Structural Scene Analysis and Content-based Image Retrieval Applied to Bone Age Assessment

Benedikt Fischer[a], André Brosig[a], Thomas M. Deserno[a,1], Bastian Ott[b], Rolf W. Günther[b]

[a] Department of Medical Informatics,
Aachen University of Technology (RWTH), 52057 Aachen, Germany
[b] Department of Diagnostic Radiology, RWTH Aachen, 52027 Aachen, Germany

ABSTRACT

Radiological bone age assessment is based on global or local image regions of interest (ROI), such as epiphyseal regions or the area of carpal bones. Usually, these regions are compared to a standardized reference and a score determining the skeletal maturity is calculated. For computer-assisted diagnosis, automatic ROI extraction is done so far by heuristic approaches. In this work, we apply a high-level approach of scene analysis for knowledge-based ROI segmentation. Based on a set of 100 reference images from the IRMA database, a so called structural prototype (SP) is trained. In this graph-based structure, the 14 phalanges and 5 metacarpal bones are represented by nodes, with associated location, shape, as well as texture parameters modeled by Gaussians. Accordingly, the Gaussians describing the relative positions, relative orientation, and other relative parameters between two nodes are associated to the edges. Thereafter, segmentation of a hand radiograph is done in several steps: (i) a multi-scale region merging scheme is applied to extract visually prominent regions; (ii) a graph/sub-graph matching to the SP robustly identifies a subset of the 19 bones; (iii) the SP is registered to the current image for complete scene-reconstruction (iv) the epiphyseal regions are extracted from the reconstructed scene. The evaluation is based on 137 images of Caucasian males from the USC hand atlas. Overall, an error rate of 32% is achieved, for the 6 middle distal and medial/distal epiphyses, 23% of all extractions need adjustments. On average 9.58 of the 14 epiphyseal regions were extracted successfully per image. This is promising for further use in content-based image retrieval (CBIR) and CBIR-based automatic bone age assessment.

Keywords: Computer-Aided Diagnosis (CAD), Bone Age Assessment, General Methodology, Scene Analysis, Content-based Image Retrieval (CBIR), Segmentation, Evaluation, Graph Matching

1. INTRODUCTION

Bone age assessment based on hand radiographs is a frequent and time-consuming task for radiologists. The two most commonly used methods are based on image comparison. In the method of Greulich and Pyle [1], the radiologist compares all bones of the hand to radiographs in the standard atlas. In the method of Tanner and Whitehouse (TW3) [2], a certain subset of bones is examined. There are different scoring systems available considering the carpal bones; radius, ulna and short bones; or a combination of these.

Different approaches have been taken to fully or partially automate bone age assessment. Martin-Fernandez et al. apply a set of rotations and translations to the input image in order to register it to a set of reference images with a known age [3]. Here, the entire radiographs are compared. Pietka et al. have developed a system for the extraction of epiphyseal/metaphyseal regions [4]. A set of horizontal step-wedge functions is applied to find the fingers and the phalangeal tips. Central axes are detected for three fingers, and the epiphyseal gaps are identified along the axis. Regions of interest (ROI) are extracted and used for a more detailed analysis of the epiphyseal area. In the method of Park et al. [5] the step-wedge functions are replaced by another heuristic approach for ROI detection. Nonetheless, the

[1] Corresponding author: Prof. Dr. Thomas M. Deserno, né Lehmann, Department of Medical Informatics, Aachen University of Technology (RWTH), Pauwelsstr. 30, D - 52057 Aachen, Germany, email: deserno@ieee.org; web: http://irma-project.org/deserno, phone: +49 241 80 88793, fax: +49 241 80 33 88793.

Medical Imaging 2009: Computer-Aided Diagnosis, edited by Nico Karssemeijer, Maryellen L. Giger
Proc. of SPIE Vol. 7260, 726004 · © 2009 SPIE · CCC code: 1605-7422/09/$18 · doi: 10.1117/12.811632

weak points of such approaches are: (i) reliable ROI extraction, where error rates up to 10% are reported [5], (ii) registration to the references, and (iii) combination of matching results to obtain a final age suggestion.

2. MATERIALS AND METHODS

Our approach is based on the Image Retrieval in Medical Applications (IRMA) framework (http://irma-project.org) that supports content-based access to large medical image repositories [6]. In this framework, global, local, and structural features are supported to describe the image, a ROI, or a constellation of ROIs, respectively [7].

