

Object-based modeling, identification and labeling of medical images for content-based retrieval by querying on intervals of attribute values

Christian Thies, Tamara Ostwald, Benedikt Fischer, and Thomas M. Lehmann
Department of Medical Informatics
Aachen University of Technology (RWTH), Aachen, Germany

ABSTRACT

The classification and measuring of objects in medical images is important in radiological diagnostics and education, especially when using large databases as knowledge resources, for instance a picture archiving and communication system (PACS). The main challenge is the modeling of medical knowledge and the diagnostic context to label the sought objects. This task is referred to as closing the semantic gap between low-level pixel information and high level application knowledge. This work describes an approach which allows labeling of a-priori unknown objects in an intuitive way.

Our approach consists of four main components. At first an image is completely decomposed into all visually relevant partitions on different scales. This provides a hierarchical organized set of regions. Afterwards, for each of the obtained regions a set of descriptive features is computed. In this data structure objects are represented by regions with characteristic attributes. The actual object identification is the formulation of a query. It consists of attributes on which intervals are defined describing those regions that correspond to the sought objects. Since the objects are a-priori unknown, they are described by a medical expert by means of an intuitive graphical user interface (GUI). This GUI is the fourth component.

It enables complex object definitions by browsing the data structure and examining the attributes to formulate the query. The query is executed and if the sought objects have not been identified its parameterization is refined.

By using this heuristic approach, object models for hand radiographs have been developed to extract bones from a single hand in different anatomical contexts. This demonstrates the applicability of the labeling concept. By using a rule for metacarpal bones on a series of 105 images, this type of bone could be retrieved with a precision of 0.53 % and a recall of 0.6%.

Keywords: image analysis, content modeling, image interpretation, object extraction, content description

1. INTRODUCTION

Retrieval from medical image databases depends on formalizing task specific knowledge.¹ Not only querying entire images but also for well defined and visually perceivable objects within an image suitable numerical features must be determined. At a regional level, object identification is only possible with cumbersome efforts at data entry time.² Each object that might be eventually of interest must be identified and labeled manually in each image, or well defined segmentation procedures must be applied to extract a specific class of objects and forming an object catalog.³ Both approaches become inefficient in clinical radiological routine, where hundreds of images from different modalities, body regions and pathologies are acquired every day, and the class and properties of sought objects are not known at data entry time.² Thus, it becomes reasonable to execute the task of object identification at query time because at this point, the physician knows exactly what he is looking for and is able to formulate his knowledge.

Correspondence: Christian Thies, Department of Medical Informatics, RWTH Aachen, Pauwelstrasse 30, 52057 Aachen, Germany E-mail: cthies@mi.rwth-aachen.de, Telephone: +49 241 8088795

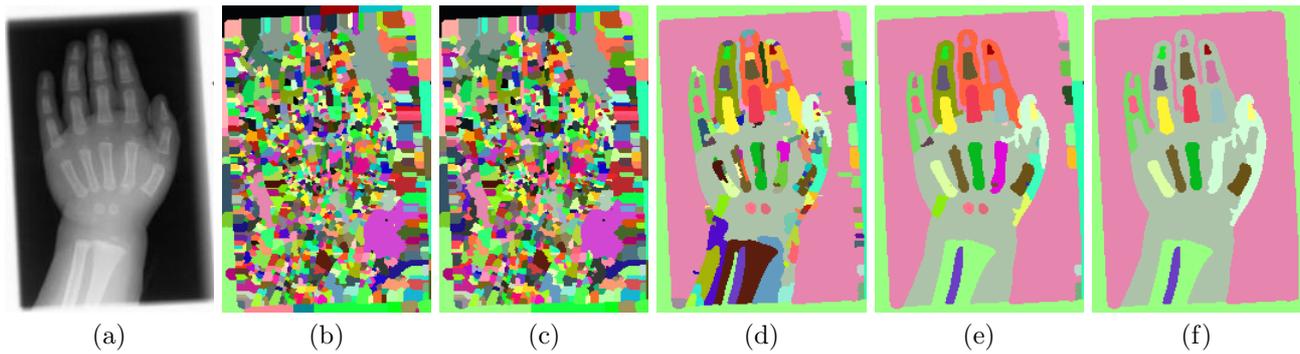


Figure 1. A hand radiograph (a) has been decomposed in regions on different scales. Scales are displayed for 0, 157, 1019, 1104, and 1130 iterations in (b), (c), (d), (e), and (f) respectively.

