Region–based volumetric medical image retrieval

Antonio Foncubierta–Rodríguez[a], Henning Müller[ab], Adrien Depeursinge[ab]

[a]University of Applied Sciences Western Switzerland (HES–SO), TechnoArk 3, CH–3960 Sierre, Switzerland;

[b]University and University Hospitals of Geneva (HUG), Department of Radiology, 4, rue Gabrielle–Perret–Gentil, CH–1211 Geneva 14, Switzerland;

ABSTRACT

Volumetric medical images contain an enormous amount of visual information that can discourage the exhaustive use of local descriptors for image analysis, comparison and retrieval. Distinctive features and patterns that need to be analyzed for finding diseases are most often local or regional, often in only very small parts of the image. Separating the large amount of image data that might contain little important information is an important task as it could reduce the current information overload of physicians and make clinical work more efficient. In this paper a novel method for detecting key–regions is introduced as a way of extending the concept of keypoints often used in 2D image analysis. In this way also computation is reduced as important visual features are only extracted from the detected key regions.

The region detection method is integrated into a platform–independent, web–based graphical interface for medical image visualization and retrieval in three dimensions. This web–based interface makes it easy to deploy on existing infrastructures in both small and large–scale clinical environments.

By including the region detection method into the interface, manual annotation is reduced and time is saved, making it possible to integrate the presented interface and methods into clinical routine and workflows, analyzing image data at a large scale.

Keywords: Content–based medical image retrieval, region–based image retrieval, 3D medical imaging, web–based interfaces, visualization.

1. INTRODUCTION

The use of medical images to assist clinicians in decision making has grown massively over the past decades and medical images are produced in vast quantities everyday.[1,2] A study estimates medical imaging to occupy 30% of world storage,[3] so management of the visual data is a task of real importance. Once image data have been used for diagnosis, reuse of images is very limited although massive amounts of knowledge are stored in the data. This limited reuse is partly due to privacy constraints[4] but also partly due to the difficulties in obtaining relevant images to queries in an effective and efficient manner. Content–Based Image Retrieval (CBIR) techniques are able to increase the re–usability potential of these images and medical cases in general, either for clinical decision support (CDS)[5–7] or for educational purposes.[8] Medical image retrieval has been an intensive area of research for the past 20 years.[1,2,9] Despite the efforts made in the field,[10,11] the clinical use of image retrieval remains underused or even unknown to many professionals.[2] To increase the reuse of the valuable knowledge stored in images daily produced in hospitals, content–based retrieval technologies need to be pushed into the clinical routine by creating new user interfaces and clinical workflows adapted to radiology routine.[12] Tools need to be able to make current work more effective and efficient and they need to deal with very large amounts of data to reduce physician overload.

Medical images contain an enormous amount of visual information, which makes it difficult and expensive to analyze them manually. In addition, abnormalities and clinically–relevant visual patterns often occur in a small region of interest of medical images,[13] which increases the need for local descriptions in contrast to global

Further author information: (Send correspondence to Antonio Foncubierta–Rodríguez)
Antonio Foncubierta–Rodríguez: E–mail: antonio.foncubierta@hevs.ch, Telephone: +41 (0)27 606–9010

descriptors. Current tomographic imaging modalities actually produce an increasing number of images so that manual work becomes more tedious than before. This also increases the risk that physician miss and important but small and localized pattern.

When local description is needed in image analysis, two approaches are commonly used: dense sampling and keypoint–based analysis. Dense sampling uses a fixed grid of points across the image where the local descriptors are computed. The decision of the sampling density is not trivial: if the grid is too dense there will be redundant descriptors and the computational cost is higher than with a sparse sampling; on the other hand, too sparse sampling has the risk to miss important information. Keypoint analysis tries to identify the optimal sampling points for computing the local descriptors. Keypoints have been widely used with various feature–detection techniques: borders, corners or salient points often based on gradients. The Scale–Invariant Feature Transform (SIFT)[14, 15] is a well–known method for automatically finding the number and location of keypoints. It is one of the most popular detectors used for 2D images (coupled with a descriptor) but there is no preferred method for local analysis in volumetric images (3D). However, SIFT is bound to the notion of point of interest and therefore the ability to determine an area or volume of interest is limited. The superpixel[16, 17] approach has been extended to supervoxels[18] and provides a region–based descriptor of the image. However, superpixels and supervoxels are defined as exhaustive partitions of the complete images and therefore do not reduce the amount of voxels to be analyzed. In order to achieve this, a second step that evaluates the meaningfulness of each superpixel or supervoxel would be required. Figure 1 shows examples of the detections made by the SIFT 1(a) and the superpixels 1(b) approaches

