[On getting theory and practice straight:]
Wagner's music is much better than it sounds.
Mark Twain.

Chapter 4

Comparison of Edge-based and Ridge-based Registration of CT and MR Brain Images.

J.B. Antoine Maintz
Petra A. van den Elsen
Max A. Viergever (1996a),
Medical Image Analysis,
slightly modified

Abstract

In modern medicine, several different imaging techniques are frequently employed in the study of a single patient. This is useful, since different images show complementary information on the functionality and/or structure of the anatomy examined. This very difference between modalities, however, complicates the problem of proper registration of the images involved, and rules out the most basic approaches –like direct grey value correlation– to achieve registration. The observation that some common structures will always exist is supportive of the statement that registration may be feasible using edges or ridges present in the images. The existence of such structures defined in the binary sense is questionable, however, and their extraction from images requires a segmentation by definition. In this paper we propose to use fuzzy edgeness *and* ridgeness *images, thus avoiding the need for segmentation and using more of the available information from the original images. We will show that such fuzzy images can be used to achieve accurate registration. Several ridgeness and edgeness computing operators were compared. The best registration results were obtained using a gradient magnitude operator.*

4.1 Introduction

Medical imaging techniques in common use today show very different aspects of the anatomy examined. For example, CT[1] shows mostly information on dense matter, while MRI[2] shows information on softer tissue types. Both modalities clearly show anatomical morphology, while SPECT[3] and PET[4] show functional aspects of the anatomy. When several imaging modalities are used in a single patient's case, correct *registration*, i.e., determining the transformation to bring one of the acquired images into agreement with the other(s), may facilitate correct diagnosis and/or treatment. Registration often is the first of two steps of an integration process, the second being image *fusion* (integrated or combined display), which mainly concerns the proper visualization of useful image information. In this paper we concentrate on the registration step. In particular, we focus on the registration of CT and MR brain images. This type of registration is useful, for example, in radiation therapy planning, where CT is used for dose calculations, while the lesion to be treated is often best seen on MR images (Chen & Pelizzari 1989). Another example for its use is skull base surgery. The delicacy of this type of surgery requires maximum knowledge of the anatomy involved, which can be supplied by integrating CT bone structures and MR soft tissue contrast images (Ruff, Hill, Robinson & Hawkes 1993).

Medical image registration can be divided into extrinsic registration methods based on artificial marking devices, and intrinsic registration methods using patient related image properties (van den Elsen, Pol & Viergever 1993, Viergever, van den Elsen & Stokking 1992). The method described here falls into the latter category. Registration algorithms using patient related image properties maximize a similarity measure between two images. This similarity may apply directly to the original grey value images (van den Elsen et al. 1994), to statistical voxel similarity measures, to feature images derived from the original images, or to objects defined in the initial or the derived feature images. Maximizing the similarity of the initial images will be useful in particular when two images of the same modality are to be registered. In multi-modality image registration, however, the physical realities of the two images can be quite different, which may call for statistical similarity based (Collignon, Vandermeulen, Suetens & Marchal 1995, Collignon, Maes, Delaere, Vandermeulen, Suetens & Marchal 1995, Studholme et al. 1995*b*, Woods et al. 1993, Hill et al. 1994),

[1] Computed Tomography
[2] Magnetic Resonance Imaging
[3] Single Photon Emission Computed Tomography
[4] Positron Emission Tomography

feature based or object based registration. Features used in image registration are, for example, edges (Borgefors 1988) and ridges (Monga et al. 1992, Guéziec & Ayache 1992, van den Elsen, Maintz & Viergever 1992, van den Elsen 1993, Thirion 1994, Liu et al. 1994). Object based registration may, e.g., be based on surface definitions (Levin et al. 1988, Pelizzari et al. 1989, Hemler, Sumanaweera, van den Elsen, Napel & Adler 1995) or anatomical landmarks (Maguire et al. 1991, Hill, Hawkes, Crossman, Gleeson, Cox, Bracey, Strong & Graves 1991, Lemoine, Barillot, Gibaud & Pasqualini 1991). Object based image registration has the disadvantage that the objects must be defined first, which is a high-level image processing task that often proves difficult, and may introduce errors for complex images. The use of low-level differential geometric features for image registration is attractive, but requires the careful choice of operators that produce sufficiently similar feature images when applied to multi-modal images. Note that such a similarity (within this context) is not always immediately obvious by visually comparing images. The optimal similarity with respect to specific computer vision measures does not necessarily coincide with the 'visual' optimum, as shown in our experiments.

When CT and MR brain images are depicted as intensity landscapes, the skull forms a ridge in the CT image, and a negative ridge (trough) in the MR image. If a 'ridgeness' extracting operator is applied to these images, the resultant feature images show remarkable similarity when compared visually (van den Elsen, Maintz & Viergever 1992). Moreover, since the skull is a virtually undeformable structure, its ridge/trough is ideally suited for registration purposes. Edgeness images of CT and MRI brain scans often have less visual similarity than ridgeness images, because the edge extraction generally produces a number of structures in either image which are not matched by similar structures in the companion image. Enough similarity, however, is present to furnish a good registration, as will be shown. It is the aim of this paper to compare the quality of registration using ridgeness and edgeness operators, combined with grey value cross-correlation of the feature images for registration.

Feature images can be extracted by means of differential operators. Conventional differentiation is ill-posed in the sense of Hadamard since we are dealing with digital, sampled images rather than smooth mathematical functions. Well-posed differentiation is possible, however, by convolving images with derivatives of Gaussians, as is explained in section two. This section also deals with the necessity of the operators being invariant under rigid transformations. In section three we define ridgeness and edgeness measuring differential operators, and include some examples of ridgeness and edgeness images. A cross-correlation based hierarchical registration algorithm is proposed in section four. In section five, the operators are applied to CT/MRI registration algorithm proposed,

and the results are reviewed. In section 6 the results are discussed, and some conclusions are drawn.

