Journal of Computing and Information Technology - CIT 6, 1998, 2, 165–179

Voxel-Based Similarity Measures for Medical Image Registration in Radiological Diagnosis and Image Guided Surgery

Thorsten M. Buzug and Jürgen Weese

Philips Research Laboratories, Division Technical Systems, Hamburg, Germany

Registration of images is a key technique for numerous medical applications from diagnosis to image guided therapy. In the present paper we focus on gray-value based registration methods. Of special interest is the so-called similarity measure which must be optimized. It is the intention of the paper to emphasize the fact that a successful registration requires a similarity measure that is carefully chosen with respect to the underlying medical application. Two single-modality examples are presented in the paper, i.e. the diagnostic tool of digital subtraction angiography and an intervention guidance under X-ray fluoroscopic control. We discuss new similarity measures, i.e. the class of one-dimensional histogram based measures and the pattern intensity, designed for these applications and compare them with frequently used measures like the cross-correlation function, cross-structure function and deterministic sign change criterion. It is demonstrated that the results obtained with the new measures are superior to the results obtained with the latter mentioned ones.

Keywords: Image registration, single modality, similarity measure, pattern intensity, energy, digital subtraction angiography, image guided surgery.

1. Introduction

For diagnosis and therapy it is often necessary to merge information from different images. These images may be of the same modality, but acquired at different times to judge and control the success of an intervention. They may show a patient before and after injection of contrast agent as in the case of digital subtraction angiography (DSA) or they may represent different modalities such as MRI, CT, PET etc. and provide complementary information about a patient. In every case, the images must be registered to obtain morphologically correct image overlays, because the patient has usually been differently positioned in the imaging devices or slightly moved during an image acquisition session.

Registration may be done point-based, surface-based or voxel-based (Van den Elsen, 1993). In the latter case, a similarity measure is evaluated for a given transformation between the image coordinate systems and registration is performed by optimizing the similarity measure with respect to the parameters of the transformation. Especially for multi-modality registration, i.e. registration of images obtained from different acquision techniques, it is very difficult to define a suitable similarity measure leading to a proper registration result. Major advances have been made recently in that context with the application of mutual information (Collignon et al., 1995a/b). But even in the case of single-modality registration, well-known similarity measures such as the cross-correlation function (Pratt, 1974, Rosenfeld and Kak, 1982), cross-structure function (Schulz–Dubois and Rehberg, 1981, Svedlow et al., 1978) and correlation coefficient (Hua and Fram, 1993) fail in various applications. Though the images represent the same modality they may differ because of the injection of contrast agent, surgical instruments or implants visible in one but not in the other image, anatomical differences due to an intervention or various other reasons. To cope with such problems,

further similarity measures such as deterministic sign change (Venot and Leclerc, 1984, Venot et al., 1984, 1988), histogram-based measures (Buzug et al., 1997) and pattern intensity (Weese et al., 1997a) have been developed.

There is, however, no similarity measure which is in general superior to the others. It depends on the properties of the images and on application requirements such as computation time which measure performs best for a specific application. In this contribution we derive for two medical applications — registration for digital subtraction angiography image enhancement and 2D/3D registration for navigation support in image guided surgery — the requirements for similarity measures and discuss to what extent the above mentioned measures are suitable.

In the following section the medical applications are presented and the requirements for the similarity measures are derived from typical images and application requirements. In the third section the similarity measures are briefly introduced and the capabilities as well as limitations with respect to the considered applications are discussed. The fourth section includes examples illustrating the most important points. The conclusions are summarized in the fifth section.

2. Medical Applications

Two X-ray fluoroscopy-based medical applications are described and requirements for the similarity measure are derived in the following subsections. The first application is digital subtraction angiography (DSA) which is a well-known tool in vessel diagnosis (Chilcote et al., 1981). The second application is registration of 2D X-ray fluoroscopies with 3D CT images. The registration result can be used to support navigation when placing pedicle screws in spine surgery (Lavallee et al., 1996) or a stent in the treatment of abdominal aortic aneurysms according to the Transfemoral Endovascular Aneurysm Management (TEAM) procedure which is considered in the EC project EASI (Buurman, 1996).

2.1. Digital Subtraction Angiography Image Enhancement

Digital subtraction angiography (DSA) is a standard diagnostic tool for the examination of blood vessels. For this method X-ray images are taken from a patient before and during injection of a radio-opaque contrast agent through a catheter. The first image of such a sequence, the mask image, is not affected by the contrast agent. This image is subtracted from a contrast image, i.e. an image acquired during injection of the contrast agent, to visualize the vessels. Usually the patient moves slightly which leads to disturbing artifacts in the subtraction image.

In clinical routine the mask image is manually shifted with respect to the contrast image to perform a rough patient motion compensation and to reduce the motion artifacts. To improve this procedure an algorithm working automatically in a region-of-interest defined by the physician is desired. Such an algorithm must be very robust, highly reliable and at least as fast as the manual correction which is performed within a few seconds. Otherwise, it won't be acceptable in clinical routine. In addition, more complex transformations than a shift may be applied to further increase the image quality.

A suitable algorithm for DSA image enhancement consists, for example, of the following steps (Buzug et al., 1996):

1. Subdivision of a user-defined region-of-interest into a set of quadratic, disjunct templates.
2. Computation of the template shift in the mask image with respect to the contrast image. In that way a motion vector field is obtained.
3. Approximation of the motion vector field by an affine transformation using singular value decomposition (Press et al., 1990).
4. Subtraction of the affinely corrected mask image from the contrast image.

As an example, Fig. 1 shows a mask image (a), contrast image (b), uncorrected subtraction image (c) and enhanced subtraction image (d) of the head. The region-of-interest is drawn as a rectangle and blown-up for better visualization of the motion vector field which is indicated as white lines attached to the template centers. Thanks to the affine motion compensation, most