Besides the problem of reducing the number of voxels to be analyzed, there are other challenges that need to be addressed. Local description is a fuzzy term that raises the obvious question of *how local is local enough*? Existing key point methods such as SIFT already provide an idea of the area of influence of the keypoint, using size and orientation features as shown in Figure 1(a). In this article we propose that instead of obtaining a single keypoint, the detection works on a regional basis, providing a region of interest based on the actual patterns of the image with no predefined shape.

In this paper we present a novel region of interest detector for medical images that is able to first reduce the amount of data that needs to be analyzed and second provide locally relevant regions that allow for region–based CBIR . These characteristics make the technique suitable for general purpose CBIR[19] but also for many specific applications. The region detector is integrated into a web–based retrieval and visualization interface that maximizes the possibility to be adopted in clinical practice by using a platform–independent visualization toolkit that is easy to deploy on any existing infrastructure.

The rest of the paper is organized as follows: Section 2 explains the methods and materials used with specific focus on the region of interest detector, Section 3 presents the web–based interface for medical image retrieval as a result of the integration of the region detector and Section 4 discusses the conclusions and future work.

2. MATERIALS AND METHODS

Medical images contain a large amount of visual information. However, most important is often to capture with great detail the visual appearance of small regions of interest or areas with anomalies. This creates a need for local description of the images, allowing computer–based approaches to understand the images not as a whole but as one or several independent regions.

In this work, a key–region detector with the following features is developed:

- Multi–scale detection of *saliency* of regions in a SIFT–like fashion. Region detection is based on a wavelet pyramid, providing meaningful regions at various scales. Non–salient regions can therefore be discarded from analysis, reducing the amount of voxels to be analyzed and improving the description time.
- Geodesic mathematical morphology based on wavelet coefficients is used to determine the region extent.
- Mathematical definition of the region detector is dimensionality–independent. The algorithm can detect regions not only in 3D but in any type of signal from 1 to N dimensions.

The region–detection procedure is described in details in the following subsections.

(a) Keypoint detection used by SIFT.[15] Keypoints are displayed as vectors indicating location, scale and orientation.

(b) Superpixel partition of an image.[16]

Figure 1. Methods for keypoint detection and local analysis.

2.1. Multi–scale difference of Gaussians

Medical images are often defined in a set of heterogeneous physical magnitudes: inter–slice distance, spacing and voxel width change among different acquisition protocols. In order to have a stable ground for retrieval and image comparison, every image is first resampled to obtain 1 millimeter per side cubic voxels.

The volumetric image is filtered at various scales in a wavelet–based framework, using the difference of Gaussians as mother wavelet. This is shown in Equations 1 to 3.

Definition 1. *Let $I(\boldsymbol{x})$ be a n–dimensional image indexed by the coordinates $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and j a non–negative integer value. The difference of Gaussians wavelet at scale $s = 2^j$, $I_\psi^s(\boldsymbol{x})$ is defined as:*

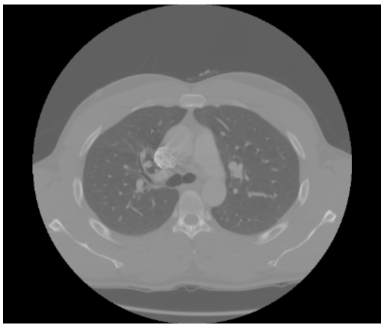
$$\sigma_1 = s;\ \sigma_2 = 1.6\sigma_1 \tag{1}$$

$$\psi_s(\boldsymbol{x}) = \frac{e^{-\frac{|\boldsymbol{x}|^2}{2\sigma_1^2}}}{\sigma_1^n\sqrt{(2\pi)^n}} - \frac{e^{-\frac{|\boldsymbol{x}|^2}{2\sigma_2^2}}}{\sigma_2^n\sqrt{(2\pi)^n}}. \tag{2}$$