However, this concept requires a method for object modeling which is applicable during routine image evaluation without supervision of an image processing expert. Current approaches either support only one type of objects or they are highly integrated and must be reconfigured by an expert.^{4,6,7} The main challenge is the mapping of medical knowledge in a formal language on a sufficiently high level of abstraction.⁸ On behalf of this, a concept of separating image partitioning and content description is applied that allows an user friendly knowledge representation.⁹

For image partitioning, a generalized causal multiscale decomposition is used. Herein the entire visual information of an image is provided without considering its content.¹⁰ Result of this computation is a hierarchically organized region representation. There are approaches for region labeling based on image decomposition in image archives but they only extract single scale information, which is not of sufficient accuracy for medical applications and, which does not permit complex object modeling.¹¹

The selection of relevant regions for content description is performed on the value ranges of their descriptive attributes such as roundness, mean grey-value, or size. On these attributes, searching intervals are defined by the user to select distinct objects. Those queries form the object model and their results are used to label the content of an image. Query formulation is done in a data mining feedback cycle where the user composes a hypothesis on an image gets the result for this query and decides whether it is correct or it must be reformulated.¹² Thus the heuristic aspect of learning object description is integrated in the concept. Main component of the approach is a GUI for browsing the data structure as well as formulating and verification of the query.

This work describes the concept of object-based modeling and identification and its evaluation based on labeling hand radiographs that are taken from radiological routine. Hand radiographs have been chosen because they are used in a variety of medical routine examinations and hereby they are acquired following a standardized protocol. For instance, in maturity determination of children, the size, number and distribution of the carpal bones is verified.¹³ In gout diagnosis, the bone deformation is determined,¹⁴ and of course bone fractures are detected from x-ray images.

2. HIERARCHICAL ANALYSIS OF IMAGE CONTENT

2.1. Image Decomposition

Object identification is based on a generalized hierarchical image decomposition, which is executed fully automated at image entry time. In this work, a region growing algorithm was applied which starts with a complete watershed segmentation of the input image.¹⁰ From those segments, merge candidates are determined where the difference between homogeneity and contrast along the neighboring contour pixels is below a predetermined threshold. This procedure yields a multiscale representation of all visually plausible image regions (Fig. 1).

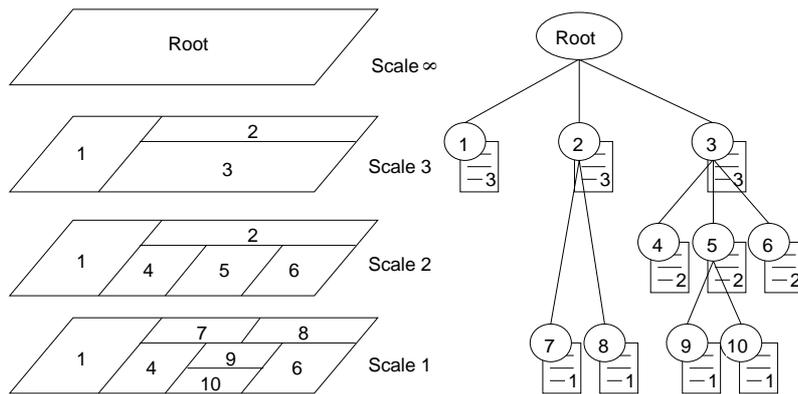


Figure 2. The region hierarchy describes the inter-scale inclusion relation for each region between the scales. All edges in coarser scales have corresponding edges in finer scales (left). This causality leads to a tree structure (right), where each region is represented as a node (circles) with an attached feature vector (boxes). The edges represent a region's inclusions. For a better overview, the neighbor adjacencies are not displayed in this figure.

2.2. Data Structure

The resulting data structure represents the topological adjacencies of regions by a region adjacency graph (RAG)⁵ as well as the inclusion relations induced by the multiscale approach. Therefore, it is called hierarchical attributed RAG (HARAG). In this graph, each image partition (i.e. region) is represented by a graph node. In the concept of the proposed image labeling method it is assumed, that the hierarchical data structure contains every possible region that may be sought. From this point of view, object identification by a human examiner becomes the identification of appropriate image regions. This closes the semantic gap between low-level pixel information and high-level knowledge on image content. The formal representation of low-level information is an attribute vector. It is linked to a region node describes region properties such as size, roundness, texture, etc. (Fig. 2).

2.3. Region Attributes

Constant features are static information that is persistently stored within the HARAG file. It encloses the size of a region, the region pixels themselves as well as the adjacent and hierarchically included regions. Shape, grey value, and texture features are computed from this basic information. To obtain normalized and consequently comparable values relative attributes are computed. They put the image dimensions in relation to the absolute values. The current set of attributes has been taken from literature and focuses on shape properties. If necessary, the implementation is efficiently extensible by new attributes and attribute classes. The following set of features is available.