Epiphyseal regions as well as the carpal bone area are located indirectly via the phalanges, which are segmented based on a multi-scale region merging approach. Based on a set of 100 reference images from the IRMA database, a so called structural prototype (SP) is trained in a preliminary step. In this graph-based structure, the 14 phalanges and 5 metacarpal bones are represented by nodes, and location, shape, as well as texture parameters are modeled with Gaussians. Accordingly, the Gaussians describing the relative positions, relative orientation, and other relative parameters between two nodes are associated to the edges. Thereafter, segmentation of a new hand radiograph is done in several steps (Fig. 1):

(i) a multi-scale region merging scheme is applied to extract visually prominent regions and yields an inclusion-hierarchy graph (hierarchical attributed region adjacency graph, HARAG) [8, 9, 10], used for training of SP and segmentation of images not seen in training phase;

(ii) a graph/sub-graph matching to the SP identifies a subset of the 19 bones;

(iii) carpal bone regions that are possibly missing due to segmentation problems are estimated from the SP yielding a complete scene reconstruction, and

(iv) the relevant epiphyseal regions are determined and extracted.

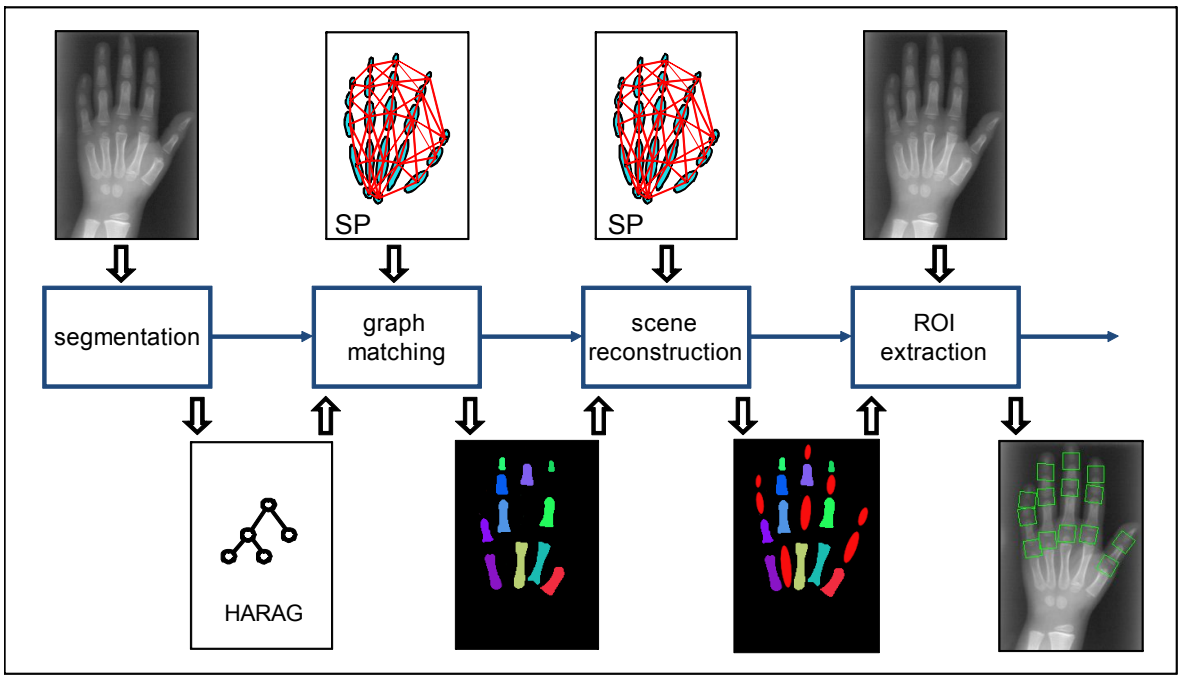


Figure 1: Structural scene analysis.
Segmentation of a previously unseen image results in a graph (HARAG), where phalanges and metacarpal bones are represented by nodes. They are extracted by graph/sub-graph-matching to the structural prototype (SP). Missing nodes are reconstructed with SP information, and the 14 relevant epiphyseal ROIs are extracted.

The benefit of this method is that the segmentation is robust against false detections, since the ROI is located from the prototype, which must be anchored on the image by only a small number of reliable bone matches.

For further enhancement of the previous achievements of steps (i) to (iii), the training of the structural prototype has been adjusted and the relative neighborhood intensity (RNI) is introduced as a new valuable feature. This extends the global, local, and structural features of the IRMA framework with an appearance-based method to the specific needs of automatic bone-age assessment.

2.1. Step-wise approach to ROI extraction

Preliminary step: Structural Prototype graph training

The training of the structural prototype graph is based on the procedure described in [10]: There, training images are transformed into HARAGs by a hierarchical region merging process. The segmented regions in the HARAG that are relevant for the prototype, e.g. metacarpal and phalangeal bones, are manually labeled and their regional and relational features are statistically analyzed. During the training process, a feature subset selection is performed by sequential backward elimination. Afterwards, for each object class (the particular bone), the best regional and relational feature set is known.

