

Hierarchical feature clustering for content-based retrieval in medical image databases

Thies C.^a, Malik A.^a, Keysers D.^b, Kohnen M.^c, Fischer B.^a, Lehmann T.M.^a

^aDepartment of Medical Informatics, Aachen University of Technology, Germany

^bChair of Computer Science VI, Aachen University of Technology, Germany

^cDepartment of Diagnostic Radiology, Aachen University of Technology, Germany

ABSTRACT

In this paper we describe the construction of hierarchical feature clustering and show how to overcome general problems of region growing algorithms such as seed point selection and processing order. Access to medical knowledge inherent in medical image databases requires content-based descriptions to allow non-textual retrieval, e.g., for comparison, statistical inquiries, or education. Due to varying medical context and questions, data structures for image description must provide all visually perceivable regions and their topological relationships, which poses one of the major problems for content extraction. In medical applications main criteria for segmenting images are local features such as texture, shape, intensity extrema, or gray values. For this new approach, these features are computed pixel-based and neighboring pixels are merged if the Euclidean distance of corresponding feature vectors is below a threshold. Thus, the planar adjacency of clusters representing connected image partitions is preserved. A cluster hierarchy is obtained by iterating and recording the adjacency merging. The resulting inclusion and neighborhood relations of the regions form a hierarchical region adjacency graph. This graph represents a multiscale image decomposition and therefore an extensive content description. It is examined with respect to application in daily routine by testing invariance against transformation, run time behavior, and visual quality. For retrieval purposes, a graph can be matched with graphs of other images, where the quality of the matching describes the similarity of the images.

Keywords: Content-based image retrieval, region growing, content description, region adjacency graph, multi-scale approach

1. INTRODUCTION

Archiving of visual information in medicine is not only important for medical and legal documentation but is also a valuable source of knowledge for comparison, statistical inquiry, and education. Fields of application are differential diagnosis, therapy control, or real world cases as educational examples. A central claim of image retrieval in medical applications (IRMA) is the independency of image modality, image content, and diagnostic question.¹ This is important because in clinical routine and especially in diagnostic radiology, a large variety of images regarding numerous questions is processed. The physicians need an integrated software covering all kinds of applications, which requires a general content-based image retrieval (CBIR) approach.² Moreover, system interaction has to be reduced to a minimum. Above all, manual low level parameterization of image processing must be avoided. With these applications and constraints, a CBIR system requires efficient data structures for storing, indexing, and retrieving images by contextual content dependent means. One of the central operations and main challenges in the image retrieval process is content-based image comparison. Consequently, a formal content description is needed, which contains all visually relevant image information and can be efficiently compared to other images.

There are two major constraints to formal image content description. On the one hand, it must be as general as possible to be valid for a variety of images. On the other hand, it has to be as specific as possible to include all information contained in an image. Alphanumeric descriptions as the most flexible mean of complex information

Christian Thies, e-mail: cthies@mi.rwth-aachen.de, phone: +49 241 8088795, fax: +49 241 803388790, Dept. of Medical Informatics, Aachen University of Technology, Pauwelsstr. 30, 52057 Aachen, Germany

storage and comparison are of limited use to medical applications since they differ from physician to physician.² Moreover, the information has to be extracted automatically from the image.

Appearance based approaches for measuring image similarity such as the tangent distance have been found to be a powerful and reliable tool for image comparison.³ However, they can hardly be applied in a generalized medical routine application since each new specification of image content has to be trained by a set of manually classified sample images. This is not practicable in the above described radiologic setting.

To overcome the duality of generalization and specialization of content description, many approaches invoke global feature values such as histograms or position independent region/edge counts⁴⁻⁶ Those features are extracted automatically from the image. With respect to a feature vector images can be efficiently compared via distances in the high dimensional feature space. However, this excludes local information but in medical applications, findings depend on the topological arrangement of objects. For instance a femur fracture in an x-ray image can be visible as a black gap between two white rectangular shapes. Therefore, spatial representation of image regions is necessary. This leads directly to the problem of spatial queries and image segmentation.^{7,8}

Interactive image segmentation is one way to define regions of interest for which appropriate feature vectors can be extracted. Afterwards CBIR is performed by comparing these vectors to identical regions in other images.⁹ In medical applications, this approach yields good results for special tasks, e.g., as shown by the ASSERT system on high resolution computed tomography archives of lung cancer images. However, the approach can hardly be generalized.¹⁰ The same holds for other systems,^{11,12} automated segmentation without any assumption on sought objects is required. Thus, an image has to be partitioned in visually perceivable regions. If the partitioning is sufficiently generalized, similar objects are represented identically in different images. A promising approach is persecuted by the BLOBWORLD system. Here, pixels with similar visual information such as homogeneity or texture are clustered. The resulting clusters are represented by their second area moment, mean feature value, and variance and represent local image content. In BLOBWORLD, retrieval becomes the finding of similar cluster descriptors by their Euclidean distance.¹³

However, all of those approaches only extract mono-scalar image descriptions without considering the multi-scale property of visual information. For example, in a frontal radiography of a skull, the skull itself may be of interest for a certain investigation, while for another task, the eye holes lying within the skull are regarded. Thus, the data structure has to reflect the skull containing the eye holes that lay left and right of the nose respectively.

