Efficient Similarity Search on 3D Bounding Box Annotations

Hans-Peter Kriegel, Marisa Petri, Matthias Schubert, Michael Shekelyan, Michael Stockerl

Database Systems Group, Ludwig-Maximilians-Universität München, Munich, Germany

ABSTRACT

Searching for similar image regions in medical databases yields valuable information for diagnosis. However, most of the current approaches are restricted to special cases or they are only available for rather small data stores.

In this paper, we propose a fast query pipeline for 3D similarity queries on large databases of computed tomography (CT) scans consisting of minimum bounding box annotations. As these box annotations also contain background information which is not part of the item that was actually annotated, we employ approximate segmentation approaches for distinguishing between within-object texture and background texture in order to correctly describe the annotated objects. Our method allows a compact form of object description. In our framework, we exploit this advantage for enabling very fast query times.

We have validated our method on data sets of 111 and 1293 bounding box lesion annotations within the liver and other organs. Our experiments show a significant performance improvement over previous approaches in both runtime and precision.

Keywords: content-based image retrieval, pattern recognition, segmentation, runtime

1. INTRODUCTION

Content-based image retrieval (CBIR) and similarity search in particular can be a valuable tool for radiologists or other physicians in the clinical routine. Finding similar patient cases based on visual correlations facilitates the comparison of a current patient's status with earlier patient histories. Additionally, an up-to-date retrieval system can enable the identification of appropriate ongoing studies for a query patient. CBIR in medicine needs to address a number of technical problems like varying image quality and characteristics or legal barriers.[1,2] Furthermore, there are practical obstacles like a slowly-growing consumer acceptance.[3]

An additional constraint on any CBIR application is the high annotation effort. Image annotations are especially time-consuming if they involve the exact outline of a three-dimensional object within a volume, for example the slices of a CT scan.

Current CBIR approaches are therefore either restricted to small datasets of exact 2D annotations[4,5] or they approximate a larger set of 3D annotations as minimum bounding boxes or rectangles (MBRs),[6,7] instead of using more precise exact mesh or mask annotations. An advantage of MBRs is that they are very convenient for visual examination because any 3D region of interest (ROI) is usually loaded and displayed in a bounding box. However, a drawback of MBRs is that they also contain irrelevant background information which might disturb the majority of conventional image descriptors. In Ref. 7, this problem is dealt with by representing a box annotation as an unordered sequence of 2D image descriptors and by using a flexible but expensive distance measure. Though this approach yielded effective similarity search, the descriptors and the similarity measures are rather complex.

In this paper, we introduce more compact 3D image descriptors on MBR annotations which aim at excluding interfering background information without requiring an explicit 3D segmentation of the targeted ROI. A great variety of segmentation algorithms has been developed in the past years – especially for our use case of liver

Corresponding authors: M.P.:   E-mail: petri@dbs.ifi.lmu.de,   Telephone: +49 89 2180 9329
               M.Sch.: E-mail: schubert@dbs.ifi.lmu.de,   Telephone: +49 89 2180 9328

lesions.[8] Our method deliberately generates an ROI descriptor without the necessity of a full segmentation step and thus, it works much faster. We define a discriminative feature vector based on the intensity distributions of the observed Hounsfield Unit values in the target segments and non-target segments.

The rest of the paper is structured as follows: Section 2 gives an overview on related work, Section 3 describes the proposed CBIR method and Section 4 presents our experimental evaluation.

2. RELATED WORK

In recent years, CBIR in medical databases was mostly restricted to special cases like 2D skin lesions[9] or cervicographic images.[10] These specialized systems are usually used for decision support, i.e. they try to summarize their results in a way equivalent to a classification system distinguishing a limited set of cases. In Ref. 11 for instance, Wang et al. present an alternative classification system for lesion tissue classification using image patches via support vector machines. A recent summary over retrieval-based classification systems is listed in Ref. 2.

However, the classical idea of image retrieval has as well advanced into the field of liver lesions: Napels et al.[4] and Seifert et al.[7] describe similarity retrieval on liver lesions. In Ref. 4, the electronic representation of a liver lesion is derived from exact 2D segmentation annotations of the lesion's center slice, and it is focused rather on semantic ROI features than on automatically-generated image features, as elaborated in Ref. 12. A recent extension using only automatic image features still requires the exact segmentation of the lesion as it specifically targets the lesion's shape.[5] The query approach of Ref. 7 uses actual 3D annotations, however due to the high annotation overhead of exact lesion segmentations, we only generated a rough lesion representation by outlining bounding boxes. Since these bounding boxes are not restricted to the target lesions, they also contain background information of the liver or any bounding organs, bones or neighboring air.