CONSTANT FEATURES

Number of Sons
 Number of Neighbors
 Contour Length
 Size
 Absorption Iteration
 Creating Iteration
 Number of Holes
 Minimal Depth
 Maximal Depth
 Relative Size
 Relative Contour Length

SHAPE FEATURES

Centroid X
 Centroid Y
 Relative Centroid X
 Relative Centroid Y
 Height of Ferret Box
 Width of Ferret Box
 Relative Height of Ferret Box
 Relative Width of Ferret Box
 Extent of Ferret Box
 Relative Height of Bounding Box
 Relative Width of Bounding Box
 Extent of Bounding Box
 Eccentricity of Bounding Box
 Orientation
 Principal Axis Radius
 Auxiliary Axis Radius
 Relative Principal Axis Radius
 Relative Auxiliary Axis Radius
 Eccentricity
 Radius of Fitting Circle
 Relative Radius of Fitting Circle
 Extent of Circle
 Form factor
 Roundness
 Compactness

GREY VALUES

Mean Gray value
 Relative Mean Gray value
 Variance
 Entropy

TEXTURE FEATURES

Contrast
 Variance of Cooccurrence Matrix
 Differential Moment
 Entropy of Cooccurrence Matrix
 Homogeneity of Cooccurrence Matrix
 Correlation of Cooccurrence Matrix
 Correlation1 of Cooccurrence Matrix
 Correlation2 of Cooccurrence Matrix

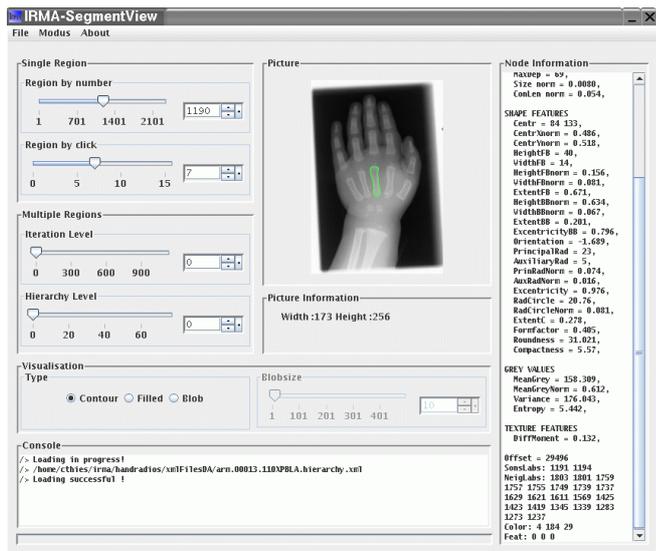


Figure 3. The GUI consist of browsing tools (left) a display area (middle) and the attribute inspector (right). Furthermore a logging window for system messages is provided (bottom left, bottom middle).

2.4. Graphical User Interface

To extract objects from an HARAG, relevant regions must be sought by the examiner. Since HARAGs consist of multiple layers, the main functionality of the user interface is displaying and changing multiscale layers to examine all image partitions. Besides the visual representation of the image regions the user interface permits examination of their formal attributes by clicking on them. The GUI has been realized as a JAVA applet to avoid installation of a client software, which is typically not applicable at hospital workstations. It consists of four sections to support browsing of a HARAG (Fig. 3). In the left column there are sliders for HARAG browsing. The middle section is the display window. On the right side, the attributes of the currently selected region are displayed. Below the left and middle column, a console for logging information is situated. This provides an intuitive interface for actual object modeling from hierarchically decomposed images.

3. FORMULATING AN OBJECT MODEL

The hierarchical image decomposition combined with the feature vectors establish a search space for object extraction. Herein a search pattern becomes the actual object model. The latter is a combination of intervals that are selected for distinct attributes. This search pattern is similar to a database query.

3.1. Designing a query

Such a query is designed in a data mining cycle by formulating and verifying a hypothesis on attributes and intervals for a relevant image region. For this purpose, the user browses a HARAG by its layers until a region that corresponds to the object of interest is identified. By clicking the region, the browser displays its attribute vector (Fig. 3). Based on the displayed values and the visual impression, the user makes an educated guess on the relevant attributes. They are entered in a dialog window which offers selection of attributes and setting of limits that reflect the values from the attribute vector of the region (Fig. 4).

As an example metacarpal bones form ellipse-shaped regions with a high aspect ratio of the principal and auxiliary axis. They are of similar size and yield a certain degree of roundness, which reflects the ratio of size and length of the region contour. Thus, one might guess those three attributes as relevant. The appropriate limits for these attributes are obtained from the values of a sample region (Fig. 5).

