An overlap invariant entropy measure of 3D medical image alignment

C. Studholme\textsuperscript{a, *}, D.L.G. Hill\textsuperscript{b}, D.J. Hawkes\textsuperscript{b}

\textsuperscript{a}Department Diagnostic Radiology and Electrical Engineering, Yale University, BML 332, 333 Cedar Street, New Haven 06511, USA
\textsuperscript{b}Computational Imaging Science Group, Radiological Sciences, UMDS, Guys Hospital, London SE1 9RT, UK

Received 31 October 1997; in revised form 29 April 1998

Abstract

This paper is concerned with the development of entropy-based registration criteria for automated 3D multi-modality medical image alignment. In this application where misalignment can be large with respect to the imaged field of view, invariance to overlap statistics is an important consideration. Current entropy measures are reviewed and a normalised measure is proposed which is simply the ratio of the sum of the marginal entropies and the joint entropy. The effect of changing overlap on current entropy measures and this normalised measure are compared using a simple image model and experiments on clinical image data. Results indicate that the normalised entropy measure provides significantly improved behaviour over a range of imaged fields of view.

Keywords: Multi-modality; 3D medical images; Registration criteria; Information theory; Entropy; Mutual information; Normalisation

1. Introduction

Recently there has been active research into the use of voxel-similarity-based measures of multi-modality medical image alignment [1,2]. Such an approach avoids the task of identifying corresponding anatomical structures in two different modalities, which is difficult to automate for a broad range of clinical data. The use of image similarity measures, especially those derived from information theory, have been shown to allow fully automated alignment in a number of important clinical applications [3,4]. In this paper we are concerned

*Corresponding author.
and from this a set of related measures including mutual information and normalized mutual information are developed and compared. The following sections then compare the behaviour of joint entropy, mutual information and the normalized measure to varying image statistics in their region of overlap. This is achieved using a simple image model and also by examining the behaviour of an automated registration algorithm using these measures as the initial misalignment and field of view of the images is varied.

2. Information measures and image alignment

In imaging the head using a 3D medical imaging device we form an image of measurements of a physical property of material within the patient. Such properties delineate regions of material, for example grey matter, white matter, bone or tumor tissue, within the patient. Different modalities are commonly used in clinical diagnosis and therapy planning which record different anatomical or physiological properties within the patient, and can therefore delineate different structures of clinical interest (for example to distinguish between tumor and non-tumor tissues).

In order to accurately relate different structures of interest delineated by the two modalities we need to align the images into a common coordinate system. In practice these modalities will delineate both common and complementary structure within the images. To define alignment we wish to make use of any shared regions delineated by the two modalities.

2.1. Information measures

The concept of uncertainty can provide a useful basis for developing ideas about information theory [10]. If we examine the probabilities of values occurring in an image, some values are rare and some are common. In the context of image alignment, information measures such as joint entropy, mutual information, and normalized mutual information are useful for quantifying the similarity and alignment of images.
Fig. 2. The information content of a binary symmetric channel: This can be equated to the information content of a binary image consisting of foreground ($m_1$) and background ($m_2$) where $p(m_1) = 1 - p(m_2)$.

predicting what value a voxel has we can form an estimate of the uncertainty of our guess at the value, given the observed distribution of probabilities. If all probabilities are equal, uncertainty in guessing what value a given voxel might have is largest. If there is only one value in the scene then the uncertainty in guessing the value of a given voxel would be zero.

If we learn the value of a measurement (pixel or voxel) that we were very uncertain about (i.e. it would be difficult to guess) then we gain a large amount of information. Conversely, if we learn a value that would be easy to guess because it has a high probability of occurrence, then we only gain a small amount of information. If we have a set of values (for example the set of voxel values occurring in an image) we can then look at the average amount of information provided by the whole set. Essentially, we can say an image with similar numbers of each voxel value contains more information than an image where the majority of voxels have the same value.

