

Scene Analysis with Structural Prototypes for Content-Based Image Retrieval in Medicine

Benedikt Fischer^{*}, Michael Sauren, Mark O. Güld, Thomas M. Deserno
Department of Medical Informatics, RWTH Aachen University of Technology,
Pauwelsstr. 30, D-52057 Aachen, Germany

ABSTRACT

The content of medical images can often be described as a composition of relevant objects with distinct relationships. Each single object can then be represented as a graph node, and local features of the objects are associated as node attributes, e.g. the centroid coordinates. The relations between these objects are represented as graph edges with annotated relational features, e.g. their relative size. Nodes and edges build an attributed relational graph (ARG). For a given setting, e.g. a hand radiograph, a generalization of the relevant objects, e.g. individual bone segments, can be obtained by the statistical distributions of all attributes computed from training images. These yield a structural prototype graph consisting of one attributed node per relevant object and of their relations represented as attributed edges. In contrast to the ARG, the mean and standard deviation of each local or relational feature are used to annotate the prototype nodes or edges, respectively. The prototype graph can then be used to identify the generalized objects in new images. As new image content is represented by hierarchical attributed region adjacency graphs (HARAGs) which are obtained by region-growing, the task of object or scene identification corresponds to the problem of inexact sub-graph matching between a small prototype and the current HARAG. For this purpose, five approaches are evaluated in an example application of bone-identification in 96 radiographs: Nested Earth Mover's Distance, Graph Edit Distance, a Hopfield Neural Network, Pott's Mean Field Annealing and Similarity Flooding. The discriminative power of 34 local and 12 relational features is judged for each object by sequential forward selection. The structural prototypes improve recall by up to 17% in comparison to the approach without relational information.

Keywords: Object identification, graph matching, classification, neural net, optimization problem

1. INTRODUCTION

The diagnostically most important information in medical images is rarely deduced from the whole image, but is restricted to specific regions of interest (ROIs)¹. While global features describing the complete image have been successfully used to detect the image category according to imaging modality, direction of acquisition, anatomical region or biosystem², they are inadequate for local analysis. Yet many applications require the analysis of local regions of interest, e.g. the detection of fractures or tumours, while the rest of the image is mainly important for orientation.

Due to the high inter- and intra-patient variability of the ROIs in medical images, it is infeasible to define precise, standard values for their attributes¹. For example, the local attributes of the metacarpal bones in radiographs are not discriminative enough to allow a distinction between the fingers. This still holds even if the images are acquired with the hands in the same orientation, with identical resolutions and captured with the same modality, showing the same part of the wrist or arm. Even for a human observer it is impossible to identify a single finger bone without knowledge of its surrounding. However, this becomes easier when relations to the other bones (position, size, etc.) are known. Hence, in applications where multiple objects need to be identified, and if the constellation of the objects is characteristic of that application, this fact can be used as additional information for the identification of the single objects^{3,4}. Once identified, these individual objects can be analyzed in more detail and with highly specialized features and methods. A structural prototype graph therefore aims to support their identification by generalizing their *structure* through the statistical description of the single objects as well as their relations between each other.

^{*} Corresponding author: Benedikt Fischer, Department of Medical Informatics, Aachen University of Technology (RWTH), Pauwelsstr. 30, D-52057 Aachen, Germany, email: bfischer@mi.rwth-aachen.de; phone +49 241 80 85174; fax +49 241 80 33 85174; web: <http://irma-project.org>

