is weighted by the probability of occurrence. The resulting entropy term is the *average* amount of information to be gained from a certain set of events.

In line with Hartley's entropy, we can also view Shannon's entropy as a measure of uncertainty. The difference is that Shannon's measure depends not only on the number of possible messages, but also on the chances of each of the messages occurring. When all messages are equally likely to occur, the entropy is maximal, because you are completely uncertain which message you will receive. When one of the messages has a much higher chance of being sent than the other messages, the uncertainty decreases. You expect to receive that one message and in most cases you will be right. The amount of information for the individual messages that have a small chance of occurring is high, but, *on average*, the information (entropy/uncertainty) is lower. As a hypothetical example, let us assume a 1-yr old child uses the words "mummy," "daddy," "cat," and "uh-oh." If the child uses all words almost as frequently, with a slight preference for "mummy," the respective percentages of times the words are used could be 0.35, 0.2, 0.2, and 0.25. The entropy of the child's language is then $-0.35 \log 0.35 - 0.2 \log 0.2 - 0.2 \log 0.2 - 0.25 \log 0.25 = 1.96$. Some time later, the vocabulary may have expanded and changed to ("mummy" 0.05), ("daddy" 0.05), ("cat" 0.02), ("train" 0.02), ("car" 0.02), ("cookie" 0.02), ("telly" 0.02), and ("no" 0.8). Now one word is dominant and the entropy of the language has dropped to 1.25. There is less uncertainty about which word the child will utter. Whatever you ask, the answer is almost certainly "no."



(a) (b) (c)

Fig. 1. Example of a feature space for (a) a CT image and (b) an MR image. (c) Along the axes of the feature space, the gray values of the two images are plotted: from left to right for CT and from top to bottom for MR. The feature space is constructed by counting the number of times a combination of gray values occurs. For each pair of corresponding points $(\mathbf{x}, \mathbf{y})$, with $\mathbf{x}$ a point in the CT image and $\mathbf{y}$ a point in the MR image, the entry $(I_{\mathrm{CT}}(\mathbf{x}), I_{\mathrm{MR}}(\mathbf{y}))$ in the feature space on the right is increased. A distinguishable cluster in the feature space is the stretched vertical cluster, which is the rather homogeneous area of brain in the CT image corresponding to a range of gray values for the MR image.

The Shannon entropy can also be computed for an image, in which case we do not focus on the probabilities of letters or words occurring, but on the distribution of the gray values of the image. A probability distribution of gray values can be estimated by counting the number of times each gray value occurs in the image and dividing those numbers by the total number of occurrences. An image consisting of almost a single intensity will have a low entropy value; it contains very little information. A high entropy value will be yielded by an image with more or less equal quantities of many different intensities, which is an image containing a lot of information.

In this manner, the Shannon entropy is also a measure of dispersion of a probability distribution. A distribution with a single sharp peak corresponds to a low entropy value, whereas a dispersed distribution yields a high entropy value.

Summarizing, entropy has three interpretations: the amount of information an event (message, gray value of a point) gives when it takes place, the uncertainty about the outcome of an event and the dispersion of the probabilities with which the events take place.

III. Mutual Information

A. History

The research that eventually led to the introduction of mutual information as a registration measure dates back to the early 1990s. Woods *et al.* [5], [6] first introduced a registration measure for multimodality images based on the assumption that regions of similar tissue (and, hence, similar gray values) in one image would correspond to regions in the other image that also consist of similar gray values (though probably different values to those of the first image). Ideally, the ratio of the gray values for all corresponding points in a certain region in either image varies little. Consequently, the average variance of this ratio for all regions is minimized to achieve registration.

Hill *et al.* [7] proposed an adaption of Woods' measure. They constructed a *feature space*, which is a two-dimensional (2-D) plot showing the combinations of gray values in each of the two images for all corresponding points. Fig. 1 shows an example

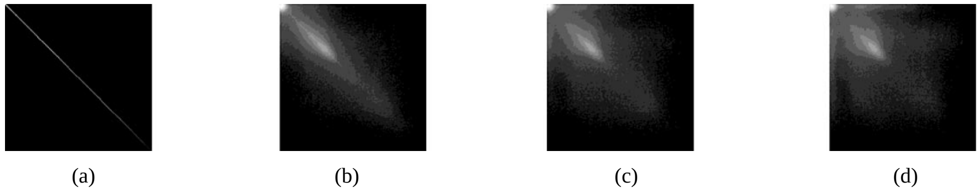


Fig. 2. Joint gray value histograms of an MR image with itself. (a) Histogram shows the situation when the images are registered. Because the images are identical, all gray value correspondences lie on the diagonal. (b), (c), and (d) show the resulting histograms when one MR image is rotated with respect to the other by angles of 2°, 5°, and 10°, respectively. The corresponding joint entropy values are (a) 3.82; (b) 6.79; (c) 6.98; and (d) 7.15..

of such a feature space for a magnetic resonance (MR) and a computed tomography (CT) image. The difference with Woods' method is that instead of defining regions of similar tissue in the images, regions are defined in the feature space. These regions are based on the clustering one finds in the feature space for registered images.

The feature space (or joint histogram) changes as the alignment of the images changes. When the images are correctly registered, corresponding anatomical structures overlap and the joint histogram will show certain clusters for the gray values of those structures. For example, the cluster in the top left corner of the histogram in Fig. 1 is the combination of background in both images. As the images become misaligned, structures will also start overlapping structures that are not their anatomical counterparts in the other image. Consequently, the intensity of the clusters for corresponding anatomical structures will decrease and new combinations of gray values emerge such as skull and brain or skin and background. This will manifest itself in the joint histogram by a dispersion of the clustering. Fig. 2 contains several histograms of an MR image with itself for different rotations of one image with respect to the other. Clearly, the histogram shows increasing dispersion as the misregistration increases.

Using this characteristic of the joint histogram of two images, measures of dispersion emerged, to use for image registration. Hill *et al.* [8] proposed the third-order moment of the joint histogram, which measures the skewness of a distribution. Both Collignon *et al.* [9] and Studholme *et al.* [10] suggested to use entropy as a measure of registration. As we have explained in Section II, entropy measures the dispersion of a probability distribution. It is low when a distribution has a few sharply defined, dominant peaks and it is maximal when all outcomes have an equal chance of occurring. A joint histogram of two images can be used to estimate a joint probability distribution of their gray values by dividing each entry in the histogram by the total number of entries. The Shannon entropy for a joint distribution is defined as

$$-\sum_{i,j} p(i,j)\log p(i,j). \tag{4}$$

By finding the transformation that minimizes their joint entropy, images should be registered.

Once entropy, a measure from information theory, had been introduced for the registration of multimodality medical images, another such measure quickly appeared: mutual information. It was pioneered both by Collignon *et al.* [11], [12], and by Viola and Wells [13]–[15]. Applied to rigid registration of multimodality images, mutual information showed great promise and within a few years it became the most investigated measure for medical image registration.

B. Definition

Most books on information theory ([16]–[18], for example) discuss the notion of mutual information. The definition of the term, however, can be presented in various ways. We will next treat three frequently used forms of the definition, because more than one is used in the literature. All three forms are identical; each can be rewritten into the other two[1]. Each form of definition, however, explains the relation to registration in a different way. We will describe mutual information for two images, as used in image registration, and not in a general sense.

The first form of definition we discuss is the one that best explains the term "mutual information." For two images A and B, mutual information I can be defined as

$$I(A,B) = H(B) - H(B\,|\,A) \tag{5}$$

where $H(B)$ is the Shannon entropy of image B, computed on the probability distribution of the gray values. $H(B\,|\,A)$ denotes the conditional entropy, which is based on the conditional probabilities $p(b\,|\,a)$, the chance of gray value b in image B given that the corresponding voxel in A has gray value a. When interpreting entropy as a measure of uncertainty, (5) translates to "the amount of uncertainty about image B minus the uncertainty about B when A is known." In other words, mutual information is the amount by which the uncertainty about B decreases when A is given: the amount of information A contains about B. Because A and B can be interchanged, $I(A,B)$ is also the amount of information B contains about A. Hence, it is *mutual* information. Registration is assumed to correspond to maximizing mutual information: the images have to be aligned in such a manner that the amount of information they contain about each other is maximal.

The second form of definition is most closely related to joint entropy. It is

$$I(A,B) = H(A) + H(B) - H(A,B). \tag{6}$$

This form contains the term $-H(A,B)$, which means that maximizing mutual information is related to minimizing joint entropy. We have described above how the joint histogram of two images' gray values disperses with misregistration and that joint entropy is a measure of dispersion. The advantage of mutual information over joint entropy *per se*, is that it includes the entropies of the separate images. Mutual information and joint entropy are computed for the overlapping parts of the images and the measures are therefore sensitive to the size and the contents

[1]We conjecture that the identity of the three definitions only holds for the Shannon entropy and we invite anyone to prove of refute this hypothesis.

TABLE I
CLASSIFICATION SCHEME FOR MUTUAL-INFORMATION-BASED REGISTRATION METHODS

Method	**Application**
Preprocessing	Modalities
Measure	monomodality
entropy	multimodality
normalization	modality to model
spatial information	modality to physical space
Transformation	Subject
rigid	intrasubject
affine	intersubject
perspective	model
curved	
Implementation	Object
interpolation	
pdf estimation	
optimization	
acceleration	
Image dimensionality	Number of images
2D/2D	2
3D/3D	> 2, with known inter-image geometry
2D/3D	> 2, with unknown inter-image geometry

of overlap. A problem that can occur when using joint entropy on its own, is that low values (normally associated with a high degree of alignment) can be found for complete misregistrations. For example, when transforming one image to such an extent that only an area of background overlaps for the two images, the joint histogram will be very sharp. There is only one peak, that of background. Mutual information is better equipped to avoid such problems, because it includes the marginal entropies $H(A)$ and $H(B)$. These will have low values when the overlapping part of the images contains only background and high values when it contains anatomical structure. The marginal entropies will thus balance the measure somewhat by penalizing for transformations that decrease the amount of information in the separate images. Consequently, mutual information is less sensitive to overlap than joint entropy, although not completely immune.

The final form of definition we discuss is related to the Kullback–Leibler distance, which is defined as $\sum_i p(i)\log(p(i)/q(i))$, for two distributions p and q. It is a measure of the distance between two distributions. Analogous to the Kullback–Leibler measure, the mutual information of images A and B is defined as

$$I(A,B) = \sum_{a,b} p(a,b)\log\frac{p(a,b)}{p(a)p(b)}. \tag{7}$$

The interpretation of this form is that it measures the distance between the joint distribution of the images' gray values $p(a,b)$ and the joint distribution in case of independence of the images, $p(a)p(b)$. It is a measure of *dependence* between two images. The assumption is that there is maximal dependence between the gray values of the images when they are correctly aligned. Misregistration will result in a decrease in the measure.

C. Properties

Mutual information has the following properties [16].

1) $I(A,B) = I(B,A)$.
   It is symmetric; otherwise it would not be *mutual* information. However, although it is a logical property in theory, mutual information is not symmetric in practice. Implementational aspects of a registration method, such as interpolation and number of samples, can result in differences in outcome when registering A to B or B to A.
2) $I(A,A) = H(A)$.
   The information image A contains about itself is equal to the information (entropy) of image A.
3) $I(A,B) \leq H(A), I(A,B) \leq H(B)$.
   The information the images contain about each other can never be greater than the information in the images themselves.
4) $I(A,B) \geq 0$.
   The uncertainty about A cannot be increased by learning about B.
5) $I(A,B) = 0$ if and only if A and B are independent.
   When A and B are not in any way related, no knowledge is gained about one image when the other is given.

IV. SURVEY OF LITERATURE

For our survey of the different aspects of mutual-information-based registration, we have defined a classification scheme, which builds on earlier schemes for medical image registration in general [19], [20]. The mutual information scheme is presented in Table I.

The main subdivision of the classification is in aspects concerning the *method* and those concerning the *application*.

In addition, the classes *image dimensionality* and *number of images* are defined. The elements of these two subclasses can concern purely the application, but they can also necessitate an adaptation of the method. They are therefore treated separately. The class "Method" can be further subdivided into *preprocessing*, *measure*, *transformation*, and *implementation*. Preprocessing entails any image processing to prepare or improve the images for registration. Typical examples are filtering to remove noise, extraction of regions of interest and isotropic resampling. The measure will obviously be (based on) mutual information, but differences are possible, characterized by the choice of entropy, by normalization and by adaptations to incorporate spatial information. The transformation can be classified as either rigid (rotations and translations only), affine (rigid plus scaling and shearing), perspective (affine without preservation of parallelism of lines) or curved. Implementation is an important category, because the choice of method for interpolation, probability distribution function (pdf) estimation, optimization and acceleration can have a substantial influence on the registration results.