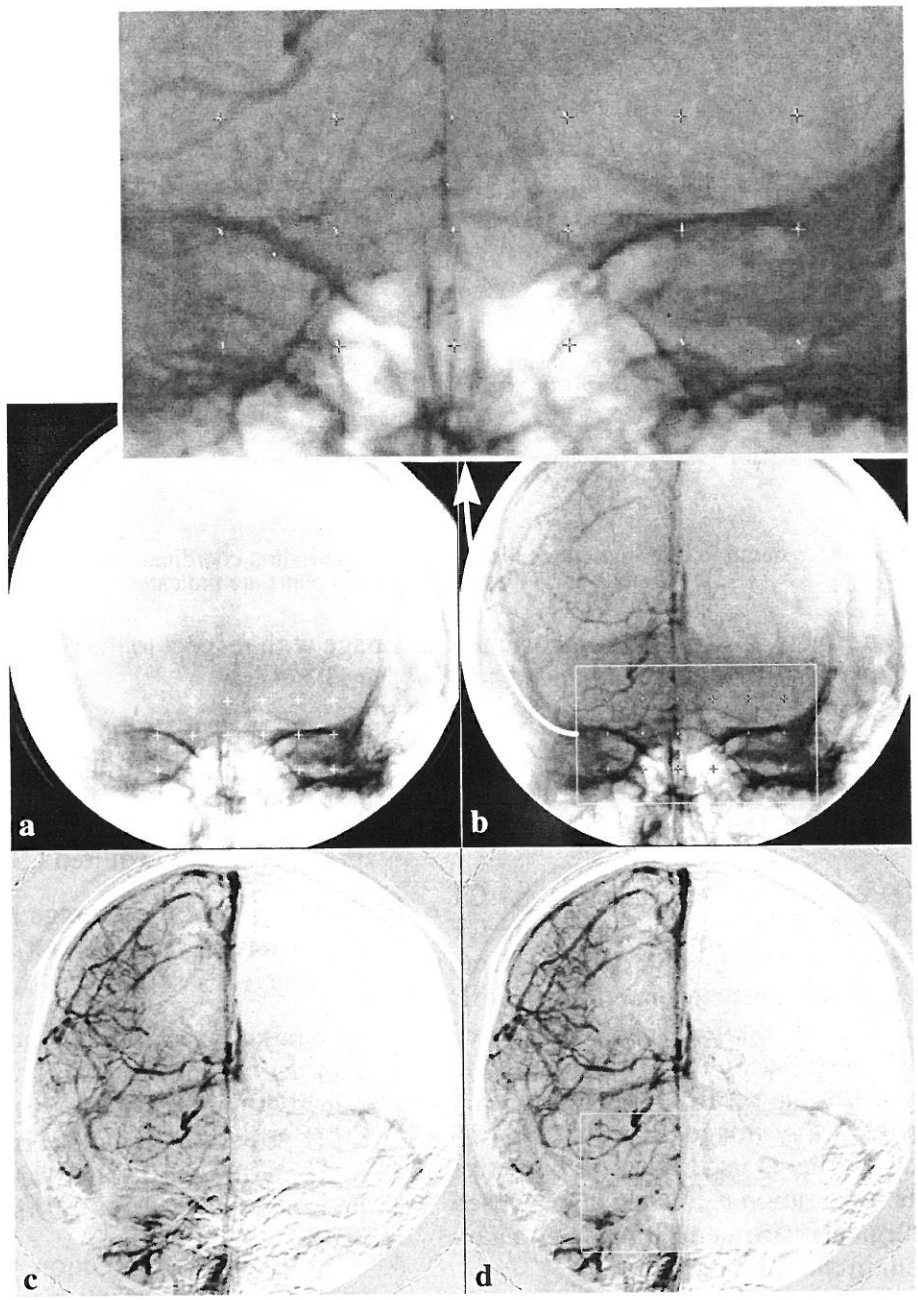

Fig. 1. Mask (a) and contrast (b) image of head fluoroscopies. A motion-vector field is estimated via template matching. As a result a set of homologous point landmarks are obtained and indicated as white and dark crosses in the mask and contrast image, respectively. The motion — visualized with white lines attached to the crosses in the contrast image (see blown-up region) is compensated inside a user-defined region-of-interest indicated by white rectangles. Figure (c) shows the manual-shift corrected subtraction result to demonstrate the motion artifacts and figure (d) shows the affinely corrected subtraction result.

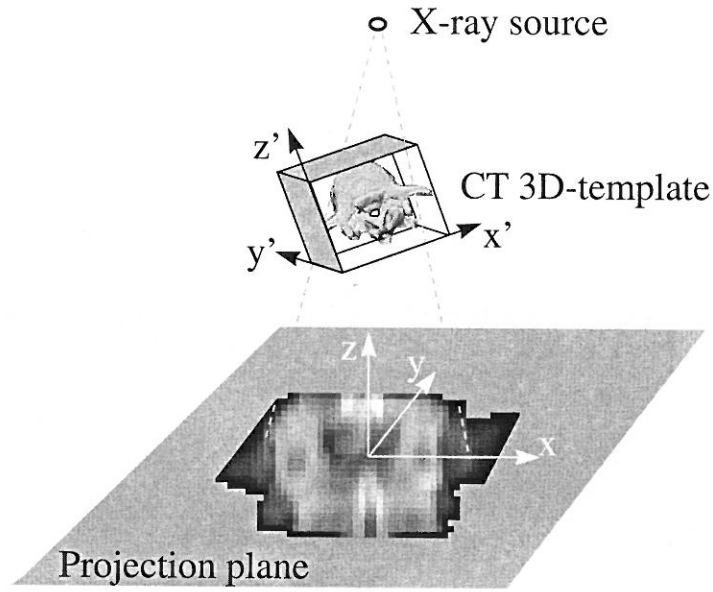

Fig. 2. Setup for the generation of pseudo projections. The corresponding coordinate systems of the segmented CT 3D-template as well as the projection plane are indicated.

of the cloudy artifacts, visible in the uncorrected subtraction image, vanish in the enhanced subtraction image.

2.2. 2D/3D Registration for Navigation Support in Image Guided Surgery

Given a 2D X-ray fluoroscopy and a 3D CT image, the location and orientation of the CT image with respect to the geometry of the X-ray imaging device can be determined by 2D/3D registration. The geometrical setup is indicated in Fig. 2. The registration result can be used to superimpose information from the CT or CTA image onto the X-ray image. If a suitable calibrated X-ray device is used for intra-operative imaging, the registration result can also be used to display surgical instruments tracked by a navigator system in the CT image.

For this purpose a novel registration method has been proposed recently (Weese et al., 1997a). The basic idea of this method is to use a rigid structure as e.g. a vertebra for registration and to rub out all structures due to the object-of-interest in the X-ray fluoroscopy using pseudo projections of the CT volume with the object-of-interest. In the case that all structures vanish, the correct location and orientation of the CT image with respect to the X-ray device has been found.

The corresponding 2D/3D registration algorithm consists of the following steps:

1. Segmentation of the object-of-interest in the pre-operatively acquired CT image.
2. Subtraction of the average gray-value of the tissue around the object-of-interest from the CT image.
3. Computation of pseudo projections taking into account only the CT 3D-template obtained by segmentation (Fig. 2). In that way a 2D projection template is obtained showing only the gray-value variation due to the presence of the object-of-interest.
4. Scaling of the projection template's gray-values and their subtraction from the X-ray image. For a proper gray-value scaling and the correct location and orientation of the CT image, the structures in the X-ray projection corresponding to the object-of-interest will vanish, and overall there will be less structures visible after subtraction.
5. Calculation of an appropriate similarity measure characterizing the 'structuredness' of