There have been many plausible functions proposed to express this concept of information. The most commonly used measure of information which satisfies these requirements is the Shannon–Wiener entropy measure [10]. Here, the average information supplied by a set of $i$ symbols whose probabilities are given by $\{p_1, p_2, \ldots, p_i\}$, can be expressed as,

$$H(p_1, p_2, \ldots, p_i) = - \sum_{s=1}^{i} p_s \log p_s.$$  

(1)

This forms the basis for a large body of work originally termed “communication theory” developed to study the performance of communication systems.

In terms of image processing the measure $H(\cdot)$ can be applied directly to the occurrence of values at points within an image, to form a first-order estimate of the entropy of the image.

2.2. Information measures from two views of a scene

Given a pair of views of a scene $m(x)$ and $n(x)$ provided by different modalities, we can find the sets of values $m \in M$ and $n \in N$ occurring in them. For these we can estimate the marginal probabilities $p[m]$ and $p[n]$ of individual values occurring and then calculate the average information (marginal entropy) for each image:

$$H(M) = - \sum_{m \in M} p[m] \log(p[m]),$$  

(2)

$$H(N) = - \sum_{n \in N} p[n] \log(p[n]).$$  

(3)

If a given value $m$ in a modality delineates a region of material within the image, then we can relate the probability of that value, $p[m]$, directly to the size of that region divided by the total imaged volume. The simplest case is where we have a binary scene with foreground and background values. This can be equated to a binary symmetric channel as illustrated in Fig. 2. Information content is maximised when the area of the two regions are the same.

For a more realistic scene, if we have a large number of different anatomical regions delineated by distinct values, the more similar the probabilities, the more difficult it is to guess a value, and the greater the average entropy of the image.
2.3. Joint entropy and image alignment

In aligning different modalities we wish to use any shared regions in the two images to define the registration. If regions are delineated by both modalities, but the images are misaligned, then in the combined image these regions will be at least partially duplicated. For example, when a transaxial slice through the head is mis-aligned, there may be four eyes and four ears. As the images are brought into alignment the duplication of features is reduced and the combined image is simplified, as shown in Fig. 3.

When two views of a scene are combined by a transformation which maps from points in one view to the other, the joint probability distribution tells us how often pairs of values occur together. The information content of this combined scene can be evaluated in a way similar to that of a single image by forming an estimate of the joint entropy,

\[ H(M, N) = - \sum_{m \in M} \sum_{n \in N} p(m, n) \log(p(m, n)), \]

from the joint probability distribution of values in the region of overlap of the two images.

If we look again at the plot for the binary image (Fig. 2), this shows that the function \( H(m_1, m_2) \) reaches a minimum when one value is highly probable (either foreground or background). By minimising the joint entropy of the two images, we are seeking a transformation where it is easiest to guess what pair of values a voxel will have in the image overlap. The use of joint entropy as a measure of multi-modality alignment has been investigated independently in two recent papers [6,7].

It is important to note here that in using joint entropy as a measure of alignment we are making the assumption that the co-occurrence of the most probable values in the two images is maximised at registration. For many clinical images of the brain from MR, PET and CT this is a valid assumption. If there are large regions in both images, but at true alignment their overlap is not maximised, then joint entropy may not be minimised at registration.
As we vary rotational alignment ($T_1$ and $T_2$) of two images ($m(x)$ and $n(y)$) of a circle we would expect any measure not to favour a particular alignment. Here though direct measures of information such as the joint entropy and mutual information are influenced by the change in image statistics, as the proportion of background in the region of overlap varies. Overall estimate $T_2$ has most similar foreground and background probability and therefore greatest uncertainty and information content.

### 2.4. Overlap statistics

If we take a closer look at the use of joint entropy as a measure of alignment, we can see in practice that there is a problem. In bringing two images into alignment we are interested in essentially comparing two alignments to decide which is better. When deriving a measure of alignment from images, any measure of information content in their overlap will be a function of the structures in the two images falling within this overlap. By minimising joint entropy we are simply trying to find the overlap which contains least information, not necessarily the most corresponding information. What we need to do is to relate changes in the value of the joint entropy $H(M, N)$ to the marginal entropies of the two images $H(M)$ and $H(N)$ derived from their region of overlap.