The first aspect of the class "Application" is the type of *modalities* it concerns. The images can be of the same kind (monomodality), acquired by different techniques (multimodality), a type of model can be involved (a phantom or atlas, for example) or images are registered to physical space. By the latter we mean registration of previously acquired images to a person, as is used for image-guided surgery or radiotherapy treatment. *Subject* denotes whether images of a single person are involved, which is called *intra*subject registration, or images of different persons, *inter*subject registration, or whether images of a person are matched to a model. Finally, the anatomy that the registration focuses on is what we term *object*.

Of *image dimensionality* we have found instances of 2-D/2-D, 3-D/3-D, and 2-D/3-D registration in the literature. The *number of images* involved in the registration is usually 2, but registration of more than two images has been described in a number of publications. The latter aspect can be further divided into registration problems where the transformations between several images are known and only a single transformation is to be found or problems that require several transformations.

In the following, we will discuss most categories of our scheme. Some aspects are so dominant in the literature (for example, 3-D/3-D registration), that we will only review the exceptions. We have also taken the liberty to focus on the aspects that *we* find most interesting. After the classification scheme, we will consider a number of comparison studies.

A. Preprocessing

Several techniques of processing images before registration have been described. The most common preprocessing step is defining a region [21], [22] or structures [23]–[38] of interest in the images to exclude structures that may negatively influence the registration results. Other processing techniques reported include low-pass filtering to remove speckle in ultrasound images [33], [38]–[40] and thresholding or filtering to remove noise [41], [42]. Blurring is also applied to correct for differences in the intrinsic resolution of the images [34], [43]–[45]. Intensity inhomogeneities in images are corrected in several methods, both for MR images [29], [46] and for portal images [47]. Some methods resample the images isotropically, to achieve similar voxel sizes in all image dimensions [21], [35], [43], [44], [48], [49], others resample to obtain similar voxel sizes in the images to be registered [22].

B. Measure

Obviously, in a literature survey on mutual-information-based image registration, the measure in question will always be mutual information. However, when using a definition of mutual-information-based on entropy, different definitions of entropy can be chosen. Furthermore, several adaptations of mutual information have been proposed: normalization with respect to the overlapping part of the images and inclusion of spatial information.

A few recently proposed methods do not adapt the mutual information measure, but cannot be considered standard implementations either. Butz and Thiran [50] compute feature images (gradients) of which the mutual information is calculated. Nyúl *et al.* [51] evaluate the mutual information of "scale images": the value of a voxel is the radius of the largest sphere which is centred at the voxel and which falls completely within a single object.

1) Entropy: By far the most common measure of entropy in the papers in this survey is the Shannon entropy [4]. Rodriguez and Loew [52], [53] use the Jumarie entropy [54]. The Jumarie entropy is defined for one-dimensional (1-D) signals and resembles a normalized version of Shannon entropy, applied not to a probability distribution, but to function value differences of neighboring samples. In [52], 2-D images are registered. The authors define the Jumarie entropy of a 2-D image on the gradient magnitude of pixels. The joint Jumarie entropy is defined on the gray value difference of corresponding pixels, which presumably makes the measure less suitable to registration of multimodality images. Ioannides *et al.* [55] use the Rényi entropy of order 4, although not for image registration, but for comparison of brain activity during different tasks. The Rényi entropy of order 2 is employed by Pompe *et al.* [56] to measure the strength of dependence between 1-D respiratory and cardiac signals.

2) Normalization: The size of the overlapping part of the images influences the mutual information measure in two ways. First of all, a decrease in overlap decreases the number of samples, which reduces the statistical power of the probability distribution estimation. Second, Studholme *et al.* [43], [44] have shown that with increasing misregistration (which usually coincides with decreasing overlap) the mutual information measure may actually *increase*. This can occur when the relative areas of object and background even out and the sum of the marginal entropies increases, faster than the joint entropy. Studholme *et al.* proposed a *normalized* measure of mutual information [44], which is less sensitive to changes in overlap

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

They found a distinct improvement in the behavior of the normalized measure for rigid registration of MR-CT and MR-PET (positron emission tomography) images.

Collignon [11] and Maes [57] have suggested the use of the *entropy correlation coefficient* (ECC), another form of normalized mutual information. NMI and ECC are related in the following manner: $\mathrm{ECC} = 2 - 2/\mathrm{NMI}$.

Normalized mutual information was used in a large number of studies [24], [26], [30], [32], [36], [37], [42], [48], [58]–[77].

An upper bound of mutual information was derived by Skouson *et al.* [78].

3) Spatial Information: A drawback of mutual information as it is commonly used, i.e., based on the Shannon entropy, is that the dependence of the gray values of neighboring voxels is ignored. The original Shannon entropy [4] *does* include a dependence of preceding signals, but the definition used in practice is the one for independent successive signals. The assumption of independence does not generally hold for medical images. Incorporating the dependence of the gray values of neighboring voxels, what we term the spatial information of the images, could improve registration.

As mentioned, Rodriguez and Loew [52] employ the Jumarie entropy, which considers the gray value differences of neighboring voxels in an image. Studholme *et al.* [79] compute the mutual information of two images together with a labeling of one of the images. Voxels with identical gray values can then be differentiated when they belong to different regions. The use of a cooccurrence matrix has been put forth by Rueckert *et al.* [80]. The cooccurrence matrix of distance d of an image is a 2-D histogram giving the frequencies of two gray values in the image being distance d apart. Rueckert *et al.* show the effect the method has on curved registration of MR images. Another method of incorporating spatial information is to combine mutual information with a measure based on the gradients at corresponding points. The measure seeks to align gradient vectors of large magnitude as well as of similar orientation [69], [81]. A slightly adapted version of the measure is used by Lötjönen and Mäkelä for curved registration [82].

C. Transformation

The transformation applied to register the images can be categorized according to the degrees of freedom. We define a *rigid* transformation as one that includes only translations and rotations. Although in the literature, rigid transformations are sometimes allowed to include scaling, we classify such transformations as *affine*. An affine transformation can furthermore include shearing. This type of transformation maps straight lines to straight lines and preserves the parallelism between lines. The *perspective* transformation differs from the affine transformation in the sense that the parallelism of lines need not be preserved. It is usually applied in 2-D/3-D registration. No instances of "true" perspective transformation were encountered. All methods using a perspective transformation limited the optimization to the rigid-body or affine parameters; the projective parameters were kept fixed. The final class consists of *curved* transformations, which allow the mapping of straight lines to curves.

1) Rigid: Translations and rotations suffice to register images of rigid objects. Examples include registration of bone or of the brain when neither skull nor dura has been opened. Rigid registration of images based on mutual information has been applied in a large number of papers [11], [12], [21]–[23], [27], [31], [35], [41], [43]–[45], [47], [49], [51], [53], [57], [63], [71], [83]–[101]. Rigid registration is also used to approximately align images that show small changes in object shape (for example, successive histological sections [102], [103] and serial MR images [24], [26]) or small changes in object intensity, as in functional MR time series images [93], [104].

2) Affine: The affine transformation preserves the parallelism of lines, but not their lengths or their angles. It extends the degrees of freedom of the rigid transformation with a scaling factor for each image dimension [25], [32], [58], [105], [106] and, additionally, a shearing in each dimension [13], [28], [38], [99], [107], [108]. In [109], [110] an affine registration with nine degrees of freedom is performed to correct calibration errors in the voxel dimensions. Holden [110] furthermore measures the *relative* scaling error between scans. Shekhar and Zagrodsky [33] compare registration of ultrasound images using transformations of increasing complexity (rigid, rigid with uniform scaling, rigid with nonuniform scaling and fully affine).

3) Curved: Curved registration methods can differ on several aspects. The mutual information measure can be calculated globally, on the entire image, or locally, on a subimage. Smoothness of the deformation can be achieved in different ways and the deformation can be either free-form (any deformation is allowed) or guided by an underlying physical model of material properties, such as tissue elasticity or fluid flow. Besides these aspects, methods can also differ in smaller, implementational details, but such differences will not be discussed.

Meyer *et al.* [39], [111]–[116] compute the mutual information measure globally. The deformation is determined by thin-plate splines through a number of control points, which are initialized by the user, but are adapted automatically. The number of control points defines the elasticity of the deformation. Apart from registration of 3-D multimodality images, the method was applied to warp a slice into a volume, including out-of-plane deformations [117]. Also computing both measure and deformation globally is Horsfield [118], who uses a third-order polynomial to nonuniformly correct MR images for eddy current distortion.

Other methods compute the mutual information globally, but find the deformation on a local scale. A grid of control points is defined to determine the deformation, usually in a multiresolution manner. The points of the grid are moved individually, defining local deformations. Transformations in between control points are propagated by linear interpolation [29], [119], Gaussian kernels [120] or other symmetrical, convex kernels [82], [121]. Rueckert *et al.* [74], and Studholme *et al.* [37], [76] calculate B-splines through the control points, which have a local region of influence (as opposed to thin-plate splines). A similar method is employed in [72], [73], [122]. The effect of the choice of transformation (rigid, affine or curved) on registration of MR breast images was studied by Denton *et al.* [62]. The method by Rueckert was adapted to allow for rigid structures within deformable tissue by Tanner *et al.* [77] through fixation of intercontrol point distances. A nonuniform deformation grid of active and passive control points is described in [75]. Applications of the method include propagation of segmentations [42], [61] and the construction of a statistical deformation model [123].

Contrary to the previous methods which compute mutual information globally, some methods compute the mutual information measure for subsets of the images [30], [46], [60], [65], [124], [125]. A problem with local computation of mutual information is that the results can suffer from the small number of samples. Usually, relatively large subimages are required, which prohibits deformations on a very small scale. Several adaptations have been proposed to overcome this problem. Likar and Pernuš [66] define local probabilities as a weighted combination of the probability distribution of a subimage and the global distribution. Maintz *et al.* [126] compute a conditional probability distribution of intensities in one image given intensities in the other image, based on a global joint histogram. Using the conditional distribution, translations of subimages are computed. Finally, Rueckert *et al.* [80] enhance the power of locally computed measures by including spatial information, in the form of cooccurrence matrices.

Hermosillo and Faugeras [127] compare global and local computation of both mutual information and the correlation ratio [128]. Schnabel *et al.* [129] propose a validation method for curved registration, which is demonstrated on mutual information.

Most methods ensure smoothness of the deformation field, by filtering of the vector field (e.g., [29], [64]–[66], [130]) and/or by regularization terms to constrain local deformations (e.g., [29], [30], [36], [74], [75], [82], [125], [127]). Rohlfing and Maurer [73] incorporate a regularization term that prevents compression of contrast-enhanced structures.

To the best of our knowledge, there are only two papers on inclusion of physical models of tissue deformation in mutual-information-based curved registration methods. Both Hata *et al.* [124] and Hermosillo and Faugeras [127] use a model of an elastic solid material for regularization of the deformation.

D. Implementation

The importance of the implementation of a mutual-information-based method should not be underestimated, since implementational decisions can have a large influence on the registration results. The main choices involve interpolation, estimation of the probability distributions and optimization. Additionally, one may choose to improve the speed of registration. Zhu and Cochoff [101] study the influence of several implementation choices, viz. optimization method, interpolation method, number of histogram bins and multiresolution approaches. The choice of implementation remains a matter of debate. An optimal implementation has not been agreed on, partly because all aspects of the implementation interact. For instance, one cannot compare optimization methods without taking the other aspects into account, because these influence the smoothness of the function to be optimized.

1) Interpolation: When transforming points from one image to another, interpolation is usually required to estimate the gray value of the resulting point. In this section, we focus on interpolation *during* the registration process, which is applied numerous times and which, consequently, necessitates a tradeoff between accuracy and speed. In addition, interpolation is required to yield a final, registered image. Since this task is performed only once, speed is less of an issue and a different choice of interpolation method (e.g., a higher order method) may be more appropriate.

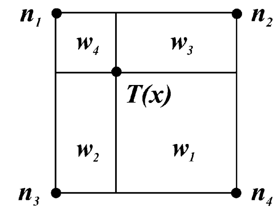


Fig. 3. Interpolation weights; the areas w_i for 2-D linear interpolation.

The most popular technique of interpolation is linear interpolation, which defines the intensity of a point as the weighted combination of the intensities of its neighbors. The weights are linearly dependent on the distance between the point and its neighbors, as shown in the 2-D example in Fig. 3. A handful of papers report the use of nearest neighbor interpolation (assigning the gray value of the spatially closest neighbor), often for speed [32], [47], for comparison to other interpolation methods [57], [101] or for the initial testing of a novel idea [52].

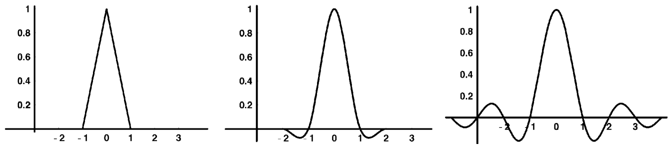
An interpolation method specifically designed to create joint histograms of intensities is partial volume interpolation, introduced by Collignon [12]. It uses the weights of linear interpolation, but not to compute a weighted intensity and update a single histogram entry, like linear interpolation. It uses the weights for fractional updates of the histogram entries corresponding to a transformed point and each of its neighbors. Effectively, this creates smoother changes of the joint histogram for varying transformations and hence a smoother registration function[2]. The method has been adopted by several others [33], [57], [70], [88], [96], [125].