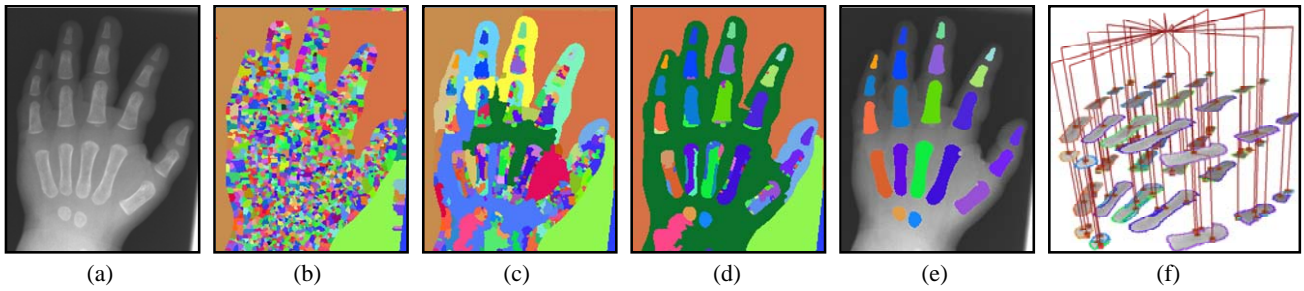


Figure 1: Selected steps of the partitioning process for a skeletal hand radiograph: (a) original image, (b) start of merging after initial watershed segmentation, (c) to (d) intermediate merging steps, (e) bone regions contained in merging hierarchy, (f) Subgraph of the complete HARAG concerning only the bone regions and the background (root of the graph).

In earlier works, the general representation of images of unknown content by the partitioning of images into hierarchical graphs has been introduced⁵. The used partitioning scheme is based on bottom-up region merging and is able to preserve almost arbitrary granularity within the image. The merging process (Fig. 1), which is based on local, inter-regional, and hierarchical distances, yields an inclusion hierarchy for the regions. As it also records the adjacency relations between regions as well as their local features, the resulting data-structure is consequently termed hierarchical attributed region-adjacency graph (HARAG).

Since new images of unknown content as well as the structural prototype are both represented by graphs, the task of object identification is equivalent to the problem of inexact sub-graph matching for which many established solutions exist⁶. The most limiting factor for the choice of graph matching algorithms is the size of the graphs to be matched. The benefit of the generic representation of unknown image content by HARAGs comes at the expense of rather large graph sizes. In the evaluation application, even small hand radiographs of maximum size 256x256 pixels lead to HARAGs between 1,000 and 6,000 nodes. The prototype graph, on the opposite can be kept very small, as it contains only the relevant objects, e.g. the hand bones. For a node-to-node matching this still yields a huge search space, prohibiting the use of brute force algorithms for optimal solutions, demanding approximating algorithms.

In this paper, a variety of graph matching algorithms is evaluated in an example application belonging to the field of computer-aided diagnosis. For certain scenarios, the real age of a patient needs to be compared to his/her bone-maturity. The single hand bones develop in individual order and together with the ossification, they can be analyzed by complex predefined schemes to estimate the bone age⁷ (Fig. 2). All approaches are compared to each other by their recall concerning the identification of the single bones which is of great help in the bone age determination. Furthermore, the impact of regarding the relational structure is evaluated by comparison to using only local features.

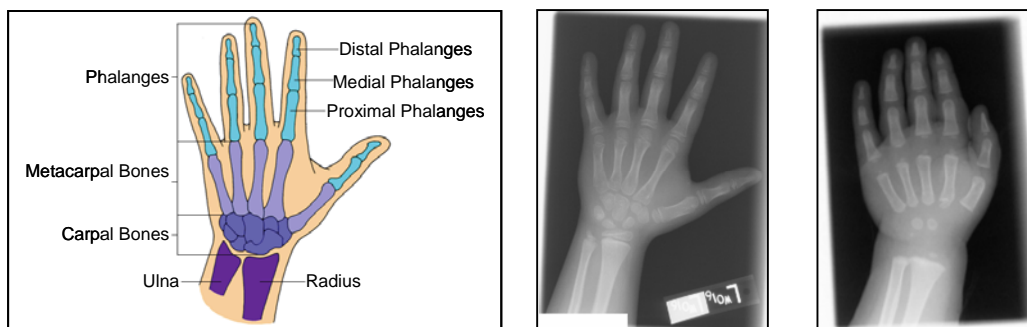


Figure 2: Taxonomy of bony elements of a human hand (left); two skeletal radiograph examples acquired in clinical routine (middle and right). Note the variety in position, exposure, and form of relevant bones characterizing the age of the patient.