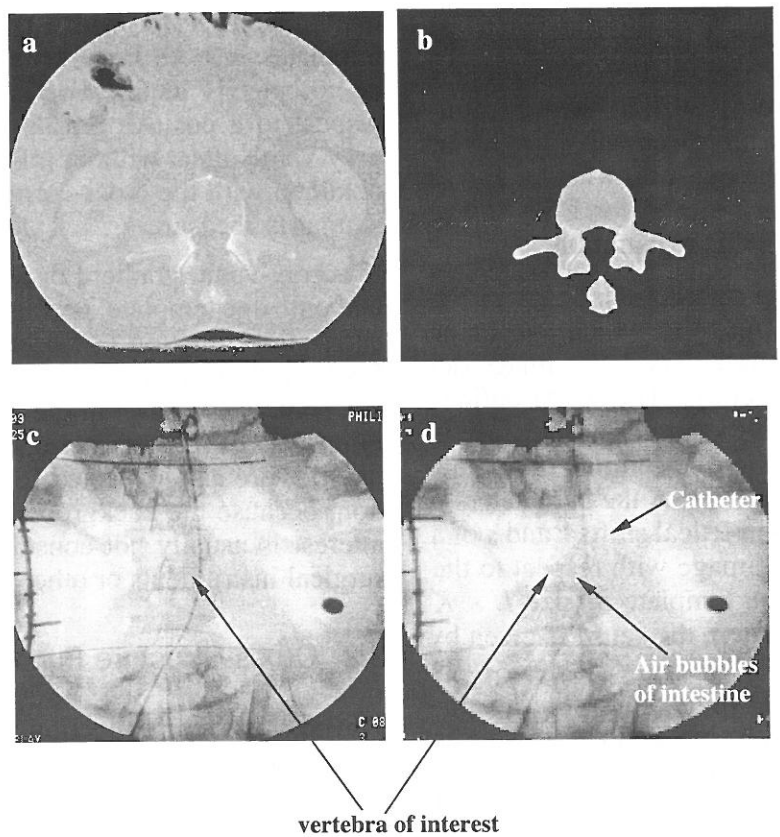

Fig. 3. Slice of CT image showing a vertebra slice (a) and segmented vertebra (b). The intra-operatively acquired vertebra fluoroscopy (c) as well as subtraction result (d) are illustrated.

the X-ray image after subtraction of the gray-value scaled projection template.

6. Optimization of the 'structuredness' with respect to the gray-value scaling and the parameters characterizing location and orientation of the CT image. Within each iteration of the optimization, steps 3 to 5 must be repeated.

Pre-processing of the CT image in steps 1 and 2 can be done prior to the intervention where time is not crucial. Registration itself is done in steps 3 to 6 and takes about 60–90 s (WEESE et al., 1997a) which is nearly acceptable for an intervention. However, most time is spent in the computation of the pseudo projection from the CT 3D-template and time for the evaluation of the similarity measure can be neglected.

As an example images of a TEAM procedure are used. Fig. 3a shows a slice of a CT image with a vertebra which is the object-of-interest, Fig. 3b a projection template, Fig. 3c an intra-operative X-ray image, and Fig. 3d the intra-operative X-ray image with the registered projection template subtracted. After registration all structures due to the vertebra-of-interest vanish, indicating proper registration. There are, however, structures left. Air bubbles in the intestine as well as a catheter are clearly visible in the area of this vertebra. This means that the projection template and the corresponding template in the X-ray image differ not only by a gray-value scaling and a gray-value offset representing a homogeneous background. There are further structures overlaid to the template in the X-ray image and for a reliable and accurate registration result it is most important that the similarity measure is robust with respect to such structures and not disturbed by them.

3. Similarity Measures

In the following subsections cross-correlation function, structure function, correlation coefficient, deterministic sign change, histogram-based measures, mutual information and the pattern intensity are introduced. The properties of these measures as well as their capabilities and limitations are discussed. Particular attention is paid to the question whether and to what extent they can be used for DSA image enhancement and 2D/3D registration.

Within the following subsections $g_1(x, y)$ denotes the reference image to which the image $g_2(s, y : \vec{p})$ is registered. The latter image depends on the parameters $\vec{p}$. The set M defines the points over which the similarity measure is evaluated and # M denotes the number of points in this set. In the case of DSA the parameters $\vec{p}$ are the horizontal and vertical shifts r and s of a template in the mask image with respect to the contrast image. For a template of size $L \times K$ with its corner at (x_0, y_0), the set M is given by $\{(x, y)|x_0 \leq x < x_0 + L \wedge y_0 \leq y < y_0 + K\}$. In the case of 2D/3D registration, the parameters $\vec{p}$ describe gray-value scaling as well as location and orientation of the CT image used to generate the pseudo projection. The set M contains the pixels of the pseudo projection for which the corresponding X-ray intersects the object-of-interest.

3.1. Cross-Correlation Function

The cross-correlation function (Pratt, 1974, Rosenfeld and Kak, 1982) is defined as

$$G(\vec{p}) = \frac{1}{\#\mathrm{M}} \sum_{(x,y)\in M} (g_1(x, y) - m_1) \times \times (g_2(x, y; \vec{p}) - m_2), \quad (1)$$

where

$$m_1 = \frac{1}{\#\mathrm{M}} \sum_{(x,y)\in M} g_1(x, y), \qquad m_2 = \frac{1}{\#\mathrm{M}} \sum_{(x,y)\in M} g_2(x, y; \vec{p}) \quad (2)$$

are the average gray-values in the area of the template. By maximizing this function similar templates can be found. If a constant gray-value is added to one of the images, the maximum remains unchanged, because the average gray-value is subtracted from each image. The location of the maximum also remains unchanged if one image is scaled with a positive factor. The templates in both images need, therefore, not be completely identical. They can differ with respect to a positive gray-value scaling and a gray-value offset without influencing the result obtained with the cross-correlation function as similarity measure.

If a gray-value gradient or a structure is overlaid onto one template, problems with the cross-correlation function can occur. In the case of DSA, difficulties have been observed if there is a contrasted vessel in the template (Hua and Fram, 1993, Fitzpatrick et al., 1987, 1988). Difficulties are also expected for 2D/3D registration, because the background of the object-of-interest is usually not constant and may show surgical instruments or other organs.

3.2. Cross-Structure Function

The cross-structure function, first mentioned in (A. N. Kolmogorov, 1941), is defined by

$$S(\vec{p}) = \frac{1}{\#\mathrm{M}} \sum_{(x,y)\in M} (g_1(x, y) - g_2(x, y; \vec{p}))^2 \quad (3)$$

and must be minimized for registration. Registration with this function can be considered as a Least-Squares method (PRESS et al., 1990) where $g_1(x, y)$ are the data, $g_2(x, y; \vec{p})$ the model and $\vec{p}$ the parameters which are fitted. Good results are, therefore, expected in the case of identical templates and every dissimilarity degrades the result. Compared to the cross-correlation function, results obtained with the cross-structure function are sensitive to a gray-value offset and a gray-value scaling in one of the templates.

In the case of DSA, similar difficulties as for the cross-correlation function are expected if there is a contrasted vessel in the template. In order to use the cross-structure function for 2D/3D registration, an additional parameter representing the gray-value of the background must be introduced and fitted. Nevertheless, in clinical practice the background of the object-of-interest may not be constant and overlaid structures should cause problems.