Maes [92] introduced partial intensity interpolation. This method calculates a weighted average of the neighboring gray values, identical to linear interpolation. Then, however, two histogram entries (those corresponding to the floor and the ceiling of the weighted average) are updated by a fractional amount.

Thévenaz and Unser [97] are advocates of higher order interpolation methods. They suggest cubic interpolation, particularly in multiresolution methods. Cubic spline interpolation is also used in [104] and [122]. Fig. 4 shows 1-D interpolation kernels for linear, cubic and sinc interpolation. Sinc interpolation is the ideal kernel in theory, but it is impractical for two reasons: 1) the images are expected to be band-limited, which medical images rarely are and 2) the width of the kernel is infinite. The cubic kernel has a larger extent than linear interpolation and is, therefore, more expensive to compute, but does approximate the sinc kernel better. The influence of the order of the interpolation method is studied by Netsch *et al.* [93].

A serious problem with interpolation is that it can cause patterns of artefacts in the registration function. When the grids of

[2]The registration measure as a function of transformation.



(a) (b) (c)

Fig. 4. Several 1-D interpolation kernels: (a) linear; (b) cubic; and (c) sinc (truncated).

two images can be aligned for certain transformations, no interpolation is required for such transformations. Because interpolation influences the value of the registration measure, the absence of interpolation—at grid-aligning transformations—can cause a sudden change in the value of the measure, resulting in a pattern of local extrema. The occurrence of such patterns has been noted in several publications [11], [92], [131]. In [132], the different patterns created by linear and partial volume interpolation are extensively studied. Holden [110] describes the existence of artefacts for both mutual information and the ratio image uniformity [133] measures when using linear interpolation and proposes low-pass filtering as a solution. Likar and Pernuš [66] try to overcome the severe artefacts in the registration functions of subimages, either by a random resampling of the image grids or by including the probability distribution of the entire images. Chen and Varshney [134] employ a *generalized partial volume* interpolation method, which is identical to partial volume interpolation using a higher order kernel instead of a first-order one. Interpolation artefacts deserve serious attention, not only because they can cause misregistrations, but also because they prohibit subvoxel accuracy.

2) Probability Distribution Estimation: The most straightforward way to estimate the joint probability distribution of intensities in two images is to compute a joint histogram of intensities. Each entry $h(a, b)$ in the histogram denotes the number of times intensity a in one image coincides with b in the other image. Dividing the entries by the total number of entries yields a probability distribution. The probability distributions for each image separately are found by summing over the rows, resp. columns, of the histogram. This method is chosen in the majority of papers [12], [22], [32]–[36], [38], [40], [44], [45], [47], [51], [52], [57], [58], [66], [72], [74], [77], [88], [91], [101], [106], [107], [112], [116], [120], [121], [125], [126], [134]–[136]. Camp and Robb [59] propose a method that better distributes the entries across all histogram bins.

Another frequently used method of distribution estimation is Parzen windowing. Given a set S of n samples, the probability $p(x)$ of x occurring is the sum of the contributions of each sample s from S to $p(x)$. The contributions are functions of the distance between s and x. This results in the following definition of the probability of x given a sample S

$$P(x, S) = \frac{1}{n} \sum_{s \in S} W(x - s).$$

The weighting function W is a Gaussian function in most implementations described in the literature [83], [96], [99], [100], [108], [119], [127]. Other choices are double exponential functions [137] and splines [97], [122].

Whereas the simple histogram method places a spike function of value 1 at the bin corresponding to s and updates only a single bin, Parzen windowing places a kernel at the bin of s and updates all bins falling under the kernel with the corresponding kernel value.

3) Optimization: The registration measure as a function of transformation defines an n-dimensional function, with n the degrees of freedom of the transformation. The optimum of this function is assumed to correspond to the transformation that correctly registers the images. Unfortunately, the registration function is generally not a smooth function, but one containing many local maxima. The local maxima can have two different causes. Some represent a local good match of the two images. Others are imperfections inherent to the implementation, for example, local maxima can occur as a result of interpolation or because of changes in the overlapping part of the images. Local maxima in the registration function can be reduced, among other things, by improving implementation choices (e.g., a higher order interpolation method), by filtering the images to reduce noise or by increasing the bin size of the intensity histogram. Because of the existence of local maxima, the choice of optimization routine has a large influence on the results of the registration method, particularly on the robustness of the method with respect to the initial transformation.

A second important property of the registration function that influences the choice of optimization method is the capture range of the optimum [1], [2], [35], [44]. For intensity-based registration measures, it is possible that a large misregistration of two images results in a higher value of the measure than the correct transformation. The desired maximum may not be the global maximum of the search space and only part of the search space leads to the desired maximum. This has two consequences for optimization of the registration function. First of all, an optimization started outside the capture range of the desired maximum has little chance of leading to a correct registration of the images. Second, probabilistic optimization routines such as some multistart methods and genetic algorithms, may prove to be less suitable for optimization of the mutual information measure, because they can move outside the capture range. The extent of the capture range depends on the registration measure and on image properties (e.g., modality, contents, field of view) and cannot be determined *a priori*.

We will only mention some characteristics of the optimization methods. Detailed descriptions can be found in general works on optimization techniques such as [138] and [139] and in the papers cited or references therein.

A popular method is Powell's routine, which optimizes each transformation parameter in turn. It does not require function derivatives to be calculated, but is relatively sensitive to local optima in the registration function [11], [21], [34], [51], [52], [57], [59], [65], [66], [70], [88], [96], [101], [108], [140], [141].

Equally popular is the Simplex method, which does not require derivatives either, but, contrary to the previous method, considers all degrees of freedom simultaneously [22], [32], [33], [38]–[40], [64], [68], [85], [89], [101], [105], [111], [116], [118], [135], [141]. It is not known for its speed of convergence.

Plattard *et al.* [47] use a combination of the Powell and Simplex methods, whereas Kagadis *et al.* [28] combine Powell and a genetic algorithm. Jenkinson and Smith [107] propose an optimization routine that extends Powell's method with initialization and a multistart technique.

Rodriguez and Loew [53] combine Powell with topographical global optimization. This involves a graph structure with the nodes denoting points in the search space and the arcs pointing in the direction of nodes with lower function values. In this manner, the influence zones of local maxima can be determined and a number of local maxima is selected based upon the graph to start optimizations from.[3]

Although being one of the simplest optimization techniques, hill-climbing optimization was shown to produce good results in a multiresolution scheme, with the step size of the hill-climbing method decreasing as the image resolution increased [30], [43], [44].

Methods that do require function derivatives (whether mathematically derived or numerically estimated) are gradient ascent [36], [37], [45], [72]–[74], [99], [100], [121], [127], [140]–[142], quasi-Newton methods [122], [141] and the method by Levenberg–Marquardt [97], [141]. Exact expressions for the gradient of mutual information are derived in [141].

A method little used in image registration is simulated annealing, which has the seemingly paradoxical property of sometimes taking a step in the "wrong" direction (i.e., toward a smaller function value when the goal is maximization) [91], [94], [106]. This move is allowed occasionally to make escapes from local maxima possible. Equally uncommon are genetic algorithms [50], [60], which are based on the survival-of-the-fittest principle of combining current elements and selecting the best of the new elements.

An unconventional approach of finding the optimal transformation is employed in [41]. Template matching of subimages is used to define a set of corresponding points (the center points of the subimages), based upon which a rigid transformation is determined.

To improve the chances of finding the global optimum of the registration function, Chen and Varshney [134] compute the mutual information both of the entire images and of four subimages, assuming that when the global mutual information is maximum, this should also hold for subimages. Zagrodsky *et al.* [38] use the mutual information value of three intensity histograms of different bin widths simultaneously to find the optimal transformation.

Optimization is often performed in a multiresolution manner, as this is expected to decrease the sensitivity of the method to local maxima in the registration function. The term multiresolution can be used with respect to the images, in the sense that the images are down- or upscaled to a number of resolution levels [21], [29]–[32], [35], [36], [44], [46], [48], [51], [71], [82], [97], [100], [101], [120], [127], [140], [141]. Multiresolution can also apply to the deformation grid of curved registration methods [29], [30], [36], [37], [46], [60], [64], [65], [72], [74]–[76], [82], [94], [121], [122], [125].

[3]We have adapted the description of the method to apply to function *maximization.*

Holmes *et al.* [140] compare two optimization methods together with several other aspects such as subsampling and thresholding to extract objects. The most extensive comparison of optimization methods for mutual-information-based image registration, including multiresolution implementations, can be found in [141].

4) Acceleration: Apart from improving the behavior of a method with respect to local maxima in the registration function, multiresolution schemes can also improve the speed of an algorithm. A rough estimate of registration is found in relatively little time using downsampled images, which is subsequently refined using images of increasing resolution. Registration at finer scales should be faster as a result of a reasonable initial estimate. In [70], simple equidistant subsampling, both with and without Gaussian blurring of the images, is compared for registration of MR, CT, and PET images. Similarly, Zhu and Cochoff [101] compare subsampling both with and without averaging of gray values. Maes *et al.* [141] study the behavior of a large number of optimization methods in combination with multiresolution approaches. Rohlfing and Maurer [73] decrease the computational demand by selectively refining the deformation grid, based on a local entropy measure. Rohde *et al.* [121] base the selective refinement on the gradient of the registration function, assuming that a large gradient is likely to denote a mismatched area. Similarly, Schnabel *et al.* [75] label selected control points as passive, based either on a segmentation of the image or local statistical measures. Mattes *et al.* [122] combine a hierarchical refinement of the deformation grid with a hierarchical degree of Gaussian blurring of the images before registration.

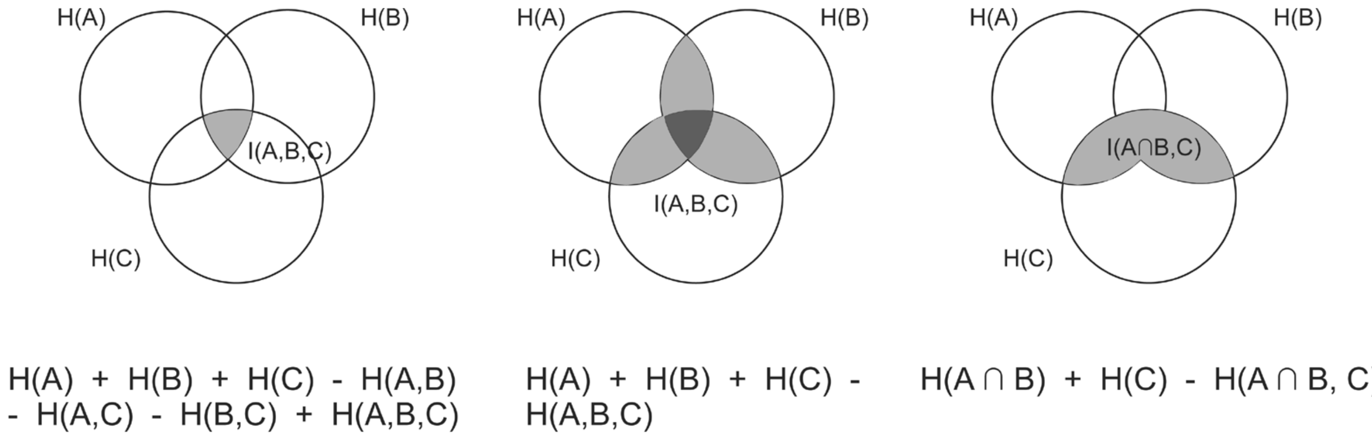
Several authors replace costly calculations by lookup tables. Sarrut and Miguet [143] use lookup tables to avoid several computations for each voxel, such as the calculation of the weights of interpolation. Meihe *et al.* [144] speed up the Parzen windowing process using lookup tables for the Gaussian functions. Zöllei *et al.* [45] employ *sparse* histogramming, i.e., using a small number of samples.

E. Image Dimensionality

The majority of papers treats registration of 3-D images. We will next discuss the exceptions: two 2-D images or a 2-D and a 3-D image.

1) 2-D/2-D: The difficulty with 2-D images is that the number of samples usually is substantially smaller than with 3-D images. This can result in a less reliable estimation of the probability distributions. Good results have been reported nonetheless. The choice for 2-D images is often guided by the application [34], [47], [65], [66], [87], [95], [103], [106], [108], [112], [119], [145]. Other times 2-D images are chosen for initial testing of a novel idea, frequently with the intention of extension to three dimensions [52], [60], [111], [131].