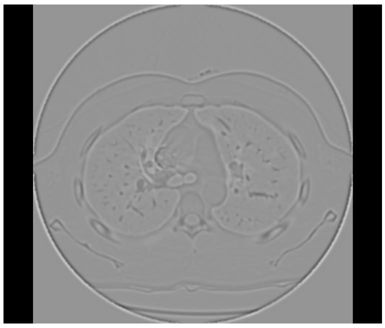
Then, the wavelet coefficients of an image $I(\boldsymbol{x})$ at scale s are obtained as:

$$I_\psi^s(\boldsymbol{x}) = I(\boldsymbol{x}) * \psi_s(\boldsymbol{x}). \tag{3}$$

Figure 2 shows an example of the result of a CT scan filtered at various wavelet scales.



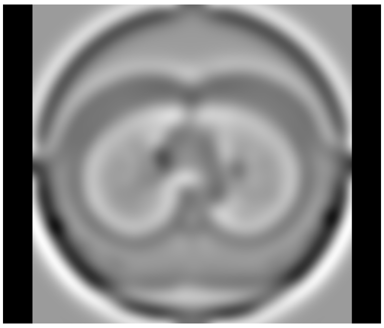
(a) Original image, resampled to 1mm–sided voxels.



(b) Wavelet coefficients image, at scale s=1.

(c) Wavelet coefficients image, at scale s=2.

(d) Wavelet coefficients image, at scale s=4.



(e) Wavelet coefficients image, at scale s=8.

(f) Wavelet coefficients image, at scale s=16.

Figure 2. Slice of a CT scan of the lungs and the results filtered at increasing wavelet scales.

The wavelet transform can be seen as a bank of band–pass filters. At each wavelet scale s, a range of frequencies is selected by the filter. Strong responses are produced when the image contains patterns in this specific frequency range. Small details, borders, blobs and ridges are contained in high frequencies (lower scales) and large objects have higher responses in lower frequencies as seen in Figure 2. Therefore, the aim of the multi–scale difference of Gaussians wavelet filtering is to be able to highlight *salient objects* of various sizes.

2.2. Geodesic detection of regional local extrema

The wavelet coefficient image highlights strong changes at a given scale with either positive or negative values. Values close to zero mean that the image does not contain large changes. Figure 3 shows how the local extrema of the wavelet images are related to salient objects for a given scale in comparison to the values of the original image, where saliency is not directly related to the gray level values but to its gradient.

Therefore, a way of computing the extrema is needed. Rather than identifying the position of the local maxima and minima, a geodesic algorithm is used to obtain the *region of influence* of the extrema. This provided a set of key–regions that correspond to the objects in terms of shape and size. The geodesic algorithm is performed as follows to each wavelet coefficient image at scale s:

1. In order to identify the regional minima and their extent, the difference of Gaussians image is considered a grayscale geodesic map. The *holes* (regions with gray values smaller than their neighbors) are filled using the fill hole algorithm.[20] As a result the new geodesic map at scale s, $I^s_{FILLED}(\boldsymbol{x})$, contains flat areas where the image contained a valley or hole.

2. By substracting the filled version of the image to the difference of Gaussian image, a grayscale map of the regional minima I_{MIN} can be obtained, i.e. a grey level image containing only the valleys and holes of the image.
$$I^s_{min}(\boldsymbol{x}) = I^s_{FILLED}(\boldsymbol{x}) - I^s_{\psi}(\boldsymbol{x})$$

3. The regional minima map is binarized, setting the voxel value to 1 only for those voxels in the holes of the difference of Gaussians.
$$B^s_{min}(\boldsymbol{x}) = \begin{cases} 0 \text{ if} I^s_{min}(\boldsymbol{x}) = 0 \\ 1 \text{ otherwise} \end{cases}$$

4. This process is repeated with the dual version of the algorithm, called grind peak,[20] and thus obtaining a binary map of the regional maxima $B^s_{max}(\boldsymbol{x})$.