3.2. Executing a query

A query hypothesis represents an intuitive model of the sought objects and forms a search pattern. By transforming the intervals into relational predicates, the pattern is formalized and becomes computable. Such a

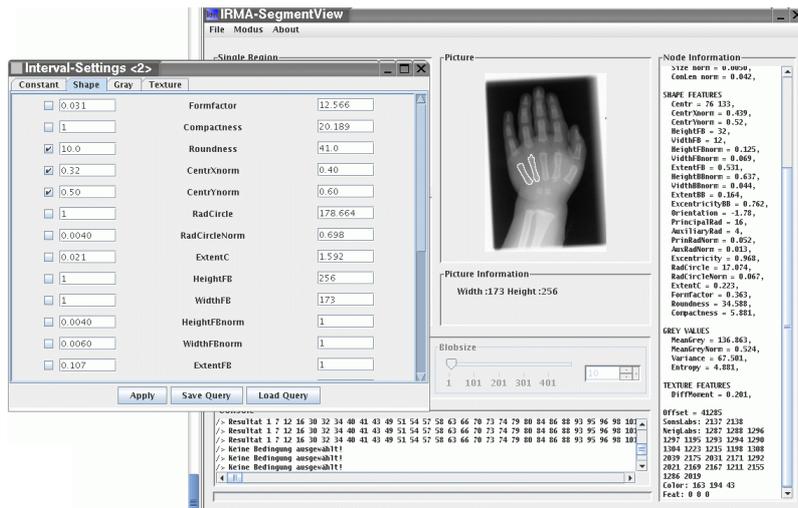


Figure 4. The interval settings dialog provides the query interface where the selection of attributes and the setting of the intervals is performed. It is used simultaneously to the main HARAG dialog. Once a query is defined, it is stored via the save query button, where the name of the stored file represents an object label.

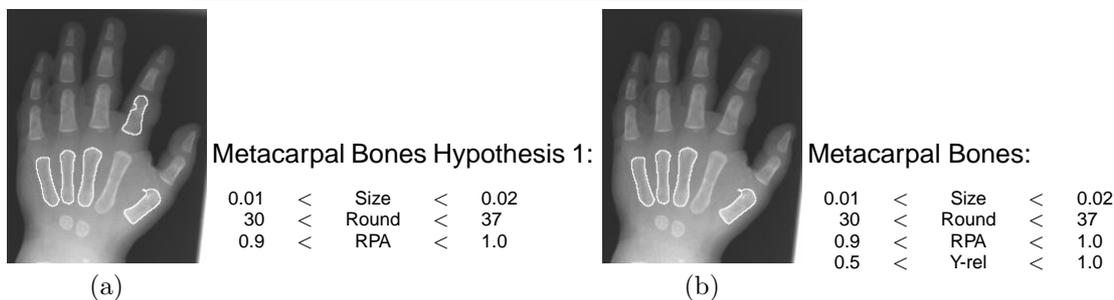


Figure 5. The hypothesis for metacarpal bones intuitively uses relative size (Size), roundness (Round) and ratio of principal and auxiliary axis radius (RPA) as attributes, which yields an additional phalanx (a). This is avoided by assuming that metacarpal bones are situated in the lower image part defined by centroid Y (Y-rel), (b). The missing metacarpal bone was not available in the HARAG.

formula consists of an attribute name $A \in \{\text{Size, Roundness, Variance, ...}\}$, a relation $R \in \{<, >, =\}$ the logic AND-operation \wedge , and the limit $l \in \mathbb{R}$. Thus the formula

$$R(A_1, l_1) \wedge \dots \wedge R(A_n, l_n) \quad (1)$$

is verified for each attribute A_n in each region. Because of the inclusion relation of the multiscale representation the computation follows a top-down sequence in the HARAG which ensures great speed and a minimum of checks for each pattern. The result of a query is a list of those regions in the current HARAG that fulfill the logical formula. For verification of the result the regions are displayed in the browser window. The data mining cycle of query design and execution, is iterated until the user retrieves the sought objects with respect to their actual occurrence in the HARAG 5. According to the query the objects can now be labeled. Once a query is formalized it can be applied to each image of a database. The resulting regions are stored for post-processing such as image series evaluation.

4. EXPERIMENTS

The experiment aims at illustrating the principle of object queries by labeling bones in hand radiographs from different contexts. For this purpose bones are extracted for an anatomic and a functional labeling from the same hand radiograph.