#### 2.4.1. Mutual information

If we assume that one image falls completely within the other, for all plausible transformation estimates under consideration, then the marginal entropy of the smaller image, say $H(N)$, will be fixed, while the joint entropy $H(M, N)$ and the marginal entropy $H(M)$ will vary as the transformation estimate and the overlap vary. In such a case, in minimising joint entropy $H(M)$, we may choose a transformation estimate which simply minimises the marginal entropy $H(M)$, and we can end up selecting the overlap of image $m(x)$ where the marginal entropy is minimum.

A plausible approach in this case is to maximise the mutual information or relative entropy between the two images. This was proposed independently by Viola and Wells [5] and Collignon et al. [8] as a measure of multimodality alignment. It was initially derived in communication theory as a measure of information shared between the transmitter and receiver at either end of a communication channel. Simply, this expresses the difference between joint entropy and the sum of the marginal entropies,

$$I(M; N) = H(M) + H(N) - H(M, N).$$

A useful way of visualising the relationship between these entropies is provided by a Venn diagram as shown in Fig. 4. Here the size of the circles represents the value of the particular entropy. The overlapping areas represent the mutual information or relative entropies.

#### 2.4.2. An overlap invariant measure

Both mutual information and joint entropy were developed in communication theory, as direct quantitative measures of information. For any given channel they will be dependent on the information transmitted. Fundamentally, we wish to define a measure which can be used to compare two transformation estimates. These estimates may correspond to significantly different regions of overlap within both imaged volumes. The amount of both the “transmitted” and “received” information may vary between the two transformations simply because of the change in overlap. Such a case is illustrated by the Venn diagrams of Fig. 5. Here, when imaging a uniform intensity circle any measure of alignment should remain constant with respect to rotation, since there is no image structure to define rotational alignment. In practice, the
area of the background region, and therefore the probability of the background values, will be a function of overlap and therefore alignment. This will be reflected in a change in the joint entropy and mutual information. Minimising joint entropy seeks a transformation for which either the background or foreground regions predominate in the image overlap and therefore favours the solution on the left of Fig. 5.

Conversely for mutual information, as we vary alignment the marginal and joint entropies in the overlap are identical,

\[ H(M) = H(N) = H(M, N) \]  (6)

and

\[ I(M; N) = H(M) + H(N) - H(M, N) = H(M) \]  (7)

and so maximisation of mutual information seeks a solution where the marginal entropies are maximised, where the foreground and background values are equally probable. This corresponds to the solution on the right of Fig. 5.

This problem can be related back to the derivation of the measures of correlation and correlation coefficient for single modality image matching. Correlation may vary because of alignment and also because of local image statistics. The correlation coefficient [11], being normalised with respect to the local signal, provides a better measure of alignment.

2.4.3. Normalising mutual information for overlap invariance

In order to provide overlap invariance we need a measure independent of changes in the marginal entropies \( H(M) \) and \( H(N) \) of the two images in their region of overlap. A direct approach to normalisation proposed in this paper is to evaluate the ratio of the joint and marginal entropies

\[ Y(M; N) = \frac{H(M) + H(N)}{H(M, N)} . \]  (8)

This we refer to as Normalised Mutual Information. Here, any change in uncertainty of the image values and, therefore, the marginal entropies, will not result in a change in the alignment measure. Maximisation of \( Y(M; N) \) seeks a transformation where joint entropy is minimised with respect to the marginal entropies.
3. The image overlap problem: a simulation

3.1. Image overlap: a simple model

In this section we shall examine the behaviour of measures to the misalignment of an image where there is both a variation in overlap statistics and a true point of alignment to be detected by the measure. We want a simple model where there are two parameters, one controlling misalignment and one controlling the field of view to vary the overlap statistics. In this case for simplicity we will neglect any differences in modality.