The new algorithm introduced in this paper builds a cluster hierarchy by successively pooling adjacent regions into bigger ones. This provides the necessary multiscale representation of all visually perceivable image regions in a tree-like data structure preserving their topological arrangement. Exclusively unifying existing clusters, the challenging problem of causality in scale space is solved implicitly.¹⁴ The algorithm operates without parameterization and fulfills the demands of generalization and specificity. The focus of this work lies on the algorithm and its visually sensible results. Actual integration of the data structure into CBIR systems is proposed but not further evaluated.

This paper is structured as follows. At first the data structure is introduced (Sec. 2) where the clustering (Sec. 2.1), visually significant features (Sec. 2.2), description of the algorithm (Sec. 2.3), and its complexity (Sec. 2.4) is treated. This is followed by experiments on a set of characteristic images (Sec. 3) and their results (Sec. 4). The discussion focuses on the application of the data structure for CBIR (Sec. 5). Finally, conclusions are drawn (Sec. 6).

2. GRAPH REPRESENTATION OF IMAGE CONTENT

Content description is based on information extraction. For image retrieval approaches, content has to be extracted in an all-enclosing manner because it is previously unknown which image aspect will be requested by the user. For that, an information-enclosing data structure is computed on the basis of a hierarchical clustering algorithm. Based on pixel features such as the gray value, the algorithm is performed similar to region-growing but without seed pixels or other constraints.

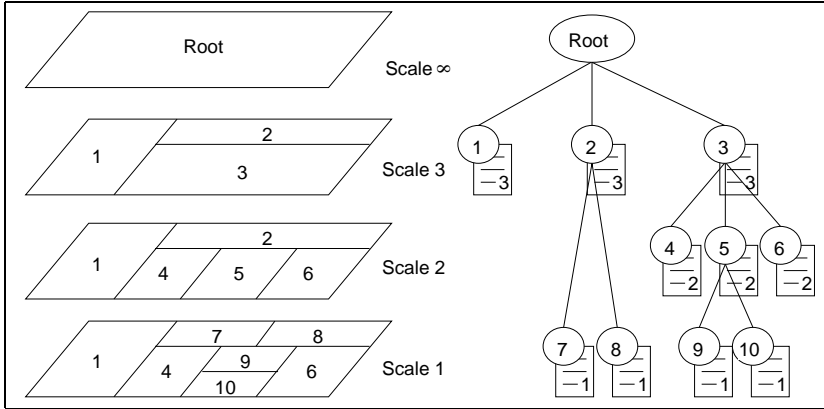


Figure 1. The region hierarchy describes the inter-scale inclusion relation for each region between the scales. All edges in coarser scales have corresponding edges in finer scales (left). This causality leads to a tree structure (right), where each region is represented as a node (circles) with an attached feature vector (boxes). The edges represent a region's inclusions. For a better overview the neighbor adjacencies are not displayed in this figure.

2.1. A region adjacency graph by hierarchical clustering

Aim of the clustering is a data structure containing all visually significant regions by fulfilling the claimed multi-scale criterion for complete information extraction.¹⁵ Assuming that image regions describe visual information in an appropriate way and that each pixel carries image information, a complete partitioning of the image is required. In this case, clustering of the pixel set means processing of its power set to find the appropriate partitions. Since a single level cannot contain all degrees of details, the multiscale approach results in several levels containing fractions or unions of such partitionings (Fig. 1). The first layer of the data structure consists of the image pixels and the preselected feature vectors for each pixel position. Based on the 4-neighborhood on the pixel grid, each pixel is regarded as an attributed graph node and its four neighbors as adjacent nodes. The pairwise Euclidean distance between the feature vectors of two neighbored nodes serves as a measure of similarity. It is assumed that visually perceivable regions consist of similar components. Therefore, the smallest Euclidean distance in the entire image is determined and all of the pixel pairs with this distance in their 4-neighborhood are merged. During the merge step, new regions are created. They can be described by the mean of all participating feature vectors. Those new regions are father nodes to the contained pixels, and can be seen as connected components. Each element that is connected to a pixel outside of the connected component produces an adjacency for the new region. New regions hold new adjacencies and require an updating of their similarity. A continuous iteration of the determination of the smallest distance, the adjacency merging, and the update step lead to the whole image as the final cluster. Protocolling the merge process as a graph results in a tree-like data structure, where image regions are stored as entities described by their pixel positions. Additionally, adjacent regions are represented as a list of neighbors, succeeding subregions as a list of sons, and enclosing regions as father nodes reflect the local topology of the regions. The pixel set refers to the positions in the original image from which complex descriptors can be extracted. Suitable descriptors are the mean or variance of feature vectors as well as regional shape parameters or area moments. When using clustering approaches on the pixel grid, the challenging problem of causality in multiscale image decompositions is circumvented because the cluster hierarchy induces a trivial causal inclusion of subsequent regions. From a theoretical point of view, the partitioning process as a causal recombination of connected components means operating on a complete lattice.¹⁶