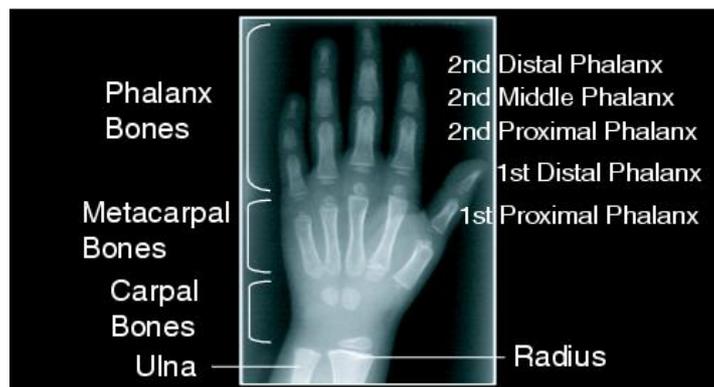


Figure 6. The taxonomy of hand bones represents anatomic knowledge that is used for image labeling.

4.1. Labeling of Single Anatomic Objects

The extraction rule generation follows the general anatomic nomenclature for hand bones (Fig. 6). The bone structure is well defined, and six classes of bones have been identified. These classes are (a) the second distal phalanx, (b) second middle phalanx, (c) second proximal phalanx, (d) metacarpal, (e) carpal, and (f) radius. Each of the classes should be represented by a distinct query which represents the knowledge on the sought objects. The search patterns for the anatomically represented bones are developed by browsing the HARAG. The data-mining loop is iterated as described in Section 3 until the sought objects are extracted. The experiment is conducted on an image, of 169x250 pixels that had been transformed into an HARAG with 5227 nodes. For the defined six classes, the appropriate regions are extracted which correspond to the anatomically defined labels.

4.2. Formulation of A-priori Knowledge

Fingers are anatomic objects that consist of several single bones, which are typically aligned vertically in a hand radiograph. The three medial fingers are nearly of similar size and topology of the single limbs. They differ in their relative position, which is the contextual knowledge to form a query hypothesis. Intuitively the horizontal position should be the most discriminative attribute when selecting these three fingers. Thus an object model that mainly varies the “centroid x” attribute is designed in this experiment. This illustrates the general validity of this approach for modeling object descriptions.

4.3. Automated Object Identification

The automated application of a query rule for image labeling on a sample of the IRMA archive is examined. For this purpose a set of 105 hand radiographs arbitrarily obtained from clinical routine was selected from the IRMA database. The radiographs consisted of approx. 250 x 150 pixels resulting in approx. 2300 nodes for each HARAG. Where HARAG sizes range from min. 958 to max. 6323 nodes with a mean value of 2356 regions.

From those HARAGs a human observers selected 372 regions that corresponded to metacarpal bones. This set of region labels served as a ground truth to which the automated object extraction was compared. The object model was defined by testing and refining it against five different HARAGs. It was applied to the 100 remaining HARAGs and the resulting region lists were compared to the region sets as defined for the ground truth.

5. RESULTS

5.1. Labeling of Single Anatomic Objects

Results are presented in Figure 7. The query rules reflect the visual properties of the regions.

The rule for second distal phalanx (a) extracts small regions of relative sizes between 0.07 and 0.15 percent of the image size. Roundness ranges from 0 to ∞ , and is limited between 10 and 30 since regions are triangle shaped