3.3. Correlation Coefficient

Considering registration with the cross-structure function as a Least-Squares method, a gray-value scaling G and a gray-value offset m_0 can be treated within this concept. For that purpose these parameters are inserted in Eq. (3) leading to

$$s(\vec{p}; G, m_o) = \frac{1}{\#\mathrm{M}} \sum_{(x,y)\in M} (g_1(x,y) - Gg_2(x,y;\vec{p}) - m_0)^2. \quad (4)$$

Minimization with respect to the gray-value scaling G and the gray-value offset m_0 can be done in a Least-Squares sense (see e.g. Press et al., 1990). In that way the equation (5)

$$s(\vec{p}) = \frac{1}{\#\mathrm{M}} \sum_{(x,y)\in M} (g_1(x,y)-m_1)^2 - \frac{\left(\sum_{(x,y)\in M} (g_1(x,y)-m_1)(g_2(x,y;\vec{p})-m_2)\right)^2}{\#\mathrm{M} \sum_{(x,y)\in M} (g_2(x,y;\vec{p})-m_2)^2} \quad (5)$$

is obtained. The first term in this equation is the gray-value variance of the template in the reference image $g_1(x,y)$. The second term represents the reduction of the gray-value variance due to registration of $g_2(x,y;\vec{p})$. The ratio of both terms

$$C^2(\vec{p}) = \frac{\left(\sum_{(x,y)\in M} (g_1(x,y)-m_1)(g_2(x,y;\vec{p})-m_2)\right)^2}{\sum_{(x,y)\in M} (g_1(x,y)-m_1)^2 \sum_{(x,y)\in M} (g_2(x,y;\vec{p})-m_2)^2} \quad (6)$$

is the square of the well-known correlation coefficient $C(\vec{p})$. The correlation coefficient is frequently used as a similarity measure in medical imaging (see e.g. Hua and Fram, 1993) and shows essentially the same characteristic properties as the cross-correlation function.

3.4. Deterministic Sign Change

The deterministic sign change (Venot and Leclerc, 1984, Venot et al., 1984, 1988) is defined in a 3-step procedure:

I. Addition of a periodic pattern to one of the images:

$$\tilde{g}(x,y) = g_1(x,y) + \delta \text{ if } x+y \text{ is even}$$
$$\tilde{g}(x,y) = g_1(x,y) - \delta \text{ if } x+y \text{ is odd,} \quad (7)$$

where δ is the so-called pattern depth. The pattern depth δ may be chosen in the same order of magnitude as the standard deviation of the noise in the images and it has been shown that there is no advantage in selecting the pattern depth more than 2 times larger (Venot and Leclerc, 1984).

II. Subtraction of images:

$$d(x,y;\vec{p}) = \tilde{g}_1(x,y) - g_2(x,y;\vec{p}). \quad (8)$$

III. Evaluation of criterion:

$$D(\vec{p}) : \text{number of sign changes in the template of the difference image } d(x,y;\vec{p}) \text{ scanned line by line} \quad (9)$$

This similarity measure is based on the idea that in the case of misregistration most parts of the periodic pattern are hidden whereas in the case of registration all structures except for the periodic pattern vanish in the difference image $d(x,y;\vec{p})$. The deterministic sign change does, however, not only work in the case of two identical templates, but also if part of the periodic pattern is hidden by an additional structure as e.g. a contrasted vessel. It can also be applied if one of the templates shows a gray-value offset. To compensate for this, the average gray-values in the area of the templates can be subtracted.

There are, however, some problems related to the application of the deterministic sign change. Since image statistics vary largely with image position, type of examination, and acquisition hardware, it may be necessary to introduce the pattern depth and, as mentioned above, the gray-value offset as dynamically adaptable parameters (K. J. Zuiderveld et al., 1989a/b). Furthermore, this measure is not a smooth function and not differentiable with respect to the parameters $\vec{p}$, because it is based on counting. For that reason many optimization methods cannot be used (Venot et al., 1988), and optimization is difficult or, at least, very time-consuming.

As indicated above, the deterministic sign change can cope with additional structures overlaid to one template and should hence be suitable for DSA image enhancement as well as for 2D/3D registration. In the case of DSA a horizontal and a vertical shift have to be determined which is a solvable optimization problem, because, in principle, a full search in the parameter space is possible. Nevertheless, with those optimization strategy it is hard to meet clinical requirements concerning computation time. In the case of 2D/3D registration where six parameters describing location and orientation of the CT image and the gray-value scaling must be adjusted, reliable and fast optimization is, however, nearly impossible.

3.5. Histogram-Based Similarity Measures

One-dimensional histogram-based similarity measures have been proposed in the context of DSA image enhancement (Buzug et al., 1996). They are defined in a 3-step procedure:

I. Subtraction of images:

$$d(x, y; \vec{p}) = g_1(x, y) - g_2(x, y; \vec{p}). \quad (10)$$

II. Calculation of gray-value histogram:

p_k : fraction of pixels in the template of the difference image $d(x, y; \vec{p})$ with gray-value k. These fractions depend on the parameters . (11)

III. Evaluation of criterion:

$$M(\vec{p}) = \sum_k f(p_k) \quad (12)$$

using a function $f(p)$ which is strictly convex for $p \geq 0$. Suitable functions are, for example,

$$f_1(p_k) = p_k \log p_k, \text{ (Information)} \quad (13)$$
$$f_2(p_k) = p_k^2, \text{ (Energy)} \quad (14)$$
$$f_3(p_k) = -\sqrt{p_k}, \text{ (negative branch of square-root)} \quad (15)$$

The starting point for histogram-based measures is the observation that in the case of registration the histogram is peaky, whereas misregistration leads to gray-value dispersion and a much broader histogram. $M(\vec{p})$ measures gray-value clustering and it has been proven that for all strictly convex functions f, $M(\vec{p})$ increases for increasing gray-value clustering (Buzug et al., 1997). In the case of maximum dispersion, i.e. if all gray-value fractions p_k have the same value, $M(\vec{p})$ has its global minimum.

For identical templates the histogram shows one large peak in the case of registration. A gray-value offset in one template shifts this peak, but does not influence dispersion or clustering and the measure $M(\vec{p})$ remains, therefore, unchanged. An overlaid structure, as e.g. a vessel, will cause another peak in the histogram, but will not significantly increase gray-value dispersion, if the overlaid structure has a relative constant gray-level. For that reason, one-dimensional histogram-based measures are assumed to be suitable for DSA image enhancement as well as for 2D/3D registration. However, problems are expected if e.g. a strong gradient is overlaid to one template, because a gradient may cause a considerable gray-value dispersion.

3.6. Mutual Information

Mutual information has been introduced recently as a similarity measure for 3D multimodality image registration. This measure is based on the so-called feature space or scatter plot $p(g_k, g_l)$ which are two-dimensional histograms describing the frequency of voxels with gray-values g_k and g_l at corresponding spatial positions in the images to be registered. Given the scatter plot, mutual information can be evaluated according to

$$I(\vec{p}) = \sum_{k,l} p(g_k, g_l) \log\left(\frac{p(g_k, g_l)}{p_1(g_l) p_2(g_k)}\right). \quad (16)$$

where $p_1(g)$ and $p_2(g)$ represent the frequency of voxels with gray-value g in each of the images (Collignon et al., 1995a/b).