2.2. Image features

Finding visually perceivable partitions in a two-dimensional image is formalized as finding connected components in a planar graph, where the connectivity criterion is the visual significance. The most simple criterion is homogeneity of intensity, which is not always sufficient to reflect visual comprehension. A variety of features has been proposed which can roughly be classified into regional aspects such as texture or local aspects such as contrast.^{11,20} Each of these features correlates with a specific perception. Thus, a sensible combination of features is the first step in the proposed image decomposition.¹⁷ As an example, sometimes textured areas with high intensity dynamics can be hardly considered as homogenous or specifically separated from other regions by contrast. They differ from neighboring regions mainly by change of polarity which is not caught by the intensity homogeneity. Since image features depend on modality or question to the image content, their extraction is the first step in the processing chain for constructing an appropriate hierarchy. In the IRMA system, this is done

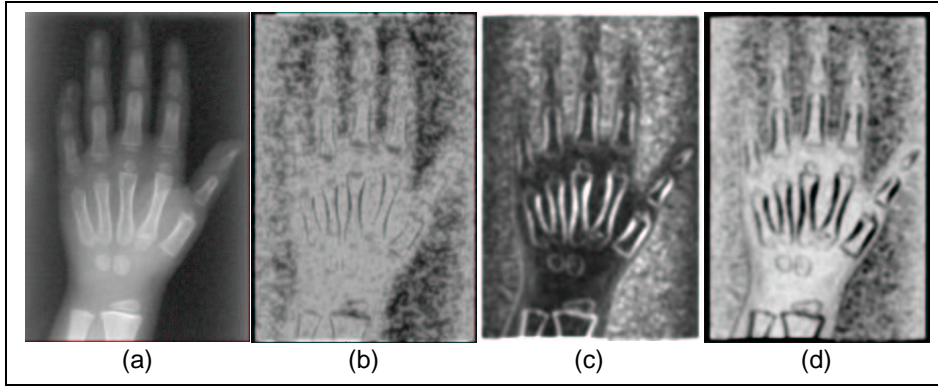


Figure 2. Texture features may be used to emphasize characteristic visual impressions. An image (a) is analyzed by its polarity (b), contrast (c), and anisotropy (d).

by a preprocessing layer called categorization to determine the image modality, orientation, body region, and biological system.¹

Furthermore, the necessity of a multiscale approach requires a feature computation for the finest scale. In other words, pixel-based features are required. Local features such as contrast only take their direct neighbors into account and, therefore, they are strictly local. In contrast, texture is a regional feature that characterizes the intensity distribution in a neighborhood with respect to its local replications. It is therefore a multiscale feature itself. However, it can be used to classify a pixel with respect to texture properties in its neighborhood such as anisotropy or polarity.¹⁸ By now, three textural measures are integrated into the IRMA system, which have already been successfully applied in the BLOBWORLD approach¹³ (Fig. 2):

- contrast, i.e., the identification of maximal grey value difference;
- polarity, i.e., the main direction of texture; and
- anisotropy, i.e., the degree of ordered directions in the actual environment.

Each of these features is computed from the eigenvalues of a local gradient of a pixel and reflects local texture characteristics. By convolution of each pixel with templates of increasing size, the local information is correlated to its environment on different scales. The value is considered as stable and selected as the pixel's feature value if the actual polarity value for a pixel does not differ by more than two percent¹³ between the preceding or succeeding scale. The three different features combined with the intensity form a four-dimensional feature vector for each pixel that is used to characterize its visual significance.

2.3. Implementation

From a theoretical point of view, topological feature clustering is a precise and powerful partitioning technique. However, designing efficient algorithms is a non-trivial task preliminary with respect to the processing sequence of the neighboring pixels. Actually, an implemented algorithm can process pixel vectors only sequentially. Initializing a merge process means selecting one pixel at a time and computing the Euclidean distance to its neighbors. Consequently, there is always a preference for the first neighbor. This fundamental problem underlies all clustering algorithms. Two effects occur depending on the counter strategy. Either some large regions consume all the others without visual significance or restrictive rules leave to many small regions leading to an over-segmentation. Both problems can be dampened by heuristic convergence criteria,¹⁹ but the dependency of initial processing direction remains. A brute force approach, where only those regions with the smallest overall distance are merged in each step, is the only solution to overcome this effect. This is necessary to achieve independence with respect to rotation and translation of the image objects.

Our algorithm is initialized by regarding each pixel as a cluster with exactly four adjacencies, namely the neighbors in the 4-neighborhood on the pixel grid. In the first step, the smallest Euclidean distance between