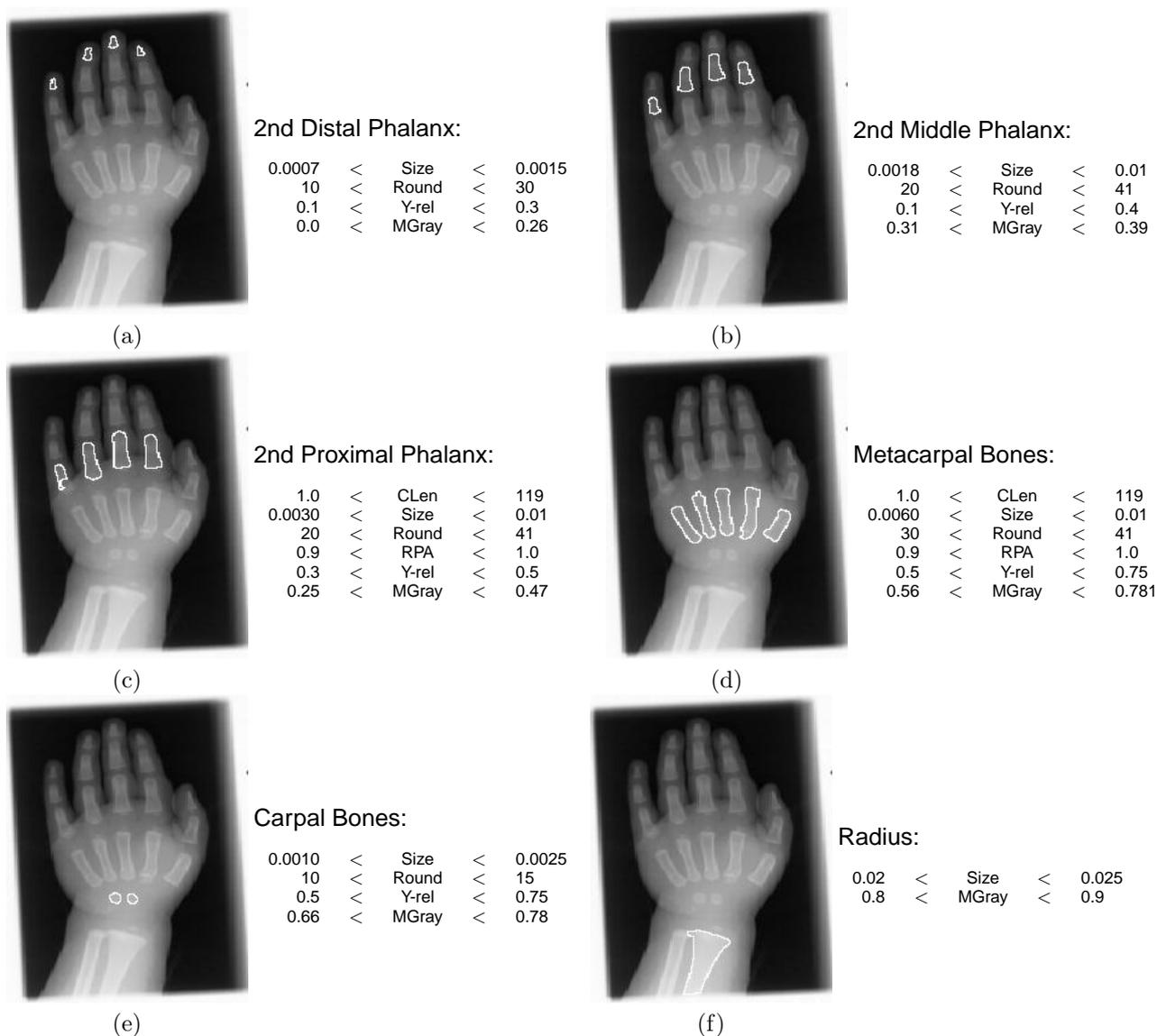


Figure 7. Different anatomic parts of fingers in handradiographs can be classified in (a) distal, (b) medial, and (c) proximal phalanxes. Further structures are the (d) metacarpal and (e) carpal bones. The radius (f) is already a bone of the arm. The attributes are abbreviated as contour length, relative size, roundness, ratio of principal and auxiliary axis radius, centroid-Y, and relative mean gray value by CLen, Size, Round, RPA, Y-rel, MGray respectively.

and therefore yield a good roundness value. Besides size the y-position is an intuitive measure that is restricted to the upper third part of the image. Since those bones do not absorb as much energy as the other parts of the hand the gray-value is also in the lower quarter of the intensity range ($0 < \text{gray-value} < 0.26$).

The rule for second middle phalanx (b) considers a higher lower limit of sizes of 0.0018 and a larger value for roundness than the rule for (a) to reflect the larger and more elongated regions. By considering y-positions down to a fraction of 0.4 of the image height the lower localization of the bones is considered. The regions are brighter since the tissue absorbs a larger part of the x-ray energy.

Bones from the second proximal phalanges (c) are nearly twice as large as the second middle phalanx thus the lower size limit is raised to 0.003 they are even more elongated than the second distal phalanx thus the principal axis radius lies in the upper ten percent. The y-positions is located between the upper third (0.3) and middle