2. METHODOLOGY

2.1 Structural Prototype Graph

The prototype graph represents objects by nodes and object-relations by edges. In contrast to the HARAG, where the node attributes are instances of local features, e.g. the coordinates of the center of gravity, the attributes of the prototype graph nodes represent the statistical distribution of such local features. The edges in the HARAG are binary, i.e. hierarchy (inclusion) or adjacency between two regions are stored, but not quantified and no other relations are used. In the prototype graph, the relational attributes are again statistical distributions of (relational) features, e.g. the relative orientation of the regions corresponding to the connected graph nodes. In a first approach, unimodal Gaussian distributions have been assumed for the features.

The complete process for synthesizing the structural prototype graph is illustrated in Figure 3: In order to obtain the means and standard deviations of the features, these are obtained from manually labeled segments within the HARAGs of training images. For the labeling, the user clicks onto a pixel in the image and scrolls through the regions in which this pixel is contained in the hierarchy. When the correct region is shown, the region number and label are recorded. The labeled nodes of the HARAG are then extracted and additional local features as well as the relational features beyond the attributes in the HARAG are computed. After all HARAGs have been processed in this manner, the distributions of the assembled local and relational features are calculated and used as attributes for prototype graph's nodes and edges, respectively.

The similarity of nodes and edges of the prototype to other graphs is computed by Mahalanobis distance with an additional class-dependent weight for each dimension:

$$g_k(x) = \exp\left(-\frac{1}{2} \sum_{d=1}^{D_k} a(k)_d \frac{(x_d - \mu(k)_d)^2}{\sigma^2(k)_d}\right) \quad (1)$$

where x denotes the feature value, k the object class, e.g. metacarpal bone of the left thumb, D_k the dimensionality of class k , i.e. the number of features used, d corresponds to the current dimension, i.e. feature, $\mu(k)_d$ and $\sigma^2(k)_d$ stand for mean and variance of the d -th feature of class k . The class-dependent weights $a(k)_d$ are adjusted on the training data by

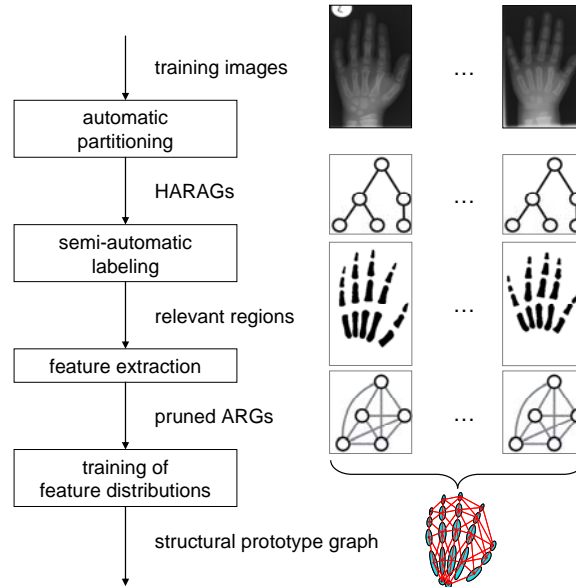


Figure 3: The required training process for synthesis of the structural prototype. Processing steps are shown inside boxes, input and output next to the edges. All steps except for the labeling of the relevant regions in the HARAGs are performed fully automatically.

sequential forward selection so that the classification rate of the corresponding feature set is maximized. For this purpose, the feature sets are built up beginning with the (according to the classifier) best single feature, increasingly adding the next single feature which enhances the classification the most, until no such feature exists anymore. Once a feature has been added, it is not removed at later stage.

In summary, the synthesized structural prototype graph

- represents only objects (classes) relevant to the specific application, one node per object
- all objects are set in relation to each other
- each object has an individual set of local features assigned
- each relation has an individual set of relational features assigned
- mean and standard deviation are recorded for each feature

2.2 Graph Matching

New image material is at first transformed into a HARAG, as this allows for a generic representation independent of resolution or specialized feature sets. On the contrary, the structural prototypes are specialized for certain objects with specific feature sets for each object and object relation. The HARAG therefore needs to be transformed into a corresponding attributed relational graph (ARG). The complete transform process until the actual object identification is illustrated in Figure 4: After the automatic partitioning of the image, the image content is represented by a HARAG. In the next step, the local feature sets of the HARAG and all objects of the prototype graph are compared. Missing or superfluous local features are added to or deleted from the HARAG, respectively, yielding an enriched HARAG.