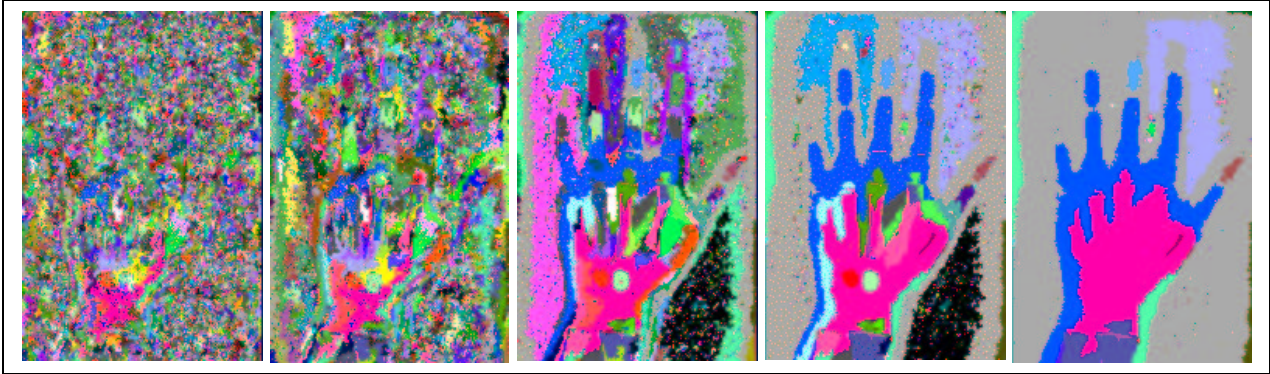


Figure 3. Fine scales hold small regions that are merged successively to larger ones (left to right). Since merging means inclusion, the edges of regions in coarser scales have a causal representation in finer scales.

pixel pairs is determined by scanning all adjacencies. In the second step, all adjacent regions with this value are merged. It is important to compute the transitive closure of adjacencies with the same distance to avoid wrong merging sequences on rotation of image objects. In the third step, the cluster descriptors are recalculated by computing the mean value of all contained pixel values. Finally, the merged clusters are integrated as the next level on top of the current cluster layer by updating the adjacency lists of the new clusters. This process is iterated until only one last cluster exists that represents the complete image. The algorithm can be summarized as follows:

1. initialize each pixel as a region with its 4-neighborhood adjacencies as neighbors;
2. compute distances between top level neighbors, insert adjacency in sort list if distance is same as or below the actual smallest value, update actual smallest value if distance is smaller;
3. if the last adjacency pair has not been visited, mark as visited, merge both clusters, add old clusters as sons, and remove last pair from list;
4. compute the transitive closure for both clusters with respect to the actual smallest distance value. If an adjacency equals the smallest value, merge regions into new region from Step 3 and mark processed clusters as visited;
5. repeat Steps 3–4 until no adjacency equals the smallest distance value;
6. update the new cluster descriptors and adjacencies to the old top level clusters; and
7. repeat Steps 2–6 until a merge results in the complete image as a cluster.

Figure 3 illustrates the computational process of the algorithm. To exemplify the iteration, computational data for the sample image of Figure 2 (a) is indicated. The radiograph image of a hand has a size of $155 \times 227 = 35185$ pixels. It is partitioned after 11179 iterations and the region adjacency graph yields 65452 regions. The sample scales as computational results of iterations 3100, 6600, 9800, 10900 11000 (from left to right) illustrate the evolution of visual perceivable partitionings. The computation took 130 minutes on a Pentium PC of 1 GHz. For better visualization, the regions are displayed in pseudo colors and not by the mean gray value.

2.4. Complexity and performance

The algorithm has a worst case complexity of $O(n^3)$ (n number of pixels) because of the two nested loops and the computation of the transitive closure in the inner loop. In the worst case, the inner loop (Steps 3–4) must process half of the region pairs leading to a transitive closure computation for all of their neighbors which also depends on the number of regions n . Planarity of the adjacency graph limits the number of possible adjacencies

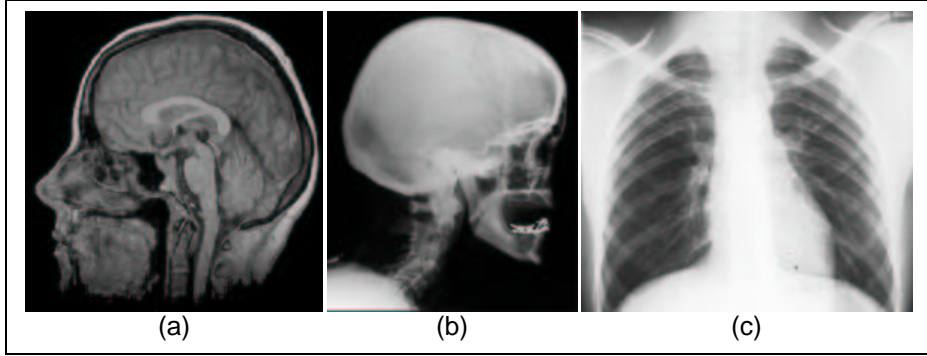


Figure 4. In the context of the IRMA project,¹ a large database of images from diagnostic radiology was collected. The partitioning algorithm is tested on the three displayed examples, a sagittal MRI of a head (a), a panoramic radiograph of the skull (b), and a common chest radiograph (c).

e to $e \leq 3n - 6$. However, this still depends on the number of clusters n . Thus, the inner loop causes n^2 steps. And since at least one merge takes place, the overall number of possible merges is $2 \cdot n$, which leads to an upper bound of n for the outer loop (Steps 2–6). In total, this leads to a runtime complexity of $O(n^3)$. But worst case complexity hardly occurs on real world images, because only few merges are necessary with respect to the smallest value. Thus, the inner loop reduces to minor fractions of n steps. Alternatively, homogenous regions are merged in one pass of the transitive closure computation leading to a smaller total number of necessary merges.