This similarity measure is motivated by the observation that registration goes along with clustering in the scatter plot and mutual information quantifies this clustering. There have also been defined a variety of other functionals to characterize the scatter plot (Haralick et al., 1973,

Bro–Nielsen, 1997) and some of them may also be suitable for quantifying clustering. Especially, all strictly convex weighting functions introduced in the previous section lead to useful similarity measures. The concept of mutual information is similar to that of one-dimensional histogram-based similarity measures, but more general because it can be applied to multimodality image registration.

Mutual information has been applied with great success to 3D image registration of different modalities as e.g. CT, MR and PET (Collignon et al., 1995a/b). However, in the case of DSA image enhancement and 2D/3D registration difficulties arise, because both are 2D problems and templates with less than 500 pixels may be used. Compared to that, the scatter plot has a huge amount of bins as e.g. 65536 for images with a gray-value resolution of 8 bits, and the scatter plot will only sparsely be occupied. This leads to a poor statistics, and therefore, the mutual information as well as other scatter plot-based measures are not subsequently considered.

3.7. Pattern Intensity

With the registration becoming better, corresponding structures in the images should erase each other in the difference image and generally the number of structures should decrease. To characterize the 'structuredness' in the difference image $d(x, y; \vec{p})$, a small value may be assigned to points in the neighborhood of structures such as gray-value edges or lines and a large value to points in the areas showing only little gray-value variation.

A suitable quantity leading to the desired effect is the pattern intensity (Weese et al., 1997a):

$$P_{R,\sigma}(\vec{p}) = \sum_{x,y} \sum_{(x-v)^2+(y-w)^2} \frac{\sigma^2}{\sigma^2+(d(x,y;\vec{p})-d(v,w;\vec{p}))^2}. \quad (17)$$

It depends on two parameters. The parameter R defines the size of the neighborhood in which gray-value variations are taken into account. The parameter σ is the sensitivity defining whether a gray-value variation is considered to be a structure or not. Evaluation of the pattern intensity is time-consuming, because for each pixel in the template of the difference image the gray-values of a number of neighboring pixels have to be considered. The pattern intensity depends, therefore, on the spatial distribution of the gray-values in the difference image. This is an important difference compared to the similarity measures mentioned above which give a value that is not related to the spatial gray-value distribution.

The pattern intensity is not affected by a gray-value offset in one template and should also work if structures are overlaid onto one template. Nevertheless, it is not suitable for DSA image enhancement, because evaluation of the pattern intensity is very time consuming and computational efficiency is highly important for this application. In the case of 2D/3D registration time for evaluation of the similarity measure is less important, because a major part of the computation time is spent during generation of the projection template.

4. Discussion and Examples

In this section it is shown that voxel-based single modality registration is a challenging problem if there are additional structures overlaid onto one of the images. In this case well-known similarity measures as e.g. the cross-correlation function fail and measures especially adapted to cope with such additional structures must be applied. For DSA image enhancement the cross-correlation function, the cross-structure function, deterministic sign change and histogram-based measures are taken into consideration and the influence of a contrasted vessel is illustrated. In connection with 2D/3D registration the correlation coefficient, the entropy of the gray-value histogram and the pattern intensity are evaluated.

4.1. DSA Image Enhancement

For the DSA algorithm outlined in the second section the choice of an appropriate similarity measure used to calculate the motion vector field is most crucial for success. It is not sufficient that the similarity measure works for the templates which are nearly identical in mask and contrast image. As the physician will always choose a region-of-interest containing vessels,

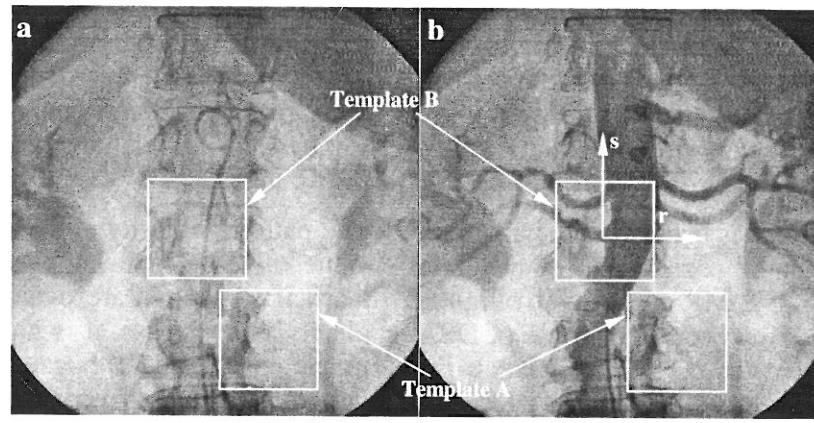

Fig. 4. Mask (a) and contrast image (b) of the abdomen. The aorta, superimposed to the vertebra, is strongly contrasted by injection of the contrast agent in (b). Two (128x128)-pixel templates are indicated in the images. Template A is not affected by contrast agent. Hence, nearly identical templates can be found in the mask and contrast image (indicated by arrows). For template B the location has been chosen to include large gray-value dissimilarities due to the contrasted aorta.

it is most important that the similarity measure leads also to the correct shift for templates which are overlaid by a vessel and strongly affected by the injection of contrast agent.

Fig. 4 shows a typical DSA image pair of the abdomen. In the mask image part of the vertebra column and a catheter inserted in the aorta but not the aorta itself are visible. In the contrast image the strongly contrasted aorta is superimposed to the vertebra column. Two templates have been selected to demonstrate the effect of additional structures on the quality of the similarity measures. They have a size of (128×128) pixels and are indicated by white boxes in both images. Template A is not affected by injection of contrast agent and nearly identical templates can be found in mask and contrast image. For template B the location has been chosen to include part of the contrasted aorta. This template is considerably different in mask and contrast image, because approximately one half of it is covered by the contrasted aorta. Fig. 5 and 6 (corresponding to templates A and B in Fig. 4) show the dependence of the cross-correlation function, the cross-structure function, the deterministic sign change criterion and three histogram-based measures (information, energy and (negative)square-root) on the shift values (r, s) for templates A and B, respectively. The maximum of the similarity measures is expected at $(r, s) = (0, 0)$, because mask and contrast images have been manually shift-corrected. The range of the shifts is $r, s \in [-15, 14]$.

Fig. 5 illustrates that the cross-correlation as well as the cross-structure function work for Template A, but both functions fail completely for Template B (Fig. 6). In the latter case they have only a bump but no optimum around $(r, s) = (0, 0)$. This shows that a contrasted vessel can have a considerable influence on these measures which makes them inappropriate for DSA image enhancement. A further discussion of the cross-correlation function, correlation coefficient, cross-structure function and similar functions like the sum of absolute differences in the context of DSA can be found in (Svedlow et al., 1978, Fitzpatrick et al., 1989, Tran and Sklansky, 1992).

