A Novel Similarity Measure to Evaluate Image Correspondence

Robert Marti
Division of Computer Science
University of Portsmouth, UK
robert.marti@port.ac.uk

Reyer Zwiggelaar
Division of Computer Science
University of Portsmouth, UK
reyer.zwiggelaar@port.ac.uk

Caroline Rubin
Breast Screening Unit
Royal South Hants Hospital
Southampton, UK

Abstract

We have developed a novel similarity measure to evaluate image correspondence. Our method is based on the mutual information between images. The main difference to other mutual information approaches is that we incorporate spatial information using grey-level co-occurrence matrices, leading to a more general measurement. We have used this technique to evaluate two registration algorithms (local affine transformation and thin plate splines) applied to a dataset of mammographic images.

1. Introduction

In many computer vision applications, a comparison between images is needed (i.e. medical imaging, aerial images, image retrieval) but, most of the time, this comparison can not be done by direct subtraction of images. Instead, a registration algorithm is used, which applied to a set of images, deforms the target image trying to match the reference image. Once alignment between the images has been achieved, a direct subtraction can be done, leading to smaller errors. This approach provides a way to detect objects, which appear only in a subset of the images. Many different methods have been proposed in the literature (see [3] and [9] for a general overview).

In this specific field, as in many others in computer vision, there is a need to evaluate the accuracy and robustness of a proposed algorithm for a specific application. In this paper we propose a novel evaluation technique, based on the work developed by Bello and Colchester[1], which uses mutual information to compute a similarity measure between images. In our approach, we have extended the mentioned work by using spatial information provided by grey-level co-occurrence matrices. We have used this technique to evaluate the accuracy of two registration algorithms (local affine transformation [3] and thin plate splines [2]). Such techniques can be used for the detection of abnormal structures in mammography by comparing mammographic images from the same patient obtained at different times. The structure of this paper is as follows. In Section 2 we will discuss image registration techniques, concentrating on local affine and thin plate splines methods. Similarity measures are discussed in Section 3 which is followed by evaluation results and discussion in Section 4. Finally, some concluding remarks can be found in Section 5.

2. Image Registration

The main purpose of image registration is to recover a deformation in corresponding images. Therefore, this process is shown as a mapping between two images (I_1 and I_2),

$$I_2(x', y') = I_1(f_x(x, y), f_y(x, y)) \tag{1}$$

where f_x, f_y are the deformation functions which map spatial coordinates (x, y) in I_1 to coordinates (x', y') in I_2.

We can classify registration algorithms taking different aspects into account. First, complexity of deformation, refers to what kind of deformation our algorithm can deal with. For example, a rigid body deformation only takes into account rotation and translation. An affine transformation extends a rigid body deformation by adding scaling and shearing capabilities. The most complex deformation is a non-rigid transformation, where pixels can move independently. Another classification of algorithms refers to whether the deformation parameters are computed from the whole image (global) or from a subset of pixels (local). Although local algorithms are computationally costly, they give better results dealing with non-uniform deformations. We describe two local registration algorithms, both of which use landmark (control) points to obtain the deformation parameters.

2.1. Local Affine Transformation

This technique applies an affine transformation which, as mentioned above, permits rotation, translation, scaling and shearing [3]. With this kind of deformation some geometrical constraints are preserved (i.e. parallelism), but not scale nor direction. The general 2-D affine transformation is defined as:

$$f_x(x,y) = t_x + a_x x + a_y y \quad f_y(x,y) = t_y + b_x x + b_y y$$

where t_x, a_x, a_y, t_y, b_x and b_y are the affine transformation parameters. In this paper we use a local approach, which solves the above equation for each pixel using its three closest control points.

2.2. Thin Plate Splines

This method, proposed by Bookstein [2], deals with local non-rigid transformations. It uses thin plate splines to solve mapping between 2D points by minimising a measure of bending energy of the thin plate surface. This method computes image mapping by solving the equations

$$f_x(x,y) = t_x + a_x x + a_y y + \sum_{i=1}^{N} w_i U(|P_i - (x,y)|)$$
$$f_y(x,y) = t_y + b_x x + b_y y + \sum_{i=1}^{N} w_i U(|P_i - (x,y)|)$$

where $t_x, a_x, a_y, t_y, b_x, b_y$ are the deformation parameters, N is the number of control points, P_i is the i^{th} control point, (x,y) is any point and U is the thin plate surface equation described in [2].

3. Similarity Measure

As explained in the introduction, having such a huge number of registration algorithms (we have only described two), we have to be able to select the most suitable technique for a particular application. In this section we analyse a novel similarity measure based on mutual information [1]. Mutual information has proven to be very efficient in measuring spatial correspondence between objects and it has been used as a basis of multi-modal registration algorithms [8] [6]. Our approach computes the mutual information of grey-level co-occurrences between images. Grey-level co-occurrence matrices (GLCM) are widely used in computer vision where they have obtained satisfactory results as texture classifiers in different applications [4] [5]. By using GLCM we obtain more information about similarity between two images as GLCMs incorporate spatial information.