2) 2-D/3-D: Registration of 2-D and 3-D images is regularly applied to find the correspondence between the operative scene and a preoperative image. Viola and Wells [13], [99], for example, devised a method of using mutual information to register 2-D video images to a model of a 3-D object (usually based on an MR or a CT image). Other papers in this area include [86], [142]. Bansal *et al.* [83] register 2-D portal images to a preoperative CT in order to verify the position of the patient with respect



(a) (b) (c)

Fig. 5. Different definitions of the mutual information (shaded areas) of three images (a)–(c). The dark gray color in (c) signifies that the area is counted twice. The circles denote the entropy of an image; joint entropy is the union of circles.

to the radiotherapy treatment plan. They propose an iterative approach, which switches between segmenting the images based on the current registration and registering the images based on the current segmentation. Other papers on registration of 2-D portal images and CT are [47], [96]. Zöllei *et al.* [45] and Kim *et al.* [89] register CT and 2-D fluoroscopy images to verify patient position.

Kim *et al.* [135] correct for motion in functional MRI (fMRI) acquisitions by registering fMRI slices into a 3-D anatomical MR scan. In [117], out-of-plane deformation of the slices is introduced. Calibration of an ultrasound probe using 2-D/3-D registration is described by Blackall *et al.* [58], registering 2-D B-mode ultrasound images to an MR volume to allow reconstruction of a 3-D ultrasound image.

A comparison of six intensity-based registration measures, for registration of a 2-D fluoroscopy image to a CT volume, has been made by Penney *et al.* [31].

F. Number of Images

Commonly, two images are involved in the registration process. However, in certain situations several images of a scene are to be registered or a series of images taken at different times needs to be compared.

When more than two images are employed, two types of registration problems can be distinguished: with known and with unknown inter-image geometry. In the first case, the transformations between several images are known and only a single transformation has to be determined. In the second case, no knowledge about the transformation between individual images is available and multiple transformations are to be found to transform the images to a common coordinate system.

1) More Than Two Images, With Known Inter-Image Geometry: An example of the first type is the problem of determining the position of a 3-D object amidst a number of 2-D images of the object, taking from different, known, angles. Several different mutual-information-based solutions have been proposed for this problem. One could simply sum the measures of each 2-D image and the 3-D image [45], [83], [86], [89] or combine the intensity correspondences for each 2-D image and the 3-D image in a single joint histogram [96]. Clarkson *et al.* [142] have compared three methods of combining measures (adding the measures of individual 2-D images and the 3-D image, alternating between separate optimizations or creating a single 2-D histogram of intensity correspondences in each of the 2-D images and the 3-D image). This was applied to registration of 2-D video images and a CT volume.

A similar problem, that of registering an MR volume to a set of 2-D ultrasound images, is tackled by Blackall *et al.* [58] by gathering the corresponding intensity pairs for each slice and the volume into a single joint histogram. Pagoulatos *et al.* [68], on the other hand, optimize the sum of the mutual information of each ultrasound slice and the MR volume.

Another example of a multidimensional registration problem that requires only a single transformation is given by Andersson and Thurfjell [146], who register two “series” of images (one series consisting of two differently weighted MR images and the other of a PET transmission and an emission scan), using a higher dimensional joint intensity histogram. Boes and Meyer [111] also propose to use higher-dimensional mutual information to register two images, using a third image for additional information (which is assumed to be in register with one of the other two images). Studholme *et al.* [79] use higher dimensional mutual information to include a segmentation of an image in the registration to another image.

An interesting question is how to define higher dimensional mutual information. In textbooks and theoretical essays on generalized (i.e., higher dimensional) mutual information [17], [147], [148], the definition of the measure for three images corresponds to Fig. 5(a). In this Venn diagram notation, the shaded area denotes the mutual information between images A, B and C. A property of this definition is that it is not necessarily nonnegative [17]. In the medical image registration literature a different definition has been proposed [79], [111]

$$I(A,B,C) = \sum_{a,b,c} p(a,b,c) \log \frac{p(a,b,c)}{p(a)p(b)p(c)}$$

which can also be written as

$$I(A,B,C) = H(A) + H(B) + H(C) - H(A,B,C).$$

This corresponds to the shaded area in Fig. 5(b) (the darker area is counted twice). This definition is nonnegative, contrary to the previous definition. However, it does not define the mutual information of three images as one would expect, namely as the information that is shared between all three images. The definition depicted in Fig. 5(c) is a slightly different application. Studholme *et al.* [79] and Lynch *et al.* [48] use the mutual information of the *union* of two images (a bivalued image) together with a third image.

2) More Than Two Images, With Unknown Inter-Image Geometry: One instance of a registration problem from the second class (requiring more than one transformation) is described by Lynch *et al.* [48], who register three images. They circumvent the problem of having to optimize several transformations simultaneously by first registering two images. The third image is registered to the previous two using a 2-D intensity distribution for the registered images, which results in a higher dimensional mutual information measure.

Krücker *et al.* [113] register several 3-D ultrasound scans, acquired under different angles, to form a compounded 3-D image with a better signal-to-noise ratio (SNR). The first scan is used as the reference scan to register all subsequent scans to.

Images of a patient that have been taken over a period of time need registration to study changes through time. Usually, the first recorded image acts as a reference to which all subsequent images are registered [24], [26], [108], [119], although sometimes another image is more suitable as a [72]. Kim *et al.* [135] correct for patient motion during the acquisition of fMRI time series by registering each slice into an anatomical volume.

G. Modalities

Mutual information has been applied to a wide variety of image modalities. These can be subdivided into applications of monomodality images, of multimodality images, of an image and a model, and of an image to physical space (e.g., using intraoperative images of a patient).

1) Monomodality: Even though, when first introduced, one of the main advantages of mutual information was its capability to register multimodality images, the measure has also been shown to be well suited to registration of images of the same modality. The following is a brief overview of modalities found in the literature. Although MR images can have very different characteristics for different scanning protocols, we have classified all MR registration problems as monomodality.

MR: Registration of MR images has been described in many publications [21], [51], [52], [149], often for curved transformations [26], [27], [29], [36], [37], [46], [62], [64], [72], [74], [77], [80], [94], [111], [114], [117], [120], [121], [123], [124]. A study by Holmes *et al.* [140] includes matching of MR-T1 and MRA images. Netsch *et al.* [93] register time series of perfusion MR. Some first results of registering interventional MR images can be found in [23], [67], [150]. Time series of fMRI images require registration to detect changes in brain function [93]. Changes in brain anatomy are studied in [24], [42]. Furthermore, registration is needed to map the functional information onto an anatomical MR scan [30], [71], [76], [125]. Kim *et al.* [135] register individual fMRI acquisitions to an anatomical image to correct for patient motion. Freire and Mangin [104] register fMRI images to correct for patient motion and they study the sensitivity of several registration measures to activated areas. Registration to correct for eddy current distortions in diffusion-weighted images is described in [108], [118]. Hill *et al.* [151] employ a curved registration method for intraoperative brain shift measurement. A rigid registration method is used to estimate brain motion with respect to the cranium as a result of patient position [63]. Studholme *et al.* [37] estimate tissue deformation of the brain after electrode implantation. Liu *et al.* [152] compare their proposed measure for extracting the midsagittal plane to mutual information.

CT: In the mutual information literature, CT is usually combined with other modalities and few monomodality cases have been reported. Extraction of the midsagittal plane is described by Liu *et al.* [152]. Martens *et al.* [153] use registration of pre- and postoperative CT images to validate pedicle screw placement, whereas Bergmans *et al.* [154] validate root canal treatment. An unusual application is described by Król *et al.* [91] who use registration to find suitable locations for bone grafting.

SPECT: Holmes *et al.* [140] compare mutual information with a measure similar to Woods' measure [5]. Radau *et al.* [32] compare normalized mutual information with two other measures for the registration of single photon emission computed tomography (SPECT) images to an atlas, created by averaging of a number of SPECT images. The performance of several measures for registration of ictal and interictal images is reported in [85]. Registration of transmission images to achieve alignment of the corresponding emission images is described by Van Laere *et al.* [105].

PET: In Holmes' comparison of mutual information and Woods' measure, PET–PET registration is one of the modality combinations described [140].

US: Meyer *et al.* [39] use mutual information to match breast ultrasound images, whereas Zagrodsky *et al.* [38], [40] register two series of cardiac images. Shekhar and Zagrodsky [33] study the effect of median filtering, number of histogram bins and interpolation method on the smoothness of the registration function of cardiac ultrasound images. Krücker *et al.* [113] form a 3-D ultrasound image with better SNR by registering several 3-D scans, which were acquired under different angles.

Microscopy: Registration of histological sections has been reported both using rigid transformations [102], [103], [145] and curved ones [65], [66].

X-Ray: Sanjay–Gopal *et al.* [95] compare mutual information and the correlation coefficient for registration of intrasubject mammograms. Plattard *et al.* [47] register both 2-D portal images and portal to x-ray images to verify the position of the patient with respect to previous radiotherapy treatment sessions.

Various: Ritter *et al.* [106] apply mutual information to the registration of retinal images, acquired by a fundus camera. Another paper on retinal images is the one by Butz and Thiran [50], who maximize the mutual information of the gradient images. Baker *et al.* [119] register series of electrophoresis images (images of protein, separated based on their isoelectric charge) in order to simplify segmentation. Sjögreen *et al.* [34] register emission and transmission scintillation images of an entire body.

2) Multimodality: Mutual information has been studied for many combinations of modalities.

MR-CT: A popular combination, and one of the earliest described, is registration of MR and CT images [11], [41], [43], [44], [49], [50], [53], [57], [69], [92], [97], [100], [103], [126], [130], [140]. An interesting category are the papers that report on registering what are commonly known as the "RREP" or "Vanderbilt" images [41], [44], [57], [69]–[71], [97], [103], [136], [141]. These images are publicly available and an accurate gold standard is known [155] (although it has been suggested that registration by mutual information may in some cases yield more accurate results than the gold standard [92]). This is one of the few sets of images that allows direct comparison of the accuracy of different methods.

MR-PET: A variety of applications of MR-PET registration has been recounted [11], [29], [41], [43], [44], [57], [69], [79], [92], [97], [100], [116], [140], [146]. The RREP images mentioned above also include MR and PET image pairs.

MR-SPECT: Comparisons between mutual information and other measures for registration of MR and SPECT images are made in [84], [94], [140]. Other publications on the subject of MR-SPECT matching are [48], [98], [101], [103].

MR-US: Registration of ultrasound images to other modalities using mutual information is a relatively unexplored field. Roche *et al.* [156] study the possibilities of using mutual information to register ultrasound to anatomical MR images, while Slomka *et al.* [22] match ultrasound with MRA images. Blackall *et al.* [58] use registration of ultrasound and MR images to calibrate the ultrasound probe. Because ultrasound images can be acquired relatively easily during a procedure, Pagoulatos *et al.* [68] report initial results of registering anatomical MR and ultrasound images with the intent of using the method for image-to-physical-space registration.

CT-PET: Both Erdi *et al.* [25] and Mattes *et al.* [122] register CT and PET *transmission* images of the thorax to achieve fusion of CT and PET *emission* images. CT-PET registration of the thorax is furthermore described by Meyer *et al.* [116].

CT-SPECT: Meyer *et al.* [116] also registered CT and SPECT images, now focusing on the abdomen, as did Koral *et al.* [90]. Kagadis *et al.* [28] compare a surface-based and a mutual-information-based registration routine.

CT-Various: CT has been registered using mutual information to several other modalities, such as 2-D video images [86], [142], 2-D fluoroscopy images [31], [45], [89] and portal images [47], [83], [96].

Microscopy: Flynn *et al.* [87] match stained histological sections with the corresponding radioluminographs (RLG) of the sections. Kim *et al.* [112] warp a histological section to a video image taken before slicing.

3) Modality to Model: By a model we denote any kind of simplified or processed image. A model can be a simulated image [42], [61], [107], [115] or a segmentation [120], [157]. Another possibility is an average image or a statistical model composed of several images [29], [32], [105].

4) Modality to Physical Space: A previously acquired image of a person can be registered to the actual person, via an intraoperative image. This is what we term "registration to physical space." A common application in radiotherapy is the verification of patient position with respect to a treatment plan based on a previously acquired image. Usually, this involves registration of a pretreatment CT to portal [47], [83], [96] or fluoroscopy images [45], [89].

Registration to physical space is also required in image-guided treatment, for transferring the information of a pretreatment image (and any treatment plans based upon the image) to a patient on the operating table. Preoperative images are registered to intraoperatively acquired images, such as video images from the operation microscope [86], [142], ultrasound images [68], fluoroscopy images [31], PET transmission images [25] or interventional MR images [23], [67], [150].

A final application is tracking of a person's movements. Viola and Wells [13], [99] achieve this through registration of a 3-D model to video images.

H. Subject

The subject in the images to be registered can be the same (intrasubject registration), can differ (intersubject registration) or one of the images can be a model. Intersubject registration based on mutual information is a highly relevant topic, because it can form the basis for methods such as tissue segmentation (e.g., [42], [158]), bias field correction in MR images (e.g., [159]) and analysis of images of groups of subjects (e.g., [160], [161]). Only a small percentage of the references deal with intersubject registration, which is partly because intersubject registration using mutual information has only recently gained more attention and partly because we have not included all papers that use registration as one of several steps in an application, but that focus on the application.