The deterministic sign change shows a better result than cross-correlation and cross-structure functions, because it is more robust with respect to vascular structures in one template. Nevertheless, for Template B with the overlaid aorta this similarity measure is very rough. This complicates optimization which may be trapped in a local maximum. The reason for this effect is that a major part of the periodic pattern is hidden by the overlaid structure and that, consequently, only a relatively small number of sign changes is observed, which leads to poor statis-

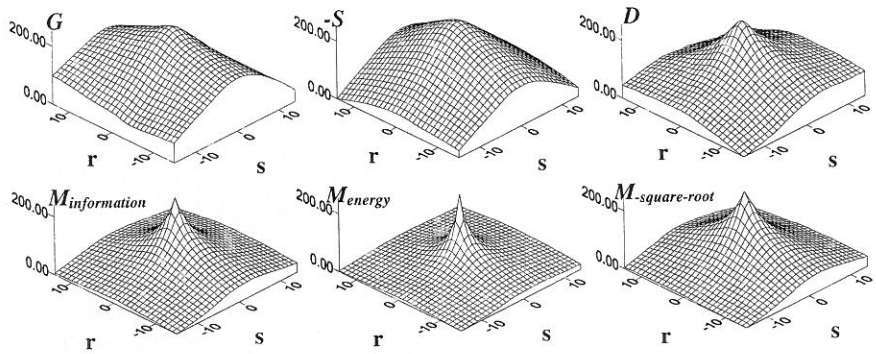

Fig. 5. For the Template A of Fig. 4, six different objective functions are shown, i.e. the cross-correlation function (G), the cross-structure function (S), the deterministic sign change criterion (D), and the three histogram-based measures information, energy and (negative) square root. All measures shown in the range of shift values $r, s \in [-15, 14]$ are normalized to $[0, 255]$. The expected maximum at $(r, s) = (0, 0)$ can be found in all measures.

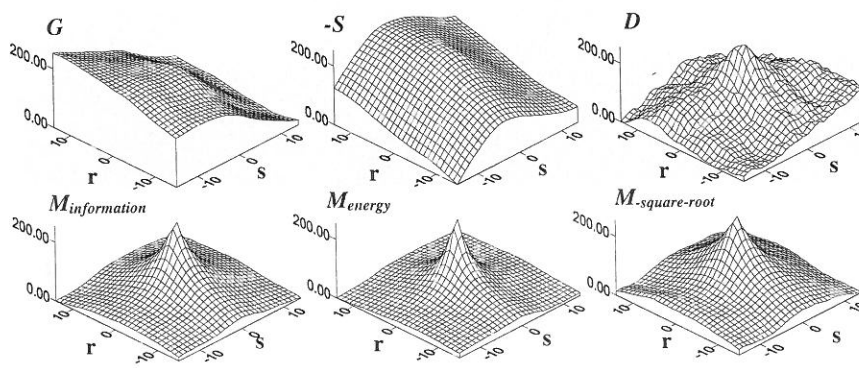

Fig. 6. For the template B of Fig. 4, six different objective functions are shown, i.e. the cross-correlation function (G), the cross-structure function (S), the deterministic sign change criterion (D), and the three histogram-based measures information, energy and (negative) square root. All measures shown in the range of shift values $r, s \in [-15, 14]$ are equally normalized to $[0, 255]$. The expected maximum at $(r, s) = (0, 0)$ can be found only in the deterministic sign change criterion and the histogram-based measures. The latter functions additionally show a desired smooth surface.

tics. For further discussion on the deterministic sign change in comparison to the sum of absolute differences and cross-correlation see (Fitzpatrick et al., 1987).

The similarity measures based on strictly convex weighted histograms lead to good results for Template A as well as for Template B. Information, energy and (negative) square-root have a significant maximum at the expected position of $(r, s) = (0, 0)$. The similarity measure plots are smooth and show a relatively large attractive basin of approximately ± 15 pixels. This holds for Template A as well as for Template B. The shape of the maximum differs slightly for the different functions. The energy which has the smallest slope for $p \to 0$ yields the narrowest peak, whereas the negative square-root, which has the steepest slope for $p \to 0$, shows the broadest peak. For the DSA application the energy is preferred as similarity measure, because it can be evaluated very fast and integer arithmetic can be used.

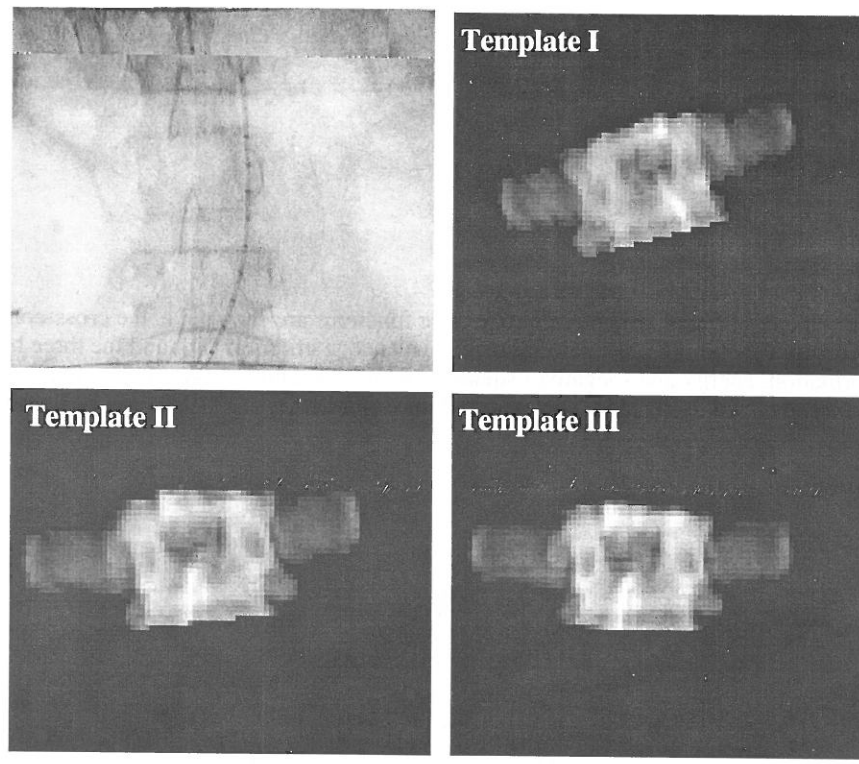

Fig. 7. Fluoroscopy of the vertebra and three templates derived from pseudo projections of the segmented 3D vertebra. Template I: Optimal parameters obtained from correlation coefficient, Template II: Optimal parameters obtained from entropy, and Template III: Optimal parameters obtained from pattern intensity.

4.2. 2D/3D Registration

In order to assess suitability of similarity measures, it must be checked whether they have a pronounced maximum for parameters corresponding with a good registration. In addition, the similarity measure should show a reasonable convergence range enabling reliable optimization. In the case of 2D/3D registration it is difficult to illustrate these properties by using plots of the similarity measure, as it was done in the previous subsection for DSA image enhancement, because the six parameters characterizing location and orientation of the CT image as well as the projection template's gray-value scaling must be estimated. For that reason, registrations using various different start values have been performed.