In order to minimize the computational complexity of the graph matching later on in the process, HARAG nodes which are very unlikely to match any prototype node are removed. For this purpose, equation (1) is used to compute the local similarity to all prototype objects, i.e. for all k in (1). A user-defined threshold is then applied to remove all nodes without likely matching partners from the HARAG.

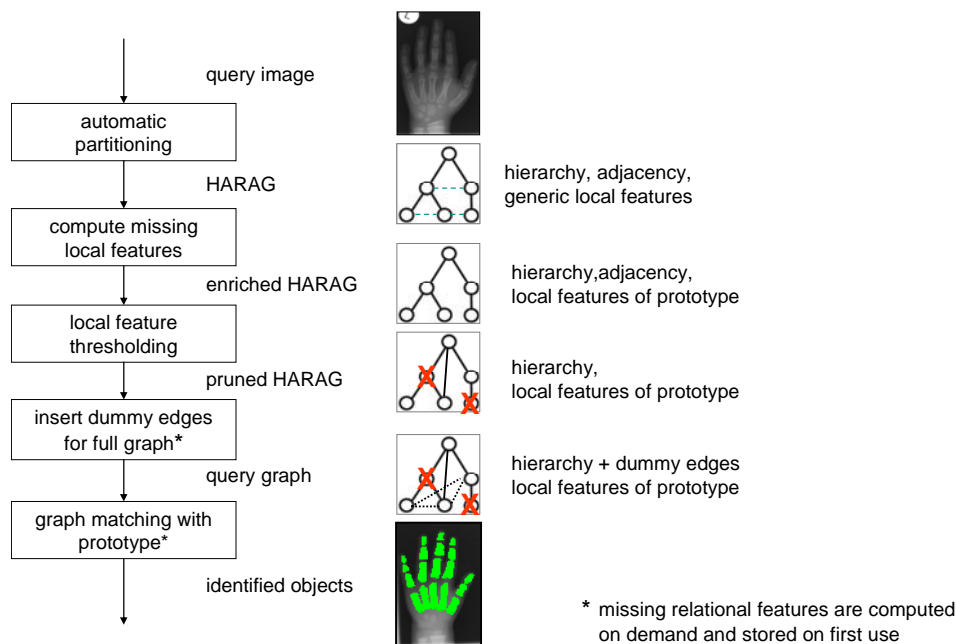


Figure 4: Overview of the identification process of the objects in the structural prototype graph in new images. Actions are shown inside boxes, input and output next to the edges.

The pruned HARAG is afterwards transformed into a query graph, by pair-wise insertion of missing edges between all nodes. Since the computation of the full relational feature set for a full graph is rather expensive, the relational computation is done on demand during the graph matching process. Then, just the relational features needed for the current comparison, i.e. the relational feature set of the concerned prototype edges, are computed and stored at first use. For recurring computations, the stored features can be re-used. Depending on the used computer architecture, this constitutes a trade-off between computation- and memory access-time. After the transform from pruned HARAG to query graph, the different approximation schemes for the graph matching can be applied.

The following approaches were chosen for evaluation, offering a good variety from established solutions for similar optimization problems from various application fields: the Nested Earth Mover's Distance⁸, as a variant of the Earth Mover's Distance known from image retrieval^{9,10}, the Graph Edit Distance (GED)¹¹, itself a well-respected graph matching version derived from string matching, a Hopfield neural network (NNGM)¹², originally applied to the matching of chemical molecules¹³, Pott's Mean Field Annealing¹⁴, a deterministic version of the standard artificial intelligence approach Simulated Annealing¹⁵, as well as Similarity Flooding¹⁶ having been introduced for relational database schema matching.