The model used represents alignment of 2D coronal slices through the brain by images of a half circle imaged with different fields of view, as illustrated in Fig. 6. Both images have an extent of 1 unit vertically. Horizontally, the first image, for simplicity, has an infinite extent, while the other has a limited extent determined by the parameter $F_{OV}$. A half circle is placed at the centre of the rectangular field of view. We consider in this model the response of different measures to the change in in-plane rotational alignment $\theta$, around the centre of the field of view (and circle).

To evaluate the measures we need to calculate the area of overlapping regions in the two images as a function of the rotational alignment $\theta$ and the field of view parameter $F_{OV}$. This is described in detail for this model in Appendix A.

3.2. Response of entropy measures to varying field of view

Using this simple model, response of measures can be evaluated directly for different rotational misalignment and field of view. Plots for rotation $\theta$ between $-30^\circ$ and $30^\circ$, and field of view parameter $F_{OV}$ between 0.5 and 3 are shown in Figs. 7–9. As can be seen, the response of the joint entropy measure is affected considerably by the
variation in field of view. As the field of view increases the value of the joint entropy increases and the peak at alignment is flattened. Going back to the axioms for information theory (Section 2.1), minimisation of joint entropy favours overlapping images with a large field of view, if the background is already larger than the foreground. In this case as the proportion of background increases, a value is more probably background than foreground and therefore the entropy falls.

A worse response with respect to field of view is exhibited by mutual information as shown in Fig. 8. The behaviour of mutual information is more complex than that of joint entropy, but can be simplified by considering the three components of $dI(M;N)/d\theta$ separately:

$$\frac{dI(M;N)}{d\theta} = \frac{dH(M)}{d\theta} + \frac{dH(N)}{d\theta} - \frac{dH(M, N)}{d\theta}. \quad (9)$$

As we move away from registration, joint entropy may increase so $dH(M, N)/d\theta > 0$. But, as the overlap falls due to misalignment, $H(M)$ and $H(N)$ may not necessarily fall. In this case the “background” region in the image is truncated as the images misalign. If this makes the background and foreground regions more similar in size then, as in Fig. 2, the marginal entropies $H(M)$ and $H(N)$ will increase. So, if this increase is such that as we move away from registration,

$$\frac{dH(M)}{d\theta} + \frac{dH(N)}{d\theta} \geq \frac{dH(M, N)}{d\theta}, \quad (10)$$

then mutual information does not provide a useful measure of alignment. In our model, the field of view increases to a point where this applies and the response to rotational alignment changes from a maximum to a minimum.

The normalised mutual information measure (Fig. 9) provides a response virtually independent of the field of view parameter $F_{OV}$, with a much sharper exponential relationship between field of view and measure. Because here we are dividing the marginal entropies by the joint entropies, any increase in the marginal entropies will be counteracted by the change in the joint entropy.

4. Response to varying field of view

4.1. Experimental data

The problem of recovering alignment is a function of both starting estimate and field of view of the image. In these experiments MR (geometrically corrected), CT and PET image data with an accurately known independent estimate of registration from the Vanderbilt Retrospective Registration Project [12] was used. Registration estimates derived from bone implanted markers in this data (removed from the images prior to experimentation) provide an accurate estimate of rigid registration between the modalities against which successful and failed registrations can be judged. The initial data was acquired to include a stereotactic frame and as a result has a significantly large transaxial field of view.

All data was acquired with contiguous slices (without a slice gap or overlap). The CT images had a sampling resolution of $0.65 \times 0.65 \times 4.0$ mm with 512 × 512 pixels in plane, with between 28 and 34 slices. The PET images had a sampling resolution of 2.6 mm in a plane with 158 mm slices. The spatial resolution of these was appreciably less, with a full width half maximum of around 15 mm. The MR data had a nominal sampling resolution of around 1.25 mm in-plane with 256 × 256, and between 20 and 26 4 mm thick slices. These images were corrected to account for MR scaling errors, resulting in slightly modified voxel dimensions. A number of the images contained significant volumes of space occupying lesions, which were visible on one or more of the acquisitions.
4.2. Registration algorithm

Prior to evaluation and optimisation of transformation parameters, all image pairs were first resampled to a common cubic sampling resolution using tri-linear intensity interpolation to increase the sampling rate and Gaussian blurring to reduce the sampling rate as necessary in each dimension.

