Classification of Medical Images using Non-linear Distortion Models

Daniel Keysers, Christian Gollan and Hermann Ney

Lehrstuhl für Informatik VI, Computer Science Department
RWTH Aachen University, D-52056 Aachen, Germany
Email: {keysers,gollan,ney}@cs.rwth-aachen.de

Abstract. We propose the application of two-dimensional distortion models for comparisons of medical images in a distance-based classifier. We extend a simple zero-order distortion model by using local context within the compared image parts. Vertical and horizontal image gradients as well as small sub images are used as local context. Taking into account dependencies within the displacement field of the distortion by using a pseudo two-dimensional hidden Markov model with additional distortion possibilities further improves the error rate. Using the methods presented in this work, the previous best error rate of 8.0% on the used medical data could be considerably reduced by about one third to 5.3%.

1 Introduction

Classification of medical images is a fundamental step in different applications as e.g. within a medical image retrieval system [1]. Due to the high variability of medical image data it is important to use appropriate models in the classification process [2]. We describe a classification method for medical images based on non-linear image distortion models that considerably improves classification results with respect to other known methods on a database of medical radiographs. The most important improvement that leads to a significant reduction of the error rate is the inclusion of feature vectors using image gradient and image parts as local context instead of the value of one pixel only.

A large variety of methods for classification of medical images is discussed in the literature. A number of these have also been evaluated on the data used in this work, the RWTH Aachen University IRMA (image retrieval in medical applications) database. Best results on these data were achieved using a statistical model incorporating various techniques that cope with the inherent variability of the data [2]. Other techniques like the use of cooccurrence matrices or the Euclidean nearest neighbor yielded higher error rates when applied to this task. The statistical approach with a model of variability (distorted tangent distance) obtained an error rate of 8.0% on the used database of radiographs.

We propose the use of a distortion model for classification, which is connected to the field of image registration by the inherent optimization or matching process. The topic of image registration is a wide research area especially in the

domain of medical image processing and elaborate techniques exist. The fundamental difference to the matching methods that result from an image distortion model is the objective of the matching process: In distortion modeling for classification the aim is to compensate only those deformations that leave the class unchanged. Deformations that change the class are unwanted in the matching, i.e. the emphasis is placed on discrimination between classes. On the other hand in image registration it is usually known that the images are from the same class (e.g. the same body region of the same patient) and the best matching is sought in order to determine the differences between the two images.

2 Methods

The classification methods used here are based on the well-known nearest neighbor classifier. The main contribution is the use of a distance within this classifier that effectively takes into account image distortions. In the asymmetrical matching process, always all of the test image pixels are explained by the reference image pixels.

In previous work [2] two distortion models were found to be especially appropriate for medical images: tangent distance for global deformations and a local zero-order image distortion model (IDM) allowing for small pixel displacements. This model allows to match a pixel of a test image to the best fitting counterpart in the reference image within a small region. This zero-order model does not take into account dependencies between neighboring pixels and the minimization involved in the matching process is therefore computationally inexpensive. On the other hand, tangent distance (which is not used in this work) copes with global affine and brightness transformations computationally efficiently. A further efficient method to enhance classification of medical images is the use of a pixel distance threshold limiting the local pixel-wise distance to a maximum value.

We take into account local dependencies in the matching using two methods:

Local image context. The local context within the images can be represented by using local neighborhoods of e.g. 3×3 or 5×5 pixels in the matching process and for the calculation of the distances. Additionally, the image gradient in horizontal and vertical direction as computed by a Sobel filter can be used to effectively model the local image structure.

Dependencies between displacements. The two-dimensional dependencies can be taken into account by restricting the possible pixel mappings with respect to the mappings of neighboring pixels. The chosen restrictions should ensure monotonicity ('no crossings') and continuity ('no holes') of the pixel displacement field. If complete two-dimensional dependencies are taken into account the matching problem is NP-complete [3] and also known approximation algorithms are computationally expensive. The dependencies can therefore be relaxed in one of the dimensions: e.g. the vertical displacements of pixels of neighboring image columns are not taken into account. This approach results in a pseudo two-dimensional hidden Markov model (P2DHMM) [4]. We propose to extend this model by additionally allowing distortions of each pixel mapping from the

Table 1. Restrictions on the deformation by the different deformation models. IDM: image distortion model; P2DHMM: pseudo 2-dimensional hidden Markov model; P2DHMDM: pseudo 2-dimensional hidden Markov distortion model.