Earth Mover's Distance

The EMD has been used for image retrieval and analysis as a metric between two distributions, e.g. color histograms^{8,9,17}. It is based on the well-known simplex algorithm¹⁸ for to the transportation problem. Here, the optimization problem is defined as

$$\max\{c^T x \mid Mx = b, x \geq 0\}$$

with $M \in \mathbb{R}^{m \times n}$ a real-valued matrix, $c \in \mathbb{R}^n$ a target function vector and $b \in \mathbb{R}^m$ a vector of constraints. The goal is to find a solution for the above linear equation system, maximizing the value $c^T x$. For the computation of the EMD, a distance matrix for all possible matching hypotheses is computed which then serves as input for the simplex algorithm. This speeds up the iterations of the simplex, since access to single graph elements is required only when building the distance matrix, instead of frequent re-computations during its iterations. Instead of directly using signatures⁸ for input, we use a real-valued distance matrix, which is calculated from the signatures and the distance measure.

An application of the EMD for graph matching is known as a nested EMD (NEMD)⁷. In a first step, the EMD is computed for each pair of nodes, taking into account node and edge attributes. The results of these computations then are used as input for an "outer" EMD, i.e. the elements of the distance matrix used are now the distance matrices of the inner EMD. The outer EMD therefore serves to estimate the overall similarity between the compared graphs and leads to a final node matching.

Graph Edit Distance

The GED is based on the Levenshtein distance for the comparison of strings and has been adapted for graphs⁸. The GED computes the costs to transform one graph into another. The transformation is achieved by edit operations, i.e. deletion and insertion of nodes or edges and label substitution. Each edit operation is assigned a pre-defined cost. The matching is optimal if the sum of all edit costs is minimal. In order to avoid exploring the complete search space, an additional heuristic is used: Beginning with an empty partial matching, in each iteration step the edit operation 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.

Neural Network Graph Matcher

The NNGM¹¹ employs a Hopfield net¹⁹ to solve an optimization problem. A fully connected neural Hopfield net is used for the minimization of an energy function, guaranteeing the net to convergence of the net to a minimum¹⁹. The NNGM has been adapted from an earlier approach featuring a Hopfield net for the structural comparison of chemical molecules¹².

The Hopfield net is built by an association graph (AG) between prototype and query graph¹¹. Each AG node represents one matching hypothesis. A pre-filtering is used to include only pairs of similar nodes in the AG. The local 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 local or relational distributions. The process iterates until a stable state is reached.

Pott's Mean Field Annealing

PMFA is a variant of the simulated annealing technique¹². The aim is to approximate the stochastic simulated annealing by a deterministic approach leading to better run-times and less sensitivity to parameter settings. For the application, an AG and a Hopfield net as described above is used. The energy function is re-written as an integral and evaluated at its saddle point. In the classical Ising model, where neurons have a binary state (active or inactive), many invalid solutions exist, where a neuron η_{ij} with source node p_i is active for more than one target node q_j . The usage of so-called Potts-neurons constrains the solution space such that the state vector is binary containing only a single component set to active. This reduces the solution space from 2^k to k for each AG-node. The state of a neuron is computed with the help of mean-field equations for the potential and output of a neuron, leading to the method's name. By iterative solving of these equations a gradient descent in the energy topology is performed.

Similarity Flooding

The key idea of SF¹⁴ is that for two nodes of different graphs to be similar, their neighborhoods in each graph must also be similar. The method has been derived for matching relational database schemata, yet its approach is generic and also suitable for the application at hand. The approach also uses an association graph (AG) to describe possible node matchings. Each node in the AG again stands for a matching hypothesis of two nodes, attributed by their direct similarity comparison. The nodes are connected by edges with propagation coefficients controlling the spread of the similarity to neighboring nodes. The spreading of similarity is performed iteratively until stability is reached or after a fixed amount of iterations. The AG is synthesized by pair-wise adding two adjacent nodes of each graph: Two nodes p and p' of a first graph P are inserted as matching nodes (p,q) and (p',q') , if the corresponding nodes q and q' of the second graph are similar and their relations (p,p') and (q,q') are also similar.