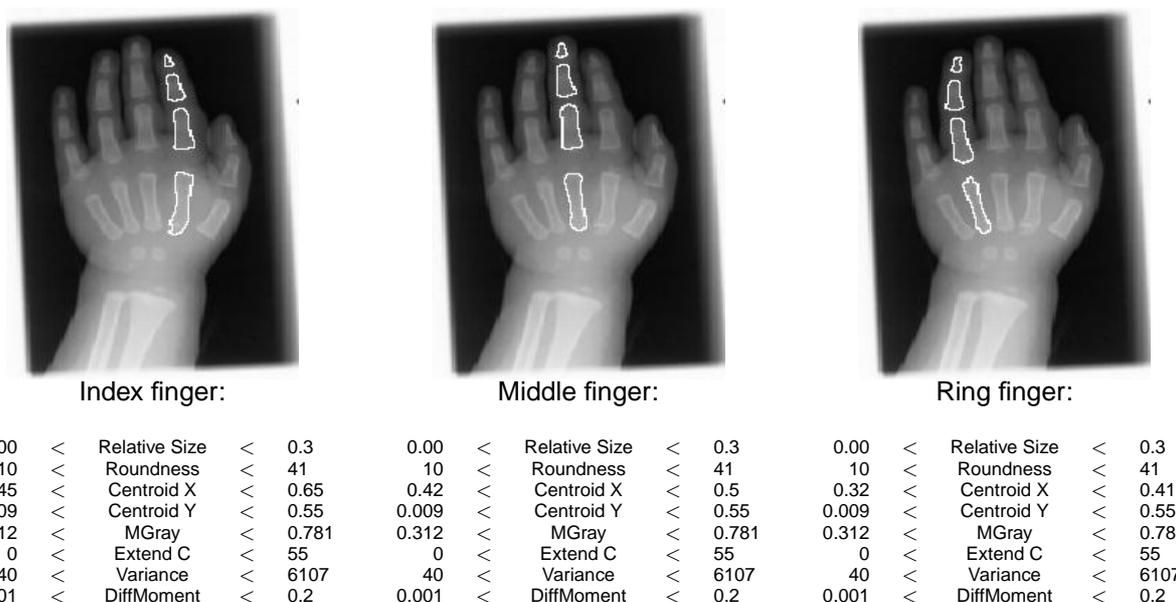


Figure 8. Intuitive adoption of the normalized y-positions provide appropriate results for three fingers, where each rule represents an adequate object model for the same image.

(0.5) of the image height. The gray-value limits are slightly elevated.

Metacarpal bones (d) are again of double size than proximal phalanges (c) thus size is considered as at least 0.006. Intervals for roundness and principal axis radius remain unchanged since both classes are of similar shape. Metacarpal bones are typically located in the lower middle quarter of the image reflected by $0.5 < Y\text{-rel} < 0.75$, and gray-values become brighter on behalf of the thicker hand tissue.

Carpal bones (e) are a little bigger than the distal phalanges (a) and somewhat smaller than the middle phalanges (b), which is reflected by the restriction of size to 0.001 and 0.0025. They are basically round thus roundness is considered between 10 and 15. This specification allows the same interval for the y-location as the metacarpal bones (d). The latter are differentiated by their disjoint roundness interval. However the lower gray-value boundary is risen to 66 percent, to avoid extraction of similar shaped but darker regions.

Radius has been extracted by considering size and gray-value. Since it is the largest and brightest bone in the entire image only size and gray-value is considered. The size is reflected by the interval between 2 and 2.5 percent of the image size and the gray-value lies within 80 and 90 percent to avoid selection of the white shutter fields at the image borders.

5.2. Formulation of A-priori Knowledge

The functional labeling for the ring finger, middle finger, and index finger yielded the three types of fingers and the attributes are the same. The interval ranges are designed to extract the bones of the hand by the same heuristic approach as in the last experiment. To formulate the a-priori knowledge where the attribute ranges had to be adopted for the x-positions as predicted.

Once the middle finger was extracted the ring- and middle finger have been labeled via the translation of the x-Coordinate limits.

5.3. Automated Object Identification

The query was generated by selecting metacarpal bones from five different images, where PAR denotes the principal axis radius.

0.005	<	Relative Size	<	0.025
0.4	<	Centroid Y	<	0.75
-2.2	<	Orientation	<	-0.8
30	<	Roundness	<	55
13.0	<	PAR	<	39.0
0.48	<	Excentricity BB	<	1.0
0.2	<	Formfactor	<	2.0
4.0	<	Compactness	<	7.0
0.3	<	MGray	<	0.85

This manually generated query obtained 225 of the available 372 regions, which makes a recall of 0.6 % and a precision 0.34 %.

6. DISCUSSION

The experiments for single object extraction (Sec. 4.1) and knowledge based modeling (Sec. 4.2) yielded comprehensible object-models. It is possible to extract single bones as well as complex combinations from an image, which represents domain specific knowledge. Since the attributes and interval boundaries have been determined heuristically they do reflect the visual comprehension of the objects. The obtained selection is not sound and unique since there may exist other attribute combinations and intervals that extract the same regions. Thus several models are obtainable for a single query context. However this does not restrict the applicability of the approach since different observers always formulate differing content descriptions.