In contrast to the previously described prototype generation, where the objects were labeled in the HARAG, now manually labeled reference images are used. The labeling is achieved by manually drawing the contour of the relevant bones and assigning a label to it, instead of assigning labels to automatically obtained segments. For each relevant image object, i.e. the 19 metacarpal and phalangeal bones, a unique label is defined. Additional labels are defined for the background and for overlapping image objects, e.g. due to superposition of bones in the radiographs. The label images have been created for all images. For the training image set, the labels are used for segmentation. For the test image set, they serve as ground truth in the evaluation process.

Now, the segmentation for the HARAG synthesis for each training image is performed on the manually segmented label images instead of using the automatically obtained segmentation. As only the relevant regions are contained, the resulting HARAGs are very small and the region delineation is precise. All feature values, such as mean gray intensity, size and shape descriptions, are still computed on the original radiographs, but using the region contours from the label image. After all HARAGs have been created, the statistical analysis to obtain mean and standard deviation of the regional and relational features is again performed as it has been described in [10].

Step (i): Graph matching

When the SP graph is to be compared to a new image, at first the new image is transformed into a HARAG using the standard hierarchical region growing of the IRMA framework [8]. The matching between the resulting HARAG and the prototype is then achieved by graph matching algorithms. For this purpose, we apply the Similarity Flooding approach as well as an evolutionary graph matcher [10, 11]. The matching result then provides the identification of (a subset of) the hand bones in the HARAG of the new image (Fig. 2a). The identified regions may not show all desired bones if the missing bones did not survive all distance/similarity computations in the matching process, e.g. if certain features are too far from the mean value of its prototype class.

Step (iii): Scene reconstruction

If the local and relational feature values of an image region do not sufficiently match the prototype, then the corresponding object is not considered in the final matching result. For the image objects not included in the matching result the geometrical features necessary for the extraction of the epiphyseal ROI (e.g. center of gravity, orientation and length of the main axis) have to be estimated. For all matched image objects, the top and bottom points of the main axis are computed from these features. The same is done for all prototype nodes using the mean feature values for the corresponding image objects observed during training. The top and bottom points of the matched image regions and their respective centers of gravity now form a point cloud. The corresponding points derived from the prototype can then be registered to the points derived from the matching resulting in a RST transformation with nonuniform scales [12] that yields a least squares fit of the corresponding points (Fig. 2b). The orientation of the hand in the image is assumed to be upright. The transformed top, center, and bottom points of the prototype nodes that correspond to unmatched image objects can then be used to compute an estimate for the necessary feature values of these objects.

Step (iv): Epiphyseal ROI extraction

In order to locate the epiphyseal ROI, a-priori knowledge about their position with respect to the identified bones is used (e.g. the distal phalangeal ROI of the index finger is located between the distal and medial phalanges of the index finger). For each pair of identified image regions that are adjacent to an epiphyseal ROI, we consider the centers of gravity, the length, and the orientation of their main axes from their respective feature vectors (Fig. 2c).

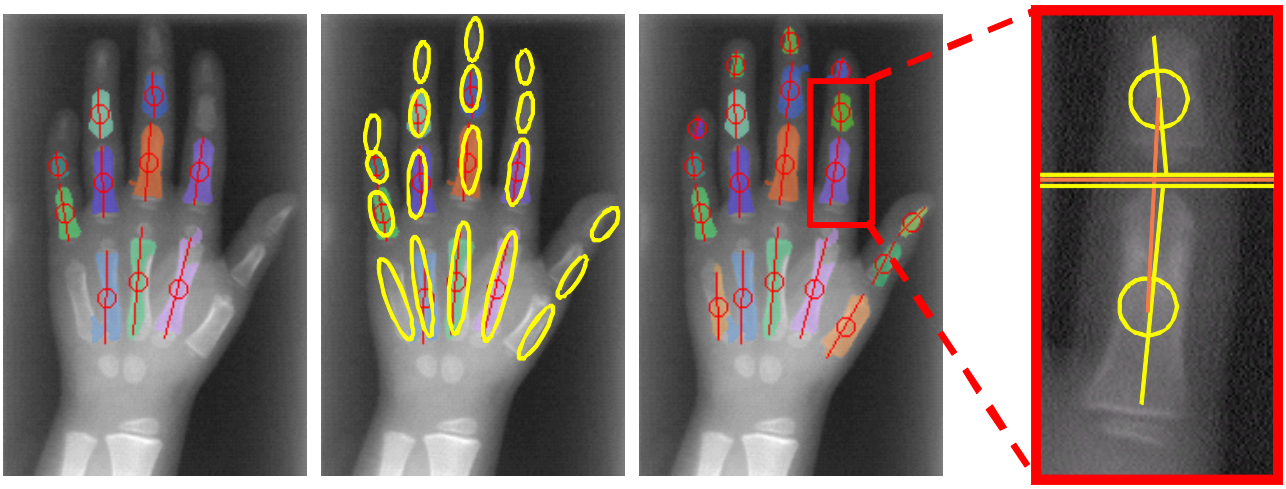
First, the lower end point of the main axis of the upper bone and the upper end point of the main axis of the lower bone adjacent to the epiphyseal region are computed. Then, a horizontal line parallel to the x-axis of the image is computed at the average y-position of these points (Fig. 2d). The epiphyseal region is then assumed to be centered at the intersection of this line with the line connecting the centers of gravity of the lower and upper adjacent bone. The orientation of the ROI is assumed to be the same as the line connecting the centers of gravity of the two adjacent bone regions.