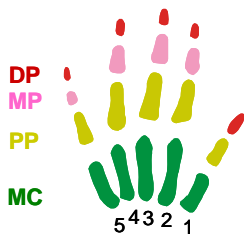
2.3 Experiments

Aim of the experiments was to assess the benefit of our new concept for structural scene analysis in content-based image retrieval. For this purpose the application of hand bone identification was chosen as it is

- comparably simple to implement,
- relevant, as it is supportive for medical routine (bone maturity estimation), and
- a ground truth can be easily established.

The different approximation methods had to identify 19 hand bones in each of 96 hand radiographs. In our evaluation, only the meta-carpal bones to the upper phalange have been used. The carpal bones were omitted in the conducted experiments, as their number and constellation change significantly with age and for a reliable identification, separate structural prototype graphs would have to be used for multiple age groups for which not enough image material would have been available.

The HARAG segmentations led to 96 graphs with a total of 221,615 regions. The HARAGs contained between 960 and 6,000 nodes, on average a graph consisted of 2,300 nodes. The ground truth was established by manually labeling the bones for each image if they had been segmented, resulting in 1,290 labels. This means, that for certain bones, no corresponding segment was contained in the HARAG due to segmentation problems. Fig. 5 displays the labels used (left) and the percentage of how often each bone class occurred in the segmentation (Tab. 1).



	5	4	3	2	1
DP	61%	80%	84%	83%	58%
MP	61%	72%	85%	85%	n.a.
PP	44%	82%	88%	86%	76%
MC	42%	36%	70%	67%	73%

Figure 5: Bone class labels refer to meta-carpal (MC) bones, proximal (PP), medial (MP), and distal phalanx (DP), and the numbering of the fingers from 1 to 5.

Table 1: Percentage of occurrences of bone class labels in the ground truth data.

Table 2: Recall in percent of ground truth by bone-group (see Fig. 1); local: no relations used, NEMD: Nested Earth Mover’s Distance, GED: Graph Edit Distance, NNGM: Neural Network Graph Matcher (Hopfield net), PMFA: Pott’s Mean Field Annealing, SF: Similarity Flooding

group	local	NEMD	GED	NNGM	PMFA	SF
distal phalanges (DP)	42.5%	48.7%	55.0%	54.4%	42.8%	62.9%
medial phalanges (MP)	64.7%	70.7%	64.0%	75.7%	64.7%	76.3%
proximal phalanges (PP)	70.4%	70.6%	74.0%	78.1%	70.4%	75.1%
metacarpal bones (MC)	64.5%	72.1%	69.9%	75.4%	64.9%	64.9%
overall	60.2%	65.0%	65.6%	70.5%	60.3%	69.8%

The prototypes featured the distributions of 34 local attributes on shape, intensity, and texture and 12 relational attributes concerning topology, size, and intensity relations. Experiments were conducted on the recall of the approximation algorithms using a six-fold cross validation. In order to judge whether the use of relational features indeed improves the matching process, one experiment was performed with the NNGM using only local features.

3. RESULTS

The numerical results of all experiments are provided in Table 2 according to the bone group, while Table 3 lists the detailed classification rates for each bone and method. The given percentages have been averaged by the six-fold cross validation. Overall, the method using only local, i.e. no relational information, classifies 60.16% correctly, the NEMD 65.0% (8% increase), the GED 65.2% (9% increase), the NNGM even 70.5% (17% increase), the PMFA 60.5% (0.2% increase), the SF 69.8% (16% increase).

The computation time for the prototype-generation was 15 seconds per prototype. The classifications experiments for local features, NEMD, GED, NNGM, PMFA, and SF took on average about 20, 21, 18, 30, 5, and 5 seconds per graph, respectively. Runtimes were taken on a 2.4 GHz Intel Xeon PC with 2 GB RAM.