In addition to ensuring equivalent sampling resolution, the images were filtered to provide an equivalent spatial resolution (impulse response). This is important because modalities such as PET may have a very broad impulse response (say 8 mm), whereas other modalities such as MR have a much narrower response (say 1 mm). We wish to examine the co-occurrence of measurements in the two modalities which occur both at a corresponding location and, from an equivalent region of material (determined by the impulse response). This filtering was achieved using a Gaussian kernel to decrease the spatial resolution of the higher resolution modality to that of the lower resolution modality. This pre-processing also ensures that the marginal entropy of the initially higher resolution modality does not dominate the registration measure.

To evaluate each measure between MR and CT or PET, a discrete histogram estimate of the joint probability distribution is made using 64 intensity bins for each modality. The histogram is formed directly by incrementing the voxel count in the bin corresponding to the MR and CT or PET values, for every MR voxel with a corresponding voxel in the other modality. Corresponding values in the CT or PET are estimated using tri-linear interpolation of the neighbouring voxel values. The two marginal histograms are calculated from this joint histogram and then the joint and marginal entropies are estimated directly from the counts in these three histograms.

In order to recover rigid alignment we need to find the 6 rigid parameters which provide the optimum value of the registration criteria (maximum of mutual information and normalised mutual information or the minimum of joint entropy). For these experiments a simple multi-resolution hill climbing algorithm was used [13]. All images were further sub-sampled using a Gaussian kernel to form lower resolution versions. This produces a simple
Table 1
Number of successfully recovered alignments (as defined in the text) from 50 random starts at three levels of misalignment, as transaxial field of view is varied for MR-CT (top) and MR-PET (bottom)

<table>
<thead>
<tr>
<th>Measure</th>
<th>10 mm 10°</th>
<th>20 mm 20°</th>
<th>30 mm 30°</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR H(M, N)</td>
<td>TF1</td>
<td>TF2</td>
<td>TF3</td>
</tr>
<tr>
<td>&amp; I(M;N)</td>
<td>50</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>CT H(M, N)</td>
<td>42</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>&amp; I(M;N)</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>PET Y(M,N)</td>
<td>43</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>&amp; I(M;N)</td>
<td>27</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>PET Y(M,N)</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

Fig. 13. An example of the registration failure following optimisation of Mutual Information using the largest field of view (TF1 in Figs. 10 and 11), displayed as a CT bone contour overlayed onto MR values in three orthogonal planes of the volume. Note partial alignment of bone contour with MR, but with a significant rotational misalignment in the coronal plane similar to that investigated by the simulation in the previous section.

Gaussian pyramid where smaller objects are removed as blurring is increased resulting in a smooth function of misalignment between low resolution images at the top of the pyramid. These images containing few voxels can be used to provide an efficient initial registration estimate which is then refined using higher resolution images lower in the pyramid. In these experiments registration was started at an image sampling resolution of $6 \times 6 \times 6$ mm with an optimisation step size of 6 mm and terminated at a resolution of $1.5 \times 1.5 \times 1.5$ mm and step size of 1.5/256 mm.

4.3. Experimental design

The aim of these experiments was to examine the behaviour of the registration algorithm using the different measures when it is presented with clinical data with varying fields of view. In this case we will concentrate on recovering 6 rigid registration parameters as the transaxial field of view is varied.

Varying the in-plane field of view will determine the proportion of the image exhibiting values corresponding to air in the two modalities. In order to look at the effect this has on the registration measures, an image set (patient 5 of the Vanderbilt study) was truncated in-plane at three levels to produce three pairs of images to register. The fields of view selected (labelled TF1, TF2 and TF3) in the MR, CT and PET images are shown in Figs. 10–12.