4.2 Differentiation of images

4.2.1 Invariants

Image features obtained by means of differential operators should be independent of the choice of coordinate system; the position and orientation of the image should have no impact on the features extracted. Hence invariance under the group of orthogonal transformations (translations, rotations, reflections) is demanded. An operator that conforms to this restriction is called an orthogonal *invariant*.

We will denote by L the image luminance as a function of spatial coordinates. Subscripts denote derivation with respect to some spatial variable. We employ the Einstein summation convention, *i.e.*, the expression is summed over any index occuring twice by letting the particular index assume all spatial dimensions, e.g., $L_i L_i$ denotes $L_x L_x + L_y L_y$ in two-dimensional space. Any tensorial expression in which all indices are resolved by means of contraction (pure or by multiplication by the Kronecker tensor, δ_{ij}) or alternation (multiplication by the Lévi-civita tensor,ε_{ij}) is an invariant (Spivak 1970). All operators presented in this paper satisfy this requirement.

4.2.2 Scale space

The differentiation of any sampled signal (e.g., an image) is ill-posed in the sense of Hadamard as opposed to the generally well-posed differentiation of smooth mathematical functions. Well-posed differentiation is possible by convolving the image with derivatives of a Gaussian (Florack et al. 1992). The width of the Gaussian used introduces a new parameter, the *image scale*, σ, which extends the image dimensionality by one. The thus defined image is usually referred to as the *scale space* of the original image (Witkin 1983, Koenderink 1984). By convolving an image with derivatives of Gaussians, we can compute image derivatives that correspond to the scale of structures. The scale is naturally bounded by the *inner* scale, the finest possible resolution, usually determined

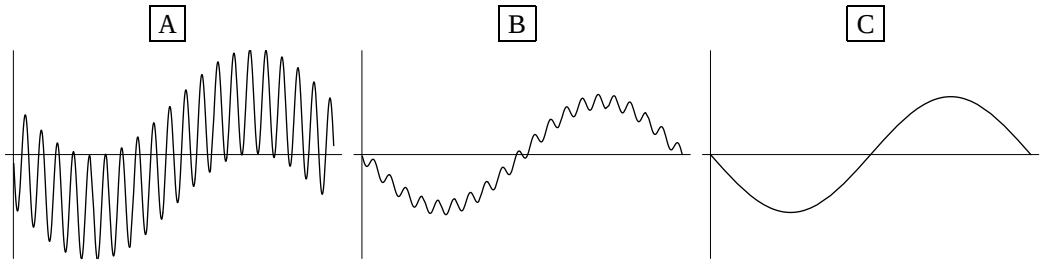


Figure 4.1 A function with distinct low scale and high scale features (A) is shown. Two convolutions of A with Gaussians of increasing width are also shown (B and C). Note that in the final picture only the large scale features of the original function remain.

by the aperture of the scanning device, and the *outer* scale, the largest possible scale, determined by the size of the entire image.

The use of the Gaussian as a convolution kernel is not mandatory. In fact, we can use any smoothing kernel to 'tune' differential operators to multi-local structures (local structures of certain spatial extent). However, upon demanding shift invariance, directional invariance (isotropy) and scale invariance, the Gaussian is the unique linear smoothing kernel (Florack et al. 1992, Florack et al. 1994). The (Gaussian) scale space $L(x, \sigma)$ of an image $L_0(x)$ is the continuous (hyper)stack of smoothed images, with the smoothing factor σ increasing as we rise in the stack. The original image rests at the bottom of the stack. The scale space can be computed using

$$L(x, \sigma) = (L_0 * G)(x, \sigma), \tag{4.1}$$

with

$$G(x, \sigma) = (2\pi\sigma^2)^{-\frac{d}{2}} \; e^{-\frac{x^2}{2\sigma^2}},$$

where G is the Gaussian kernel, x is the coordinate vector, σ is the smoothing factor, i.e., the scale, and d the number of spatial dimensions.

In figure 4.1 the notion of scale is illustrated. We show a function with a bimodal spectrum. ($f(x) = \sin(x) + \sin(20x)$, on a domain of $[0, 2\pi]$.) At a low scale, this function has 20 maxima and 20 minima –precluding boundary extrema– in the domain depicted. At a sufficiently high scale, the function has one maximum and one minimum). The use of the Gaussian (or, in fact, any smoothing kernel) for extracting these high scale extrema or to act as a low-pass filter is widely known and used. Scaling *and* differentiation however, are fused into a single operation within scale space: computing a derivative of L is equivalent to replacing

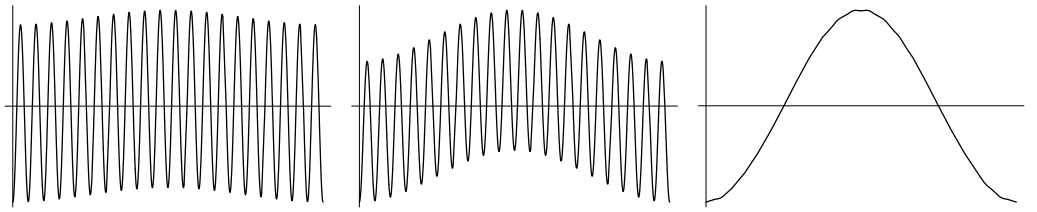


Figure 4.2 Three scaled first derivatives with increasing scale of the function in the previous figure. Note that in the final picture only the large scale features of the original function determine the resultant derivative. (Because the amplitude of the images varies largely, the images have been rescaled for display purposes.)

the Gaussian G with its appropriate derivative in the convolution operation:

$$(L_{i_1 \ldots i_n})(x,\sigma) = (L_0 * G_{i_1 \ldots i_n})(x,\sigma), \tag{4.2}$$

where subscripts i_j denote the order of derivation with respect to the spatial variables; $i_j \in \{\mathbf{x},\mathbf{y},\mathbf{z}\}$, $n \in \mathbb{N}^+$, $j = 1 \ldots n$. This property is easily verified: suppose $f = f(x)$, where f represents a sampled signal on a bounded domain (say, $f(x) = \text{Ш}(x)f'(x)$, Ш being the Shah or replicating symbol, and f' the actual signal) with the Fourier transform defined in the limiting case, and suppose $g = g(x)$, where g is some well-behaved smooth kernel, and consider $\frac{d}{dx}(f * g)$. In theory, this last expression equals $\frac{df}{dx} * g$ as well as $f * \frac{dg}{dx}$, since the Fourier transform of all three equals $i2\pi s\mathcal{F}(s)\mathcal{G}(s)$,where calligraphic letters $(\mathcal{F},\mathcal{G})$ denote Fourier transforms. The latter form is the only one well-posed, however. Figure 4.2 illustrates scaled differentiation.