3.1. Mutual Information

Mutual information (MI) [1], based on information theory concepts, measures the amount of information we can obtain from one random variable (A) relative to another random variable (B). This measure is related to the entropy of the variable distribution and it is determined by,

$$C(A,B) = \sum_{i,j} p_{AB}(i,j) \log \frac{p_{AB}(i,j)}{p_A(i)p_B(j)} \tag{2}$$

where p_A, p_B are the marginal distribution of the random variables A and B, respectively, and p_{AB} is their joint distribution. Applied to images, each image is identified as one random variable, therefore we can determine the similarity of one image to another by using equation 2. In this case, probabilities are obtained from normalised marginal and joint histograms. As the above measure is dependent on the amount of image overlap, we propose to use normalised mutual information [8] as it is independent on such overlap and is given by,

$$\overline{C(A,B)} = \frac{C(A,B)}{\sum_i p_A(i) \cdot \log \frac{1}{p_A(i)}} \tag{3}$$

3.2. Including Spatial Information

The MI similarity measure, previously discussed, is based on the grey-level histograms of both images. As histograms do not provide any spatial information about the pixels of an image, we believe the use of another metric (instead of histogram information) which will take spatial information into account results in a more robust similarity measure. We propose to use grey-level co-occurrence matrices (GLCM), which are defined by the grey-level co-occurrence between pixels within an image taking a particular translation into account. Each position (i,j) of the GLCM contains the number of pixels with intensities i and j, where the pixels are separated by translation t. Usually, more than one GLCM is computed using various translations t.

We propose to calculate the amount of MI between images using their GLCMs instead of using their histogram distribution. To indicate the validity of this approach, lets consider the following example: Figure 1 we have three images. Using (a) as the reference image, we want to measure the similarity between images (a) and (b) and between (a) and (c). By visual inspection, we would say that (b) is more similar to (a) than the image (c). If we calculate the amount of MI using intensity histograms, we obtain the same value of 0.012189 for (a)-(b) and (a)-(c). This is because there is no overlap between objects so the joint histogram remains constant in both cases. On the contrary, using our GLCM approach combining co-occurrence matrices using various translations we obtain a similarity measure which is not i dependent on such overlap. For (a)-(b) comparison we obtain a measure of 0.183060 while for (a)-(c) the value is 0.037711, which indicates that (b) is more similar to (a) than (c).

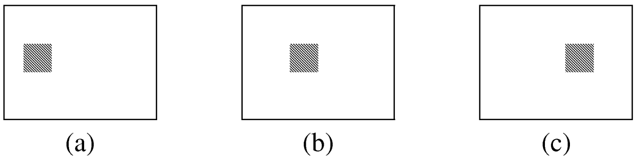

Figure 1. Example of local image structure to demonstrate the validity of the proposed similarity approach (see main text for explanation)

The proposed measure is based on calculating the MI of several GLCM of both images. A GLCM calculated from two images is a four dimensional matrix where each position (i, j, k, l) contains the number of pixels which intensities are (i, j) from the reference image and (k, l) from the other image, in addition the distance between pixels i and j and between pixels k and l is the same and defined by a translation parameter t. From this joint GLCM we can obtain the marginal GLCM, e.g. for the first image,

$$G_A(i,j) = \sum_{k,l}^{K,L} G_{AB}(i,j,k,l)$$

where G_{AB} is the joint GLCM and G_A is the marginal GLCM of the image A (with a similar description for the second image). The MI measure is obtained as a four dimensional extension of equation 3

$$C_{GLCM}(A,B) = \frac{\sum_{i,j,k,l} p_{AB}(i,j,k,l) \log \frac{p_{AB}(i,j,k,l)}{p_A(i,j)p_B(k,l)}}{\sum_{i,j} p_A(i,j) \log \frac{1}{p_A(i,j)}} \tag{4}$$

In order to achieve a measure which is overlap independent we combine MI measures from various GLCM having different translation parameters. The final measure is obtained by calculating the mean of each GLCM.

Figure 2 shows the MI measure between identical images as a function of translation in x on one of the images, which contain a structure similar to the one in figure 1. As stated above, MI using joint intensity histogram remains constant when there is no overlap. Using GLCM with different translations (5 different translations in the example of figure 2) we obtain a wider measure, even when no overlap exists. The measure range is related to the number of translation parameters.

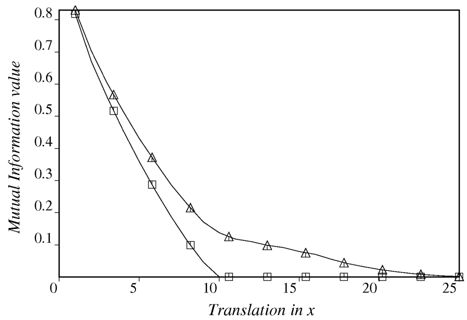

Figure 2. Similarity measure related to translation in x using MI joint histogram (□) and GLCM (△) using 5 different translations