Some of the intersubject registration methods include a model and those have been treated in Section IV.G.3. Hellier *et al.* [162] compare five measures for intersubject registration of MR brain images. Studholme *et al.* [36], [160] use a single image as a reference for intersubject registration in cohort studies of patients. Rohde *et al.* [121] register MR images of different patients using a curved transformation. Rangarajan *et al.* [145] register sets of sulcal points sets of different individuals. Finally, Rueckert *et al.* [123] register images of different patients in order to create a statistical deformable model of brain anatomy.

I. Object

The object in medical image registration is the part of the anatomy involved. We have found a varied list of objects, which is summarized in this section.

Brain: A large part of the literature of mutual-information-based registration concerns head or brain images [11], [12], [14], [24], [26]–[30], [32], [35]–[37], [41]–[44], [47], [49]–[53], [57], [64], [67], [69]–[71], [76], [80], [84], [86], [88], [92]–[94], [97]–[101], [104], [107] [108], [111], [115]–[117] [123], [125], [126], [135], [136], [140]–[142], [150], [160], [162].

Thorax/Lungs: Moving downwards we come to the thorax and the lungs, which have been the central theme in a small selection of papers [25], [89], [116].

Spine: Penney *et al.* [31] use images of a realistic spine phantom and add structures (soft tissue, stents) from clinical images to assess the performance of several intensity-based registration measures. Martens *et al.* [153] apply registration to preand postoperative CT images to validate pedicle screw placement.

Heart: Zagrodsky *et al.* [38], [40] register two series of cardiac ultrasound images, with each series of images a sequence of heart cycles. In [33], several adaptations to the method are made to allow affine registration.

Breast: Registration of breast images has been described for various imaging modalities, in particular, MR [46], [62], [74], [77], [114], [149], X-ray [95] and ultrasound [39], [113].

Abdomen/Liver: Several papers have been published on registration of abdominal images [52], [90], [116] and of the liver [23], [72].

Pelvis: As registration of pelvic images is quite a challenging task, almost all references given propose some adaptations (whether large or small) to the standard method of mutual-information-based matching [21], [79], [83].

Tissue: Histological sections of tissue are the object of registration in a number of studies [65], [66], [87], [102], [112], [145].

Various: Retinal images are registered by Ritter *et al.* [106] and by Butz and Thiran [50]. Lynch *et al.* [48] align two MR images and a SPECT image of the knee. MRA and power Doppler ultrasound images of carotid bifurcations are registered by Slomka *et al.* [22] and the resulting transformation is used to register B-mode ultrasound and MRA. Sjögreen *et al.* [34] register whole-body scintillation images. Dental CT images are registered by Bergmans *et al.* [154].

J. Comparison Studies

By the term "comparison study" we mean all papers written with the sole intention of comparing several different registration measures and not papers that primarily intend to present a new method (which often includes a comparison to other methods). Admittedly, the dividing line is thin. Naturally, all studies include mutual information.

Studholme *et al.* apply three measures to rigid registration of MR and CT brain images [49] and five to MR and PET images [35]. A number of measures based on a joint intensity histogram is compared by Bro–Nielsen [163], for registration of MR and CT images. By far the most extensive and the most important comparison study was performed by West *et al.* [155]. It originally comprised 16 methods, but has been extended substantially since. It has the advantage that the registrations were done by the research groups themselves. The accuracy of the methods for rigid registration of clinical CT-MR and PET–MR images pairs was established relative to a method based on bone-implanted markers. In [164], the performance of a number of the methods in the study is compared, after subdivision into surface-based and intensity-based methods. Penney *et al.* [31] study 2-D/3-D registration of fluoroscopy and CT images using six measures. A phantom is used, but the robustness of the measures with respect to differences in image content is studied by extending the phantom images with soft tissue structures and interventional instruments from clinical images. Brinkmann *et al.* [85] study three measures for registration of ictal and interictal SPECT, using phantom, simulated and clinical images. One manual and four automated methods are compared by Flynn *et al.* [87]. They apply the methods to registration of radioluminographs and histological sections, focusing on accuracy (versus markers). Nikou *et al.* [94] adapt two existing measures by including robust estimators and compare all four measures to mutual information. Monomodality registration of MR images and multimodality registration of MR and SPECT images is studied with regard to accuracy (versus a manual solution), robustness with respect to starting estimate and the presence of nonbrain structures. Barnden *et al.* [84] compare the accuracy of five methods to register SPECT and MR images against that of skin fiducials. Mutual information is outperformed by two methods designed specifically for registration of functional and anatomical images. Holden *et al.* [27] compare eight measures for registration of 3-D MR time series of the brain. The property under scrutiny is consistency, which is measured by registering images in triangles (A to B, B to C and C back to A) and calculating the deviation of the composite of the three transformations to the identity transformation. Carrillo *et al.* [23] apply one manual and four automated methods to matching of differently weighted MR images (including contrast enhanced images). The accuracy (versus anatomical landmarks) and the robustness (with respect to, e.g., field of view and starting estimate) were investigated. Van Laere *et al.* [105] describe the performance of three measures for registration of SPECT transmission images. In [81], [136], mutual information is compared to other dependence measures from information theory, in particular, f-information measures. Mutual information is a member of this class of measures, which are all potential registration measures. Freire and Mangin [104] study the performance of six measures on registration of fMRI images, focusing on their sensitivity to activated areas. Nonrobust measures can give rise to erroneous activations in the analysis of the images. Otte [30] compares two measures for curved registration of fMRI to anatomical MR data. Radau *et al.* [32] investigate the sensitivity of three registration measures to (simulated) defects in SPECT images. Four measures for registration of MR and SPECT images are validated by Grova *et al.* [88]. The SPECT images are simulations, derived from the MR images. Hellier *et al.* [162] evaluate intersubject registration of MR brain images for five similarity measures. The transformation considered is curved except for the mutual information method, which employs a rigid transformation. Four intensity-based measures are evaluated by Sarrut and Clippe for registration of 2-D portal and a 3-D CT image [96]. Two methods for CT-SPECT registration, one based on surfaces and one on mutual information, are compared by Kagadis *et al.* [28]. Zhu [165] shows that mutual information is a special case of cross entropy. Several other cases are deduced (such as conditional entropy) which are suitable registration measures. Combinations of the measures are compared for rigid registration.

To conclude, we present a list of papers that we did not consider true comparison studies, but that do contain comparisons between mutual information and other measures [12], [29], [44], [46], [51]–[53], [64], [69], [89], [91], [93], [95], [102], [103], [107], [127], [128], [130], [152], [156].

V. Discussion

Over the past seven years, a lot of understanding has been gained about mutual information as an image registration measure. It is not an easy measure to understand: the underlying process of how misregistration influences the probability distribution is difficult to envisage. How it influences the relation between joint and marginal distributions is even more mystifying. In contrast, minimizing the distance between corresponding points, for example, is a much easier concept to grasp. However, extensive experimenting, applying, and comparing of the measure has given a good deal of insight into the strengths and weaknesses of mutual information.

From the diversity of modalities and objects found in the literature, it is clear that mutual information lives up to its reputation of being a generally applicable measure. For numerous clinical applications it can be used without need for preprocessing, user initialization or parameter tuning. On the other hand, from the conclusions of certain comparison studies [31], [84], [85] and from the interest in adaptations of the measure [50]–[52], [69], [79], [80] it can be inferred that mutual information may not be a universal cure for all registration problems. For instance, better results with other measures have been reported for registration of serial images which show relatively large changes [31], [85], for extraction of the midsagittal plane of the brain in MR images [152] and for curved registration of MR brain images [130]. Furthermore, it may turn out that mutual information is not the optimal measure for images of thin structures (e.g., retinal images) or for the combination of MR and ultrasound images [156].

What we have learnt from past research is that normalization of mutual information with respect to image overlap is a useful adaptation of the measure. It has been shown by quite a number of different methods that curved registration based on mutual information is viable, although the best way to set about it is yet unclear. We have seen that the choice of interpolation method influences both accuracy and smoothness of the measure. Several options for estimation of the probability distributions have been proposed, while large numbers of optimization routines have been investigated. The question remains, however, how best to implement a mutual-information-based method. That certain options are more promising than others has been shown, but the optimal choice also depends on the interaction between the various aspects of the implementation. For example, a higher order interpolation method will most likely yield a smoother registration function, which reduces the need for a highly complex, yet robust, optimization technique. The best implementation will always be a balance between time constraints and the demands of the application. Naturally, comparing the different implementations proposed is a problem because of the different applications, the different test sets and sometimes also because of a lack of detail described. A huge step forward has been the introduction of the RREP data sets, with which a large number of registration methods has already been compared. However, only the *accuracy* of the participating methods can be studied, as it is unlikely that anyone will submit results that are evidently incorrect by visual inspection. An interesting observation from the RREP study is that the methods by Maes *et al.* [57], Studholme *et al.* [44], Thévenaz and Unser [97] and Viola and Wells [100], although very differently implemented, all yield comparable results with respect to accuracy.

The challenges ahead lie, for example, in the field of curved registration. As far as we know, only two of the curved registration methods reported explicitly include a physical model of deformation. For many applications more than just a regularization term will be required to achieve physically realistic (let alone correct) deformations. Another interesting topic is the registration of three images (or more). This is a problem in subtraction SPECT, for example, where two SPECT images may need to be registered with an anatomical scan. All the papers on registration of three images either assume two of the images are already in register or this is achieved by first registering two images and then the third. How to optimize two different transformations simultaneously and whether there is a single global optimum to this problem is another question. Challenging also is the field of intraoperative registration, including patient position verification in radiotherapy and correction for tissue deformation, which usually requires fast matching to an image of relatively poor quality and also entails deformations. Relatively little research has as yet gone into intersubject registration, as well as certain combinations of modalities. Ultrasound, to name one of the most challenging, poses a serious problem for registration, because of the difference in imaging physics. It is based on tissue *transitions*, which results in a strong dominance of edges in the resulting images. A final example of an area demanding further research is the question how to "correct" the assumption of Shannon entropy that the gray values of neighboring voxels are uncorrelated. In other words, how to include the images' spatial information.

From the continuing interest in the measure it can be deduced that mutual information will not be abandoned in the near future. It is already a successful registration measure for many applications and it can undoubtedly be adapted and extended to aid in many more problems.

Appendix
Hartley Entropy

Hartley wanted a measure that increases linearly with length. Furthermore, he assumed that given messages of length n_1 and n_2 from s_1 and s_2 numbers of symbols, respectively, if $s_1^{n_1} = s_2^{n_2}$, i.e., the number of possible messages is equal, then the amount of information per message is also equal. *Ergo*

$$H = Kn$$
$$s_1^{n_1} = s_2^{n_2}.$$

His deduction of the definition of entropy is as follows:

$$\begin{aligned}
K_1 n_1 &= K_2 n_2 \quad \{n_x = {}^{s_x}\!\log s_x^{n_x}\} \\
K_1 \, {}^{s_1}\!\log s_1^{n_1} &= K_2 \, {}^{s_2}\!\log s_2^{n_2} \quad \{{}^{s_x}\!\log s_x^{n_x} = \log s_x^{n_x} / \log s_x\} \\
K_1 \frac{\log s_1^{n_1}}{\log s_1} &= K_2 \frac{\log s_2^{n_2}}{\log s_2} \quad \{s_1^{n_1} = s_2^{n_2} \rightarrow \log s_1^{n_1} = \log s_2^{n_2}\} \\
K_1 / \log s_1 &= K_2 / \log s_2.
\end{aligned}$$

The final equality holds only when $K_x = c \log s_x$, with c an arbitrary constant that should be equal for all K_x. It can therefore be omitted and $K = \log s$ results.

References

[1] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, Eds., *Medical Image Registration*. Boca Raton, FL: CRC, 2001.

[2] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Phys. Med. Biol.*, vol. 46, no. 3, pp. R1–R45, 2001.

[3] R. V. L. Hartley, "Transmission of information," *Bell Syst. Tech. J.*, vol. 7, pp. 535–563, 1928.

[4] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, p. 379–423/623–656, 1948.

[5] 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, no. 4, pp. 620–633, 1992.

[6] R. P. Woods, J. C. Mazziotta, and S. R. Cherry, "MRI-PET registration with automated algorithm," *J. Comput. Assist. Tomogr.*, vol. 17, no. 4, pp. 536–546, 1993.

[7] D. L. G. Hill, D. J. Hawkes, N. A. Harrison, and C. F. Ruff, "A strategy for automated multimodality image registration incorporating anatomical knowledge and imager characteristics," in *Information Processing in Medical Imaging*, H. H. Barrett and A. F. Gmitro, Eds. Berlin, Germany: Springer-Verlag, 1993, vol. 687, pp. 182–196. *Lecture Notes in Computer Science*.

[8] D. L. G. Hill, C. Studholme, and D. J. Hawkes, "Voxel similarity measures for automated image registration," in *Visualization in Biomedical Computing*, R. A. Robb, Ed. Bellingham, WA: SPIE, 1994, vol. 2359, pp. 205–216. *Proc. SPIE*.