Finally, for each located epiphyseal ROI a rectangular sub-image of fixed size can be extracted in the correct orientation (using quadratic interpolation). For images scaled to a height of 256 pixels an extraction window of 20x30 pixels was found to be appropriate in order to fully contain the epiphyseal regions. These sub-images can then be used for database queries following the query by example paradigm [13] in content-based image retrieval (CBIR) [14].

2.2. Relative neighborhood intensity

In our previously published studies, only the features of the relevant image objects themselves or relations between them have been computed. Due to computational considerations, this disregarded information about further regions. For example, for the 19 relevant objects of the metacarpal and phalangeal bones, there exist on average 2,000 regions on images with a bounding box of 256 pixels. Obviously, the consideration of all combinations between relevant and irrelevant regions (and possible feature set compositions) is computationally prohibitive.

Yet the saliency of the bones for human observers is significantly due to the contrast to its surrounding tissue. Thus, a new computationally feasible feature, the relative neighborhood intensity (RNI), is introduced to determine the relative gray value of a region with respect to its surrounding. The neighborhood of a region is defined here as the set of pixels that do not belong to the region itself and for which a path to a pixel of the region of length smaller than a constant n exists (Fig. 3). A path between two pixels is defined here as a set of steps of in horizontal, vertical or diagonal direction in the 8-pixel neighborhood. This set of pixels can be found by performing n steps of morphological dilation on the region pixels and subtracting the region pixels afterwards. For images scaled to 256 pixels of height, n=3 was found to



(a) (b) (c) (d)

Figure 2: Example of processing pipeline results
(a): Graph matching results
(b): Least squares fitting of mean prototype features of the structural prototype (yellow), bone regions (various colors)
(c): Main axes and bone region centers (red); reconstructed regions
(d): Estimation of location and orientation of an epiphyseal ROI by matched neighboring bone regions (enlarged); main axes of ROI region with vertical center (orange).

be appropriate.

The relative neighborhood intensity then computed as the difference between the normalized mean gray value of the region and the normalized mean gray value of its neighborhood:

$$\mathrm{RNI}(region) = \mathrm{mean}_{\mathrm{gray\_norm}}(region) - \mathrm{mean}_{\mathrm{gray\_norm}}(neighborhood)$$

where the normalized mean gray value is normalized by subtracting the global minimum gray value found in the image and dividing by the difference between the maximum and the minimum gray value of the image:

$$\mathrm{mean}_{\mathrm{gray\_norm}}(region) = \frac{\mathrm{mean}_{\mathrm{gray}}(region) - \min_{\mathrm{gray}}(image)}{\max_{\mathrm{gray}}(image) - \min_{\mathrm{gray}}(image)}$$

The normalized gray value of the surrounding is computed accordingly.

Thus, the RNI is normalized to a range between -1 and 1, where values greater than 0 correspond to a bright region on a dark neighborhood. The usefulness of the new feature is evaluated together in combination with the overall matching quality. Overall, evaluations are performed for the quality estimation of the segmentation, the graph matching, and – most important – the desired epiphyseal ROI extraction.

3. RESULTS

The training of the structural prototypes and evaluation of the segmentation was based on 100 images randomly selected from clinical routine at the Aachen University Hospital. The graph matching and epiphysis extraction is evaluated using a subset of 137 images of Caucasian boys compiled from the digital hand atlas of the University of Southern California, Los Angeles, CA, USA (www.ipilab.org/Baaweb). This accounts for a true separation of the training and the test set. Not only are the two sets distinct, but they additionally are obtained by other imaging devices. To reduce computational complexity, all images were scaled to a height of 256 pixels preserving the aspect ratio. Once the epiphyseal regions have been located, the actual extraction of the ROIs was performed on the original image.

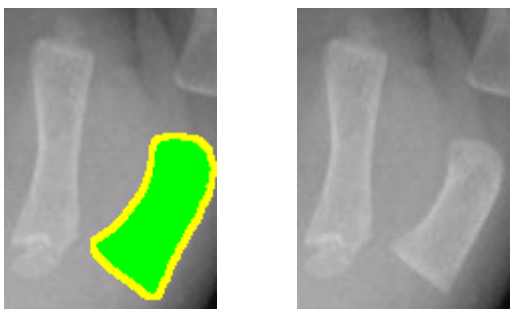
The evaluation is first performed quantitatively (Sections 3.1 to 3.3) and then some illustrative examples of the ROI extraction given (Section 3.4).