In order to handle this problem, we are using a quasi-segmentation step in our feature generation approach. Therefore, we also point out a small sample of existing segmentation techniques in the field of liver lesions: Bellon et al.[13] describe a semi-automated watershed approach in 2D. Soler et al.[14] present a deformable model-based 3D liver segmentation approach, which also outlines lesions in order to improve the quality of the organ segmentation, whereas Taieb et al.[15] proposed a semi-supervised, iterative growth algorithm based on Gaussian intensity estimates. Militzer et al.[8] use a probabilistic boosting tree for detecting and segmenting two types of liver lesions. As our focus is not the exact segmentation of lesions but an acceptable separation of the lesion foreground from the background of the given box annotations, we did not yet test whether any of these methods were better suited for our query process than the method described later on.

3. METHOD

The method being described in this paper aims at the retrieval of similar ROIs in volume scans. Thus, our method needs to derive compact and descriptive feature representations for the annotated ROIs which are stored in a database. Now, given a query ROI and its descriptor, a query ranks the ROIs in the database with respect to the similarity to the query ROI. Since usually only the k most similar ROIs are of interest, our method supports top-k ranking queries which only retrieve the k most similar results. In the following, we will propose two new descriptors for representing 3D ROIs being contained in rectangular annotation boxes. Afterwards, we will sketch how the new descriptors are employed for the retrieval of similar ROIs from the database.

3.1 A Descriptor for ROIs in Box Annotations

In this paper, we aim to define compact and discriminative image descriptors for 3D MBR annotations which are well suited to find similar lesions. An example annotation of a lesion is depicted in Figure 1(a).

In Ref. 7, a first solution to this problem is introduced. The most simple image characteristic of a 3D annotation is its actual size in millimeters. The first lesion representation SIZE therefore consists of the extension of the lesion's bounding box in all three dimensions, i.e. a vector (w_x, w_y, w_z) as in Figure 1(b). A second approach is the use of 2D image descriptors (grey value histograms HIST and Haralick texture features[16] HAR). These descriptors are derived from the 2D rectangles on each slice corresponding to the 3D bounding box, thus forming a collection of feature vectors $A = \{a_1, a_2, \ldots, a_{10}\}$ of a previously unknown size like in Figure 1(c). To

deal with the effect of disturbing background noise and the matching problem of differently-sized annotations, the method employs flexible multi-instance distance measures. The best results were observed by using histogram features HIST with the Sum of Minimum Distances (SMD):

$$d_{\mathrm{SMD}}(A, B) = \frac{1}{2} \cdot \left(\frac{1}{|A|} \sum_{a_i \in A} \min_{b_j \in B} d_{\mathrm{man}}(a_i, b_j) + \frac{1}{|B|} \sum_{b_j \in B} \min_{a_i \in A} d_{\mathrm{man}}(b_j, a_i) \right) , \tag{1}$$

where the feature sets A and B represent two lesion annotations consisting of d-dimensional feature vectors $a_i, b_j \in \mathbb{R}^d$. To compare two instances, the Manhattan Distance $(d_{\mathrm{man}}(a_i, b_j) = \sum_{k=1}^{d} |a_{i,k} - b_{j,k}|)$ is employed.

In spite of a feasible performance in matters of descriptiveness, the used image representation and the pursued retrieval approach are very expensive w.r.t. both data storage and distance complexity as there are multiple image descriptors per ROI. In order to compare two ROIs A and B, it is necessary to compute the distance $d_{\mathrm{man}}(a_i, b_j)$ for each pair of instances making a distance computation quadratic with respect to its time complexity. In addition, high-dimensional slice-wise 2D feature descriptors are prone to over-fitting in case of very small box annotations.
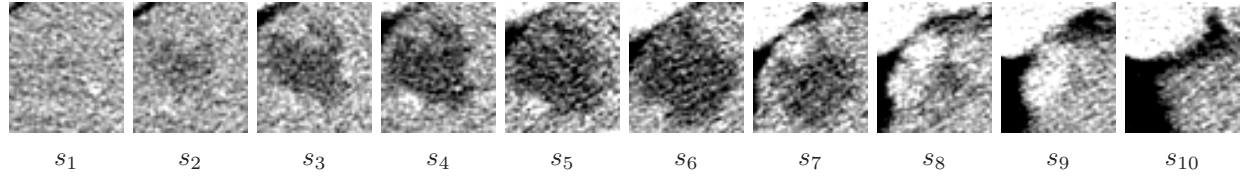
To avoid these disadvantages, our new method compares ROIs by representing them as a single vector which can be compared in linear time with any L_p metric, like the Euclidean or the Manhattan Distance. Furthermore, since the dimensionality of the resulting description vector is comparable to the description of a single slice-wise descriptor in the multi-instance model the space complexity is also considerably smaller. In the following, we will introduce two new descriptors for representing 3D ROIs.