4. Evaluation

As explained in the introduction, we will use our similarity measure as a way to evaluate the accuracy of two registration algorithms described in Section 2. Our aim is to evaluate which method gives better overall results in a concrete application, such as breast cancer diagnosis. A way to help radiologists in the detection of mammographic abnormalities is to compare mammographic images of a patient, either taken sequently in time or using right and left mammograms. Before dealing with this problem, and because both methods use landmark points, we study the relationship between the number of control points and the accuracy of the results. Figure 3 shows evaluation results of both registration methods related to the number of control points used to recover the deformation. Such evaluation has been carried out using four different similarity measures: normalised cross-correlation [7], entropy of the difference image [7], mutual information [1] and our co-occurrence approach. A higher value of the similarity measure denotes a more accurate algorithm, except for the entropy measure where a lower value indicates the same effect. It can be seen from the measures that a number of control points of 16 or 17 achieves maximum accuracy. Results obtained are coherent because a low number of control points can not deal with the complex deformation of the breast, while with a high number of control points the system becomes unstable and no better results are obtained. Figure 4 shows evaluation measures obtained with our approach. From this graph we conclude that thin plate splines obtain overall better results compared to local affine transformation.

The main drawback of our approach is the high computational cost we have to pay for calculating the co-occurrence between images. We have studied the possibility of decreasing such cost following different ways. Firstly, and straightforward, is to reduce the number of grey-levels of our co-occurrence matrix. This would decrease the num-

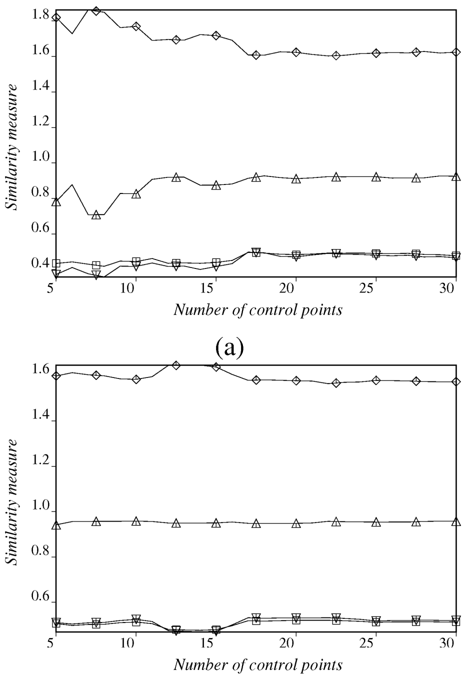

Figure 3. Evaluation of (a) local affine registration and (b) thin plate splines related to the number of control points, using ◇: Entropy of the difference image, △: Normalised cross-correlation □: co-occurrence based MI and ▽: histogram based MI

ber of elements of the GLCM. But this has the handicap of decreasing the accuracy at the same time. Another way would be to eliminate values from the GLCM which do not provide valuable information for our measure, like small or zero values. Finally, we believe the use of dimension reduction methods (like principal component analysis) could be used.

5. Conclusions

We have developed a novel similarity approach, based on mutual information and grey-level co-occurrence matrices. As shown in Section 3, our method takes into account in image spatial information which is omitted by the method of Bello and Colchester [1]. However, we have to pay a highly computational cost for calculating the 4-D joint GLCM. Future work should include a way to reduce computational cost without loose of accuracy. In addition, we plan to study the effects of translation parameters when obtaining the GLCM information related to image contents.

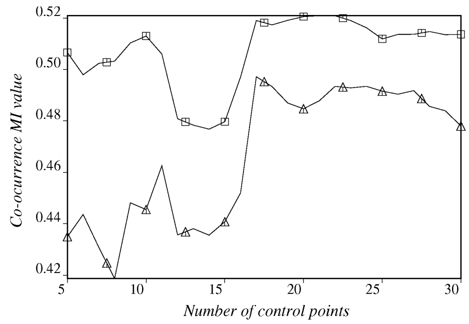

Figure 4. Comparison between local affine (△) and thin plate splines (□) registration related to the number of control points, using co-occurrence based MI

Acknowledgements

This work was supported by EPSRC grant GR/M53387, "Multi-Modality Mammography".

References

[1] F. Bello and A. Colchester. Measuring global and local spatial correspondence using information theory. 1^{st} *International Conference Medical Image Computing and Computer-Assisted Intervention*, Cambridge, USA:964–973, 1998.

[2] F. Bookstein. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[3] L. Brown. A survey of image registration techniques. In *ACM Computing-Surveys*, volume 24, pages 325–376, New York, 1992. ACM Press.

[4] M. Haralick. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804, 1979.

[5] V. Kovalev and M. Petrou. Multidimensional co-occurrence matrices for object recognition and matching. *Graphical Models and Image Processing*, 58(3):187–197, 1996.

[6] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[7] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes. A comparison of similarity measures for use in 2D-3D medical image registration. 1^{st} *International Conference Medical Image Computing and Computer-Assisted Intervention*, pages 1153–1161, 1998.

[8] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.

[9] P. A. van den Elsen, E. D. Pol, and M. A. Viergever. Medical image matching - a review with classification. *IEEE Engineering in Medicine and Biology*, pages 26–39, 1993.