To examine how robustly different levels of initial misalignment may be recovered using the measures, sets of randomised starting estimates were produced. In these experiments 3 sets of 50 randomised transformations were used. These were derived by perturbing the 6 rigid marker based parameters with random translations and rotations of sizes (10 mm, 10°), (20 mm, 20°) and (30 mm, 30°). These were then used as starting estimates for the optimisation algorithm for each of the three measures.
4.4. Analysis of results

In looking for the successful recovery of alignment we are not interested in differentiating those estimates which are close to registration, we simply need a way of classifying and counting the “outliers” which are failures. If, following optimisation to some minimum step size, the estimate is still far from true alignment, we assume that
In order to then distinguish failures, the mean imaged patient voxel displacement from the voxel positions defined by the marker estimate was evaluated for each resulting registration estimate. The patient voxels were defined by:

- An interactively set threshold in CT or PET highlighting all none air voxels.
- The subset of these voxels which fall completely in the intersection of the MR and CT (or PET) volumes for the marker based estimate.

A registration estimate was then deemed a failure if the mean displacement of these voxels exceeded the MR slice thickness (4 mm) following optimisation down to the smallest step size. Examination of the distribution of estimates for the measures confirmed that this simple approach provided an effective differentiation between “failures” and “successes”.

4.5. Results

4.5.1. MR-CT

The number of successfully recovered misalignments, for each measure and each level of misalignment are included in Table 1. These show an appreciable difference in the behaviour of the system when using different measures as the field of view is varied. Comparison at the poorest starting estimates (30 mm and 30°) provides the strongest contrast between the measures. Minimisation of joint entropy provides robust recovery of alignment until the field of view is reduced to the smallest value (TF3). Maximisation of mutual information conversely provides good recovery of alignment for all but the largest field of view (TF1). The normalised mutual information measure $Y(M;N)$ provides overall the most robust recovery of alignment with only one failure for the largest initial misalignment and smallest field of view.

4.5.2. MR-PET

The results of the MR-PET transaxial experiments are included in Table 1. Of the entropy measures, minimising joint entropy $H(M, N)$ only recovers a significant number of alignments from the closest starting estimates. Interestingly, the results for the middle field of view (TF2) are very much worse than at the larger (TF1) and smaller (TF3) fields of view from the 10 mm and 10° starting estimates. Mutual information provides very good recovery at the two smaller fields of view but, as with the MR-CT experiment, the measure fails significantly at the larger field of view (TF1). Again, normalised mutual information provides the best overall behaviour with only 4 failures from the poorest starting estimates at the smallest field of view.

5. Accuracy over a range of clinical images

5.1. Experimental design

In order to examine the accuracy of estimates provided by the measures over a larger database of images, the
three measures were used to recover alignment between all the geometrically corrected image pairs in the Vanderbilt image database [12]. Here, we use only the original full field of view of the images and look at the accuracy of the single estimate. The starting estimate for the optimisation of the measures were the centres of the two image volumes aligned, and no residual rotations between the imaging planes. This is a typical starting estimate for automated registration in clinical use.

5.2. Results

The graphs of Figs. 14—16 show the mean imaged patient voxel displacement between the estimate for each image pair and the marker based estimate using the three measures to register all images with full field of view. The estimates provided by minimisation of joint entropy include a number of significant failures, particularly for the MR and PET image pairs. The worst performance is that provided by mutual information with the majority of estimates being complete failures, far from the marker-based estimate. This is expected, given the large field of view of all these images, and the experiments of the previous section. In particular for the MR-CT estimates it can be seen that patient image pairs 2, 4 and 7 produce failures for all but one of the proton density MR sequences. This may be a direct result of the location of the patient within the image volume together with the size of the field of view, which is identical for the three MR sequences.

The normalised measures provide MR-CT estimates close to those provided by the bone implanted markers and well within the minimum slice separation of around 4 mm. The quality of the estimate furthest from the marker estimate (the MR-T2 estimate for patient 1) is illustrated in Fig. 17. Results for the PET data are considerably worse, but given the lower sampling resolution (slice thickness 8 mm) and spatial resolution (approximately 15 mm) are clinically acceptable for many applications.