3.1. Evaluation of the segmentation

The segmentation was evaluated by comparing all regions in the HARAG to the manually drawn regions in the label image. For each region marked in the label image, the HARAG node that yields the best overhang/overlap ratio was selected. Here, the overlap of two regions r_1 and r_2 is defined as the number of pixels in the intersection of both regions:

$$\mathrm{overlap} = r_1 \cap r_2$$

The overhang is defined as the number of pixels in the symmetric difference of the regions.



(a) (b)

Figure 3: Example RNI computation; relative neighborhood intensity (RNI) for a metacarpal bone of the thumb (green) and its neighborhood (yellow).

$$\text{overhang} = (r_1 \cup r_2) - (r_1 \cap r_2)$$

For each label, the best corresponding HARAG region was evaluated in order to determine if the HARAG contains a region that sufficiently represents the object. A HARAG region was considered to sufficiently represent a region from the label image if the overhang/overlap ratio between these regions decreased below a threshold of 0.8. Based on this measure, 1696 of the 1900 metacarpal and phalangeal bone regions were sufficiently represented by a HARAG node in the set of 100 hand-radiographs from Aachen, i.e. 89 percent of the bones were segmented correctly.

3.2. Evaluation of the graph matching

Graph matching was performed using the similarity flooding graph matching algorithm and the evolutionary game theory graph matching algorithm [10,11]. During the prototype generation a feature subset selection is performed by sequential backward elimination. In order to judge the new feature RNI, experiments were conducted first without RNI, i.e. with the initial feature subset as given in [10]. Then RNI was included in the feature subset selection process, and last, the RNI feature was manually forced to survive the selection process.

A matched region was considered to be correct when the overlap to the best region available from the segmentation was greater than 60%. To ease comparisons, the quality of the graph matching is evaluated by the standard measures recall, precision, and f-measure. These are defined as:

$$\text{recall} = \frac{\text{correctly matched regions}}{\text{total regions}}$$

$$\text{precision} = \frac{\text{correctly matched regions}}{\text{total regions matched}}$$

$$\text{f}-\text{measure} = \frac{2 * \text{recall} * \text{precision}}{\text{recall} + \text{precision}}$$

The results of the graph matching and RNI performance are provided in Table 1. Forcing RNI, the best precision and f-measure of 0.78 and 0.59, respectively, is obtained using the evolutionary graph matcher.

3.3. Evaluation of the epiphyseal ROI extraction

In order to evaluate the quality of the extracted epiphyseal ROI, the coordinates of the centers of 19 epiphyseal regions have been manually marked for all images in the test set using a web interface (Fig. 4). The location of the detected ROI is evaluated by computing its Euclidean distance to the reference points and testing this against a constant distance threshold. For the images scaled to a height of 256 (while keeping the aspect ratio) a distance threshold of 6 pixels was found to lead to an evaluation similar to that of a human observer (Fig. 5).

From a total of 1,918 epiphyseal ROI in 137 images, 1,313 ROI were extracted correctly and 605 incorrectly, yielding an overall error rate of 32% with a standard deviation of 0.14. All details are provided in Table 2. A separate evaluation for the 822 distal and medial epiphyseal ROI of the 3 middle fingers (IDs 5, 6, 9, 10, 13, 14) yields 635 correct and 187 incorrect extractions from a total of 832 ROI, giving an error rate of only 23%.

Matching algorithm	**Similarity Flooding**			**Evolutionary Graph Matcher**		
RNI feature included	no	yes	forced	no	yes	forced
Total regions	2601	2601	2601	2601	2601	2601
Correct matches	1082	1200	1297	1405	1463	1223
False matches	1502	1316	1210	1090	897	343
Recall	0.42	0.46	0.50	0.54	**0.56**	0.47
Precision	0.42	0.48	0.52	0.56	0.62	**0.78**
F-Measure	0.42	0.47	0.51	0.55	**0.59**	**0.59**

Table 1: Retrieval results of the graph matching.
Shown are results for Similarity Flooding and evolutionary graph matching, each without the newly introduced feature RNI, with RNI, and with RNI forced to survive the feature subset selection. The best results for recall, precision, and f-measure are shown in bold print.