Generalization of an object-model was examined in the automated identification of metacarpal bones (Sec. 4.3). Precision and recall have been computed as a measure of quality with respect to the manually generated ground truth. A total number of 249640 regions had to be searched for 372 relevant objects. The applied interval-based classification is known for over-fitting on behalf of the parallelity to the search axis. Considering this and since the rule for region extraction was obtained heuristically on five training samples and consists of only nine attributes the recall of 60 % and the precision of 34 % can be considered as good results. Further improvement is achievable by selecting a more general classifier such a nearest neighbor computation. However this requires a new GUI for query design.

Current general medical image retrieval systems do not make use of region-based queries for object identification.² In this work, such an approach has been introduced and verified on typical medical images and questions. Now it is possible to make use of local visual information which allows a more specific and comprehensible presentation of relevance facts as well as quantitative content analysis.

A major drawback of existing approaches for general object-based retrieval is the strict formulation of content extraction.^{6,7,11} They are restricted to a fixed number of objects and do not consider visual information on different scales. Consequently the classes of objects that can be retrieved must be known a-priori at data entry time, which is not possible in medical image retrieval.

An alternative and even more intuitive approach to query design that is possible with our approach is a manual outlining of a region of interest. Based on a user drawn region, the attribute values can be computed and transformed into a query.

7. CONCLUSION

Object modeling can be done by an interval-based search in a hierarchical image partitioning. Therefore, a user-friendly GUI is most important to close the semantic gap between context knowledge and raw data representation. It also depends on the available set of attributes. Additionally, the image decomposition must yield the visually plausible regions. The actual query is efficiently computable and therefore applicable for large image archives.

By this heuristic approach different anatomic and functional context can be considered with a single application implementation, and the knowledge is stored in a reproducible fashion. However automated application does not yet yield sufficiently accurate results for reliable quantitative analysis.

By using the intuitive data-mining cycle for object modeling at query time, the object identification is performed

by the medical expert. Furthermore, it can be integrated into a classical image retrieval interface without extension of underlying concepts. The use of simple relational predicates combined with a top down evaluation strategy ensures a runtime of a query within milliseconds. Since this is work in progress, the quality of the identified queries has to be verified on a larger set of images, as well as the set of test images has to be extended systematically. However, the results obtained on the hand radiograph illustrate the usefulness of object modeling for content-based image retrieval in medical applications.

8. ACKNOWLEDGEMENT

This work was performed within the image retrieval in medical applications (IRMA) project, which is supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), grants Le 1108/4 and Le 1108/6. For further information visit <http://irma-project.org>.

REFERENCES

1. Lehmann TM, Güld MO, Thies C, Fischer B, Spitzer K, Keysers D, Ney H, Kohnen M, Schubert H, Wein BB: Content-based image retrieval in medical applications. *Methods of Information in Medicine* 2004; 43(4): 354-61.
2. Müller H, Michoux N, Bandon D, Geissbuhler A: A review of content-based image retrieval systems in medical applications. clinical benefits and future directions. *International Journal of Medical Informatics* 2004; 73(1):1-23.
3. Cootes TF, Taylor CJ, Cooper DH, Graham J: Training models of shape from sets of examples. *Procs. British Machine Vision Conference* 1992:266-275.
4. Petrakis EGM: Design and evaluation of spatial similarity approaches for image retrieval. *Image and Vision Computing* 2002 20: 59-76.
5. Sonka M, Hlavac V, Boyle R: *Image Processing, Analysis and Machine Vision*. Chapman & Hall, London, 1993.
6. Wang JZ: *Integrated Region-based Image Retrieval*. Kluwer Academic Publishers, 2001.
7. Xu Y, Duyuglu P, Saber E, Tekalp AM, Yarman-Vural FT: Object based image labeling through learning by example and multi-level segmentation. *Pattern Recognition* 2003; 36(6): 1407-23.
8. Ullman S: *High Level Vision: Object Recognition and Visual Cognition*. MIT Press, Cambridge, MA, 1996.
9. Thies C, Malik A, Keysers D, Kohnen M, Fischer B, Lehmann TM: Hierarchical feature clustering for content-based retrieval in medical image databases. *Procs. SPIE* 2003; 5032(1): 598-608.
10. 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, in press.
11. Djeraba C: Association and content-based retrieval. *IEEE Transaction on Knowledge and Data Engineering* 2003 15(1): 118-35.
12. Witten IH, Eibe F: *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufman Publishers, San Francisco, 2000.
13. Vogelsang F, Kohnen M, Schneider H, Weiler F, Kilbinger MW, Wein BB, Guenther RW: Skeletal maturity determination from hand radiograph by model-based analysis *Procs. SPIE* 2000; 3979: 294-305.
14. Pittman JR, Bross MH: Diagnosis and management of gout. *American Family Physician* 1999; 59(7): 1799-806.