Our first descriptor uses a rotation invariant shell kernel as a rough pseudo-segmentation step. We fit 3D ellipsoids of decreasing radius within the 3D bounding box of a given ROI. Regions between two succeeding ellipsoids form a shell which is described by a grey value histogram. In addition to the shells, we also include the background pixels not covered by any of the ellipsoid kernel's shells as an additional type of image information. This measure has been especially beneficial for our chosen validation procedure, where we compare feature-based similarity rankings with manually-assigned pair-wise similarity labels. The annotator is more likely going to qualify a pair of lesions to be similar if their backgrounds are of a similar appearance. The background pixels can thus be qualified as a rather valuable piece of information.

Our descriptor EHIST is therefore the concatenation of s d-dimensional grey value histograms formed from the shells defined by s concentric ellipsoids plus a d-dimensional grey value histogram formed from the remaining background pixels. In our experiments, the use of two shells and 60-dimensional grey value histograms turned out to perform best. Figure 1(d) visualizes the pixel selection kernels of a 3D ROI. The upper row displays the pixels selected as background information in 10 consecutive slices of a hypodense lesion. In the middle row, we see all pixels of the lesion's outer boundary approximation (Shell 1) and the bottom row visualizes the selected pixels of the innermost shell (Shell 2). The final feature vector of the descriptor (EHIST) has the length 180.

As second type of ROI description, we tested several more detailed segmentation approaches in order to better discern between background and foreground pixels. Our experiments showed, however, that a better segmentation does not necessarily lead to more descriptive image features. As Xu et al. reported good results using shape features for 2D applications,[5] we converted their image descriptors to 3D and tested their quality in comparison with the simple grey value histograms as used in the EHIST approach. However, the simpler EHIST features always resulted in better rankings. The reason for this effect is most likely that our segmentation method did not offer a precise enough separation of background and ROI for enabling the shape features to compete with a grey value histogram representation.

We will therefore only briefly describe the best-performing segmentation method we used for generating the new DHIST descriptors in the following. In a first step, we scale down the resolution of the annotation box to an extension of 24×24 w.r.t. width and height and a uniform height of 8. Now, each pixel in the re-scaled version is described by a texture descriptor which is derived from the corresponding pixels in the original image. Let us note that the texture descriptors can be efficiently computed in parallel by rescaling the region of interest into 10 disjunctive HU ranges. To segment the down-scaled version, we employ the clustering method GDBSCAN[17] based on a similarity measure considering spatial and textural proximity. Once a set of clusters has been found,
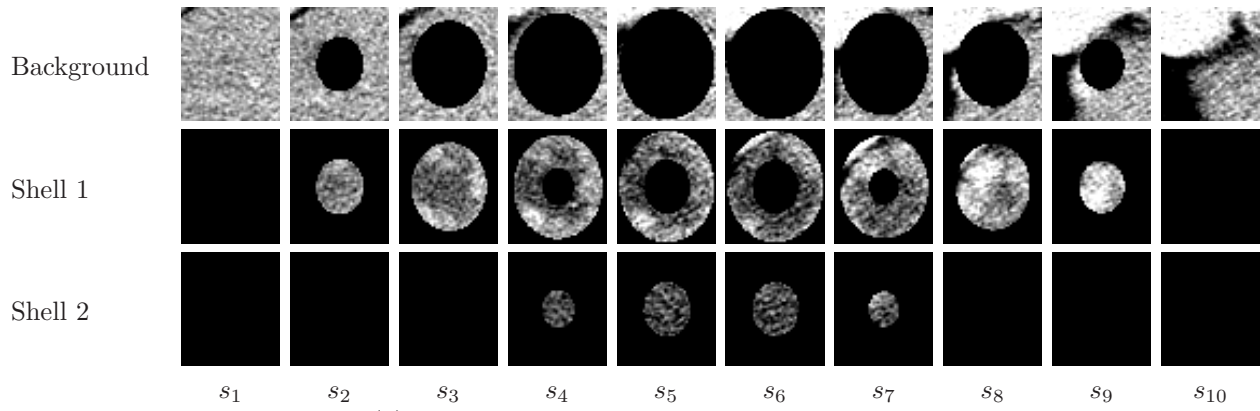
(a) Example lesion annotated as MBR in 10 slices s_i of a CT scan.