model	restrictions on $(x_{11}^{IJ}, y_{11}^{IJ})$
IDM	$x_{ij} \in \{1,\dots,X\} \cap \{i'-w,\dots,i'+w\},\ i' = \left\lceil i\frac{X}{I} \right\rceil,$ $y_{ij} \in \{1,\dots,Y\} \cap \{j'-w,\dots,j'+w\},\ j' = \left\lceil j\frac{Y}{J} \right\rceil,$ with warp range w, e.g. $w=3$
P2DHMM	$x_{1j}=1, x_{Ij}=X, y_{i1}=1, y_{iJ}=Y,$ $\exists\{\hat{x}_1,\dots,\hat{x}_I\}: \hat{x}_{i+1}-\hat{x}_i \in \{0,1,2\},$ $x_{ij}-\hat{x}_i=0,\ y_{i,j+1}-y_{ij} \in \{0,1,2\}$
P2DHMDM	$x_{1j}=1, x_{Ij}=X, y_{i1}=1, y_{iJ}=Y,$ $\exists\{\hat{x}_1,\dots,\hat{x}_I\}: \hat{x}_{i+1}-\hat{x}_i \in \{0,1,2\},$ $x_{ij}-\hat{x}_i \in \{-1,0,1\},\ y_{i,j+1}-y_{ij} \in \{0,1,2\}$

possible displacement fields [5], resulting in a pseudo two-dimensional hidden Markov distortion model (P2DHMDM).

We briefly give a formal description of the decision process: To classify a test image A with a given training set of references $B_{1k},\dots,B_{N_k k}$ for each class $k \in \{1,\dots,K\}$ we use the nearest neighbor decision rule

$$r(A) = \arg\min_k \left\{ \min_{n=1,\dots,N_k} D(A, B_{nk}) \right\},\tag{1}$$

i.e. the test image is assigned to the class of the nearest reference image. For the distance calculation the test image $A = \{a_{ij}\}, i = 1,\dots,I, j = 1,\dots,J$ must be explained by a suitable deformation of the reference image $B = \{b_{xy}\}, x = 1,\dots,X, y = 1,\dots,Y$. Here, the image pixels take D-dimensional values $a_{ij}, b_{xy} \in \mathbb{R}^{\mathsf{D}}$, where the vector components are denoted by a superscript d. We now want to determine an image deformation mapping $(x_{11}^{IJ}, y_{11}^{IJ}) : (i,j) \mapsto (x_{ij}, y_{ij})$ that results in the distorted reference image $B_{(x_{11}^{IJ}, y_{11}^{IJ})} = \{b_{x_{ij} y_{ij}}\}$. The resulting cost given the two images and the deformation mapping is defined as

$$C(A, B, (x_{11}^{IJ}, y_{11}^{IJ})) = \sum_{i,j} \sum_{\mathsf{d}} ||a_{ij}^{\mathsf{d}} - b_{x_{ij} y_{ij}}^{\mathsf{d}}||^2,\tag{2}$$

i.e. by summing up the local pixel-wise distances, which are squared Euclidean distances here. Now, the distance measure between images A and B is determined by minimizing the cost over the possible deformation mappings:

$$D(A,B) = \min_{(x_{11}^{IJ}, y_{11}^{IJ}) \in \mathcal{M}} \left\{ C(A, B, (x_{11}^{IJ}, y_{11}^{IJ})) \right\}\tag{3}$$

The set of possible deformation mappings $\mathcal{M}$ determines the type of model used. Table 1 shows the restrictions imposed on the mappings by the different models. Since the P2D-models have more restrictions on the relative positions within the displacement field, the minimization process is computationally more complex here. A pre-selection of the e.g. 100 nearest neighbors with a different distance measures like the Euclidean distance can then significantly improve the computation time at the expense of a slightly higher error rate.

Fig. 1. One image from each of the six IRMA-1617 classes: 'abdomen', 'skull', 'chest', 'limbs', 'breast', and 'spine'.
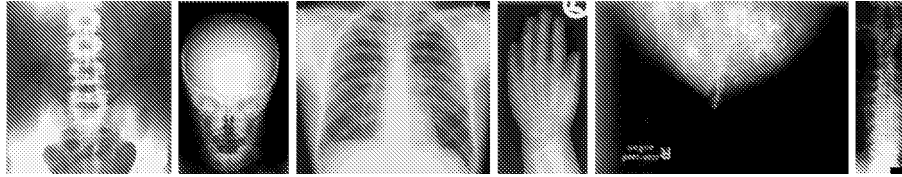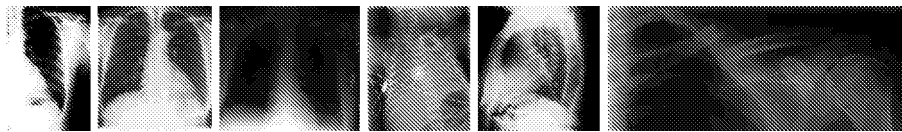
Fig. 2. Several images from class 'chest' from the IRMA-1617 database.