5. A logic OR is performed on both maxima and minima binary maps, finally obtaining a map of the regional extrema in the difference of Gaussians of the image $R^s(\boldsymbol{x})$.
$$R^s(\boldsymbol{x}) = B^s_{min}(\boldsymbol{x}) \vee B^s_{max}(\boldsymbol{x})$$

6. Each fully connected component is labelled a *key–region* for scale s: $R^s_1, R^s_2, \ldots, R^s_k$.

Figure 4 shows an example of the regions detected at the end of the process.

2.3. Region–based descriptor for similarity–based retrieval

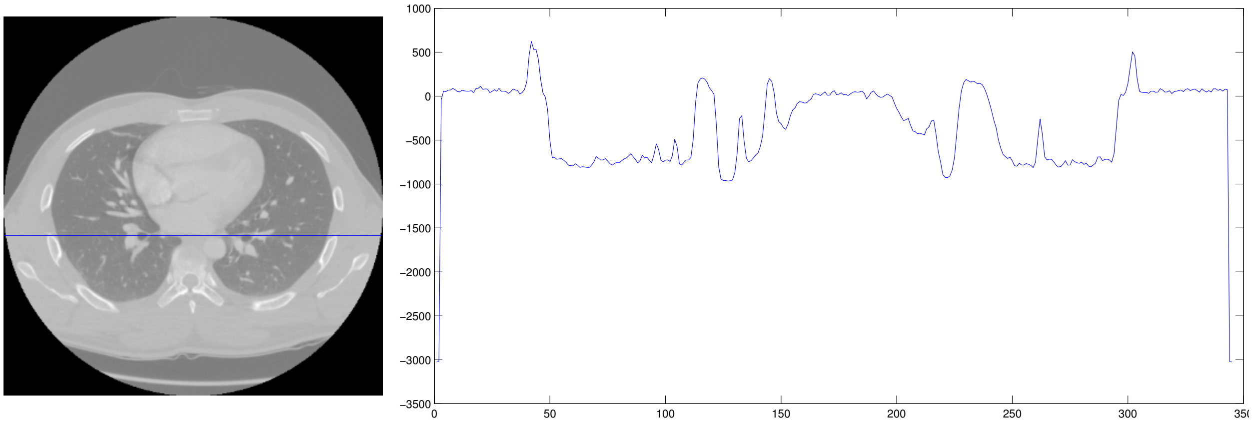
Region–based analysis enables the use of content–based retrieval that can help clinicians finding similar regions in other cases. The descriptor to be used strongly depends on the application, since similarity can be defined in terms of shape, size, texture or grey level (color) intensity.

In this work, a wavelet–based descriptor is used to characterize texture patterns contained in the images. Once the regions of interest are detected, the energies of the wavelet coefficients are used as a descriptor for similarity–based retrieval. For each region detected, the energies of the wavelet coefficients within the region are computed as a descriptor of the local texture.

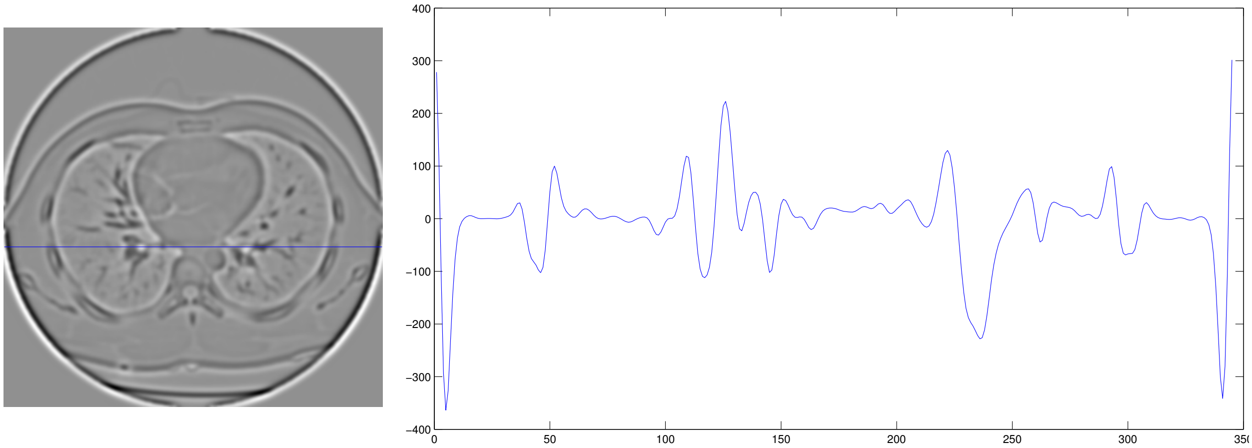
Definition 2. *Let $R_1, R_2, \ldots, R_k$ be k regions of interest. Then, an image $I(\boldsymbol{x})$ can be described by the set of k feature vectors F^I_i, $i = 1 \ldots k$ calculated as:*