For this purpose the images of Fig. 3 have been used. The CT image has $512 \times 512 \times 123$ voxels, a resolution of 0.488 mm in a slice, a slice-to-slice distance of 2 mm, and a slice thickness of 5 mm. The intra-operative X-ray projection refers to a geometry defined by an image intensifier of 12" diameter and an X-ray source 1 m above it. The projection template showing the vertebra of interest has a lower sharpness than the intra-operative X-ray projection, because of the CT slice thickness. In order to reduce this difference, the X-ray projection has been smoothed with a 5×5 pixel moving average filter. In addition, its resolution has been reduced from $512 \times 512 \times 512$ to 128×128. The resolution reduction speeds up computation of the projection template by a factor of 16 which is necessary to perform registration in a reasonable time.

Registration has been performed using the correlation coefficient, the entropy of the gray-value histogram (negative information, see Eq. 13), and the pattern intensity with parameters ($r = 3, \sigma = 10$). Optimization of the similarity measure has been done as described in (Weese et al., 1997a). Fourteen different sets of start values have been generated by adding a rotation of about 10 deg to the rotation parameters obtained by manual registration. The translation parameters have been roughly adjusted to get an overlay of the projection template with the vertebra of interest in the X-ray projection. Fig. 7 shows typical projection templates after optimization together with a magnified part of the X-ray image. For visualization the gray-values of the projection template have been inverted and scaled. Projection template I does obviously not correspond with the vertebra of interest in the X-ray image. Template II matches much better, though it is slightly rotated to the left around an axis perpendicular to the projection plane. Template III shows the best correspondence. Using this template all structures in the X-ray image caused by the vertebra of interest can be rubbed out, which indicates proper registration (Fig. 3d). The deviation of rotation parameters is about 23 deg. for templates I and III and 13 deg for template II and III.

With the correlation coefficient, results far from a good registration have been obtained. In particular, the most pronounced maximum corresponds with Template I and has the value of 0.54. On the other hand, the correlation coefficient was much lower, only about 0.42 for a reasonable registration result such as Template III. As it has been shown for DSA image enhancement, correlation is not robust with respect to overlaid structures and the misregistration can be explained by the catheter overlaid to the X-ray image and the organ behind the vertebra of interest.

The maximum found for the entropy (negative information, see Eq. 13) of the gray-value histogram has the value of 0.31 and corresponds with Template II. This is a much better result than that obtained with the correlation coefficient though there is a clear deviation from Template III which shows the best agreement. The improvement can be explained by the fact that the entropy of the gray-value histogram can cope with overlaid structures. Registration results spread, however, around a considerable area without the entropy varying much. For projection templates very similar to templates I and III, for example, values of about 0.3 have been found. This shows that entropy is not sensitive to all the parameters characterizing the location and orientation of the CT image and that robust and reliable registration cannot be done with this measure.

For the pattern intensity ($r = 3, \sigma = 10$) the most pronounced maximum had a value of 1850 and Template III was obtained. In general, optimization converged to this maximum and the standard deviation of the resulting parameters was less than 0.5 deg for the rotation, less than 0.2 mm for the translations parallel to the projection plane and about 2.5 mm for the height of the CT volume above the projection plane. In our experiments there was only one set of start values for which optimization was trapped in a local maximum significantly away from optimal registration and with much smaller pattern intensity of 640. It should also be noted that the pattern intensity was 237 and 131 for templates I and II, respectively. All results indicate that this similarity measure has a very well-pronounced maximum corresponding with a good registration and that it has a reasonable convergence range enabling reliable optimization. That the pattern intensity maximum corresponds with a good registration is supported by the first results on accuracy derived from the comparison of registration results for three different vertebrae of a spine phantom (Weese et al., 1997b).

5. Conclusions

Two medical applications requiring voxel-based registration methods have been considered. DSA image enhancement and 2D/3D registration. Both applications have been described and the requirements to be met by the similarity measures have been derived using examples of clinical images. For both applications it is important that the similarity measure is robust with respect to structures such as contrasted vessels or surgical instruments visible in only one of the images to be registered. Several voxel-based similarity measures have been described and their properties have been discussed. In particular, two similarity measures recently introduced for registration of medical images have been presented.

Examples have shown that voxel-based single modality registration is a challenging problem if there are additional structures overlaid onto one of the images. In this case well-known similarity measures as e.g. the cross-correlation function fail and measures especially adapted to cope with additional structures must be applied. The class of histogram-based measures has been shown to be robust with respect to contrasted vessels overlaid onto one of the images and is, therefore, suitable for DSA image enhancement. In addition, these measures can be evaluated very fast making it possible to satisfy clinical requirements concerning computation time. For 2D/3D registration the pattern intensity leads to good results. This measure can cope with surgical instruments or organs overlaid onto one of the images to be registered which is beneficial for clinical applications.

Acknowledgments The authors would like to thank Dr. L. J. Schultze Kool, University Hospital Leiden, for providing us with the DSA data sets and Prof. Dr. W. P. Th. M. Mali, Prof. Dr. B. C. Eikelboom and Dr. J. D. Blankensteijn University Hospital Utrecht, for providing the images of the TEAM procedure.

The work on 2D/3D registration was done within the context of the EASI project "European Applications for Surgical Interventions", supported by the European Commission under contract HC1012 in their "4th Framework Telematics Applications for Health" RTD programme. The partners in the EASI consortium are Philips Medical Systems Nederland B.V., Philips Research Laboratory Hamburg, the Laboratory for Medical Imaging Research of the Katholieke Universiteit Leuven, Utrecht University & University Hospital Utrecht, The National Hospital for Neurology and Neurosurgery in London, and the Image Processing Group of Radiological Sciences at UMDS of Guy's and St. Thomas's Hospitals in London.

The algorithms were implemented on an experimental version of the EasyVision workstation from Philips Medical Systems and we would like to thank EVM (Easy Vision Modules) Advanced Development, Philips Medical Systems, Best, for helpful discussions.

References

M. Bro–Nielsen, (1997) Rigid registration of CT, MR and Cryosection images using a GLCM framework. In Proc. of CVRMed/MRCAS'97 (J. Troccaz, E. Grimson and R. Mösges Eds.) Lecture Notes in Computer Science **1205**, pp. 171–180, Springer, Berlin.

J. Buurman and F. A. Gerritsen, (1997) European Applications in Surgical Interventions in Computer Assisted Radiology. In Proc. of CAR'96 (H. U. Lemke, M. W. Vannier, K. Inamura and A. G. Farman, Eds.) p.677–681, Elsevier, Amsterdam.

T. M. Buzug, J. Weese, C. Fassnacht and C. Lorenz, (1996) Using an entropy similarity measure to enhance the quality of DSA images with an algorithm based on template matching. In Proc. of VBC'96 (K. H. Höhne, R. Kikinis, Eds.) Lecture Notes in Computer Science **1131**, pp. 235–240, Springer, Berlin.

T. M. Buzug, J. Weese, C. Fassnacht and C. Lorenz, (1997) Image registration: Convex weighting function for histogram-based similarity measures. In Proc. of CVRMed/ MRCAS'97 (J. Troccaz, E. Grimson and R. Mösges, Eds.) Lecture Notes in Computer Science **1205**, pp. 203–212, Springer, Berlin.

W. A. Chilcote, M. T. Modic, W. A. Pavlicek, (1981) Digital subtraction angiography of the carotid arteries: A comparative study in 100 patients. Radiology **139**, 287.