Nonetheless, the average case complexity is about $O(n^2)$. For each new cluster, all of the adjacent regions must be updated in Step 6. In the worst case the number is n . Furthermore, the outer loop is passed as often as so many regions exist because on real world images, only a few adjacencies form the smallest overall distance value. Computational costs depend on the image size but above all, on the distribution of descriptive distances.

3. EXPERIMENTS

A major goal of the algorithm is to provide a visual comprehensible and formal image content description. Therefore, the resulting data structure must fulfill three basic requirements: At first, the positioning invariance of image objects is examined. Secondly, runtime is investigated since the algorithm has to process images in daily routine applications. Data entry costs are less crucial as in real time applications but it must be possible to keep up with a steady stream of images. Finally, the visual quality of the partitioning has to be evaluated because image content is described via the extracted objects and their spatial relations.

3.1. Processing invariance

An important requirement of content extraction is its invariance against image transformations. Image acquisition may cause effects of rotation, flipping, or translation of displayed body parts. For instance, secondary digitized exposures may be turned when putting them on the scanner. For this analysis, an image is chosen where mainly homogenous regions with characteristic shapes describe the visual structures. In particular, a magnetic resonance image (MRI) is selected arbitrary from the IRMA reference database (Fig. 4 (a)). This image is processed before and after rotating the image by 90 degrees.

3.2. Runtime properties

Computing of the data structure may be executed offline which leads to minor importance of the actual computation time. However, an extensive latency is not recommended if the CBIR system serves as knowledge base. As indicated in Section 2.4, runtime depends on the distribution of similarity values. Therefore, images with differing similarity values in their local adjacencies are expected to cause the longest runtime. On the one hand, radiographs are basically noisy images with dynamic similarity values for their adjacencies. On the other hand, adjacencies with the same similarity are unlikely if pixels are represented in a high dimensional feature space. Therefore, a high-energy beam x-ray image (Fig. 4 (b)) was selected randomly from the IRMA database and segmented using both, gray values only and the four-dimensional feature vector of the BLOBWORLD system.

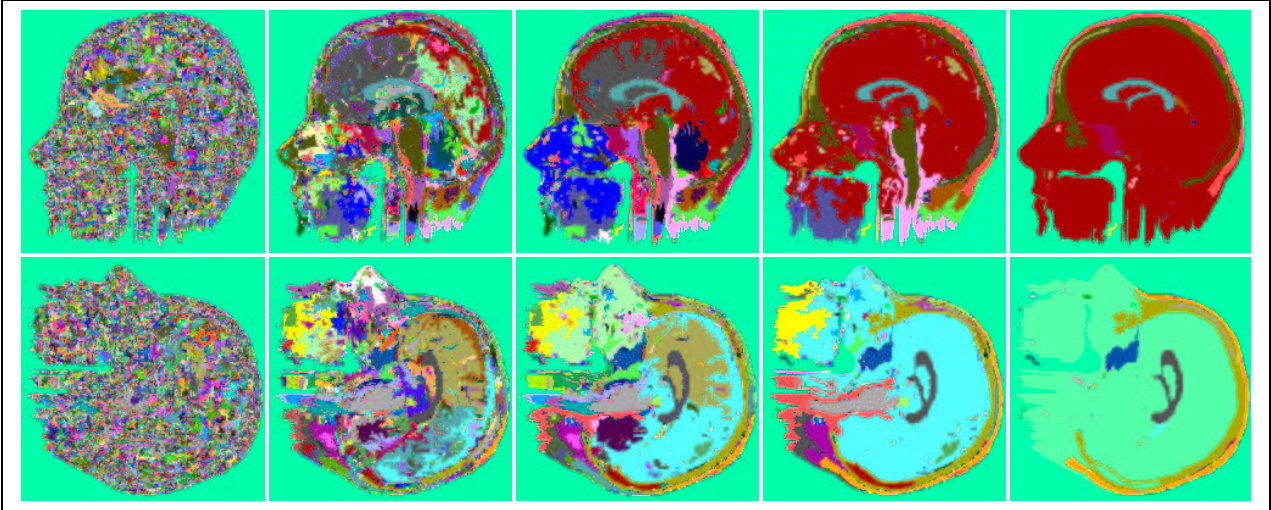


Figure 5. The sagittal MRI section of the head from Figure 4 (a) (top row) is rotated by 90 degrees (bottom row). Both rows show iterations 100, 4000, 5200, 5700, 6000 from left to right.

3.3. Visual quality

In contrast to the experiments described before, the assessment of visual relevance is of qualitative nature. We validate the algorithm's output by a human observer rating the precision of the extracted regions. The algorithm has to extract visual sensible regions from images of all kind of modalities and on all levels of detail. However, the exemplary MRI and radiograph from previous tests do not cover the great variety of medical imagery. Since anteroposterior chest radiographs are the most frequent investigation performed in a department of radiology, we chose a third radiograph of this type arbitrarily from the IRMA database for visual quality assessment (Fig. 4 (c)).