$$\begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} = \begin{pmatrix} 31.9 \\ 37.0 \\ 40.0 \end{pmatrix}$$
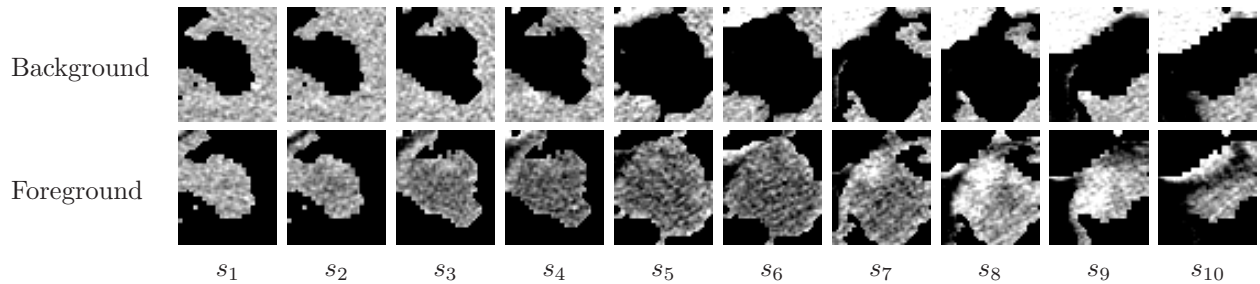
(b) SIZE descriptor: MBR width in millimeters.

$$A = \{a_1, a_2, \ldots, a_{10}\} = \left\{ \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{1,d} \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} a_{10,1} \\ \vdots \\ a_{10,d} \end{pmatrix} \right\}$$

(c) HIST and HAR descriptors: each slice s_i of (a) is described by a separate, d-dimensional feature vector a_i. They are collected as unordered set A.

(d) Ellipsoid segmentation for EHIST with $s = 2$.

(e) Density-based segmentation for DHIST.

Figure 1. Slice-wise display of an example lesion in (a) with various techniques of deriving image descriptors.

the clusters are summarized into groups describing foreground, air and noise clusters. This summarization step employs the centrality of the clusters as well as an air detection operation. The darker the texture of a cluster, the more likely is it going to consist of air and not of the annotated lesion and the more central a cluster is, the more likely is it going to contain the lesion. The pixels classified as foreground this way are then interpolated back into the original space resulting in a 3D segmentation of the original MBR into foreground and background pixels. An example segmentation is displayed in Figure 1(e).

We derive the descriptor DHIST D100 as the concatenation of 100-dimensional grey-value histograms for both the foreground and the background pixels. We additionally tested a variant DHIST D10 employing only 10-dimensional grey value histograms for both foreground and background segments. Despite the low dimensionality,

our experimental results show that this descriptor hardly suffers any loss in accuracy.

3.2 Query Processing

After introducing our methods to extract a vector description for the 3D ROI within an annotation box, we will now briefly describe the query processing system being employed on top of these descriptors. In general, the similarity search system has to maintain a database of MBR annotations being related to a volume in the PACS. For each MBR, we process the feature representation and store the triple of scan ID, annotation ID and descriptor in a dedicated index.

To search the database, our system processes example-based similarity queries, i.e. we retrieve the k most similar ROIs in the database for a given query ROI. In an application scenario, a physician might sketch an MBR around a certain type of ROI, e.g. a lesion and might afterwards pose a query to the database for the k most similar ROIs and their corresponding cases in the database.

To answer this type of query, a feature description for the query MBR has to be derived. Therefore, the effort of feature transformation might influence the query performance if the computational complexity is too extensive. Additionally, a time-consuming feature generation step increases the offline costs of inserting new annotations into the database. The descriptors being proposed in this paper are all sufficiently fast for online processing. However, more complex segmentation approaches on the original sized volumes may results in an considerable increase of the query processing times.

In the next step, the query descriptor is compared to the descriptors in the database. There are multiple similarity measures for comparing vector representations which can be employed to compute a dissimilarity score for two vector representations like the L_p metrics or the cosine coefficient. In our experiments, the Euclidean Distance i.e. the L_2 metrics displayed the best performance.