$$E_i(I, s) = \sum_{\boldsymbol{x} \in R_i} I^s_{\psi}(\boldsymbol{x})^2. \tag{4}$$

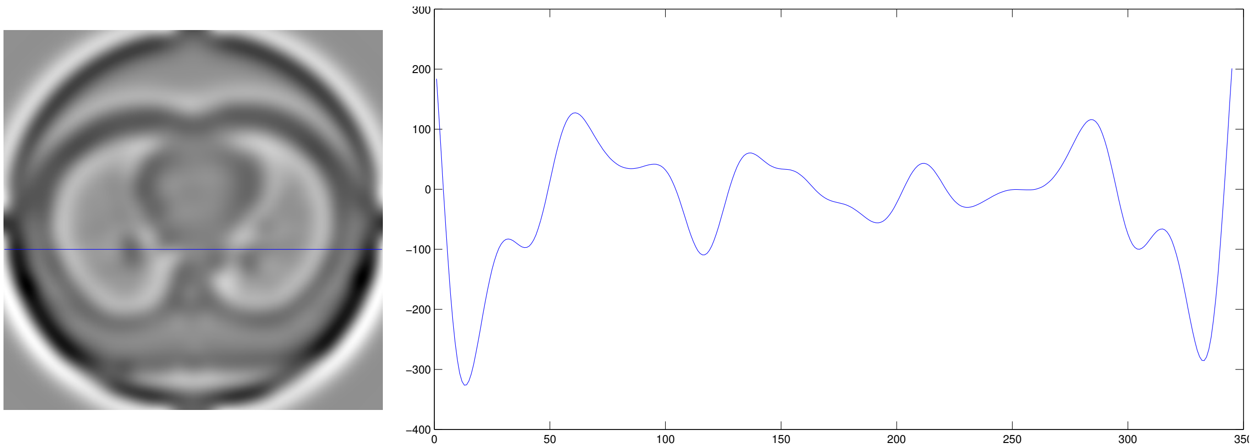
$$F^I_i = (E_i(I, 1)), E_i(I, 2), E_i(I, 4), E_i(I, 8)). \tag{5}$$



(a) Original image (left). Gray level values (right).

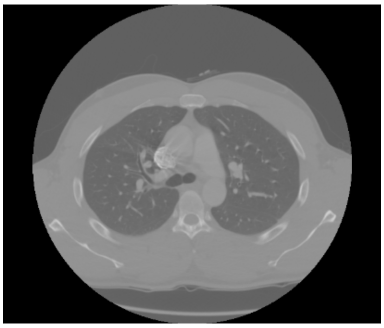


(b) Wavelet image at scale s = 2 (left). Coefficient values (right).

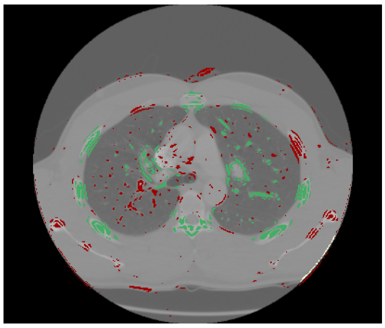


(c) Wavelet image at scale s = 8 (left). Coefficient values (right).

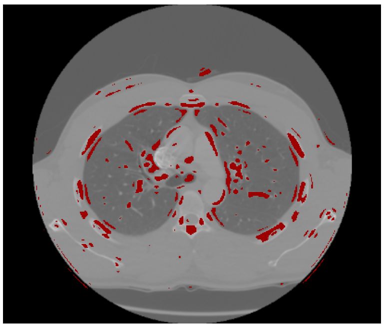
Figure 3. Examples of wavelet images and the representation of the values. Regional maxima and minima directly correspond to salient regions in the wavelet images 3(b) and 3(c) in contrast to 3(c), where saliency corresponds to large differences in gray levels.



