Structural Prototypes for Content-Based Medical Image Retrieval

Benedikt Fischer, Michael Sauren, Mark O. Güld, Ilja Bezrukov, Thomas M. Deserno
Department of Medical Informatics, RWTH Aachen University of Technology, Aachen, Germany

Abstract

The relevant contents of medical images can often be described as a composition of objects with distinct relationships. For a given setting, a generalization by the statistical distributions of regional and relational attributes can be obtained. These yield a structural prototype graph which in turn can be used for the identification of objects in new images. As new image contents are represented by hierarchical attributed region adjacency graphs, the task of object identification corresponds to the problem of inexact graph matching. For this purpose, the graph edit distance as well as a Hopfield-net are used and evaluated for the example application of bone-identification in hand-radiographs. The structural prototypes improve recall by 9% and 17%, respectively, in comparison to a traditional approach without relational information.

1 Introduction

Due to the high inter- and intra-patient variability of depicted regions of interest it is infeasible to define precise normative values for the attributes of an object-class to be identified. For example, the regional attributes of the metacarpal bones in radiographs are not discriminative enough to allow a distinction between the single fingers. Even for a human observer it is impossible to identify the corresponding finger without knowledge of a bone’s surrounding. However, this becomes easier when relations to other bones (position, size, etc.) are known.

A structural prototype graph therefore aims to describe the relevant objects (nodes) and their relations (edges) as a scene. Since only relevant objects are contained, the graph is very small. In the example application, the 19 hand bones are represented by an identical number of nodes.

During the partitioning of images with unknown content and resolution, each pixel could belong to different segments. Hence, a generic data representation is required, which preserves these ambiguities. For this purpose, an iterative region-growing concept is applied, transforming the image to a hierarchical attributed region adjacency graph (HARAG) [1]. The resulting graphs are rather large. In the example application, the HARAGs for images of 256 pixels maximum width and length lead to graphs between 1,000 and 6,000 nodes.

As prototype and image are both represented as graphs, the object identification is equivalent to the problem of inexact graph matching for which many established solutions exist [2]. Due to the huge search space, we chose two approximation algorithms: a Hopfield neural network (NNGM) [3] as well as the Graph Edit Distance (GED) [4].

2 Methods

2.1 Structural Prototype Graph

The prototype graph represents objects by nodes and object-relations by edges. Nodes and edges are attributed by the means and standard deviations of the regional and relational attributes, respectively.

Means and standard deviations are obtained for the manually labeled segments. The similarity of nodes and edges of the prototype to the HARAG is computed by the Mahalanobis distance with an additional class-dependent weight for each dimension. The weights were adjusted on the training data by sequential forward selection so that the classification rate was maximized.

2.2 Graph Matching

For the Hopfield net, an association graph (AG) between prototype and HARAG nodes is used. Each AG node represents one matching hypothesis. A pre-filtering is used to only include pairs of similar nodes in the AG. The regional similarity is used as initial potential for the AG nodes. The final matching is derived by a gradient descent optimization process. Matching hypotheses are strengthened or weakened by similarity to regional or relational distributions. The process iterates until a stable state is reached.

The well-known graph edit distance transforms one graph into another with pre-defined transform costs for substitution, deletion and insertion of nodes or edges. The matching is optimal if the sum of all transforms is minimal. In order not to cover the whole search space, a heuristic is used additionally: Beginning with an empty partial matching, in each iteration step the transform with the minimal costs is chosen to extend the partial matching. This is repeated until an end-state (full matching) is reached. To further reduce the complexity, a heuristic function estimating the costs following a selection is used.

3 Experiments

Aim of the experiments was the identification of 19 hand bones in 96 hand radiographs. The HARAG segmentations led to 96 graphs with a total of 221,615 regions and 1,000 to 6,000 nodes per graph, on average 2,300 nodes. The ground truth was established by manually labeling the bones for each image if they had been segmented, resulting in 1290 labels. **Fig. 1** displays the labels (left) and the percentage of how often each bone class occurred in the segmentation.

34 regional attributes were computed for shape, intensity, and texture. 12 relational attributes were used for topology, size, and intensity relations.

The experiments were conducted using a six-fold cross validation.