To determine the k most similar vectors in the database, the basic method is to compare the query with each database entry while maintaining a priority queue containing the k most similar entries being encountered so far. However, for large databases this simple approach might prove infeasible due to its linear runtime. Therefore, our query processing method is based on the X-tree[18] which is an extension of the R-tree being especially suitable for high-dimensional data as in our case. Though the query performance of high-dimensional index structures usually strongly decreases for feature descriptions having more than 20 dimensions, in our experiments it turned out that the X-tree was capable to improve query times even for the 180-dimensional EHIST descriptors. Let us note that this result does not contradict the well-known observations about the curse of dimensionality because there is no fixed value for the number of dimensions where an index structure does not yield any advantage compared to the basic linear search described above. Instead, query performance strongly depends on the data distribution and the internal dimensions of the employed feature space. Thus, the question of whether or not employing an index structure does make, must be examined for each particular feature description.

4. EXPERIMENTAL VALIDATION

We tested the quality of our retrieval approach on a dataset of 1293 lesions annotated as 3D MBRs on a database of 574 CT scans of 90 patients provided by the University Hospital of Erlangen for the research project THESEUS MEDICO. One example lesion is displayed in Figure 1(a).
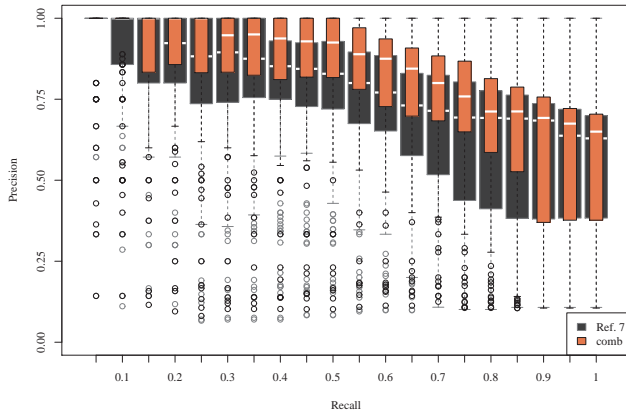
4.1 Retrieval Precision

In order to validate the precision of our method, we extracted 111 liver lesions (79 CT scans of 26 patients) for a qualitative examination. A medical expert has manually annotated those lesions with pair-wise similarity scores on a 5-step scale from 0 (completely dissimilar) to 100 (same lesion).

Table 1 shows the mean average precisions (MAP) and the average normalized discounted cumulative gain of the first 10 retrieved objects (nDCG(10))[19] of the 111 lesions tested in a leave-one out setting. As the MAP requires a relevance threshold, we considered a lesion pair to be relevant for a similarity score ≥ 50. The nDCG statistic does not require such a threshold, however, its results are comparable to those of the MAP.
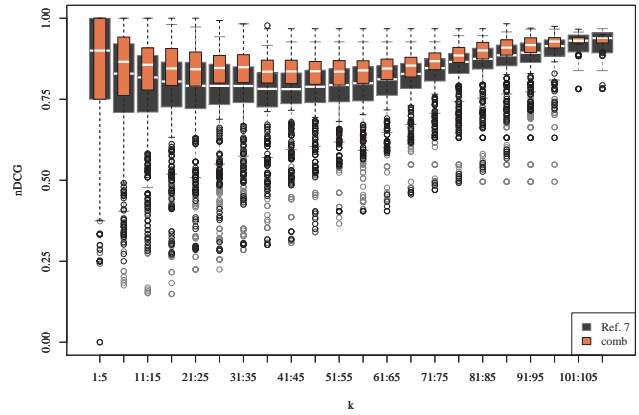
In Table 1, the rows HIST (slice-wise grey-value histograms), HAR (slice-wise Haralick features[16] retrieved from the windows of a partitioning grid of 3×3 divisions per slice) and SIZE (MBR extensions in millimeters in all

Table 1. Mean average precision (MAP) and the average normalized discounted cumulative gain[19] for the first 10 hits (nDCG(10)). The feature dimensionality d is listed in the last but one column. For the HIST and HAR features, it needs to be multiplied by the number of slices z per annotation. COMB is a combination of EHIST, DHIST D10 and SIZE.

Method	MAP	nDCG(10)	d	Feature Description
HIST	0.728	0.778	$z \cdot 150$	grey value histograms
HAR	0.653	0.717	$z \cdot 585$	Haralick texture features[16]
SIZE	0.655	0.694	3	dimension-wise MBR extension
REF. 7	0.735	0.817	$O(z \cdot 738)$	HIST : HAR : SIZE = 3 : 1 : 1
EHIST	0.781	0.817	180	ellipsoid-segmented histograms (two shells, 60 bins)
DHIST D10	0.763	0.791	20	density-clustered histograms (10 bins)
DHIST D100	0.762	0.808	200	density-clustered histograms (100 bins)
COMB	**0.788**	**0.837**	203	EHIST : DHIST D10 : SIZE = 3 : 2 : 1

(a) Precision-Recall curves of REF. 7 and COMB.