The numerical complexity of the computation of a differential image at a certain scale is reduced to mere multiplications, when all computations are done in the frequency domain:

$$(L_{i_1 \ldots i_n})(x,\sigma) = (\mathcal{F}^{-1}[\mathcal{L}_0 \cdot \mathcal{G} \cdot i^n \cdot \prod_j \omega_{i_j}])(x,\sigma), \tag{4.3}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform, and ω_{i_j} represent the spatial frequencies.

4.3 Feature measures

4.3.1 Ridgeness measures

For the differential-geometrical detection of ridge-like structures, many different schemes and mathematical ridge definitions have been proposed, some dating back for well over a century (Maxwell 1859). Koenderink showed that for some popular definitions based on the water drainage pattern of a landscape, a local ridge detector does not exist (Koenderink & van Doorn 1994). However, there are a number of geometrical invariants that approximate ridges well in a wide variety of images (Eberly et al. 1994). We have selected the so-called L_{vv} and closely related operators (van den Elsen, Maintz, Pol & Viergever 1992, van den Elsen, Maintz & Viergever 1993). A comparison with other ridge operators can be found in (Maintz, van den Elsen & Viergever 1996*b*), which also illustrates known shortcomings of this local operator.

$\boldsymbol{L_{vv}}$: The ridgeness operators used in this paper are derived from the L_{vv} operator. In this formula v is defined in a local gradient based coordinate system (v, w): $w_i = L_i$, and $v_i = \varepsilon_{ij} L_j$ in tensor notation. Therefore L_{vv} represents the second order derivative in the direction perpendicular to the local gradient. The value of L_{vv} can be computed using (Cartesian) local derivatives:

$$L_{vv} = \frac{1}{\|v\|^2}(v \cdot \nabla)^2 L = (L_y^2 L_{xx} - 2L_x L_y L_{xy} + L_x^2 L_{yy})(L_x^2 + L_y^2)^{-1}. \qquad (4.4)$$

The generalization to 3D can be found in (Maintz, van den Elsen & Viergever 1996*b*). This generalization is non-trivial, since in 3D the v direction as being perpendicular to the local gradient, needs another constraint to be properly defined.

$\boldsymbol{L_{vv}/L_w}$: L_{vv}/L_w Can also be considered a ridgeness measuring operator. This formula derives from the observation that in two dimensional images, the local gradient changes direction when crossing a ridge. Consequently, an alternative definition of ridgeness is the rate by which the gradient direction changes when moving along the v direction. Let the two-dimensional gradient orientation be denoted by $\theta = \arctan(\frac{L_y}{L_x})$. The new ridge measure then is

$$\frac{\partial \theta}{\partial v} = \frac{1}{\|v\|}(v \cdot \nabla)\theta = \frac{1}{\|v\|}(L_y \frac{\partial \theta}{\partial x} - L_x \frac{\partial \theta}{\partial y}) = \frac{2L_x L_y L_{xy} - L_y^2 L_{xx} - L_x^2 L_{yy}}{(L_x^2 + L_y^2)^{\frac{3}{2}}}, \qquad (4.5)$$

Notice that $\frac{\partial \theta}{\partial v} \equiv -\frac{L_{vv}}{L_w}$, so in fact the only difference with L_{vv} is a negation and a normalization with respect to the gradient magnitude. $-\frac{L_{vv}}{L_w}$ Often appears in

literature as the *isophote curvature*, frequently denoted by κ. The normalization of L_{vv} with respect to the gradient magnitude ($\frac{L_{vv}}{L_w}$) causes it to react stronger than L_{vv} in areas of the image where the variation in image intensity is relatively small, i.e., relatively flat areas in the intensity landscape.

$\boldsymbol{L_{vv}L_w^{\alpha}}$: Both L_{vv} and $L_{vv}L_w^{-1}$ have been identified as ridgeness measures. The notion of L_{vv} and L_{vv}/L_w as ridgeness measures can be readily expanded towards the more general formula $L_{vv}L_w^{\alpha}$, where α is bounded. The ridgeness operators used in this paper can be written in this form, with $\alpha \in \{-1, -0.5, 0\}$, although other values for α could be considered as well.

4.3.2 Edgeness measures

Well known edge measures are the gradient magnitude (L_w), and the Laplacian (L_{ii}). L_w measures the local 'steepness' of the intensity landscape, which presumably has a local maximum at an edge. It is a good detector for step edges. When edges get less steep, use of the Laplacian often gives better results. For example, in the 1D case of an edge, the Laplacian has a positive response in the convex part of the edge flank, and a negative response in the concave part. The edge locus is presumed to be at the zero crossing between these two parts.

4.3.3 Miscellaneous measures

Besides edge and ridge measures, we employed some other invariant measures, which are briefly mentioned here. Cartesian expressions are given in their 2D form.