4. DISCUSSION

The NEMD provides noticeable increases in the classification rate compared to the local method. The group of PP is the only group with no more than slight improvements to local features alone. In detail, this fact can be traced to the results from the bones PP2 and PP4, yet no explanation for this behaviour and a deeper analysis is pending. The GED shows a similar behavior. While the overall increase is roughly identical to that of the NEMD, it shows a weak point in one bone group (MP), especially for bone MP2 and MP5 where a decrease of the recall is observed. PMFA yields a disappointing performance. Even although only on bone MC4 a recall below that of the local method is given, the increases on the other bones are negligible. The results seem to indicate problems with the chosen parameters²⁰. We therefore have conducted further experiments with the weights for the influence of relational features compared to local features, start temperature, step-size for the gradient descent, the penalty factor and number of iterations per temperature, yet no better results were obtained. A computed parameter sensitivity (standard deviation of the recall values with different parameters normalized by mean recall-value) also shows a low dependency on the chosen parameters and validation experiments were successful. For the future, the PMFA will have to be tested in further medical applications. Apparently, the NNGM delivers the most consistent performance, with evident increases for all bone groups compared to the local features. It also leads to the overall best result. SF almost reaches the same overall recall, yet with increasing success from metacarpal bones to upper phalanges. The bones become smaller in this direction, with more likely matching hypotheses, as smaller regions outnumber bigger ones in the HARAG. This leads to an increased interconnection of the smaller regions in the association graph and to an increased “flow” of similarities while the bigger regions will receive input from less nodes, since the AG does not include “punishing” hypothesis as the NNGM does.

Overall, the experiments clearly demonstrate the beneficial use of structural information. With the exception of Pott’s Mean Field Annealing, which is almost identical to the local classification, all methods mainly improve the classification results. The NEMD leads to an increase of 8%, the GED to 9%, the NNGM even to 17%, the PMFA to 0.2%, and the SF to a 16% increase of the performance compared to the strict local search.

Table 3: Recall in percent of ground truth for each bone (see Fig. 1); local: no relations used, NEMD: Nested Earth Mover’s Distance, GED: Graph Edit Distance, NNGM: Neural Network Graph Matcher (Hopfield net), PMFA: Pott’s Mean Field Annealing, SF: Similarity Flooding

class	local	NEMD	GED	NNGM	PMFA	SF	class	local	NEMD	GED	NNGM	PMFA	SF
MC1	67.1%	70.0%	72.9%	74.3%	68.6%	38.6%	MP2	54.9%	62.2%	51.2%	64.6%	54.9%	64.6%
MC2	70.3%	76.6%	68.8%	78.1%	71.9%	68.8%	MP3	72.0%	82.9%	75.6%	82.9%	72.0%	85.4%
MC3	68.7%	71.6%	80.6%	79.1%	68.7%	76.1%	MP4	72.2%	79.2%	73.6%	87.5%	72.2%	87.5%
MC4	68.6%	74.3%	60.0%	74.3%	65.7%	85.7%	MP5	59.4%	56.3%	54.7%	67.2%	59.4%	67.2%
MC5	40.0%	67.5%	57.5%	67.5%	40.0%	67.5%	DP1	37.5%	51.8%	57.1%	57.1%	37.5%	58.9%
PP1	64.3%	61.9%	52.4%	69.0%	64.3%	69.0%	DP2	35.0%	43.8%	55.0%	53.8%	35.0%	57.5%
PP2	73.4%	65.8%	72.2%	74.7%	73.4%	69.6%	DP3	54.3%	55.6%	59.3%	54.3%	54.3%	66.7%
PP3	70.2%	79.8%	72.6%	75.0%	70.2%	73.8%	DP4	46.8%	45.5%	57.1%	61.0%	48.1%	70.1%
PP4	80.7%	68.7%	85.5%	85.5%	80.7%	80.7%	DP5	35.6%	47.5%	44.1%	44.1%	35.6%	59.3%
PP5	58.9%	72.6%	76.7%	82.2%	58.9%	79.5%	overall:	60.2%	65.0%	65.6%	70.5%	60.3%	69.8%

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 to be satisfactory. Even more so, as the algorithms may have classified regions similar to the real bones but not labeled in the ground truth as such.

5. CONCLUSION

Although graph matching has been used for image and object retrieval for a long time, the involved graphs have been mostly limited to a small number of nodes. Yet unknown image contents without manual segmentations require a generic data-structure resulting in large graphs with thousands of nodes such as HARAGs. The offered concept represents characteristic object constellations as a scene by the use of small structural prototype graphs. The scene analysis, i.e. also the identification of the objects the scene is composed of, can again be achieved by graph matching. Five well-known optimization techniques and distance metrics have been evaluated for this purpose in an example application. The results clearly show that the scene representation outperforms the sole search by local features. Especially with NNGM and SF two robust solutions have been found recommending the use of our framework to further applications in medical imaging.