4. RESULTS

4.1. Processing invariance

Rotation has no effect on the merge process and provides the same partitioning at each iteration. All of the objects in the rotated MRI of the head in Figure 5 (bottom) such as the brain and its substructures remain unchanged in their visual appearance on all levels of detail. Every region is extracted in the original image in the same way as in the rotated version. For example, the corpus callosum is visible from iteration 4000, and has the same shape in the original as well as is the rotated image.

4.2. Runtime behavior

Since the average case complexity lies at $O(n^2)$ and highly depends on the intensity and similarity distribution, runtimes for an image vary between 60 and 450 minutes. Experiments were conducted on a Pentium PC with 1 GHz operating with Linux and yielded the following results:

image	size	features	# of regions	# of iterations	runtime/minutes
Fig. 4 (a)	$200 \times 200 = 40000$	gray value	56763	6196	60
Fig. 4 (b)	$177 \times 212 = 37524$	gv, C, P, A	75047	37523	450
Fig. 4 (b)	$177 \times 212 = 37524$	gray value	56491	7001	150
Fig. 4 (c)	$245 \times 222 = 54390$	gray value	81335	8878	100
Fig. 2 (a)	$155 \times 227 = 35185$	gray value	65452	11179	130

As expected, the image of the skull with the defined set of pixel features, took the longest time of computation as well as the highest possible number of regions (Fig. 6). For this image every step resulted in one single merge of

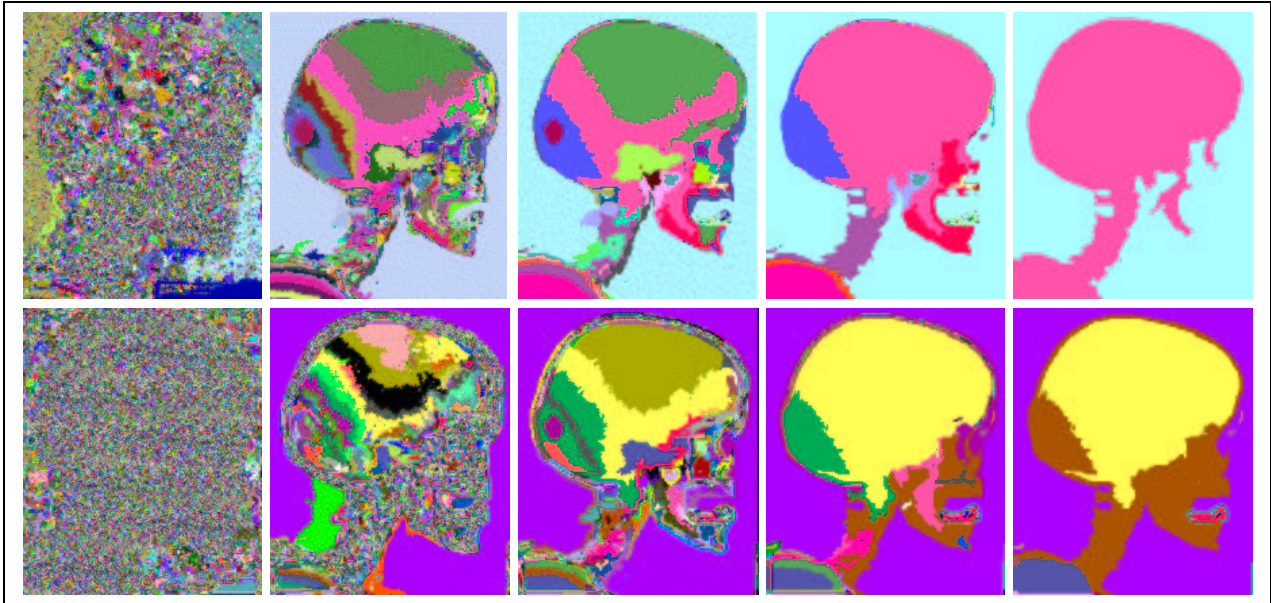


Figure 6. Computing the cluster hierarchy in the skull radiograph from Figure 4 (b) is done using the gray value only (top row) and the BLOBWORLD features (bottom row). The top row shows iterations 200, 2700, 6600, 6800, 7000, and the bottom row iterations 4800, 28700, 30000, 32300, 37500 (left to right).

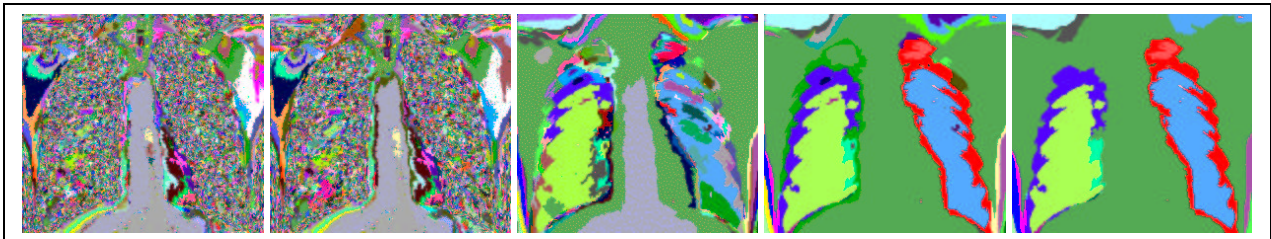


Figure 7. The scales for the chest radiograph from Figure 4 (c) are displayed from iterations 40, 1300, 5400, 8300, 8860 (left to right).

exactly two regions. Therefore, the total number of merges equals the number of pixels contained in the image. The other computations are based on the gray value only, which took significantly lower computation times. The fastest computation (60 min) was performed on the MRI of the head with its homogenous gray value structure. For this image, the smallest overall number of regions (56763) was computed.