$\boldsymbol{L_{vv}L_w^2}$: $L_{vv}L_w^2$ is a cornerness measure based on the work of Kitchen and Rosenfeld (Kitchen & Rosenfeld 1982) and Blom (Blom 1992). Note that it equals the (negated) isophote curvature times the gradient magnitude cubed: $L_{vv}L_w^2 = \frac{L_{vv}}{L_w}L_w^3$. The idea behind this detector is that the isophote curvature is extremely high at corners. As the isophote curvature reacts equally strong at 'background' structures and 'real' objects, and has a response at ridges, it is multiplied by a power of the gradient magnitude, to avoid phantom responses. This also solves the faulty detection of corners at local extrema (where $L_w = 0$). In the strictest sense, $L_{vv}L_w^2$ is categorized in our class of ridge detectors $L_{vv}L_w^{\alpha}$. However, valid values for α are bounded, and —although the exact bounds are subjective– 'cornerness measure' is a better description of $L_{vv}L_w^2$ than 'ridgeness

measure': the strong response at corners relatively suppresses the response at ridges. Some ridgeness information can be extracted from $L_{vv}L_w^2$ images by windowed display, or by enhancing the response at ridges by re-mapping the operator as $(L_{vv}L_w^2)^{(1/n)}$, where $n > 1$. Experiments have shown $n = 3$ to be a satisfying choice. In Cartesian notation, $L_{vv}L_w^2$ equals the numerator of the Cartesian expression of L_{vv}: $L_{vv}L_w^2 = L_y^2L_{xx} - 2L_xL_yL_{xy} + L_x^2L_{yy}$.

$\boldsymbol{L_{vw}}$: As $-\frac{L_{vv}}{L_w}$ equals the isophote curvature, $\frac{L_{vw}}{L_w}$ equals the flowline curvature, where the flowline is defined as the integral curve of the gradient. In each point of an image, the local flowline and isophote are perpendicular by definition. We employ L_{vw}, obtained by multiplying the flowline curvature with the gradient magnitude, thus diminishing its response in uninteresting areas. The Cartesian expression of L_{vw} equals $(L_xL_y(L_{yy} - L_{xx}) + L_{xy}(L_x^2 - L_y^2))/(L_x^2 + L_y^2)$.

$\boldsymbol{L_{ww}}$: $\frac{L_{ww}}{L_w}$ is a measure for isophote density. As with the previous expressions, we use the expression L_{ww} to reduce uninteresting responses. L_{ww} is closely connected to the Laplacian and L_{vv} ridgeness by the relation $L_{ww} = L_{ii} - L_{vv}$. The Cartesian expression of L_{ww} equals $(L_{xx}L_x^2 + 2L_{xy}L_xL_y + L_{yy}L_y^2)/(L_x^2 + L_y^2))$.

umbilicity, $\boldsymbol{L_{ij}L_{ji}}$: The umbilicity of a point can be determined by computing $\varepsilon_{ij}\varepsilon_{kl}\frac{L_{ik}L_{jl}}{L_{mn}L_{nm}} = \frac{L_{ii}L_{jj} - L_{ij}L_{ji}}{L_{kl}L_{lk}}$. The numerator equals twice the determinant of the Hessian $\det L_{ij}$, which is a measure for local ellipticity (positive value) or hyperbolicity (negative value) of a surface patch. A zero value of this determinant indicates a parabolic or planar patch. The denominator normalizes the umbilicity measure so as to be bounded by -1 and 1, which is most obvious when examining the form containing the ε tensors. The denominator can therefore be regarded as an 'unflatness'-measure. The Cartesian expression of umbilicity equals $2(L_{xx}L_{yy} - L_{xy}^2)/(L_{xx}^2 + 2L_{xy}^2 + L_{yy}^2)$.

checkerboard detector, Y-junction detector: The third and fourth order binary forms L_3 and L_4 of an image L at coordinates $\vec{x}$ are $\frac{1}{3!}L_{ijk}x_ix_jx_k$ and $\frac{1}{4!}L_{ijkl}x_ix_jx_kx_l$ respectively. The discriminants D_3 and D_4 of these forms can serve to detect Y-junctions and 'checkerboard' patterns respectively. The reader interested in a theoretical expansion of these expressions is referred to (Florack, ter Haar Romeny, Koenderink & Viergever 1993, ter Haar Romeny, Florack, Koenderink & Viergever 1991). The Cartesian expressions are: $D_3 = 6L_{xyy}^2L_{xxy}^2 - 2L_{xxx}^2L_{yyy}^2 - 8L_{yyy}L_{xxy}^3 - 8L_{xyy}^3L_{xxx} + 12L_{yyy}L_{xyy}L_{xxy}L_{xxx}$ and $D_4 = (L_{xxxx}L_{yyyy} - 4Lxxxy L_{xyyy} + 3L_{xxyy}L_{xxyy})^3 - 27(L_{xxxy}(L_{xxxy}L_{yyyy} - L_{xyyy}L_{xxyy}) + L_{xxyy}(L_{xxxy}L_{xyyy} - L_{xxyy}L_{xxyy}))^2$.

4.4 Registration method

The purpose of this paper is to register CT and MR brain images by means of geometrical image features. After having determined feature images as discussed in section 3 , we need to register the feature volumes or slices. The method we use to this end is cross-correlation of grey values. By using the grey values directly we avoid segmentation of our feature images. We seek the global optimum of the correlation value $c(t)$ of the CT feature volume L_1 and the MR feature volume L_2 over all rigid transformations t, where $c(t)$ is defined

$$c(t) = \sum_{(x,y,z)\in L_1} L_1(x,y,z)L_2(t(x,y,z)).$$

A brute force approach in which all possible values of t are investigated is computationally infeasible. We therefore resort to a multi-resolution method and a number of assumptions on the behavior of $c(t)$ to find the optimal t within an acceptable number of computational operations. A multi-resolution pyramid is created, with the original image at its bottom. The next level is formed by maximizing or minimizing each group of up to eight neighboring voxels into one voxel. The choice between maximizing and minimizing is based on the sign of the relevant response. Some operators change sign at interesting voxels, e.g., L_{ii}, and here an often arbitrary choice between maximizing and minimizing has to be made. New pyramid levels are formed as long as the largest image structures are clearly discernible. Typically, the pyramid will consist of four levels. Between the (very low resolution) top levels of the feature pyramids the optimal registration is found by optimizing the correlation value using an exhaustive search. Local extrema within a certain percentage of the strongest extremum found are passed on as search seeds to the next pyramid level, where new searches are started. As we progress further down the pyramid, the absolute search range of t (i.e., the actual range in parameter space in terms of millimeters and degrees around a certain origin) diminishes, as do the step sizes. By keeping the search range small, the number of values for t to test remain computationally feasible. To avoid the risk of a search range around a seed being too small, and a correlation optimum being missed, a hill-climbing operation may be performed after the search for a local optimum. Details on the procedure are furnished in (van den Elsen, Maintz & Viergever 1993, van den Elsen 1993). These references also explain in more detail the advantages of the correlation method over e.g. surface based methods (Levin et al. 1988, Pelizzari et al. 1989), which are computationally more attractive. These advantages include the absence of a segmentation step and decreased sensitivity to differences in the structures used for registration.