6. Discussion

In this paper we have examined the concepts behind the use of information theory to quantify multi-modality image alignment. The property of overlap invariance was identified as an important factor for clinical image alignment where misregistration can be significant with respect to the size of the overlapping image volume. Previously proposed measures are sensitive to changing image overlap statistics which can limit their ability to recover larger-scale misalignment. A normalised entropy measure is proposed to provide invariance to overlap. The behaviour of this measure with respect to image overlap statistics was then investigated and compared to current approaches using a simulation of a clinical scene. This illustrated the limitations of the un-normalised measures.

These findings were confirmed experimentally in recovering the alignment of clinically typical MR-CT and MR-PET images of the brain. These results contrasted the behaviour of the two un-normalised measures of joint entropy and mutual information with respect to varying field of view. Overall, the normalised form provides significant improvements in the robustness of automated alignment with respect to imaged field of view.

Experiments registering the full database of clinical images from a single starting estimate verify that the measure provides an accuracy acceptable for many clinical applications, over a range of image types and underlying patient anatomies.

7. Summary

Over the last few years the approach of using image similarity as a measure of multi-modality image alignment has been developed and applied to fully automated clinical image registration. Some of the most successful approaches have been those derived from information
theory; in particular, those using the joint entropy and mutual information measures. This paper is concerned with the development of such measures and proposes a simple extension of current approaches which can significantly improve the response to misalignment.

3D medical images are commonly truncated and may have a limited overlap which can vary appreciably with alignment. In practice, direct quantitative measures of information derived from the overlap of a pair of images will be affected by the local image statistics. In order to provide invariance to overlap statistics a normalised entropy based measure of registration is proposed which is simply the ratio of the sum of the marginal entropies and the joint entropy. The response of currently proposed measures are compared with those of the normalised measure using a simple image model.

Experiments comparing the measures are included on clinically typical MR, CT and PET scans of the brain for which an independent gold standard registration transformation exists. These show that as the field of view is varied both the joint entropy and mutual information measures can fail to provide a useful indication of alignment, while the normalised form remains unaffected.

Acknowledgements

The work described in this paper was undertaken with the financial support of UK EPSRC grants GR/H48170, GR/L08519 and GR/K91231 and the clinical PET centre of Guy’s and St. Thomas’ Hospitals. We are grateful to Professor Michael Fitzpatrick of Vanderbilt University, Nashville, U.S.A., for the permission to use the data collected as part of the Retrospective Registration Evaluation Project funded by the U.S. National Institutes of Health Project No. 1 R01 NS33926-01.

Appendix A. Simple image model of alignment and overlap

To evaluate the measures we need to calculate the area of overlapping regions in the two images as a function of the rotational alignment $\theta$ as shown in Fig. 6(a). We can do this by considering the non-overlapping corners of the rectangle in Fig. A.1(b) made up of the two triangles $TA$ and $TB$ shown in Fig. A.1(c). The area of these are determined by the change in the vertical position of the corner of the field of view $q$, given by

$$q = ny - 1/2,$$

where $n = \tan^{-1}(1/2/F_{OV}/2)$. For simplicity of the model, we can constrain the range of $\theta$ so that the other two corners of the rectangle (at the far left and right in Fig. A1(a)) remain inside the region of overlap. Given this we can then calculate $p$ for the triangle $TA$ using $\tan(\theta) = q/p$ which gives the area of $TA$ as,

$$v_{TA}(\theta) = \frac{pq}{2} = \frac{q^2}{2\tan \theta}. \quad (A.2)$$

Similarly the area of triangle $TB$ can be shown to be,

$$v_{TB}(\theta) = \frac{rq}{2} = \frac{(q^2 \tan \theta)}{2}. \quad (A.3)$$

From this, the overlapping area for rotation $\theta$ is given by $1F_{OV} - (v(TA) + v(TB))$ which can be rewritten as,

$$v_{\theta}(\theta, F_{OV}) = 1F_{OV} - q \left[ \tan \theta + \frac{1}{\tan \theta} \right]. \quad (A.4)$$