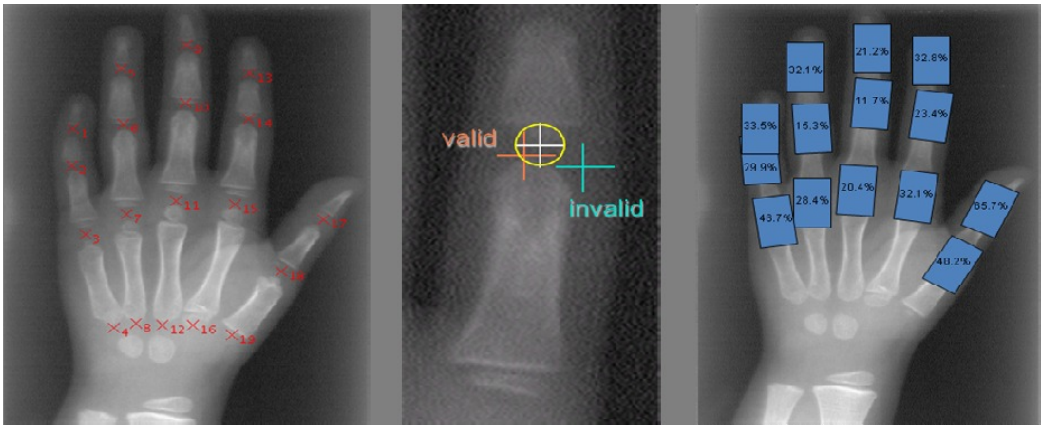
Not just the error rate is important, but also the number of images a histogram, how many images contain a certain number of correct ROIs (Table 3). On average 9.58 of 14 possible epiphyseal ROI per image could be extracted successfully. In 88% of the images, 6 or more ROI were extracted successfully.

3.4. Qualitative examples

The following images (Fig. 6, 7) show two examples for the ROI extraction. To the left an overview of the hand is shown. The main axes of the regions from graph matching result are marked in blue. The main axes of bone regions not matched in the graph matching process estimated by registration of the prototype to the matched regions are marked in yellow. The estimated ROI positions are marked in red. The extracted regions are shown again enlarged on the right.

In Fig. 6, the ROI extraction performed very well. Twelve of the 14 possible ROIs have been extracted correctly. Even the two ROIs not counted as correct since their distance to the manually set reference point is to high, roughly show the desired area. Even although the matched and reconstructed bones (blue and yellow, resp.) are not perfect, our proposed method performs sufficiently robust.

Figure 4: Web interface for setting epiphyseal reference points



(a) (b) (c)

Figure 5: Epiphyseal ROI comparison to ground truth
(a): labels of the epiphyseal ROI, set manually as ground truth
(b): valid range around ground truth (yellow), valid (orange) and invalid (blue) example points;
(c): error rates (percent) for each ROI, the absolute numbers are provided in Tab. 2.

ID	1	2	3	5	6	7	9	10	11	13	14	15	17	18	mean	std.dev.
invalid	46	41	64	44	21	39	29	16	28	45	32	44	90	66	**43.21**	**19.55**
valid	91	96	73	93	116	98	108	121	109	92	105	93	47	71	**93.79**	**19.55**
error rate	0.34	0.30	0.47	0.32	0.15	0.28	0.21	0.12	0.20	0.33	0.23	0.32	0.66	0.48	**0.32**	**0.14**

Table 3: Absolute numbers and error rate for ROIs

# valid ROIs / image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
# images	1	3	1	5	2	5	5	7	14	13	**19**	15	22	17	8

Table 3: Number of correct ROIs per image

In Fig. 7, serious problems can be observed. Here, only four of the extracted ROIs show the desired epiphyseal areas. Apparently, only 7 regions were matched correctly by the graph matcher. The finger tips were matched too low and in part next to the bone. This resulted in a poor subsequent reconstruction of the missing regions and lead to the inacceptable ROI extraction.

4. DISCUSSION

The presented works show the first results of utilizing high-level scene analysis in combination with content-based image retrieval for bone age estimation. So far, the framework has been evaluated up to the point where epiphyseal regions are extracted. The next step in this ongoing research will be to subject extracted epiphyseal regions to query by example searches and retrieve similar regions with their associated ages. The evaluation of the processing pipeline until the ROI extraction delivers a proof of concept of the application of high level scene analysis in this context.

The use of the label image for HARAG generation ensures that all relevant objects are included in each training HARAG and therefore eliminates earlier problems with the segmentation. These resulted in fewer of the relevant regions, diminishing the available data for the training of the structural prototype. In addition, the newly obtained HARAGs are smaller and further speed up the training process. The manual segmentation provided by the label images also allows a thorough quantitative analysis of the segmentation and matching performance. The evaluation of the