4.5 Application of feature measures to CT/MRI registration

4.5.1 Similarity of feature volumes

The edge (L_w and L_{ii}) and ridge operators (L_{vv}, $L_{vv}/\sqrt{(L_w)}$, and L_{vv}/L_w) all have a high degree of similarity when comparing CT and MR feature images of the brain (see figure 4.3). Of the ridge operators, (when grey values are scaled equally) L_{vv} looks most appealing because it shows less background response than the other operators. All ridge operators show a clearly visible ridge at the center of the skull. The L_w operator shows a high degree of similarity, notably the skin and skull edges. L_{ii} also shows this similarity, however, the CT pattern at the skin edge matches that of the MRI, whereas the pattern at the skull edge is reversed. Since both positive and negative responses are interesting (and occur coupled at edges), the building of the resolution pyramids can be done arbitrarily by maximizing or minimizing, as long as both pyramids are constructed the same way, and either feature image is reversed beforehand. We can then optimize the value of the correlation $c(t)$, which is maximal if the structures at the bone edge are properly aligned. It must be noted that the simultaneously occuring alignment of the skin edge has the adverse effect of lowering the correlation value. Fortunately, we have not seen this to disrupt finding the correct registration.

L_{ww} has the same ambiguity of patterns occuring both matched and reversed. It shows both edge and –at higher scales– ridge structures, and also looks promising. $L_{ij}L_{ji}$ is hard to assess visually, as it produces 'double-edge'-like structures which tend to interfere with each other. L_{vw} also poses difficulties. There is a definite similarity to the human eye, but its rapid sign changes at ridge structures makes the matching behavior hard to predict. The checkerboard, Y-junction, $L_{vv}L_w^2$ (corner), and umbilicity images show very little similarity to the human eye.

Regardless of the above observations on similarity, all operators were tested in the registration scheme described in section 4. We emphasize that we can only draw conclusions concerning the registration quality of invariant operators for our particular use of cross-correlation registration, not for integration regardless of the registration paradigm.

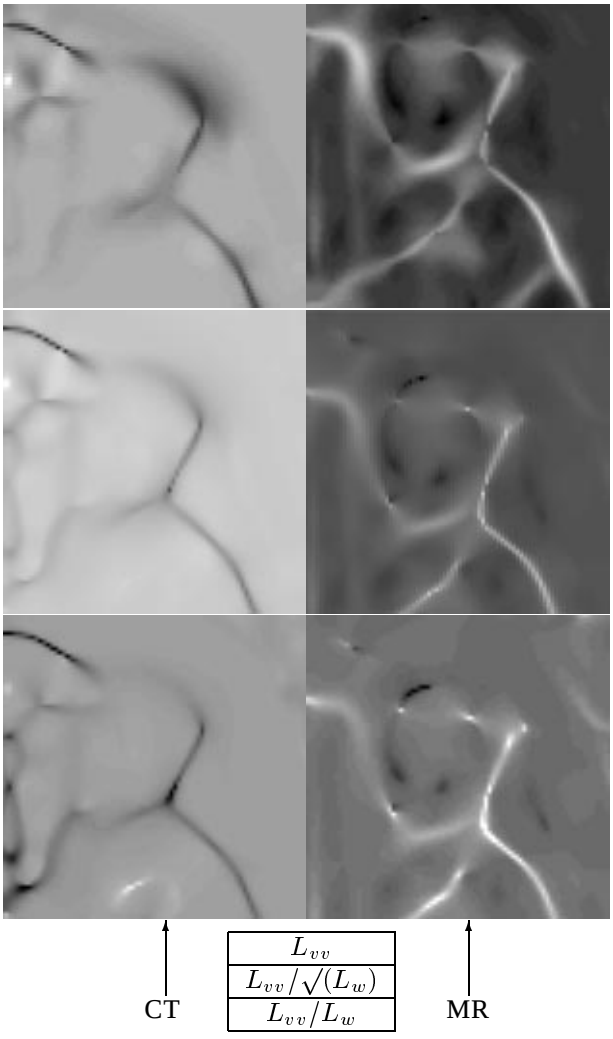


Figure 4.3 Visual comparison of the edge and ridge feature operators applied to an MR image and a registered CT image. Various scales were employed. The images show a zoom of the feature images around the right eye socket. The actual area depicted may differ between images to allow the inclusion of some interesting features. Of some images the grey values have been linearly re-mapped for display purposes.

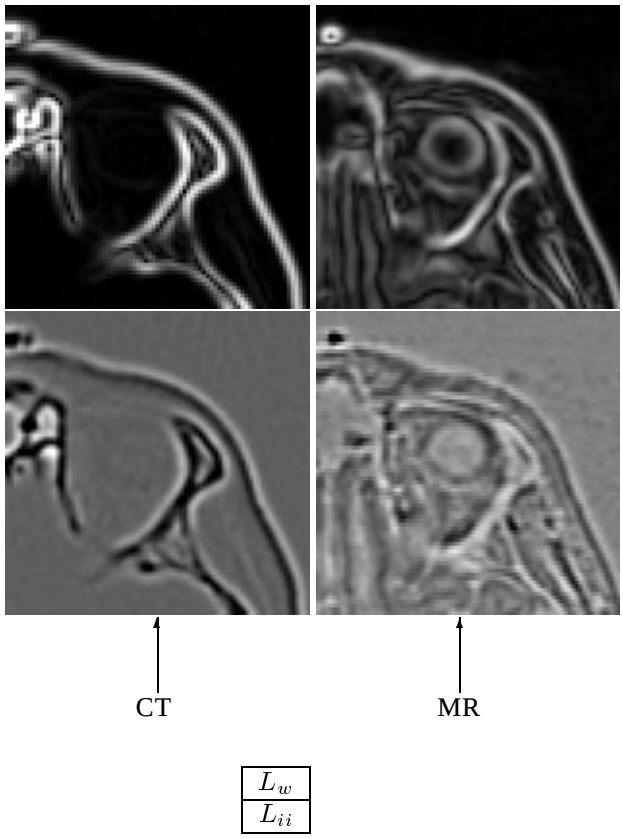


Figure 4.3 (Continued.)