[9] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal, "3D multimodality medical image registration using feature space clustering," in *Computer Vision, Virtual Reality, and Robotics in Medicine*, N. Ayache, Ed. Berlin, Germany: Springer-Verlag, 1995, vol. 905, pp. 195–204. *Lecture Notes in Computer Science*.

[10] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Multiresolution voxel similarity measures for MR-PET registration," in *Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, Eds. Dordrecht, The Netherlands: Kluwer, 1995, pp. 287–298.

[11] A. Collignon, "Multi-modality medical image registration by maximization of mutual information," Ph.D. dissertation, Catholic Univ. Leuven, Leuven, Belgium, 1998.

[12] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," in *Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, Eds. Dordrecht, The Netherlands: Kluwer, 1995, pp. 263–274.

[13] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," in *Proc. Int. Conf. Computer Vision*, E. Grimson, S. Shafer, A. Blake, and K. Sugihara, Eds. Los Alamitos, CA, 1995, pp. 16–23.

[14] P. A. Viola, "Alignment by maximization of mutual information," Ph.D. dissertation, Massachusetts Inst. Technol., Boston, MA, 1995.

[15] W. M. Wells III, P. Viola, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," in *Medical Robotics and Computer Assisted Surgery*. New York: Wiley, 1995, pp. 55–62.

[16] J. Aczél and Z. Daróczy, *On Measures of Information and Their Characterizations*. New York: Academic, 1975.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[18] I. Vajda, *Theory of Statistical Inference and Information*. Dordrecht, The Netherlands: Kluwer, 1989.

[19] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, 1998.

[20] P. A. van den Elsen, E. J. D. Pol, and M. A. Viergever, "Medical image matching—A review with classification," *IEEE Eng. Med. Biol. Mag.*, vol. 12, pp. 26–39, Mar. 1993.

[21] B. Fei, A. Wheaton, Z. Lee, J. L. Duerk, and D. L. Wilson, "Automatic MR volume registration and its evaluation for the pelvis and prostate," *Phys. Med. Biol.*, vol. 47, no. 5, pp. 823–838, 2002.

[22] P. J. Slomka, J. Mandel, D. Downey, and A. Fenster, "Evaluation of voxel-based registration of 3-D power Doppler ultrasound and 3-D magnetic resonance angiographic images of carotid arteries," *Ultrasound Med. Biol.*, vol. 27, no. 7, pp. 945–955, 2001.

[23] A. Carrillo, J. L. Duerk, J. S. Lewin, and D. L. Wilson, "Semiautomatic 3-D image registration as applied to interventional MRI liver cancer treatment," *IEEE Trans. Med. Imag.*, vol. 19, pp. 175–185, Mar. 2000.

[24] E. R. E. Denton, M. Holden, E. Christ, J. M. Jarosz, D. Russell-Jones, J. Goodey, T. C. S. Cox, and D. L. G. Hill, "The identification of cerebral volume changes in treated growth hormone-deficient adults using serial 3D MR image processing," *J. Comput. Assist. Tomogr.*, vol. 24, no. 1, pp. 139–145, 2000.

[25] Y. E. Erdi, K. Rosenzweig, A. K. Erdi, H. A. Macapinlac, Y.-C. Hu, L. E. Braban, J. L. Humm, O. D. Squire, C.-S. Chui, S. M. Larson, and E. D. Yorke, "Radiotherapy treatment planning for patients with non-small cell lung cancer using positron emission tomography (PET)," *Radiotherapy Oncol.*, vol. 62, no. 1, pp. 51–60, 2002.

[26] M. Holden, E. R. E. Denton, J. M. Jarosz, T. C. S. Cox, C. Studholme, D. J. Hawkes, and D. L. G. Hill, "Detecting small anatomical changes with 3D serial MR subtraction images," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 1999, vol. 3661, pp. 44–55. *Proc. SPIE*.

[27] M. Holden, D. L. G. Hill, E. R. E. Denton, J. M. Jarosz, T. C. S. Cox, T. Rohlfing, J. Goodey, and D. J. Hawkes, "Voxel similarity measures for 3-D serial MR brain image registration," *IEEE Trans. Med. Imag.*, vol. 19, pp. 94–102, Feb. 2000.

[28] G. C. Kagadis, K. K. Delibasis, G. K. Matsopoulos, N. A. Mouravliansky, P. A. Asvestas, and G. C. Nikiforidis, "A comparative study of surface- and volume-based techniques for the automatic registration between CT and SPECT brain images," *Med. Phys.*, vol. 29, no. 2, pp. 201–213, 2002.

[29] U. Kjems, S. C. Strother, J. Anderson, I. Law, and L. K. Hansen, "Enhancing the multivariate signal of [^{15}O] water PET studies with a new nonlinear neuroanatomical registration algorithm," *IEEE Trans. Med. Imag.*, vol. 18, pp. 306–319, Apr. 1999.

[30] M. Otte, "Elastic registration of fMRI data using Bézier-spline transformations," *IEEE Trans. Med. Imag.*, vol. 20, pp. 193–206, Mar. 2001.

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

[32] P. E. Radau, P. J. Slomka, P. Julin, L. Svensson, and L.-O. Wahlund, "Evaluation of linear registration algorithms for brain SPECT and the errors due to hypoperfusion lesions," *Med. Phys.*, vol. 28, no. 8, pp. 1660–1668, 2001.

[33] R. Shekhar and V. Zagrodsky, "Mutual information-based rigid and nonrigid registration of ultrasound volumes," *IEEE Trans. Med. Imag.*, vol. 21, pp. 9–22, Jan. 2002.

[34] K. Sjögreen, M. Ljungberg, K. Wingårdh, K. Erlandsson, and S.-E. Strand, "Registration of emission and transmission whole-body scintillation-camera images," *J. Nucl. Med.*, vol. 42, no. 10, pp. 1563–1570, 2001.

[35] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures," *Med. Phys.*, vol. 24, no. 1, pp. 25–35, 1997.

[36] C. Studholme, V. Cardenas, and M. Weiner, "Multi scale image and multi scale deformation of brain anatomy for building average brain atlases," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, Eds. Bellingham, WA: SPIE, 2001, vol. 4322, pp. 557–568. *Proc. SPIE*.

[37] C. Studholme, E. Novotny, I. G. Zubal, and J. S. Duncan, "Estimating tissue deformation between functional images induced by intracranial electrode implantation using anatomical MRI," *NeuroImage*, vol. 13, no. 4, pp. 561–576, 2001.

[38] V. Zagrodsky, R. Shekhar, and J. F. Cornhill, "Multi-function extension of simplex optimization method for mutual information-based registration of ultrasound volumes," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, Eds. Bellingham, WA: SPIE, 2001, vol. 4322, pp. 508–515. *Proc. SPIE*.

[39] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, G. L. Lecarpentier, J. B. Fowlkes, M. A. Roubidoux, and P. L. Carson, "Semiautomatic registration of volumetric ultrasound scans," *Ultrasound Med. Biol.*, vol. 25, no. 3, pp. 339–347, 1999.

[40] V. Zagrodsky, R. Shekhar, and J. F. Cornhill, "Mutual information based registration of cardiac ultrasound volumes," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 2000, vol. 3979, pp. 1605–1614. *Proc. SPIE*.

[41] L. Ding, A. Goshtasby, and M. Satter, "Volume image registration by template matching," *Image Vision Comput.*, vol. 19, no. 12, pp. 821–832, 2001.

[42] M. Holden, J. A. Schnabel, and D. L. G. Hill, "Quantifying small changes in brain ventricular volume using non-rigid registration," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 49–56. *Lecture Notes in Computer Science*.

[43] C. Studholme, "Measures of 3D Medical Image Alignment," Ph.D. dissertation, Univ. London, London, U.K., 1997.

[44] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognit.*, vol. 32, no. 1, pp. 71–86, 1999.

[45] L. Zöllei, E. Grimson, A. Norbash, and W. Wells, "2D–3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators," in *Computer Vision and Pattern Recognition*. Los Alamitos, CA: IEEE Computer Soc., 2001, vol. 2, pp. 696–703.

[46] P. M. Hayton, M. Brady, S. M. Smith, and N. Moore, "A non-rigid registration algorithm for dynamic breast MR images," *Artif. Intell.*, vol. 114, no. 1–2, pp. 125–156, 1999.

[47] D. Plattard, M. Soret, J. Troccaz, P. Vassal, J.-Y. Giraud, G. Champleboux, X. Artignan, and M. Bolla, "Patient set-up using portal images: 2D/2D image registration using mutual information," *Comput. Aided Surgery*, vol. 5, no. 4, pp. 246–262, 2000.

[48] J. A. Lynch, C. G. Peterfy, D. L. White, R. A. Hawkins, and H. K. Genant, "MRI-SPECT image registration using multiple MR pulse sequences to examine osteoarthritis of the knee," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 1999, vol. 3661, pp. 68–77. *Proc. SPIE*.

[49] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated 3-D registration of MR and CT images of the head," *Med. Image Anal.*, vol. 1, no. 2, pp. 163–175, 1996.

[50] T. Butz and J.-P. Thiran, "Affine registration with feature space mutual information," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 549–556. *Lecture Notes in Computer Science*.

[51] L. G. Nyúl, J. K. Udupa, and P. K. Saha, "Task-specific comparison of 3-D image registration methods," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, Eds. Bellingham, WA: SPIE, 2001, vol. 4322, pp. 1588–1598. *Proc. SPIE*.

[52] C. E. Rodríguez-Carranza and M. H. Loew, "A weighted and deterministic entropy measure for image registration using mutual information," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 1998, vol. 3338, pp. 155–166. *Proc. SPIE*.

[53] ——, "Global optimization of weighted mutual information for multimodality image registration," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 1999, vol. 3338, pp. 89–96. *Proc. SPIE*.

[54] G. Jumarie, "A new information theoretic approach to the entropy of non-random discrete maps relation to fractional dimension and temperature of curves," *Chaos, Solitions, Fractals*, vol. 8, no. 6, pp. 953–970, 1997.

[55] A. A. Ioannides, L. C. Liu, J. Kwapien, S. Drozdz, and M. Streit, "Coupling of regional activations in a human brain during an object and face affect recognition task," *Human Brain Mapping*, vol. 11, no. 2, pp. 77–92, 2000.

[56] B. Pompe, P. Blidh, D. Hoyer, and M. Eiselt, "Using mutual information to measure coupling in the cardiorespiratory system," *IEEE Eng. Med. Biol. Mag.*, vol. 17, pp. 32–39, Nov./Dec. 1998.

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

[58] J. M. Blackall, D. Rueckert, C. R. Maurer Jr., G. P. Penney, D. L. G. Hill, and D. J. Hawkes, "An image registration approach to automated calibration for freehand 3D ultrasound," in *Medical Image Computing and Computer-Assisted Intervention*, S. L. Delp, A. M. DiGioia, and B. Jaramaz, Eds. Berlin, Germany: Springer-Verlag, 2000, vol. 1935, pp. 462–471. *Lecture Notes in Computer Science*.

[59] J. Camp and R. Robb, "A novel binning method for improved accuracy and speed of volume image coregistration using normalized mutual information," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 1999, vol. 3661, pp. 24–31. *Proc. SPIE*.

[60] P. Castellanos, P. L. del Angel, and V. Medina, "Deformation of MR images using a local linear transformation," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, Eds. Bellingham, WA: SPIE, 2001, vol. 4322, pp. 909–916. *Proc. SPIE*.

[61] A. D. Castellano-Smith, T. Hartkens, J. Schnabel, D. R. Hose, H. Liu, W. A. Hall, C. L. Truwit, D. J. Hawkes, and D. L. G. Hill, "Constructing patient specific models for correcting intraoperative brain deformation," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 1091–1098. *Lecture Notes in Computer Science*.

[62] E. R. E. Denton, L. I. Sonoda, D. Rueckert, S. C. Rankin, C. Hayes, M. O. Leach, D. L. G. Hill, and D. J. Hawkes, "Comparison and evaluation of rigid, affine, and nonrigid registration of breast MR images," *J. Comput. Assist. Tomogr.*, vol. 23, no. 5, pp. 800–805, 1999.

[63] D. L. G. Hill, C. R. Maurer Jr., R. J. Maciunas, J. A. Barwise, J. M. Fitzpatrick, and M. Y. Wang, "Measurement of intraoperative brain surface deformation under a craniotomy," *Neurosurgery*, vol. 43, no. 3, pp. 514–526, 1998.

[64] Y. H. Lau, M. Braun, and B. F. Hutton, "Non-rigid image registration using a median-filtered coarse-to-fine displacement field and a symmetric correlation ratio," *Phys. Med. Biol.*, vol. 46, no. 4, pp. 1297–1319, 2001.

[65] B. Likar and F. Pernuš, "Registration of serial transverse sections of muscle fibers," *Cytometry*, vol. 37, no. 2, pp. 93–106, 1999.