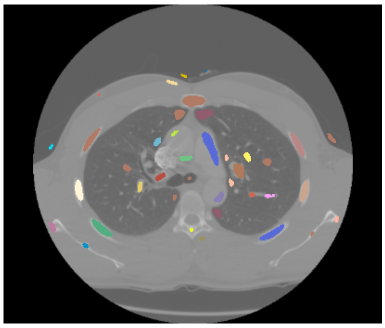
(a) Original image, resampled to 1mm–sided voxels.



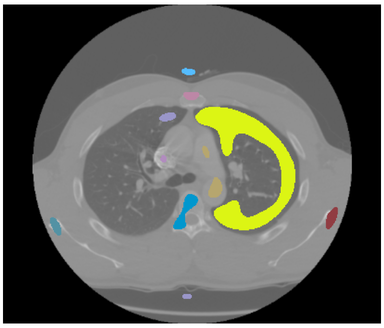
(b) Regions of interest detected at scale s=1.



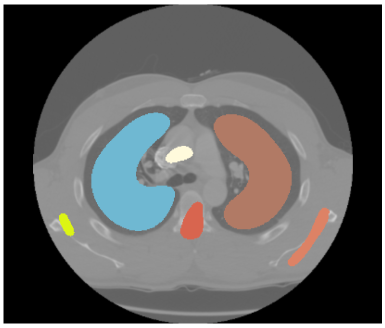
(c) Regions of interest detected at scale s=2.



(d) Regions of interest detected at scale s=4.



(e) Regions of interest detected at scale s=8.



(f) Regions of interest detected at scale s=16.

Figure 4. Examples of salient regions detected in lung CT scans at various scales.

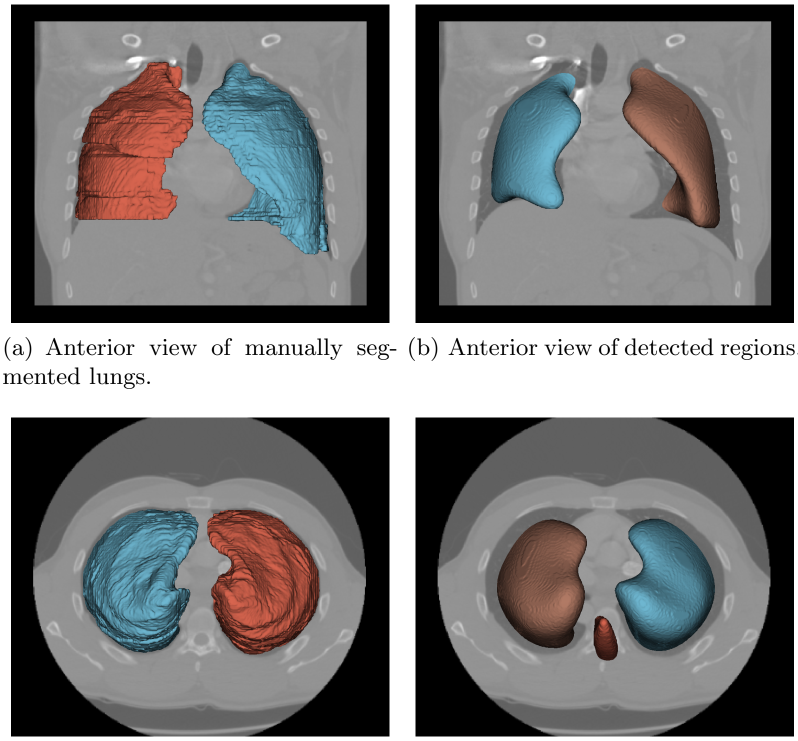
3. RESULTS

The region detector was successfully tested on medical data and meaningful anatomical regions can be detected. Figure 5 shows a comparison of regions automatically detected and a manual segmentation of the lungs. The region detector is also able to detect other regions such as the spine.

In order to provide a platform–independent retrieval interface that can be integrated seamlessly into clinical routine without affecting an existing infrastructure, a web–based interface was developed.