In addition, we can work out the overlapping and non-overlapping regions of the half circle (radius 1/4)

[Fig. A.1. Evaluating the overlapping field of view of the two images for a given rotation $\theta$ and field of view $F_{OV}$.]
shown in Fig. A.2 which are simply,

\[ v_d(\theta) = \frac{1}{2} \left( \frac{1}{4} \right)^{2 \theta} \left( 1.0 - \frac{|\theta|}{\pi} \right) \]
\[ = \frac{\pi - |\theta|}{32}, \quad -\pi < \theta < \pi, \quad \text{(A.5)} \]

\[ v_b(\theta) = \frac{1}{2} \left( \frac{1}{4} \right)^{2 |\theta|} \left( \frac{|\theta|}{\pi} \right) \]
\[ = \frac{|\theta|}{32}, \quad -\pi < \theta < \pi, \quad \text{(A.6)} \]

\[ v_c(\theta) = \frac{1}{2} \left( \frac{1}{4} \right)^{2 |\theta|} \left( \frac{|\theta|}{\pi} \right) \]
\[ = \frac{|\theta|}{32}, \quad -\pi < \theta < \pi. \quad \text{(A.7)} \]

The area of the ‘background’ region is then simply these subtracted from the overlap area,

\[ v_d(\theta, F_{ov}) = v_o - v_d - v_b - v_c. \quad \text{(A.8)} \]

From these we can then evaluate the probabilities of each region (by dividing by the area of image overlap) and therefore the joint probability distribution, as a function of the transformation parameter \( \theta \) and the field of view parameter \( F_{ov} \),

\[ P(\theta, F_{ov}) = \frac{1}{v_o} \begin{bmatrix} v_d/2 \sqrt{3} & 0 \\ v_b/2 \sqrt{3} & v_d/2 \sqrt{3} \end{bmatrix}. \quad \text{(A.9)} \]

Approaching alignment, \( \theta \to 0 \) and

\[ P(\theta, F_{ov}) \to \frac{1}{F_{ov}} \begin{bmatrix} \pi/32 & 0 \\ 0 & F_{ov} - \pi/32 \end{bmatrix}. \quad \text{(A.10)} \]

References


About the Author—COLIN STUDHOLME completed a B.Eng. in Electrical and Electronic Engineering at the University of Bradford in 1990 and an M.Sc. in Remote Sensing and Image Processing at the University of Edinburgh in 1991. He worked as a student at Marconi Communications Systems and the British Antarctic Survey. Following his MSc he worked at the Royal Signals and Radar Establishment and moved to medical image processing research at UMDS Guy’s Hospital, London, in 1992. Here he completed a Ph.D. at the University of London in 1997. He is currently doing Post Doctoral research in medical image processing at Yale, New Haven, U.S.A.

About the Author—DEREK LG HILL received his B.Sc. in Physics from Imperial College, London in 1987, his MSc in Medical Physics from the University of Surrey in 1989, and his Ph.D. in medical image processing from the United Medical and Dental Schools of Guy’s and St Thomas’ Hospitals, University of London in 1994. He is currently employed as a Lecturer in Radiological Sciences at Guy’s and St Thomas’ Medical School, London, UK.
About the Author—DAVID HAWKES graduated in physics from Oxford in 1973 and obtained his Ph.D. in dual energy CT in 1982. He has worked in medical imaging for nearly 25 yr. In recent years he has concentrated on computational aspects of medical imaging with interests in multi-modal image registration, measurement, shape analysis and augmented reality in medicine with applications in image guided surgery, interventional radiology, oncology and fetal medicine. He has authored or co-authored over 100 refereed publications. He currently heads the Computational Imaging Science Group in Radiological Sciences at GKT, King's College, Guy's Hospital, London.