4.5.2 2D registration experiments

To test the registration performance of all of the operators described, five representative registered pairs of CT and MRI transverse slices were chosen. These slices were chosen from volumes registered using skin markers (van den Elsen & Viergever 1994).

Each of the 5 operators was tested by registering the 5 slice pairs, at scales ranging from 1 to 7 mm, using 4 different artificially induced initial transformations

	initial transformation		initial maiming	
#	translation (mm)	rotation (deg)	#	power loss
1	(0,0)	0	1	0%
2	(10,10)	0	2	15.6%
3	(0,0)	15	3	55.9%
4	(15,15)	15	4	64.2%

Table 4.1 The initial transformations and initial 'maimings' of the images used in the 2D registration experiments. The last column shows the percentage of image power loss when applying the maiming, to give an indication of the severity of the maiming.

of one of the images of the registration pair, ranging from no transformation to 15 mm translations and 15 degrees rotation (see table 4.1).

After artificial transformation of the CT image of a pair, the feature operator in question was applied to both the images to be registered, and the resultant feature images were registered. Since some operators can suffer from boundary artifacts, additional experiments were performed to ensure these had no influence on the registration outcome. Ideally, the registration transformation found equals the inverse of the applied initial transformation, since the images were taken from sets registered beforehand. We must be careful to accept this transformation as a gold standard however, since the original registration will inevitably have a (small) error. Moreover, the original registration was based on 3D information, while we now register using the 2D slice information only. To make sure the original registration is an acceptable standard, it was re-assessed visually within the 2D slice, and compared visually to typical transformations found in the experiments. The original registration was visually found to be an excellent standard for all experiments.

Based on this run of experiments, the checkerboard, Y-junction, umbilicity, L_{vw}, L_{ww}, and $L_{ij}L_{ji}$ operators were discarded: in none or only very few cases a correlation optimum was found at the ideal matching transformation. The registration results based on the edge and ridge operators were very good[5] in almost all of the experiments. The L_{vv}/L_w was not used in further experiments, however, since there often were many local registration optima near the value of the optimum obtained at the correct registration transformation. Since all of these optima will have to be investigated by the registration program, the runtime increases dramatically. Because the $L_{vv}/\sqrt{(L_w)}$ and L_{vv} operator had no such problems, the L_{vv}/L_w operator was discarded.

[5] 'Very good' meaning here that shifting the registered images by a single pixel (or rotating the image so that there was a single pixel shift at the edge of the region of interest) worsened the registration when assessed visually.

deviation	index
$0 \leq . \leq 1$	1
$1 < . \leq 2$	0.75
$2 < . \leq 5$	0.5
$5 < . \leq 10$	0.25
$10 < .$	0
indeterminable	0

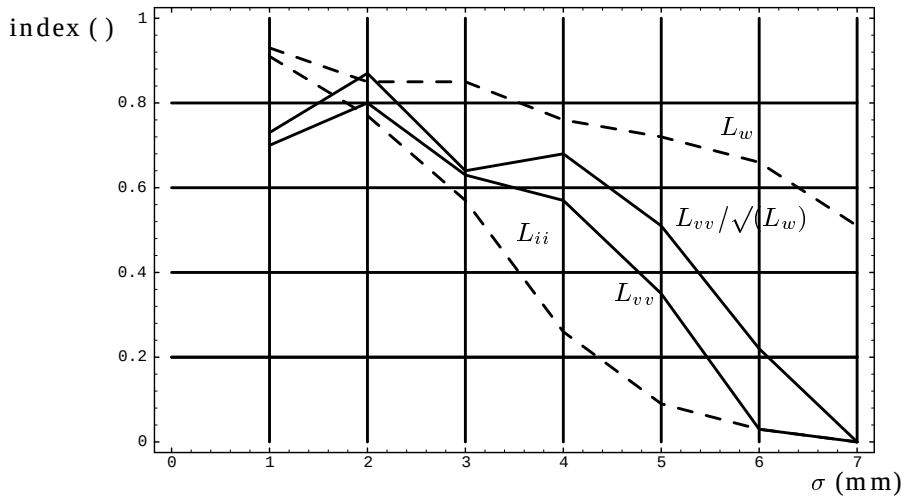
Table 4.2 Index values assigned to registations with a certain deviation from the reference match.

The remaining operators (L_{vv}, $L_{vv}/\sqrt{(L_w)}$, L_{ii}, and L_w) were tested further in three experiment runs, comprising 1680 experiments. Each run is similar to the initial run of experiments, except we now 'maim' the CT image before the start of the experiments (by cutting away specific parts of it), making correct registration harder, because there is less information available to base it on. The particular image loss induced by each 'maiming' is shown in table 4.1. To interpret the large amount of experimental results, each registration result was categorized as belonging to one of six classes of increasing registration accuracy. These classes were subsequently indexed, where the index ranged from '0' for the class of failed registrations (more than 10 pixels or degrees deviation, or no registration transformation found at all), to '1' for the class of registrations with less than 1 pixel or degree deviation from the original marker registration, as shown in table 4.2.

For each operator, the average experimental index was computed. Note that specific values of this index have no quantitative interpretation (in a sense like '0.5' means 50% of the registrations were registered adequately). Within one slice pair it can be used to compare operators though, since a higher index can be interpreted as a better registration performance. The overall performance index for each operator as a function of scale is shown in figure 4.4.

The best performance index and the scale at which this index occured is given in the table included in figure 4.4.