[66] ——, "A hierarchical approach to elastic registration based on mutual information," *Image Vis. Computing*, vol. 19, no. 1–2, pp. 33–44, 2001.

[67] C. R. Maurer Jr., D. L. G. Hill, A. J. Martin, H. Liu, M. McCue, D. Rueckert, D. Lloret, W. A. Hall, R. E. Maxwell, D. J. Hawkes, and C. L. Truwit, "Investigation of intraoperative brain deformation using a 1.5-T interventional MR system: Preliminary results," *IEEE Trans. Med. Imag.*, vol. 17, pp. 817–825, Oct. 1998.

[68] N. Pagoulatos, D. R. Haynor, and Y. Kim, "Image-based registration of ultrasound and magnetic resonance images: A preliminary study," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 2000, vol. 3979, pp. 156–164. *Proc. SPIE*.

[69] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imag.*, vol. 19, pp. 809–814, Aug. 2000.

[70] ——, "Mutual information matching in multiresolution contexts," *Image Vis. Computing*, vol. 19, no. 1–2, pp. 45–52, 2001.

[71] T. Rohlfing, J. B. West, J. Beier, T. Liebig, C. A. Taschner, and U.-W. Thomale, "Registration of functional and anatomical MRI: Accuracy assessment and application in navigated neurosurgery," *Comput. Aided Surgery*, vol. 5, no. 6, pp. 414–425, 2000.

[72] T. Rohlfing, C. R. Maurer Jr., W. G. O'Dell, and J. Zhong, "Modeling liver motion and deformation during the respiratory cycle using intensity-based free-form registration of gated MR images," in *Medical Imaging: Visualization, Image Display, and Image-Guided Procedures*, S. K. Mun, Ed. Bellingham, WA: SPIE, 2001, vol. 4319, pp. 337–348. *Proc. SPIE*.

[73] T. Rohlfing and C. R. Maurer Jr., "Intensity-based non-rigid registration using adaptive multilevel free-form deformation with an incompressibility constraint," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 111–119. *Lecture Notes in Computer Science*.

[74] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Non-rigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, pp. 712–721, Aug. 1999.

[75] J. A. Schnabel, D. Rueckert, M. Quist, J. M. Blackall, A. D. Castellano-Smith, T. Hartkens, G. P. Penney, W. A. Hall, H. Liu, C. L. Truwit, F. A. Gerritsen, D. L. G. Hill, and D. J. Hawkes, "A generic framework for nonrigid registration based on nonuniform multi-level free-form deformations," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 573–581. *Lecture Notes in Computer Science*.

[76] C. Studholme, R. T. Constable, and J. S. Duncan, "Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model," *IEEE Trans. Med. Imag.*, vol. 19, pp. 1115–1127, Nov. 2000.

[77] C. Tanner, J. A. Schnabel, D. Chung, M. J. Clarkson, D. Rueckert, D. L. G. Hill, and D. J. Hawkes, "Volume and shape preservation of enhancing lesions when applying non-rigid registration to a time series of contrast enhancing MR breast images," in *Medical Image Computing and Computer-Assisted Intervention*, S. L. Delp, A. M. DiGioia, and B. Jaramaz, Eds. Berlin, Germany: Springer-Verlag, 2000, vol. 1935, pp. 327–337. *Lecture Notes in Computer Science*.

[78] M. B. Skouson, Q. Guo, and Z.-P. Liang, "A bound on mutual information for image registration," *IEEE Trans. Med. Imag.*, vol. 20, pp. 843–846, Aug. 2001.

[79] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Incorporating connected region labelling into automated image registration using mutual information," in *Mathematical Methods in Biomedical Image Analysis*, A. A. Amini, F. L. Bookstein, and D. C. Wilson, Eds. Los Alamitos, CA: IEEE Computer Soc. Press, 1996, pp. 23–31.

[80] D. Rueckert, M. J. Clarkson, D. L. G. Hill, and D. J. Hawkes, "Non-rigid registration using higher-order mutual information," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 2000, vol. 3979, pp. 438–447. *Proc. SPIE*.

[81] J. P. W. Pluim, "Mutual information based registration of medical images," Ph.D. dissertation, Utrecht Univ., Utrecht, The Netherlands, 2001.

[82] J. Lötjönen and T. Mäkelä, "Elastic matching using a deformation sphere," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 541–548. *Lecture Notes in Computer Science*.

[83] R. Bansal, L. Staib, Z. Chen, A. Rangarajan, J. Knisely, R. Nath, and J. Duncan, "A minimax entropy registration framework for patient setup verification in radiotherapy," *Comput. Aided Surgery*, vol. 4, no. 6, pp. 287–304, 1999.

[84] L. Barnden, R. Kwiatek, Y. Lau, B. Hutton, L. Thurfjell, K. Pile, and C. Rowe, "Validation of fully automatic brain SPET to MR co-registration," *Eur. J. Nucl. Med.*, vol. 27, no. 2, pp. 147–154, 2000.

[85] B. H. Brinkmann, T. J. O'Brien, S. Aharon, M. K. O'Connor, B. P. Mullan, D. P. Hanson, and R. A. Robb, "Quantitative and clinical analysis of SPECT image registration for epilepsy studies," *J. Nucl. Med.*, vol. 40, no. 7, pp. 1098–1105, 1999.

[86] M. J. Clarkson, D. Rueckert, A. P. King, P. J. Edwards, D. L. G. Hill, and D. J. Hawkes, "Registration of video images to tomographic images by optimising mutual information using texture mapping," in *Medical Image Computing and Computer-Assisted Intervention*, C. Taylor and A. Colchester, Eds. Berlin, Germany: Springer-Verlag, 1999, vol. 1679, pp. 579–588. *Lecture Notes in Computer Science*.

[87] A. A. Flynn, A. J. Green, G. Boxer, R. B. Pedley, and R. H. J. Begent, "A comparison of image registration techniques for the correlation of radiolabelled antibody distribution with tumour morphology," *Phys. Med. Biol.*, vol. 44, no. 7, pp. N151–N159, 1999.

[88] C. Grova, A. Biraben, J.-M. Scarabin, P. Jannin, I. Buvat, H. Benali, and B. Gibaud, "A methodology to validate MRI/SPECT registration methods using realistic simulated SPECT data," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 275–282. *Lecture Notes in Computer Science*.

[89] J. Kim, J. A. Fessler, K. L. Lam, J. M. Balter, and R. K. Ten Haken, "A feasibility study on mutual information based set-up error estimation for radiotherapy," *Med. Phys.*, vol. 28, no. 12, pp. 2507–2517, 2001.

[90] K. F. Koral, S. Lin, J. A. Fessler, M. S. Kaminski, and R. L. Wahl, "Preliminary results from intensity-based CT-SPECT fusion in I-131 anti-B1 monoclonal-antibody therapy of lymphoma," *Cancer*, vol. 80, no. 12, pp. 2538–2544, 1997.

[91] Z. Król, P. Zerfass, B. von Rymon-Lipinski, T. Jansen, W. Hauck, H.-F. Zeilhofer, R. Sader, and E. Keeve, "Computer aided osteotomy design for harvesting autologous bone grafts in reconstructive surgery," in *Medical Imaging: Visualization, Image Display, and Image-Guided Procedures*, S. K. Mun, Ed. Bellingham, WA: SPIE, 2001, vol. 4319, pp. 244–251. *Proc. SPIE*.

[92] F. Maes, "Segmentation and registration of multimodal medical images: From theory, implementation and validation to a useful tool in clinical practice," Ph.D. dissertation, Catholic Univ. Leuven, Leuven, Belgium, 1998.

[93] T. Netsch, P. Rösch, J. Weese, A. van Muiswinkel, and P. Desmedt, "Grey value-based 3-D registration of functional MRI time-series: Comparison of interpolation order and similarity measure," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 2000, vol. 3979, pp. 1148–1159. *Proc. SPIE*.

[94] C. Nikou, F. Heitz, and J.-P. Armspach, "Robust voxel similarity metrics for the registration of dissimilar single and multimodal images," *Pattern Recognit.*, vol. 32, no. 8, pp. 1351–1368, 1999.

[95] S. Sanjay-Gopal, H. P. Chan, T. Wilson, M. Helvie, N. Petrick, and B. Sahiner, "A regional registration technique for automated interval change analysis of breast lesions on mammograms," *Med. Phys.*, vol. 26, no. 12, pp. 2669–2679, 1999.

[96] D. Sarrut and S. Clippe, "Geometrical transformation approximation for 2D/3D intensity-based registration of portal images and CT scan," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 532–540. *Lecture Notes in Computer Science*.

[97] P. Thévenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Trans. Image Processing*, vol. 9, pp. 2083–2099, Dec. 2000.

[98] L. Thurfjell, Y. H. Lau, J. L. R. Andersson, and B. F. Hutton, "Improved efficiency for MRI-SPET registration based on mutual information," *Eur. J. Nucl. Med.*, vol. 27, no. 7, pp. 847–856, 2000.

[99] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," *Int. J. Comput. Vision*, vol. 24, no. 2, pp. 137–154, 1997.

[100] 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, no. 1, pp. 35–51, 1996.

[101] Y.-M. Zhu and S. M. Cochoff, "Influence of implementation parameters on registration of MR and SPECT brain images by maximization of mutual information," *J. Nucl. Med.*, vol. 43, no. 2, pp. 160–166, 2002.

[102] M. Čapek and I. Krekule, "Alignment of adjacent picture frames captured by a CLSM," *IEEE Trans. Inform. Technol. Biomed.*, vol. 3, pp. 119–124, June 1999.

[103] S. Ourselin, A. Roche, S. Prima, and N. Ayache, "Block matching: A general framework to improve robustness of rigid registration of medical images," in *Medical Image Computing and Computer-Assisted Intervention*, S. L. Delp, A. M. DiGioia, and B. Jaramaz, Eds. Berlin, Germany: Springer-Verlag, 2000, vol. 1935, pp. 557–566. *Lecture Notes in Computer Science*.

[104] L. Freire and J.-F. Mangin, "Motion correction algorithms may create spurious brain activations in the absence of subject motion," *NeuroImage*, vol. 14, no. 3, pp. 709–722, 2001.

[105] K. Van Laere, M. Koole, Y. D'Asseler, J. Versijpt, K. Audenaert, F. Dumont, and R. Dierckx, "Automated stereotactic standardization of brain SPECT receptor data using single-photon transmission images," *J. Nucl. Med.*, vol. 42, no. 2, pp. 361–375, 2001.

[106] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, and P. P. van Saarloos, "Registration of stereo and temporal images of the retina," *IEEE Trans. Med. Imag.*, vol. 18, pp. 404–418, May 1999.

[107] M. Jenkinson and S. Smith, "A global optimization method for robust affine registration of brain images," *Med. Image Anal.*, vol. 5, no. 2, pp. 143–156, 2001.

[108] J.-F. Mangin, C. Poupon, C. Clark, D. Le Bihan, and I. Bloch, "Eddy-current distortion correction and robust tensor estimation for MR diffusion imaging," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 186–194. *Lecture Notes in Computer Science*.

[109] 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, no. 2, pp. 317–323, 1998.

[110] M. Holden, "Registration of 3D serial MR brain images," Ph.D. dissertation, London Univ., London, U.K., 2001.

[111] J. L. Boes and C. R. Meyer, "Multi-variate mutual information for registration," in *Medical Image Computing and Computer-Assisted Intervention*, C. Taylor and A. Colchester, Eds. Berlin, Germany: Springer-Verlag, 1999, vol. 1679, pp. 606–612. *Lecture Notes in Computer Science*.

[112] B. Kim, J. L. Boes, K. A. Frey, and C. R. Meyer, "Mutual information for automated unwarping of rat brain autoradiographs," *NeuroImage*, vol. 5, no. 1, pp. 31–40, 1997.

[113] J. F. Krücker, C. R. Meyer, G. L. LeCarpentier, J. B. Fowlkes, and P. L. Carson, "3D spatial compounding of ultrasound images using image-based nonrigid registration," *Ultrasound Med. Biol.*, vol. 26, no. 9, pp. 1475–1488, 2000.

[114] C. Meyer, J. Boes, B. Kim, and P. Bland, "Evaluation of control point selection in automatic, mutual information driven, 3D warping," in *Medical Image Computing and Computer-Assisted Intervention*, W. M. Wells, A. Colchester, and S. Delp, Eds. Berlin, Germany: Springer-Verlag, 1998, vol. 1496, pp. 944–951. *Lecture Notes in Computer Science*.

[115] C. R. Meyer, J. L. Boes, B. Kim, and P. H. Bland, "Probabilistic brain atlas construction: Thin-plate spline warping via maximization of mutual information," in *Medical Image Computing and Computer-Assisted Intervention*, C. Taylor and A. Colchester, Eds. Berlin, Germany: Springer-Verlag, 1999, vol. 1679, pp. 631–637. *Lecture Notes in Computer Science*.