Current efforts for the enhancement of our framework are directed towards improving the segmentation quality as this is crucial for the information content of the HARAGs. Besides the study of further matching algorithms, we will focus on different models for node and edge prototypes in combination with Bayesian classification. Furthermore, the structural prototypes will be synthesized independent of HARAGs by manually established ground truths. But already at present, a generic framework for scene analyses in content-based medical image retrieval is available.

ACKNOWLEDGEMENTS

The IRMA project and the related project Structural Prototypes in Radiological Routines (SPIRR) are funded by the German Research Foundation (Deutsche Forschungsgemeinschaft DFG) grants: Le 1108/4, Le 1108/6, and Le 1008/9. We also acknowledge the contributions of Ilja Bezrukov to the graph matching framework.

REFERENCES

- ¹ Tagare HD, Jaffe CC, Duncan J: Medical image databases – A content-based retrieval approach. Journal of the American Medical Informatics Association 1997; 4: 184-198
- ² Müller H, Deselaers T, Deserno TM, Clough P, Kim E, Hersch W. Overview of the ImageCLEFmed 2006 medical retrieval and annotation tasks. Lecture Notes in Computer Science 2007; 4730: 595-608

- ³ Petrakis E. Content-Based Retrieval of Medical Images. *International Journal of Computer Research* 2002; 11(2): 171-182
- ⁴ Petrakis E. Fast Retrieval by Spatial Structure in Image Databases. *Journal of Visual Languages and Computing* 2002; 13(5): 545-569
- ⁵ 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. *Proceedings SPIE* 2005; 5747:545-555
- ⁶ 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
- ⁷ Aja-Fernandez S, de Luis-Garcia R, Martin-Fernandez MA, Alberola-Lopez C. A computational TW3 classifier for skeletal maturity assessment: A Computing with Words approach. *Journal of Biomedical Informatics* 2004; 37(2): 99-107
- ⁸ Kim DH, Yun ID, Lee SU: A new attributed relational graph matching algorithm using the nested structure of Earth Mover's Distance. *International Conference on Pattern Recognition* 2004; 48-51
- ⁹ Rubner Y, Tomasi C, Guibas LJ: The Earth Mover's Distance as a metric for image retrieval. *International Journal of Computer Vision* 2000; 40(2): 99-121
- ¹⁰ Greenspan H, Dvir G, Rubner Y. Context dependent segmentation and matching in image databases. *Journal of Computer Vision and Image Understanding* 2004; 93:86-109
- ¹¹ Bunke H, Allermann G: Inexact graph matching for structural pattern recognition. *Pattern Recognition Letters* 1983; 1: 245-253
- ¹² Lappe C, Fischer B, Thies C, Güld MO, Kohlen 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
- ¹³ Schädler K, Wysotzki F. Comparing Structures using a Hopfield-style Neural Network. *Applied Intelligence* 1999; 11: 15-30
- ¹⁴ Peterson C, Söderberg B. A new method for mapping optimization problems onto neural networks. *International Journal of Neural Systems* 1989; 1: 3
- ¹⁵ Russel S, Norvig P. *Artificial Intelligence: A Modern Approach*. Prentice Hall, New Jersey, USA; 1995
- ¹⁶ Melnik S, Garcia-Molina H, Rahm E. Similarity flooding - A versatile graph matching algorithm and its application to schema matching. *Proceedings of the 18th ICDE Conference* 2002; 117-128
- ¹⁷ Grauman K, Darrell T. Fast contour matching using approximate Earth Mover's Distance. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 2004; 220-227
- ¹⁸ Hillier FS, Lieberman GJ. *Introduction to Mathematical Programming*. McGraw-Hill, New York; 1990
- ¹⁹ Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences. USA* 1984; 81: 3088-3092
- ²⁰ Jain BJ. *Structural Neural Learning Machines*. PhD-thesis. Faculty IV, Technische Universität Berlin; 2005