The interface is based on the X Toolkit (XTK)*,[21] an open source, javascript–based visualization toolkit that enables three dimensional model visualization, DICOM (Digital Imaging and COmmunications in Medicine) slice

*http://github.com/xtk/X/



(a) Anterior view of manually segmented lungs. (b) Anterior view of detected regions.

(c) Superior view of manually segmented lungs (d) Superior view of detected regions

Figure 5. Examples of salient regions detected in a lung CT scan compared to the manual segmentation of the lungs.

viewing and volume rendering.

Figure 6 shows a screenshot of the developed interface. The retrieval workflow is described as follows:

1. The user (clinician) uploads an image file as a query for content–based retrieval.
2. An external label–map can also be uploaded with the desired regions of interest, for example if the user has already generated manual annotations with third–party software.
3. Otherwise, the key–region detector produces a label map with the salient regions in the uploaded image.
4. Once the regions are generated or uploaded the clinician can select the relevant region for retrieval.
5. The retrieval engine produces a list of relevant regions in the database, grouped by patient.

The user can then load the visualization of the relevant items with the standard panning–zooming options. The interface also provides access to the complete information of the case (diagnosis, patient history, radiology report, etc.)

4. CONCLUSIONS AND FUTURE WORK

In this paper a novel *key–region* detector is presented that is able to reduce the amount of voxels to be analyzed in medical images. It provides locally relevant regions at several scales for further analysis or image retrieval. The algorithm automatically detects salient regions at various scales. This allows to extract features from all these regions at the various scales, which provides great flexibility in a retrieval applications where users might be interested in various sizes of elements.