3 Experimental results

The experimental results were obtained on the RWTH Aachen University IRMA database of 1617 secondary digital medical radiographs from the six classes 'abdomen', 'skull', 'chest', 'limbs', 'breast', and 'spine' (IRMA: image retrieval in medical applications [1]). The images were labeled by expert radiologists. They have widely differing sizes and were scaled to a common height of 32 pixels preserving the aspect ratio. One example from each of these classes is shown in Figure 1. The difficulty of this task is due to the fact that a large intra-class variability exists, as shown in Figure 2. The gray values were normalized to span the full gray level range for each image. The error rates are obtained using a leaving one out approach, i.e. each image is classified in turn using the remaining images as training data. This approach ensures that the classifier has 'never seen' the image that is tested and therefore results in a valid test error rate. Some known error rates on the IRMA-1617 database using other methods are given in Table 2 along with the results of the experiments.

Using the proposed techniques for the inclusion of local context information of image gradient (Sobel operator) and local image parts (3×3 sub images), the performance using image distortion and thresholding could be significantly improved from 9.0% to 6.6% error rate. Note that now the feature vector associated with each pixel has the dimensionality $18 = 2 \cdot 3 \cdot 3$ instead of just one value for the pixel gray value. Modeling local dependencies by using the pseudo two-dimensional hidden Markov model and the local context information of the image gradient the error rate could be reduced to 5.7%. Finally, allowing for additional deviations resulting in the P2DHMDM, the error rate could be further reduced to 5.3%. This is a remarkable relative improvement of about one third with respect to the previous best result of 8.0% that included tangent distance.

Table 2. Error rates (ER) for different methods on the IRMA-1617 corpus. NN: nearest neighbor; IDM: image distortion model; P2DHMM: pseudo 2-dimensional hidden Markov model; P2DHMDM: pseudo 2-dimensional hidden Markov distortion model.

reference	method	ER [%]
[2]	cooccurrence matrices	29.0
[2]	Euclidean 1-NN	15.8
[6]	local representations, thresholding	9.7
[2]	kernel densities, thresholding, IDM	9.0
[2]	+ tangent distance	8.0
this work	1-NN, gradients, thresholding – local image parts, IDM	6.6
	– P2DHMM	5.7
	– P2DHMDM	5.3

4 Conclusions

The proposed non-linear image distortion method achieves considerable classification improvements on the used corpus of medical radiographs. The improvements are due to the introduction of local dependencies into the distortion model and the use of context information by using image gradients and local image regions. It is interesting to see that the use of the context information alone already improves the classification considerably, especially for the simple zero-order model. The results are improved by taking into account local dependencies without restricting the possible pixel displacement fields too strictly.

Acknowledgments

References

1. Lehmann T, Wein B, Dahmen J, Bredno J, Vogelsang F, and Kohnen M: Content-Based Image Retrieval in Medical Applications: A Novel Multi-Step Approach. Procs SPIE 3972(32):312–320, February 2000.
2. Keysers D, Dahmen J, Ney H, Wein B, and Lehmann T: Statistical Framework for Model-based Image Retrieval in Medical Applications. J. Electronic Imaging 12(1):59–68, January 2003.
3. Keysers D and Unger W: Elastic Image Matching is NP-complete. Pattern Recognition Letters, 24(1–3):445–453, January 2003.
4. Kuo S and Agazzi O: Keyword Spotting in Poorly Printed Documents using Pseudo 2-D Hidden Markov Models. IEEE Trans PAMI, 16(8):842–848, August 1994.
5. Gollan C: Nichtlineare Verformungsmodelle für die Bilderkennung. Diploma thesis, Lehrstuhl für Informatik VI, RWTH Aachen University, Aachen, September 2003.
6. Paredes R, Keysers D, Lehmann T, Wein B, Ney H, and Vidal E: Classification of Medical Images using Local Representations. Procs BVM 02:171–174, March 2002.