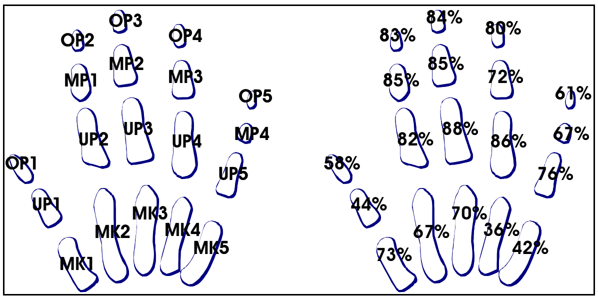

Fig. 1: Bone class labels and percentage of occurrences in the ground truth data.

4 Results

The results as shown in **Tab. 1** and **Tab. 2** clearly demonstrate the beneficial use of structural information. With the exception of the MPh bone group, where GED and regional classification are roughly the same, GED and NNGM always improve the classification results. The classification using only regional, i.e. no relational information, classifies 60.16% correctly, the GED 65.2% (9% increase) and the NNGM even 70.5% (17% increase). The computation time for the prototype-generation was 15 seconds per prototype. The classifications for all experiments took on average about 30 seconds per graph independent of the approach. Runtimes were taken on an AMD Athlon 1200.

5 Discussion

Considering that only 19 of on average 2,300 nodes, i.e. 0.8%, in the HARAGs represent hand bones, an overall classification rate of 70.5% is considered as satisfactory. Even more, as the algorithms may have classified regions similar to the real bones but not labeled in the ground truth as such. In future work, we will concentrate on improving the segmentation quality, as it is crucial for the information content of the HARAGs. Furthermore, we will investigate different models for node and edge prototypes in combination with Bayesian classification.

group	regional	GED	NNGM
MK	64,5%	69,9%	75,4%
UPh	70,4%	74,0%	78,1%
MPh	64,7%	64,0%	75,7%
OPh	42,5%	55,0%	54,4%
overall	**60,2%**	**65,6%**	**70,5%**

Tab. 1: Classification rate in percent of ground truth by bone-group (see Fig. 1); regional: no relations used, GED: graph edit distance, NNGM: Hopfield neural net

class	regional	GED	NNGM	class	regional	GED	NNGM
MK1	67,1%	72,9%	74,3%	MP2	54,9%	51,2%	64,6%
MK2	70,3%	68,8%	78,1%	MP3	72,0%	75,6%	82,9%
MK3	68,7%	80,6%	79,1%	MP4	72,2%	73,6%	87,5%
MK4	68,6%	60,0%	74,3%	MP5	59,4%	54,7%	67,2%
MK5	40,0%	57,5%	67,5%	OP1	37,5%	57,1%	57,1%
UP1	64,3%	52,4%	69,0%	OP2	35,0%	55,0%	53,8%
UP2	73,4%	72,2%	74,7%	OP3	54,3%	59,3%	54,3%
UP3	70,2%	72,6%	75,0%	OP4	46,8%	57,1%	61,0%
UP4	80,7%	85,5%	85,5%	OP5	35,6%	44,1%	44,1%
UP5	58,9%	76,7%	82,2%	**overall:**	**60,2%**	**65,6%**	**70,5%**

Tab. 2: Classification rate in percent of ground truth for each bone class; regional: without relational attributes, GED: graph edit distance, NNGM: Hopfield neural net

6 References

[1] Lehmann TM, Beier D, Thies C, Seidl T: Segmentation of medical images combining local, regional, global, and hierarchical distances into a bottom-up region merging scheme. Procs SPIE 2005; 5747:545-555

[2] Conte D, Foggia P, Sansone C, Vento M: Thirty years of graph matching in pattern recognition. International Journal of Pattern Recognition and Artificial Intelligence 2004; 18(3): 265-298

[3] Lappe C, Fischer B, Thies C, Güld MO, Kohnen M, Lehmann TM: Optimierung eines konnektionistischen Graphmatchers zum inhaltsbasierten Retrieval medizinischer Bilder. In: Tolxdorff T, Braun J, Handels H, Horsch A, Meinzer HP (Hrsg) Bildverarbeitung für die Medizin 2004, Springer-Verlag, Berlin, 2004; 338-342

[4] Bunke H, Allermann G: Inexact graph matching for structural pattern recognition. Pattern Recognition Letters 1983; 1: 245-253.