(b) nDCG[19] curves of REF. 7 and COMB.

Figure 2. Visual comparison of the approach in Ref. 7 and our new feature combination approach. The ranking evaluations are summarized via box-plots displaying the median en-boxed by the first and third quantile. The whiskers represent the farthest non-outliers.

three dimensions) display the performance of the single features we used in Seifert et al.,[7] and REF. 7 validates the proposed combination of all three representations. The above features are combined via a joint distance function, which consists in weighted distance sums of the different lesion representations. The single distance functions of REF. 7 are combined in the relation HIST : HAR : SIZE = 3 : 1 : 1. The different representations' distance ranges have been re-scaled to similar standard deviations in order to allow a fair proportion of the single features' distances.*

When examining the performance of the single features, the relatively good performance of the low-dimensional SIZE feature is striking. Our new compact 3D features, the 180-dimensional EHIST, the 20-dimensional DHIST D10 and the 200-dimensional DHIST D100 descriptors clearly perform better than the set-based image descriptors HIST and HAR, with a minor advantage of the larger DHIST D100 descriptors over DHIST D10. The combination COMB of the EHIST, DHIST D10 and SIZE descriptors is consequently superior to our previous approach REF. 7. The chosen combination weights for COMB are EHIST : DHIST D10 : SIZE = 3 : 2 : 1. A detailed comparison of the performance of the REF. 7 and COMB rankings is visualized in Figure 2. We see that both the precision recall curves and the nDCG curves of our new approach COMB are constantly better than the previous ranking REF. 7.
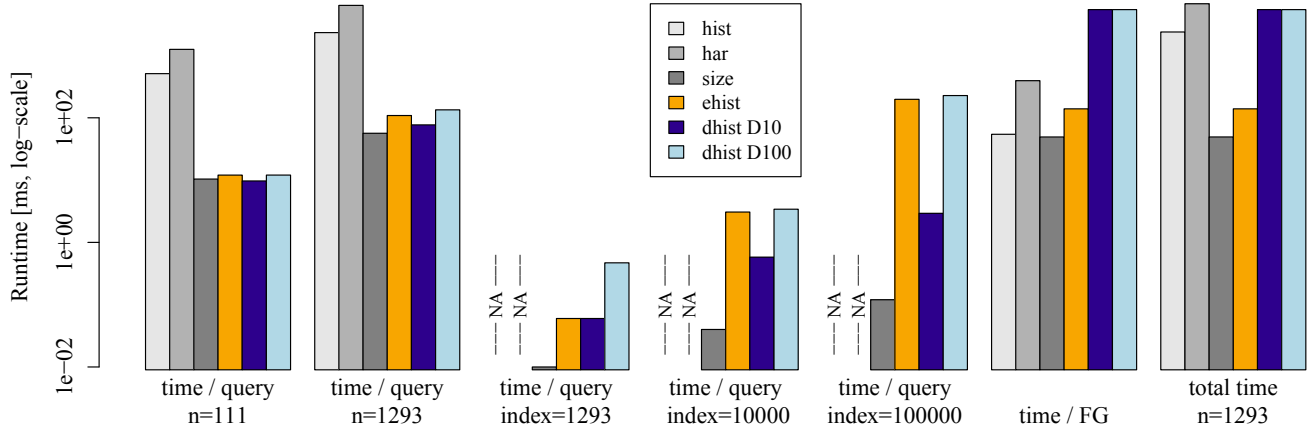
*For more details see Ref. 7.

Figure 3. Runtime evaluation of tested image descriptors. Experiments with size n specifications are measurements of database-scans, whereas experiments stating an index size show runtimes of spatial indexes. The last but one column displays feature generation runtimes (FG). All times are in milliseconds with a logarithmic scale.

4.2 Retrieval Time

We will now show that the computational cost of our approach is limited in comparison to the gain achieved by an improved descriptiveness of the resulting 3D image features. The retrieval times of the dataset of 111 lesions with similarity annotations are displayed in the first block of Figure 3 (in the following experiments, be aware of the logarithmic scale). We see that the set-based descriptors of the image features we developed in Ref. 7 are clearly more expensive than the more compact descriptors introduced in this paper. The next column displays the same effect on a larger database of 1293 lesions (which, due to its large size, could not be annotated with similarity scores). Clearly, the runtimes of 60 to 130 ms per query required by the newly-introduced features are a lot better than the runtimes of up to 6 seconds as required for the set-based descriptors HIST and HAR. Furthermore, due to the rather low dimensionality of the descriptor, it is possible to further speed-up query processing by employing a spatial index structure.