4.3. Visual quality

Visual quality can be evaluated based on all experiments. In Figure 5, all important brain structures have been detected: The corpus callosum (iteration 4000) as well as different levels of detail on the brain stem (iteration 5200, 5700) or the cortical structure (iteration 5200). In Figure 6, some visible results differ significantly as the incomprehensible merge of the maxilla with the background indicates in the last two scales of the gray value based decomposition (top row). In contrast, the feature based decomposition produces a more comprehensible partitioning (bottom row). However, visually important structures as the dental fixation in iterations 32300 and 37500 are preserved more precisely by the texture-based features. In Figure 7, significant structures as the pulmonary wings (iterations 5400 to 8860) and the sternum (iterations 40 to 8300) are extracted. The characteristic shadow of the heart on the left wing is represented as well as the rib structure (iterations 5400 to 8860).

5. DISCUSSION

The algorithm has been tested on a few examples to illustrate the principal properties of the hierarchical clustering approach. At first, the basic aspects of reproducibility via the invariance argument was examined. Unlike classical region growing approaches,²¹ the algorithm avoids a preference of initial seed regions. The adjacencies are re-sorted independently from their local appearance and all neighboring regions are taken into account. However, ensuring rotation and translation invariance requires the parallel processing of each adjacency. This avoids different merging sequences for different initial processing directions and is ensured via the transitive closure computation. Here, more experiments are necessary to validate other transformations such as scaling, rotations by different angles, and projections.

Complexity analysis has indicated that the runtime depends on the distribution of adjacency similarities. When using high dimensional pixel feature vectors, the occurrence of similar distances is unlikely. This leads to one merge per step and consequently to long computations and large data structures. Homogenous areas cause fast merging of large regions by computing the transitive closure, which leads to faster computation times and less nodes in the region graph. Sizes of radiologic exposures are about 2000×2000 pixels, which lead to unbearable runtimes for $O(n^2)$ -algorithms. Therefore, the images have to be resampled to smaller sizes while protecting their aspect ratio. Finding the appropriate image sizes is still an important task. There are also several possibilities for runtime improvement. General speed-up may be gained by using indexed similarity search and local adjacency updates. Furthermore, the algorithm can be extended by allowing merges of adjacency pairs when their similarity lies below a given threshold. This results in more merges per step and therefore reduces the number of necessary updates. By now only those pairs with the smallest similarity are merged. Since the algorithm is to be applicable to clinical settings, the processing must not hinder the physicians work. As required, the algorithm itself is parameter free and consequently no additional handicap for the physicians arises.

The visual quality of the decomposition has been tested on MRI and plain x-ray images. Thereby, the resulting data structure represented different anatomic structures and substructures on several levels of detail. This is an interesting side effect of the data structure, because searching of single objects via subtree matching requires no image registration. As long as the describing feature vector is transformation invariant the subtree representation has no orientation. On the level of pixel feature extraction for the clustering step, first experiments have shown that the set of necessary features depends on the image modality. Thus, a preprocessing step for image categorization is necessary to determine the modality. This is conform with the IRMA concept. However, the selection of appropriate features is still a challenging task. All in all, the extracted regions for the tested image modalities and varying image contents showed the flexibility of the algorithm.

Extending common mono-scalar content descriptions for image retrieval by adding multiscale information as a cluster hierarchy is a new concept. With respect to the multiscale approach, the gathering of small regions is not an over segmentation but contains possibly relevant image details. Moreover, since the data structure is used for flexible content description, it is important to keep as much information as possible.

For integration into a retrieval system the flexible data structure offers several possibilities:

1. As visual information extraction is task specific, the set of features used for the clustering process can be adapted to the needs of different image modalities or content categories. This is important, e.g., if textured structures which are not necessarily of homogenous intensity are viewed as one region.
2. The feature vector that describes a region is computed on the raw data, which allows the extension of descriptive attributes when modeling expert knowledge. Since different retrieval approaches use different features, an easy integration of well examined solutions is sensible. For example, the BLOBWORLD system extracts a fixed number of regions and describes their second area moments.¹³
3. The number of regions taken into consideration can be adopted to specific tasks by restricting the appropriate attributes. This is also important for adaption to other systems or possibly for speedup of the matching algorithms.
4. The tree structure enables fast traversal algorithms and therefore guarantees fast query operations. As large image archives consist of several thousand images, a fast query processing is necessary, but the data entry time is not critical as long as the system can keep up with the incoming material.