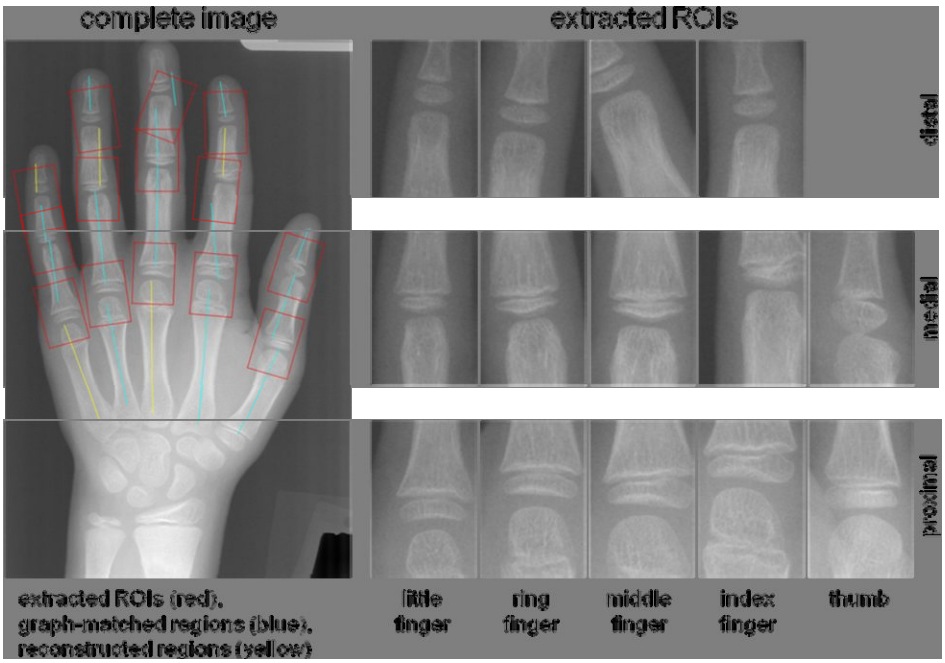


Figure 6: Example epiphyseal ROIs. Here, 12 ROIs are extracted correctly

segmentation showed that 89% of the bones were segmented correctly. The best results for the graph matcher were achieved by the evolutionary graph matcher with a precision of up to 78% and recall up to 56%. This means that 44% of the bone regions have to be reconstructed in the subsequent processing step.

On account of the high signal-to-noise ratio – on average, 3,000 nodes are contained in a HARAG of the USC data set, whereas only 19 are relevant – and considering the very high variances in the features of the structural prototype, this is deemed acceptable. If the graph matching is evaluated by 5-fold cross-validation only on the 100 images from Aachen, the values for recall, precision, and f-measure are considerably higher at 73%, 76%, and 74%. This is obviously due to the different origin of the data sets with different imaging devices, recording parameters and strategies (e.g. with or without markers, positioning of the hand, etc.).

So far, no distinct prototypes for each age group and for different anatomical poses, e.g. differently spread hands (Fig. 6/7) have been used. In combination with enhancements such as contrast equalization, correspondingly adjusted prototypes then are expected to considerably improve the retrieval results.

The reconstruction of unmatched region feature by means of a registration of the prototype using a least squares criterion is rather sensitive to outliers in the graph matching result. If, for example, the proximal phalanx of the thumb has been matched to an image region at the tip of the thumb or even on another finger or altogether outside the hand, this error will greatly distort the least squares fitting, since it contributes quadratically to the rigid transformation. To circumvent this problem, another reconstruction strategy will be implemented that uses the relational features of the structural prototype. Since the prototype is a complete graph, the relational features between each pair of objects described therein are computed during the training process. Thus, the mean relational feature values between each matched and each unmatched object can be used to estimate the required local feature values (e.g. using the mean relative distance, orientation and main axis length). This should lower the extraction error rate significantly. But even now the average number of correct ROIs per image of roughly 10 ROIs is acceptable. Gertych et al., for example, use only 6 phalangeal ROIs and carpal bones for their experiments [15]. Using only these ROIs also lowers the overall error rate of 32% to a rather acceptable level of 23%.

In the future, the ROI extraction should also improve by using ROIs of different sizes, as for example the epiphyseal area between metacarpal and proximal phalanges is much smaller than between medial and distal phalanges. At least the radius of the area in which the extraction is counted as correct or incorrect should be adjusted accordingly to allow an adequate evaluation.