With the elevated dimensionalities of the EHIST and the DHIST D100 descriptor in mind, we use an X-tree,[18] which is suited for the fast retrieval of feature vectors in medium dimensionalities. In our implementation, we set the page sizes to 2 KB for the SIZE descriptor, to 16 KB for the DHIST D10 descriptor and to 160 KB for the EHIST and the DHIST D100 descriptors, respectively. The maximum overlap fraction is set to 20 %, the minimum fanout to 14 %. In addition, for a better balance of the generated search structure we re-introduced the re-insert functionality of the original R*-tree[20] into the X-tree, with a re-insert fraction of 30 %. This index structure allows very efficient top-k queries. The three center groups of the runtime measurements in Figure 3 show the retrieval times of the lesion database of 1293 lesions and larger databases containing descriptors of 10 000 and 100 000 random box annotations. We see that even large database sizes still allow efficient query response times for SIZE and 3D histogram descriptors. However, the DHIST D10 features are obviously better suited for a fast retrieval than the larger EHIST and DHIST D100 descriptors. Let us note that there are currently no index structures for accelerating query processing on set-based descriptors like HIST and HAR except for some special cases.

Using the proposed query processing method, the actual runtime bottleneck is the feature generation of the image descriptors displayed in the *time / FG* columns of Figure 3. The final columns show the total runtimes of a query on the lesion dataset of size 1293. As expected, EHIST was the fastest descriptor following the simple SIZE descriptor. The set-based descriptors of Ref. 7 are less expensive to compute than the advanced DHIST descriptors, however, due to their inefficient retrieval times, they already reach a break-even point in the total runtime for this rather small dataset. Our new histogram-based descriptors are therefore suited best for growing medical repositories containing multiple annotations for thousands of patients in both runtime and precision.

5. CONCLUSION

In this paper, we proposed an efficient retrieval framework for CBIR on three-dimensional image annotations. We have tested two new variants of compact 3D image descriptors based on intensity histograms and could clearly outperform our previous approach of Ref. 7 in both runtime and precision. The compact size of our image descriptors allows a very efficient retrieval on larger databases in comparison to the large-dimensional features used commonly. Owing to elevated needs of privacy protection in a clinical environment, cloud-based solutions may not be embraced for some time. Even though there are ways to use spatial index structures for high-dimensional features,[21] low-dimensional objects can still be queried faster. Furthermore, our retrieval framework is both well interpretable and it can be easily integrated into larger PACS architectures. In future work, we will examine new ways of further reducing the complexity of our image descriptors and we will test machine learning methods on their suitability of improving the accuracy of existing descriptors.

ACKNOWLEDGMENTS

REFERENCES

[1] Müller, H. and Deserno, T. M., "Content-based medical image retrieval," in [*Biomedical Image Processing*], Deserno, T. M., ed., 471–494, Springer Berlin Heidelberg (2011).

[2] Akgül, C. B., Rubin, D. L., Napel, S., Beaulieu, C. F., Greenspan, H., and Acar, B., "Content-based image retrieval in radiology: Current status and future directions," *Journal of Digital Imaging* **24**, 208–222 (2011).

[3] Antani, S., Xue, Z., Long, L. R., Bennet, D., Ward, S., and Thoma, G. R., "Is there a need for biomedical CBIR systems in clinical practice?: outcomes from a usability study," in [*Proceedings of the SPIE Medical Imaging Conference 2011: Advanced PACS-based Imaging Informatics and Therapeutic Applications, Lake Buena Vista, FL, USA*], **7967**, 796708 (2011).

[4] Napel, S. A., Beaulieu, C. F., Rodriguez, C., Cui, J., Xu, J., Gupta, A., Korenblum, D., Greenspan, H., Ma, Y., and Rubin, D. L., "Automated retrieval of CT images of liver lesions on the basis of image similarity: Method and preliminary results," *Radiology* **256**(1), 243–252 (2010).

[5] Xu, J., Faruque, J., Beaulieu, C., Rubin, D., and Napel, S., "A comprehensive descriptor of shape: Method and application to content-based retrieval of similar appearing lesions in medical images," *Journal of Digital Imaging* **ePub**, 1–8 (2011). 10.1007/s10278-011-9388-8.