Evidently, L_w has the best overall performance. The edgeness detectors (L_w, L_{ii}) perform slightly better than the ridgeness detectors on their respective optimal scales. L_w is the least sensitive to changes in scale, as the graph shows. The maximum of the edge operators is a boundary maximum, which suggests that the actual maximum may occur at a scale lower than 1 mm. As computing derivatives at such low scales is infeasible on most image volumes since these scales are undersampled, scales lower than 1 mm cannot be properly examined.



operator	best index	scale
L_w	0.93	1
L_{ii}	0.91	1
$L_{vv}/\sqrt{(L_w)}$	0.87	2
L_{vv}	0.80	2

Figure 4.4 Performance index of the four feature operators as a function of scale (top), and the value of the best index that occured for each operator (bottom). In the graph, edge operators are represented by dashed lines, and ridge operators as solid lines.

4.5.3 3D registration experiments

The 3D experiments comprised correlation registrations of the L_{vv}, $L_{vv}/\sqrt{(L_w)}$, L_{ii}, and L_w feature images of high resolution 3D CT and MR brain images[6]. The optimum operator scale was derived from the 2D experiments. The registration transformation found was compared to a previously established registration based on marker methods (van den Elsen & Viergever 1994). Additionally, the same experiments were carried out on volumes of lower transverse resolution. These volumes were generated from the high resolution sets by averaging each 2 (CT) or 3 (MR) slices into one slice, thus simulating 3 mm slices. We use simulated sets instead of directly acquired sets, because an accurate registration reference standard is needed, which is difficult to check visually for low-resolution

[6]The MR data set is a transverse T1 weighted 3D/FFE set, TR=30 msec, TE=9 msec, containing 180 slices, no gap, with cubic voxels of approx. $1mm^3$, obtained on a 1.5 Tesla Philips Gyroscan S15/ACS. The CT data set is a contiguous 100 slice set, pixel size approx. 0.9 mm, slice thickness 1.5 mm, obtained on a Philips Tomoscan 350, set to 120 kV and 120 mA.

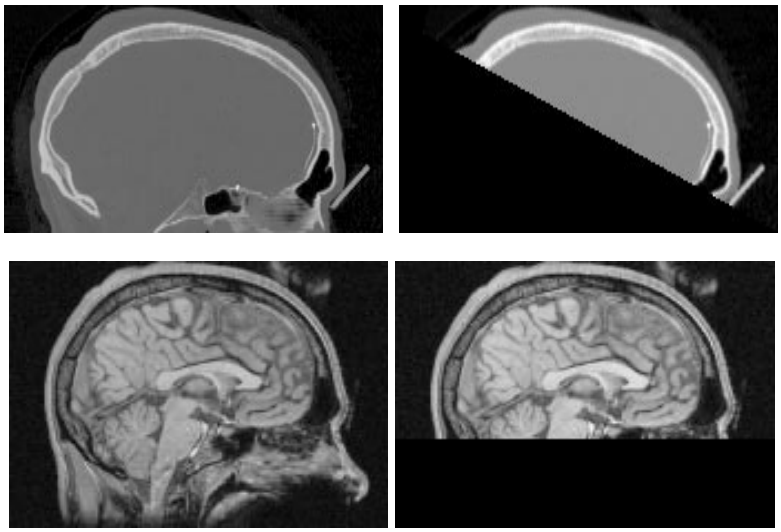


Figure 4.5 The midsagittal planes from the volumes used in the 3D registration experiments (top CT, bottom MRI). The right frames show the corresponding maimed images.

sets. Finally, we repeated all of the above described experiments on maimed sets. Both the CT and MR volumes were maimed, as indicated in figure 4.5. The power loss is 27.1% in the case of the MR volume, and 56.9% in case of the CT volume.

The results are summarized in the tables 4.3 and 4.4.

We increased the scale of the edge operator when using the low resolution sets, to maintain an appropriate scale relative to the slice thickness. If the scale were not increased, the operator is asked to supply information on an under-sampled level of detail. Table 4.3 shows the results of the 3D registration experiments.

The resultant registration transformation furthest from the reference registration was selected from the high resolution experiments (the L_{vv} registration), and subsequently applied to the original datasets. The thus found registration was compared visually to the original reference registration. This comparison was done by segmenting the bone contours –by grey value thresholding– from various transversal and sagitally and coronally reformatted slices from the registered CT volume, and overlaying these contours onto the corresponding MR slices (Maintz, van den Elsen & Viergever 1996*b*, van den Elsen, Maintz & Viergever 1992, van den Elsen, Maintz & Viergever 1993). Two independent observers concluded that the reference registration is inferior to the newly found

operator	σ	high res	maim	result						miss	#
				translation			rotation				
				x	y	z	x	y	z		
L_w	1	•		-0.02	0.04	0.44	0.45	-0.62	2.04		1
L_w	2			0.22	-0.20	-0.39	0.68	-0.17	1.37		1
L_w	1	•	•	0.22	0.98	0.44	0.00	0.28	1.37		6
L_w	2		•	0.45	0.27	-3.61	-0.22	1.17	2.04		0
L_{ii}	1	•		-0.02	0.04	1.22	0.45	-0.39	2.04		11
L_{ii}	2			0.22	0.04	1.16	0.68	0.05	2.94		0
L_{ii}	1	•	•	0.22	0.74	0.84	0.00	0.05	2.04		15
L_{ii}	2		•	0.45	0.98	-1.19	-0.44	0.05	2.04		1
L_{vv}	2	•		-0.25	0.04	1.60	0.45	-0.39	2.27		0
L_{vv}	2			-0.25	0.04	1.20	0.45	-0.17	2.04		0
L_{vv}	2	•	•	-0.25	0.04	1.60	0.45	-0.39	2.27		0
L_{vv}	2		•	∞	∞	∞	∞	∞	∞	•	(97)
$L_{vv}/\sqrt{(L_w)}$	2	•		-0.48	0.98	1.60	0.68	-0.17	2.04		0
$L_{vv}/\sqrt{(L_w)}$	2			0.45	-0.66	-1.20	1.35	0.73	4.28		1
$L_{vv}/\sqrt{(L_w)}$	2	•	•	0.22	0.98	0.83	0.23	0.05	1.82		4
$L_{vv}/\sqrt{(L_w)}$	2		•	0.22	1.21	-4.34	-0.44	0.50	0.92		16

Table 4.3 The results of the 3D registration experiments. The 'result' column shows the difference between the transformation found by feature registration and the reference transformation obtained by marker registration. The figures show the translation and rotation of the center CT volume voxel. All values mentioned are in millimeters or degrees. The obvious mismatches have a mark in the 'miss' column. The last column shows the number of extra local minima examined. If this number is bracketed, convergence of the particular correlation was too slow, and the number of minima to be examined was reduced manually.