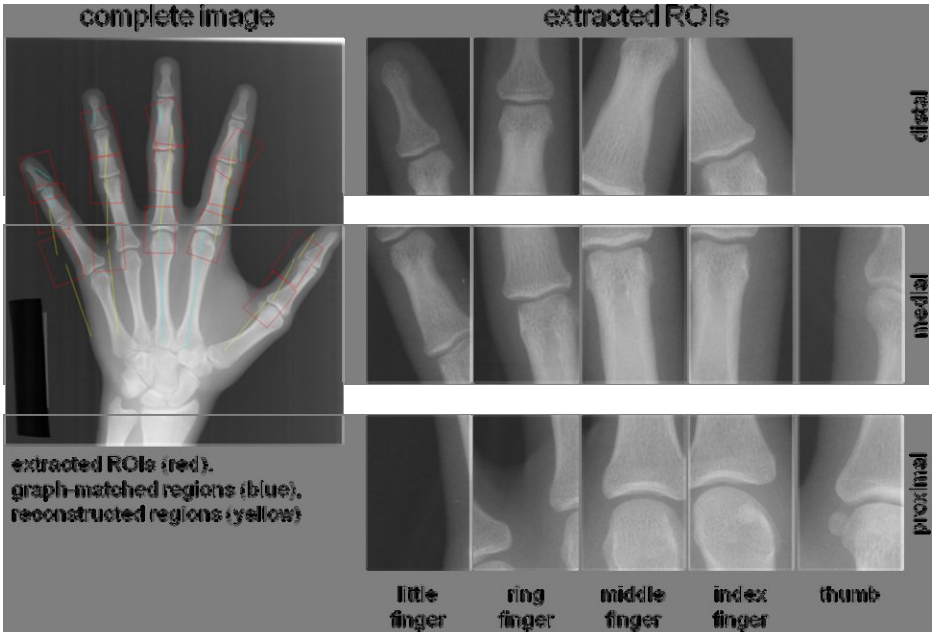


Figure 7: Example epiphyseal ROIs. Here, only 4 ROIs are extracted correctly.

Concerning the next stage in applying CBIR for age estimation, the web interface for manual marking of the epiphyseal center coordinates may also be used as a backup solution for the ROI extraction process. It offers an alternative semi automatic ROI extraction, enabling subsequent image retrieval and age estimation even if the fully automatic ROI extraction fails.

5. CONCLUSION

The proposed processing pipeline allows the application of high level scene analysis for the fully automatic extraction of epiphyseal ROIs. For this purpose, structural prototype graphs for the metacarpals and phalanges are generated and matched with automatically generated hierarchical attributed region adjacency graphs (HARAGs). Missing objects are reconstructed by a least squares approximation of the distance between prototype and the matched objects. Afterwards, the ROI extraction is achieved from the resulting reconstructed scene. The results of each involved step in the processing pipeline have been evaluated and key issues were discussed. Even although training and test images are from different institutions, the outcome of the evaluation strongly encourages further research. Besides the optimizations suggested in the discussion, future research work will focus on age estimation by application of content-based image retrieval. The extracted epiphyseal ROIs will be issued to query by example-searches, where similar ROIs will be retrieved from the image database along with their age which is also stored in the database. The retrieved ages can then be provided as a second opinion for bone age estimation.

6. ACKNOWLEDGEMENT

This research was supported (in part) by the German Research Foundation (DFG), grant no. Le 1108/9.

REFERENCES

1. Greulich WW, Pyle SI. Radiographic atlas of skeletal development of hand wrist. Stanford University Press, Stanford CA. 1971.
2. Tanner JM, Healy MRJ, Goldstein H, Cameron N. Assessment of skeletal maturity and prediction of adult height (TW3). WBSaunders, London. 2001.
3. Martin-Fernandez MA, Martin-Fernandez M, Alberola-Lopez C. Automatic bone age assessment: A registration approach. Proc SPIE. 2003; 5032: 1765–76.
4. Pietka E, Gertych A, Pospiech S, et al. Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction. IEEE Trans Med Imaging 2001; 20(8): 715-29.
5. Park KH, Lee JM, Kim WY. Robust epiphyseal extraction method based on horizontal profile analysis of finger images. Proc ISITC 2007; 278–82.
6. Lehmann TM, Güld MO, Thies C, et al. Content-based image retrieval in medical applications. Methods Inf Med 2004; 43(4): 354-61.
7. Deserno TM, Antani S, Long R. Ontology of gaps in content-based image retrieval. J Digit Imaging 2008; online-first, DOI 10.1007/s10278-007-9092-x.
8. 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. Proc SPIE 2005; 5747(1): 546-555.
9. Thies C, Schmidt-Borreda M, Seidl T, Lehmann TM. A classification framework for content-based extraction of biomedical objects from hierarchically decomposed images. Proc SPIE 2006; 6144: 559-568.
10. Fischer B, Sauren M, Güld MO, Deserno TD. Scene analysis with structural prototypes for content-based image retrieval in medicine. Proc SPIE 2008; 6914; online first, DOI: 10.1117/12.770541.
11. Fischer B, Fritsche A, Thies C, Deserno TM: Evolutionäres Graphmatching zur Handknochen-Identifikation. Procs Bildverarbeitung für die Medizin 2009, Heidelberg, Germany. In Press.
12. Skrinjar O. Point-based registration with known correspondence: Closed form optimal solutions and properties. Biomedical Image Registration 2006; : 315-21.
13. Niblack W, Barber R, Equitz W et al. The QBIC project: Querying images by content using color, texture, and shape. Proc SPIE; 1993. p. 173-87.

14. Smeulders AWM, Worring M, Santini S, Gupta A, Jain R. Content-based image retrieval at the end of the early years. IEEE Trans Pattern Anal Mach Intell 2000; 22(12): 1349-80.
15. Gertych A, Zhang A, Sayre J, Pospiech-Kurkowska S, Huang HK: Bone age assessment of children using a digital hand atlas. Computerized Medical Imaging and Graphics 2007; 31(4-5): 322-31.