[6] Criminisi, A., Shotton, J., Robertson, D., and Konukoglu, E., "Regression forests for efficient anatomy detection and localization in CT studies," in [*Proceedings of the MICCAI 2010 Workshop: Medical Computer Vision (MCV), Beijing, China*], 106–117, Springer (2011).

[7] Seifert, S., Thoma, M., Stegmaier, F., Hammon, M., Kramer, M., Huber, M., Kriegel, H.-P., Cavallaro, A., and Comaniciu, D., "Combined semantic and similarity search in medical image databases," in [*Proceedings of the SPIE Medical Imaging Conference 2011: Advanced PACS-based Imaging Informatics and Therapeutic Applications, Lake Buena Vista, FL, USA*], **7967**, 796702 (2011).

[8] Militzer, A., Hager, T., Jager, F., Tietjen, C., and Hornegger, J., "Automatic detection and segmentation of focal liver lesions in contrast enhanced CT images," in [*Proceedings of the 20th International Conference on Pattern Recognition (ICPR), Washington, DC, USA*], 2524–2527, IEEE Computer Society (2010).

[9] Baldi, A., Murace, R., Dragonetti, E., Manganaro, M., Guerra, O., Bizzi, S., and Galli, L., "Definition of an automated content-based image retrieval (CBIR) system for the comparison of dermoscopic images of pigmented skin lesions," *BioMedical Engineering OnLine* **8**(1), 18 (2009).

[10] Xue, Z., Antani, S., Long, L. R., Jeronimo, J., and Thoma, G. R., "Investigating CBIR techniques for cervicographic images," *AMIA Annual Symposium Proceedings* **2007**, 826–830 (2007).

[11] Wang, L., Zhang, Z., Liu, J., Jiang, B., Duan, X., Xie, Q., Hu, D., and Li, Z., "Classification of hepatic tissues from CT images based on texture features and multiclass support vector machines," in [*Advances in Neural Networks – ISNN 2009*], Yu, W., He, H., and Zhang, N., eds., *Lecture Notes in Computer Science* **5552**, 374–381, Springer Berlin / Heidelberg (2009).

[12] Korenblum, D., Rubin, D., Napel, S., Rodriguez, C., and Beaulieu, C., "Managing biomedical image metadata for search and retrieval of similar images," *Journal of Digital Imaging* **24**(4), 1–10 (2010).

[13] Bellon, E., Feron, M., Maes, F., van Hoe, L., Delaere, D., Haven, F., Sunaert, S., Baert, A., Marchal, G., and Suetens, P., "Evaluation of manual vs semi-automated delineation of liver lesions on CT images," *European Radiology* **7**(3), 432–438 (1997).

[14] Soler, L., Delingette, H., Malandain, G., Montagnat, J., Ayache, N., Koehl, C., Dourthe, O., Malassagne, B., Smith, M., Mutter, D., and Marescaux, J., "Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery," *Computer Aided Surgery* **6**(3), 131–142 (2001).

[15] Taieb, Y., Eliassaf, O., Freiman, M., Joskowicz, L., and Sosna, J., "An iterative bayesian approach for liver analysis: tumors validation study," in [*Proceedings of the 3D Segmentation in the Clinic MICCAI 2008 Workshop: Liver Tumor Segmentation (LTS'08)*], (2008).

[16] Haralick, R. M., Shanmugam, K., and Dinstein, I., "Textural features for image classification," *IEEE Transactions on Speech and Audio Processing* **3**(6), 610–623 (1973).

[17] Sander, J., Ester, M., Kriegel, H.-P., and Xu, X., "Density-based clustering in spatial databases: The algorithm GDBSCAN and its applications," *Data Mining and Knowledge Discovery* **2**, 169–194 (1998).

[18] Berchtold, S., Keim, D. A., and Kriegel, H.-P., "The X-Tree: An index structure for high-dimensional data," in [*Proceedings of the 22nd International Conference on Very Large Data Bases (VLDB), Bombay, India*], 28–39 (1996).

[19] Järvelin, K. and Kekäläinen, J., "Cumulated gain-based evaluation of IR techniques," *ACM Transactions on Information Systems* **20**(4), 422–446 (2002).

[20] Beckmann, N., Kriegel, H.-P., Schneider, R., and Seeger, B., "The R*-tree: An efficient and robust access method for points and rectangles," in [*Proceedings of the ACM International Conference on Management of Data (SIGMOD), Atlantic City, NJ*], 322–331 (1990).

[21] Samet, H., [*Foundations of Multidimensional and Metric Data Structures*], Morgan Kaufmann, San Francisco (2006).