6. CONCLUSION

We presented an algorithm to compute a formal data structure for image content description, which reflects the multiscale property of visual information. This data structure is developed to compare images in a CBIR system. The simple similarity criterion in the region merging process and its global consideration provide a parameter free clustering algorithm. However, reducing average case complexity is an important task. Furthermore, the race conditions between the best matching regions in the merge process is overcome. There is neither a preference for a single region nor the rasterization effect by strictly merging two regions. Avoiding initialization by seed points overcomes the problem of emphasizing specific visually outstanding image features. This offers a reproducible formal decomposition of images, which also reflects the variability of medical image material. The resulting data structure is a tree-like graph which allows polynomial-time complexity computing a matching with other image representing trees as a way to retrieve similar data structures and therefore similar images from a database.

REFERENCES

1. Lehmann TM, Wein BB, Dahmen J, Bredno J, Vogelsang F, Kohnen M: Content-based image retrieval in medical applications: a novel multi-step approach; *Procs. SPIE* 3972:312–320 2000.
2. Tagare HD, Carl Jaffe C, Duncan J; Medical image databases: A content-based retrieval approach; *Journal of Applied Mathematics and Image Analysis* 4:184–198, 1997.
3. Simard P, Le Cun Y, Denker J, Victorri B: Transformation invariance in pattern recognition – tangent distance and tangent propagation; *Lecture Notes in Computer Science*, Vol. 1524, Springer, pp. 239–274, 1998.
4. Vendrig J, Worring M, Smeulders AWM: Filter image browsing: Exploiting interaction in image retrieval; *Lecture Notes in Computer Science*, Vol. 1614, Springer, pp. 147–154, 1999.
5. Ogle VE, Stonebraker M: Chabot: Retrieval from a relational database of images; *IEEE Computer*, 28(9):40–48, 1995.
6. Brunelli R, Mich O: On the use of histograms for image retrieval; *Procs. IEEE Multimedia Systems '99 International Conference on Multimedia Computing and Systems*, pp. 143–147, 1999.
7. Chang SK, Yan CW, Dimitroff DC, Arndt T: An intelligent image database system; *IEEE Trans. on Software Engineering* 14(5):681–688, 1988.
8. Siebert A: Segmentation-based image retrieval: Storage and retrieval for image and video databases (SPIE), pp. 14–24, 1998.
9. Niblack W, Barber R, Equitz W, Flickner M, Glasman E, Petkovic D, Yanker P, Faloutsos C, and Taubin G: The QBIC project: Querying images by content using color, texture, and shape; *Procs. SPIE* 1908:173–187, 1993.
10. Shyu C, Brodley C, Kak A, Kosaka A, Aisen A, Broderick L: ASSERT – A physician-in-the-loop content-based image retrieval system for HRCT image databases; *Computer Vision and Image Understanding* 75(1/2):111–132, 1999.
11. van Ginneken B, Katsuragawa S, ter Haar Romeny BM, Doi K, Viergever MA: Automatic detection of abnormalities in chest radiographs using local texture analysis; *IEEE Trans. on Medical Imaging* 21(2):139–149, 2002.
12. Petrakis EGM, Faloutsos C: Similarity searching in medical image databases; *Knowledge and Data Engineering* 9(3):435–447, 1997.
13. Belongie S, Carson C, Greenspan H, Malik J: Color- and texture-based image segmentation using the expectation-maximization algorithm and its application to content-based image retrieval; *International Conference on Computer Vision*, Mumbai, India, pp. 675–82, 1998.
14. Lifschitz LM, Pizer SM: A multiresolution hierarchy approach to image segmentation based on intensity extrema; *IEEE Trans. on Pattern Analysis and Machine Intelligence* 12(6):529–540, 1990.
15. Metzler V, Thies C, Aach T: Object-oriented image analysis by evaluating the causal object hierarchy of a partitioned reconstructive scale-space; *Mathematical Morphology*, Talbot H, Beare R, (eds.), CSRIO Publishing, Collingwood, AUS, pp. 265–276, 2002.
16. Braga Neto U, Goutsias J: Connectivity on complete lattices: New results; *Computer Vision and Image Understanding* 85:22–53, 2002.

17. Fung PW, Grebbin G, Attikiouzel Y: Model-based region growing segmentation of textured images; *Procs. of the International Conference of Acoustics, Speech, and Signal Processing*, pp. 2313–2316, USA, 1990.
18. Haralick RM, Shanmugam K, Dinstein I: Textural features for image classification; *IEEE Trans. on Systems, Man and Cybernetics* 3(6):610–621, 1973.
19. Tilton JC: Method for recursive hierarchical segmentation by region growing and spectral clustering with a natural convergence criterion; *Disclosure of Invention and New Technology: NASA Case Number GSC 14, 328-1*, February 28, 2000.
20. Canny J: A computational approach to edge detection; *IEEE Trans. on Pattern Analysis and Machine Intelligence* 8(6):679–698, 1986.
21. Haris K, Efstratiadis SN, Magleveras N, Katsaggelos AK: Hybrid image segmentation using watersheds and fast region merging; *IEEE Trans. on Image Processing* 12(7):1684–1698, 1998.