	$\frac{L_{vv}}{\sqrt{L_w}}$	L_{vv}	L_{ii}	L_w
marker	5.27	6.00	4.21	3.97
L_w	2.29	2.33	1.18	
L_{ii}	1.46	1.89		
L_{vv}	2.66			

	$\frac{L_{vv}}{\sqrt{L_w}}$	L_{vv}	L_{ii}
L_w	6.24	2.23	3.57
L_{ii}	5.05	2.13	
L_{vv}	6.74		

Table 4.4 The maximal distances (in mm) between two corresponding voxels when transformed according to the different registration parameters found in the experiments. The distance was computed considering all voxels within a diameter of 20cm from the image center. The found value is an upper bound for the actual distance between any 2 corresponding voxels in the images. The left table shows the maximal distances between corresponding voxels for each of the edge or ridge based registrations and the marker registration. For example the '1.18' in the L_{ii} column and L_w row means that if you take a voxel from the set as transformed according to the parameters obtained from the L_{ii} based registration, it is never more than $1.18mm$ apart from the set as transformed according to the L_w based registration parameters. The right table shows the distances between the low-resolution registrations.

L_{vv} correlation registration. (This suggests that feature based registration is better than marker based registration. However, this mere observation of a few observers does not justify a definitive conclusion as regards the relative accuracy of feature based and marker based registration.) When visually comparing all high resolution/no maiming registrations in the table in the same manner, no conclusion could be drawn regarding the best registration. We may conclude that –even though owing to the inevitable intrinsic distortion of the MR image, *the* perfect rigid registration does not exist– these four registrations are clearly good approximations, and seem to be equally accurate. This is supported by table 4.4, which shows the same voxels from the high resolution sets to be never more than $2.33mm$ apart.

Table 4.3 shows that for both the low resolution sets and the maimed sets, the performance of all operators is still accurate. In the table, the last column shows the number of additional local minima examined (i.e., besides the absolute minimum found on each pyramid level) by the algorithm. This number is bracketed when the algorithm showed no or too slow convergence, e.g., too much incorrect minima remained to be examined on the high resolution pyramid levels. In this case, after the algorithm had run for a certain amount of time, all results but the best were removed from the list of minima to be examined, and the algorithm was continued. The final result is shown in the table.

When maimed low resolution sets are used, the performance of almost all operators diminishes. For the L_{vv} operator, the registration algorithm fails to converge on a single minimum. The other operators provide registrations with a less accurate z-translation. Given the poor z-resolution and the very severe maiming of the CT volume, these cannot be called poor results.

We verified the 3D results on similar datasets from two other patients. The registration results were almost identical.

4.6 Conclusion and discussion

We have tested and compared the registration of CT and MR brain images by correlation of image features, notably edge and ridge features. In 2D experiments, the registration merit of the feature operators was established using brute force experiments. Both edge and ridge operators showed good performance, with the edge operators having a slight advantage. Four operators and their optimal scale were selected for 3D registration experiments, the edge operators L_w (the gradient magnitude), and L_{ii} (the Laplacian), and the ridge operators L_{vv} and

$L_{vv}/\sqrt{(L_w)}$. The 3D registration results were accurate, visually even more accurate than the marker based results used as a reference. All operators showed proper convergence. Only when severely maimed low resolution images were employed, the L_{vv} registration failed, and the L_w and L_{ii} registrations were less accurate. The maiming of the volumes used borders the limits of normal protocol scope, however. Based on both the 2D and 3D experiments, the gradient magnitude L_w seems to be the best choice of the originally proposed operators. In the 2D experiments, it produced accurate registration results, and appeared robust under changes in scale, initial transformation and maiming of the original CT and MR images. In the 3D experiments, it led to accurate results when realistic initial CT and MR images were used. Moreover, of all operators mentioned, L_w is the fastest to compute, since only the three first order derivatives are needed.

For the registration of 3D CT and MR brain images, we propose a registration scheme using scaled L_w feature images in a hierarchical correlation algorithm. This scheme requires no interaction, and is therefore devoid of human subjectivity. Only patient-related geometrical features are used for the registration, so registration can be performed even if pre-acquisition registration accommodations have not been made. The use of patient-related features is also more patient friendly than the use of external features, since no marking devices are required to be attached to the patient's head. In this paper, the proposed scheme was shown to produce accurate registration results with the limited number of datasets employed. In future work, we hope to extend the scope of the scheme by testing it on more brain imaging protocols, and move on to other modalities, as well as images of other parts of the body. We would also like to reduce the number of computer operations required, in order to improve the clinical applicability.

4.7 Acknowledgments

This research was supported in part by the industrial companies Philips Medical Systems, KEMA, Shell Research, and ADAC, as well as by the Netherlands ministries of Education & Science and Economic Affairs through a SPIN grant, by the Murray Foundation, and by the Netherlands Organization for Scientific Research (NWO), through a travel grant. We wish to thank Dr G. Wilts, Dr L.M. Ramos, Dr P.F.G.M. van Waes and Dr F.W. Zonneveld for their efforts to supply the images. The help of Dr K.J. Zuiderveld and Dr B.M. ter Haar Romeny in the course of this work are also greatly appreciated. C. Bouma, W. Niessen, R. Stokking, and Dr Evert-Jan D. Pol are thanked for their comments on early manuscript versions.