[116] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, K. R. Zasadny, P. V. Kison, K. Koral, K. A. Frey, and R. L. Wahl, "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations," *Med. Image Anal.*, vol. 1, no. 3, pp. 195–206, 1997.

[117] B. Kim, J. L. Boes, P. H. Bland, and C. R. Meyer, "Out-of-plane non-linear warping of a slice into volume," in *Medical Image Computing and Computer-Assisted Intervention*, C. Taylor and A. Colchester, Eds. Berlin, Germany: Springer-Verlag, 1999, vol. 1679, pp. 638–645. *Lecture Notes in Computer Science*.

[118] M. A. Horsfield, "Mapping eddy current induced fields for the correction of diffusion-weighted echo planar images," *Magn. Reson. Imag.*, vol. 17, no. 9, pp. 1335–1345, 1999.

[119] M. Baker, H. Busse, and M. Vogt, "An automatic registration and segmentation algorithm for multiple electrophoresis images," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 2000, vol. 3979, pp. 426–436. *Proc. SPIE*.

[120] T. Gaens, F. Maes, D. Vandermeulen, and P. Suetens, "Non-rigid multimodal image registration using mutual information," in *Medical Image Computing and Computer-Assisted Intervention*, W. M. Wells, A. Colchester, and S. Delp, Eds. Berlin, Germany: Springer-Verlag, 1998, vol. 1496, pp. 1099–1106. *Lecture Notes in Computer Science*.

[121] G. K. Rohde, A. Aldroubi, and B. M. Dawant, "Adaptive free-form deformation for inter-patient medical image registration," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, Eds. Bellingham, WA: SPIE, 2001, vol. 4322, pp. 1578–1587. *Proc. SPIE*.

[122] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "Nonrigid multimodality image registration," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, Eds. Bellingham, WA: SPIE, 2001, vol. 4322, pp. 1609–1620. *Proc. SPIE*.

[123] D. Rueckert, A. F. Frangi, and J. A. Schnabel, "Automatic construction of 3D statistical deformation models using non-rigid registration," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 77–84. *Lecture Notes in Computer Science*.

[124] N. Hata, T. Dohi, S. Warfield, W. Wells III, R. Kikinis, and F. A. Jolesz, "Multimodality deformable registration of pre- and intraoperative images for MRI-guided brain surgery," in *Medical Image Computing and Computer-Assisted Intervention*, W. M. Wells, A. Colchester, and S. Delp, Eds. Berlin, Germany: Springer-Verlag, 1998, vol. 1496, pp. 1067–1074. *Lecture Notes in Computer Science*.

[125] P. Hellier and C. Barillot, "Multimodal non-rigid warping for correction of distortions in functional MRI," in *Medical Image Computing and Computer-Assisted Intervention*, S. L. Delp, A. M. DiGioia, and B. Jaramaz, Eds. Berlin, Germany: Springer-Verlag, 2000, vol. 1935, pp. 512–520. *Lecture Notes in Computer Science*.

[126] J. B. A. Maintz, E. H. W. Meijering, and M. A. Viergever, "General multimodal elastic registration based on mutual information," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE , 1998, vol. 3338, pp. 144–154. *Proc. SPIE*.

[127] G. Hermosillo and O. Faugeras, "Dense image matching with global and local statistical criteria: A variational approach," in *Computer Vision and Pattern Recognition*. Los Alamitos, CA: IEEE Computer Soc. Press, 2001, vol. 1, pp. 73–78.

[128] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Medical Image Computing and Computer-Assisted Intervention*, W. M. Wells, A. Colchester, and S. Delp, Eds. Berlin, Germany: Springer-Verlag, 1998, vol. 1496, pp. 1115–1124. *Lecture Notes in Computer Science*.

[129] J. A. Schnabel, C. Tanner, A. D. Castellano Smith, M. O. Leach, C. Hayes, A. Degenhard, R. Hose, D. L. G. Hill, and D. J. Hawkes, "Validation of nonrigid registration using finite elements methods," in *Information Processing in Medical Imaging*, M. F. Insana and R. M. Leahy, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2082, pp. 344–357. *Lecture Notes in Computer Science*.

[130] A. Guimond, A. Roche, N. Ayache, and J. Meunier, "Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections," *IEEE Trans. Med. Imag.*, vol. 20, pp. 58–69, Jan. 2001.

[131] P. Thévenaz and M. Unser, "A pyramid approach to sub-pixel image fusion based on mutual information," in *Proc. Int. Conf. Image Processing*, 1996, pp. 265–268.

[132] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Interpolation artefacts in mutual information-based image registration," *Comput. Vision Image Understanding*, vol. 77, no. 2, pp. 211–232, 2000.

[133] R. P. Woods, S. T. Grafton, C. J. Holmes, S. R. Cherry, and J. C. Mazziotta, "Automated image registration: I. General methods and intrasubject, intramodality validation," *J. Comput. Assist. Tomogr.*, vol. 22, no. 1, pp. 139–152, 1998.

[134] H. Chen and P. K. Varshney, "A cooperative search algorithm for mutual information based image registration," in *Sensor Fusion: Architectures, Algorithms, and Applications*, B. V. Dasarathy, Ed. Bellingham, WA: SPIE, 2001, vol. 4385, pp. 117–128. *Proc. SPIE*.

[135] B. Kim, J. L. Boes, P. H. Bland, T. L. Chevenert, and C. R. Meyer, "Motion correction in fMRI via registration of individual slices into an anatomical volume," *Magn. Reson. Med.*, vol. 41, no. 5, pp. 964–972, 1999.

[136] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "f-information measures in medical image registration," in Medical Imaging: Image Processing, M. Sonka and K. M. Hanson, Eds. Bellingham, WA: SPIE, 2001, vol. 4322. *Proc. SPIE*.

[137] J. Kim, J. W. Fisher III, A. Tsai, C. Wible, A. S. Willsky, and W. M. Wells III, "Incorporating spatial priors into an information theoretic approach for fMRI data analysis," in *Medical Image Computing and Computer-Assisted Intervention*, S. L. Delp, A. M. DiGioia, and B. Jaramaz, Eds. Berlin, Germany: Springer-Verlag, 2000, vol. 1935, pp. 62–71. *Lecture Notes in Computer Science*.

[138] J. H. Mathews, *Numerical Methods for Mathematics, Science and Engineering*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1992.

[139] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[140] D. R. Holmes III, J. J. Camp, and R. A. Robb, "Evaluation of search strategies and cost functions in optimizing voxel-based image registration," in *Medical Imaging: Image Display*, Y. Kim, Ed. Bellingham, WA: SPIE, 1996, vol. 2707, pp. 554–562. *Proc. SPIE*.

[141] F. Maes, D. Vandermeulen, and P. Suetens, "Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information," *Med. Image Anal.*, vol. 3, no. 4, pp. 373–386, 1999.

[142] 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," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 1999, vol. 3661, pp. 14–23. *Proc. SPIE*.

[143] D. Sarrut and S. Miguet, "Fast 3D image transformations for registration procedures," in *Proc. Int. Conf. Image Analysis and Processing*. Los Alamitos, CA, 1999, pp. 446–452.

[144] X. Meihe, R. Srinivasan, and W. L. Nowinski, "A fast mutual information method for multi-modal registration," in *Information Processing in Medical Imaging*, A. Kuba, M. Šámal, and A. Todd-Pokropek, Eds. Berlin, Germany: Springer-Verlag, 1999, vol. 1613, pp. 466–471. *Lecture Notes in Computer Science*.

[145] A. Rangarajan, H. Chui, and J. S. Duncan, "Rigid point feature registration using mutual information," *Med. Image Anal.*, vol. 3, no. 4, pp. 425–440, 1999.

[146] J. L. R. Andersson and L. Thurfjell, "A multivariate approach to registration of dissimilar tomographic images," *Eur. J. Nucl. Med.*, vol. 26, no. 7, pp. 718–733, 1999.

[147] H. Matsuda, "Physical nature of higher-order mutual information: Intrinsic correlations and frustration," *Phys. Rev. E*, vol. 62, no. 3, pp. 3096–3102, 2000.

[148] T. Tsujishita, "On triple mutual information," *Adv. Appl. Math.*, vol. 16, no. 1, pp. 269–274, 1995.

[149] T. Brückner, R. Lucht, and G. Brix, "Comparison of rigid and elastic matching of dynamic magnetic resonance mammographic images by mutual information," *Med. Phys.*, vol. 27, no. 10, pp. 2456–2461, 2000.

[150] A. Nabavi, P. McL. Black, D. T. Gering, C.-F. Westin, V. Mehta, R. S. Pergolizzi Jr., M. Ferrant, S. K. Warfield, N. Hata, R. B. Schwartz, W. M. Wells III, R. Kikinis, and F. A. Jolesz, "Serial intraoperative magnetic resonance imaging of brain shift," *Neurosurgery*, vol. 48, no. 4, pp. 787–797, 2001.

[151] D. L. G. Hill, C. R. Maurer Jr., A. J. Martin, S. Sabanathan, W. A. Hall, D. J. Hawkes, D. Rueckert, and C. L. Truwit, "Assessment of intraoperative brain deformation using interventional MR imaging," in *Medical Image Computing and Computer-Assisted Intervention*, C. Taylor and A. Colchester, Eds. Berlin, Germany: Springer-Verlag, 1999, vol. 1679, pp. 910–919. *Lecture Notes in Computer Science*.

[152] Y. Liu, R. T. Collins, and W. E. Rothfus, "Robust midsagittal plane extraction from normal and pathological 3-D neuroradiology images," *IEEE Trans. Med. Imag.*, vol. 20, pp. 175–192, Mar. 2001.

[153] K. Martens, K. Verstreken, J. Van Cleynenbreugel, K. Van Brussel, J. Goffin, G. Marchal, and P. Suetens, "Image-based planning and validation of C1-C2 transarticular screw fixation using personalized drill guides," in *Medical Image Computing and Computer-Assisted Intervention*, C. Taylor and A. Colchester, Eds. Berlin, Germany: Springer-Verlag, 1999, vol. 1679, pp. 860–867. *Lecture Notes in Computer Science*.

[154] L. Bergmans, J. Van Cleynenbreugel, M. Wevers, and P. Lambrechts, "A methodology for quantitative evaluation of root canal instrumentation using microcomputed tomography," *Int. Endodontic J.*, vol. 34, no. 5, pp. 390–398, 2001.

[155] J. 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, no. 4, pp. 554–566, 1997.

[156] A. Roche, X. Pennec, G. Malandain, and N. Ayache, "Rigid registration of 3-D ultrasound with MR images: A new approach combining intensity and gradient information," *IEEE Trans. Med. Imag.*, vol. 20, pp. 1038–1049, Oct. 2001.

[157] S. von Klinski, C. Derz, D. Weese, and T. Tolxdorff, "Model-based image processing using snakes and mutual information," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed. Bellingham, WA: SPIE, 2000, vol. 3979, pp. 1053–1064. *Proc. SPIE*.

[158] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based tissue classification of MR images of the brain," *IEEE Trans. Med. Imag.*, vol. 18, pp. 897–908, Oct. 1999.

[159] ——, "Automated model-based bias field correction of MR images of the brain," *IEEE Trans. Med. Imag.*, vol. 18, pp. 885–896, Oct. 1999.

[160] C. Studholme, V. Cardenas, N. Schuff, H. Rosen, B. Miller, and M. Weiner, "Detecting spatially consistent structural differences in Alzheimer's and fronto temporal dementia using deformation morphometry," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 41–48. *Lecture Notes in Computer Science*.

[161] A. Carpentier, K. R. Pugh, M. Westerveld, C. Studholme, O. Skrinjar, J. L. Thompson, D. D. Spencer, and R. T. Constable, "Functional MRI of language processing: Dependence on input modality and temporal lobe epilepsy," *Epilepsia*, vol. 42, no. 10, pp. 1241–1254, 2001.

[162] P. Hellier, C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, L. Collins, A. Evans, G. Malandain, and N. Ayache, "Retrospective evaluation of inter-subject brain registration," in *Medical Image Computing and Computer-Assisted Intervention*, W. J. Niessen and M. A. Viergever, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, pp. 258–265. *Lecture Notes in Computer Science*.

[163] M. Bro-Nielsen, "Rigid registration of CT, MR and cryosection images using a GLCM framework," in *Proceedings of Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery*, J. Troccaz, E. Grimson, and R. Mösges, Eds. Berlin, Germany: Springer-Verlag, 1997, vol. 1205, pp. 171–180. *Lecture Notes in Computer Science*.

[164] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer Jr., R. M. Kessler, and R. J. Maciunas, "Retrospective intermodality registration techniques for images of the head: Surface-based versus volume-based," *IEEE Trans. Med. Imag.*, vol. 18, pp. 144–150, Feb. 1999.

[165] Y.-M. Zhu, "Volume image registration by cross-entropy optimization," *IEEE Trans. Med. Imag.*, vol. 21, pp. 174–180, Feb. 2002.