A. Collignon, D. Vandermeulen, P. Suetens and G. Marchal, (1995a) 3D multimodality medical image registration using feature space clustering. In Proc. of CVRMed'95 (N. Ayache, Ed.) Lecture Notes in Computer Science **905**, pp. 195–204, Springer, Berlin.

A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens and G. Marchal, (1995b) Automated multimodality image registration using information theory. In Proc. of IPMI'95 (Y. Bizais, C. Barillot, R. di Paola, Eds.) pp. 263–274, Kluver, Dordrecht.

P. A. van Den Elsen, E.–J. D. Pol and M. A. Viergever, (1993) Medical image matching — A review with classification, IEEE Eng. In Medicine and Biology, 26–39.

J. M. Fitzpatrick, D. R. Pickens, J. J. Grefenstette, R. R. Price and A. E. James, (1987) Technique for automatic motion correction in digital subtraction angiography. Optical Engineering **26**, 1085–1093.

J. M. Fitzpatrick, J. J. Grefenstette, D. R. Pickens, M. Mazer and J. M. Perry, (1988) A system for image registration in digital subtraction angiography. In Information processing in medical imaging (C. N. de Graaf and M. A. Viergever Eds.) pp. 415–435, Plenum Press, New York.

J. M. Fitzpatrick, D. R. Pickens, H. Chang, Y. Ge and M. Özkan, (1989) Geometrical transformations of density images. SPIE **1137**, 12–21.

R. M. HARALICK, K. SHANMUGAM AND I. DINSTEIN, (1973) Textural features for image classification. IEEE Trans. Systems, Man, and Cybernetics, SMC-**3**, 610.

P. HUA AND I. FRAM, (1993) Feature-based image registration for digital subtraction angiography. SPIE **1898**, 24–31.

A. N. KOLMOGOROV, (1941) Local structure of turbulence in noncompressible fluid with high Reynolds number. Dokl. Acad. Sci. USSR 30, 299.

S. LAVALLEE, J. TROCCAZ, P. SAUTOT, B. MAZIER, P. CINQUIN, P. MERLOZ AND J.–P. CHIROSSEL, (1996) Computer-assisted spinal surgery using anatomy-based registration. In: Computer-Integrated Surgery (R. H. Taylor, S. Lavallee, G. C. Burdea, R. Mösges, Eds.) pp. 425–450, MIT Press, London.

W. K. PRATT, (1974) Correlation techniques of image registration. IEEE Trans. on AES, AES-**10**, 353–358.

W. H. PRESS, B. P. FLANNERY, S. A. TEUKOLSKY AND W. T. VETTERLING, (1990) Numerical Recipes in C. Cambridge University Press, Cambridge. A.

ROSENFELD, AND A. KAK (1982) Digital picture processing, 2nd ed. Academic Press, New York.

E. O. SCHULZ–DUBOIS AND I. REHBERG, (1981) Structure function in lieu of correlation function. Appl. Phys. **24**, 323–329.

M. SVEDLOW, C. D. MCGILLEM AND P. E. ANUTA, (1978) Image registration: Similarity measure and preprocessing method comparison. IEEE Trans. AES-**14**, 141–149.

L. V. TRAN AND J. SKLANSKY, (1992) Flexible mask subtraction for digital angiography. IEEE Trans. on Medical Imaging **11**, 407–415.

A. VENOT AND V. LECLERC, (1984) Automated correction of patient motion and gray values prior to subtraction in digitized angiography. IEEE Trans. on Med. Im. **4**, 179–186.

A. VENOT, J. F. LEBRUCHEC AND J. C. ROUCAYROL, (1984) A new class of similarity measures for robust image registration. Comp. Vis. Im. Proc. **28**, 176–184.

A. VENOT, J. Y. DEVAUX, M. HERBIN, J. F. LEBRUCHEC, L. DUBERTRET, Y. RAULO AND J. C. ROUCAYROL, (1988) An automated system for the registration and comparison of photographic images in medicine. IEEE Trans. Medical Imaging **7**, 298–303.

J. WEESE, T. M. BUZUG, C. LORENZ AND C. FASSNACHT, (1997a) An approach to 2D/3D registration of a vertebra in 2D X-ray fluoroscopies with 3D CT images. In Proc. of CVRMed/MRCAS'97 (J. Troccaz, E. Grimson and R. Mösges, Eds.) Lecture Notes in Computer Science **1205**, pp. 119–128, Springer, Berlin.

J. WEESE, G. P. PENNEY, T. M. BUZUG, C. FASSNACHT AND C. LORENZ, (1997b), 2D/3D registration of pre-operative CT images and intra-operative X-ray projections for image guided surgery. In Proc. of the CAR'97 (H. U. Lemke, M. W. Vannier, K. Inamura and A. G. Farman, Eds.) pp. 833–838, Elsevier, Amsterdam.

K. J. ZUIDERVELD, B. M. TER HAAR ROMENY AND W. TEN HOVE, (1989a) Fast techniques for automatic local pixel shift and rubber sheet masking in digital subtraction angiography. In: The formation, handling and evaluation of medical images (A. Todd–Pokropek and M. A. Viergever, Eds.) p. 667, Springer, Berlin.

K. J. ZUIDERVELD, B. M. TER HAAR ROMENY AND MAX. A. VIERGEVER, (1989b) Fast rubber sheet masking for digital subtraction angiography. SPIE **1137**, Science and Engineering of Medical Imaging, 22.

Received: May, 1997
Revised: April, 1998
Accepted: May, 1998

Contact address:
Thorsten M. Buzug and Jürgen Weese
Philips Research Laboratories
Division Technical Systems
Röntgenstrasse 24–26
D-22335 Hamburg
Germany
phone: ++49 40 50 78 20 55
fax: ++49 40 50 78 25 10
e-mail: {t.buzug,j.weese}@pfh.research.philips.com

THORSTEN M. BUZUG was born in Lübeck, Germany, in 1963. He received his Diplom-Physiker degree in 1989 and his Ph.D. in 1993 in Applied Physics, both from the Christian-Albrechts-University of Kiel, Germany, where he worked in the field of signal processing applied to chaotic systems. From 1993 to 1994 he had a postdoctoral position at the German Federal Armed Forces Underwater Acoustics and Marine Geophysics Research Institute, where he worked on image acquisition and processing techniques for SONAR applications. In the end of 1994 he joined the Philips Research Laboratories Hamburg, Germany. He has published numerous journal articles, conference papers and one book. He is member of the German Physical Society and the European Physical Society. Currently, he is the leader of the Philips Research Cluster Medical Image Processing.

JÜRGEN WEESE studied Physics at the University of Freiburg, Germany. After receiving his degree in 1990 he went to the Freiburger Material Research Center. He received his Ph.D. degree early in 1993 and continued his work as head of the scientific information processing group. In the end of 1994 he left the Material Research Center and joined the Philips Research Laboratories Hamburg, Germany, where he is currently working in the area of medical image processing and image guided surgery. He has published about 40 regular papers and conference articles in the area of ill-posed problems, computational physics, material research, and medical image processing.