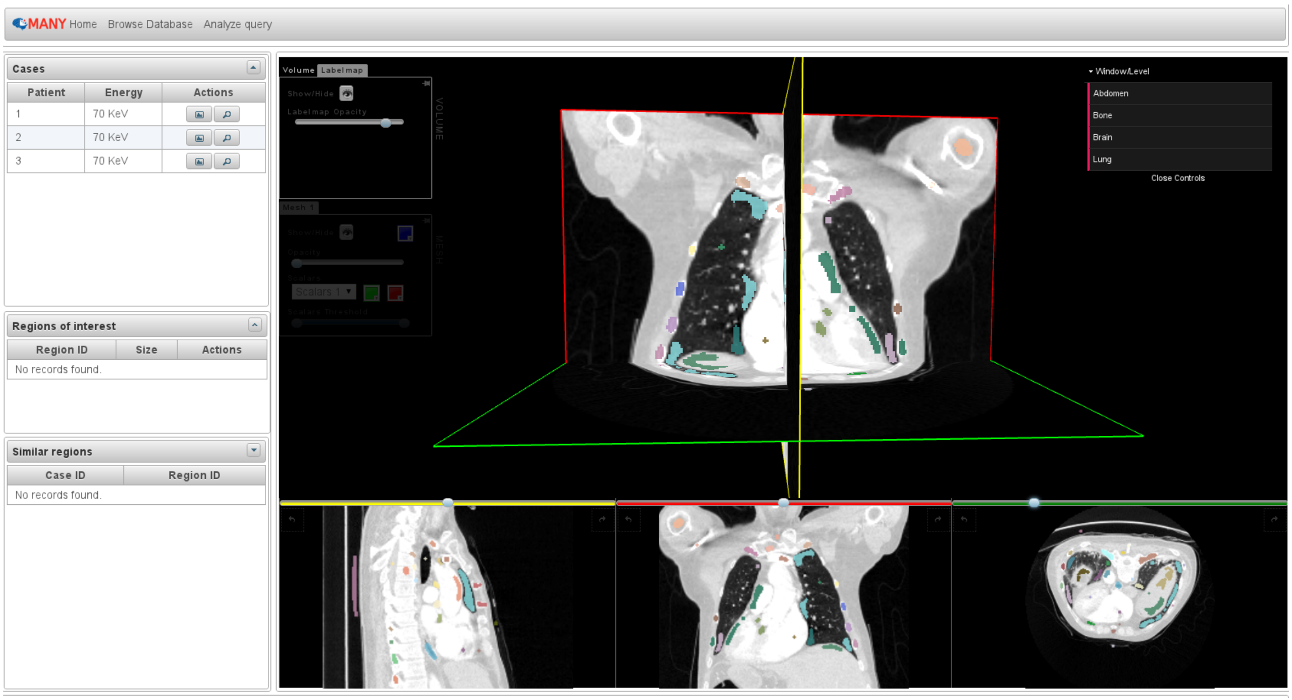


Figure 6. Screenshot of the region–based volumetric retrieval interface

The developed tool is integrated into a web–based retrieval interface for general–purpose volumetric medical image retrieval. The latter is developed using HTML and other platform–independent techniques enabling classical panning/zooming, volume rendering and slicing features that are usually only available on desktop interfaces. Region–based volumetric retrieval can improve the re–use of existing corpora of medical images in clinical practice. It allows focusing on relevant local visual information in medical images without requiring any additional user input. By integrating a novel region detector with a web–based interface and an image retrieval engine, a new tool is made available to the clinicians.

Future work includes adaptation of the user interface to smaller mobile devices and integration of text–based and multimodal queries for the retrieval of similar cases, as well as a clinical evaluation of the developed tools. 3D region detection and more importantly graphical user interfaces helping the query formulation and results display are extremely important to be able to navigate in the large and quickly growing amount of medical visual data.

5. ACKNOWLEDGMENTS

This work was supported by the European Union in the context of the Khresmoi project (grant number 257528) and the Swiss National Science Foundation (grant number 205321–130046).

REFERENCES

1. H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content–based image retrieval systems in medicine–clinical benefits and future directions," *International Journal of Medical Informatics* **73**(1), pp. 1–23, 2004.
2. C. Akgül, D. Rubin, S. Napel, C. Beaulieu, H. Greenspan, and B. Acar, "Content–based image retrieval in radiology: Current status and future directions," *Journal of Digital Imaging* **24**(2), pp. 208–222, 2011.
3. "Riding the wave: How europe can gain from the rising tide of scientific data." Submission to the European Comission, available online at `http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf`, October 2010.

4. B. Elger, J. Iavindrasana, L. L. Iacono, H. Müller, N. Roduit, P. Summers, and J. Wright, "Strategies for health data exchange for secondary, cross–institutional clinical research," *Computer Methods and Programs in Biomedicine* **99**, pp. 230–251, Sept. 2010.
5. C. E. Kahn, "Artificial intelligence in radiology: Decision support systems," *RadioGraphics* **14**, pp. 849–861, 1994.
6. C. LeBozec, M.-C. Jaulent, E. Zapletal, and P. Degoulet, "Unified modeling language and design of a case–based retrieval system in medical imaging," in *Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA)*, (Nashville, TN, USA), Oct. 1998.
7. A. A. T. Bui, R. K. Taira, J. D. N. Dionision, D. R. Aberle, S. El-Saden, and H. Kangarloo, "Evidence–based radiology," *Academic Radiology* **9**(6), pp. 662–669, 2002.
8. A. Rosset, H. Müller, M. Martins, N. Dfouni, J.-P. Vallée, and O. Ratib, "Casimage project — a digital teaching files authoring environment," *Journal of Thoracic Imaging* **19**(2), pp. 1–6, 2004.
9. H. D. Tagare, C. Jaffe, and J. Duncan, "Medical image databases: A content–based retrieval approach," **4**(3), pp. 184–198, 1997.
10. A. M. Aisen, L. S. Broderick, H. Winer-Muram, C. E. Brodley, A. C. Kak, C. Pavlopoulou, J. Dy, C.-R. Shyu, and A. Marchiori, "Automated storage and retrieval of thin–section CT images to assist diagnosis: System description and preliminary assessment," **228**, pp. 265–270, July 2003.
11. A. Depeursinge, A. Vargas, F. Gaillard, A. Platon, A. Geissbuhler, P.-A. Poletti, and H. Müller, "Case-based lung image categorization and retrieval for interstitial lung diseases: clinical workflows," *International Journal of Computer Assisted Radiology and Surgery* **7**, pp. 97–110, Jan. 2012.
12. A. Depeursinge, B. Fischer, H. Müller, and T. M. Deserno, "Prototypes for content–based image retrieval in clinical practice," *The Open Medical Informatics Journal* **5**, pp. 58–72, 2011.
13. D. Markonis, M. Holzer, S. Dung, A. Vargas, G. Langs, S. Kriewel, and H. Müller, "A survey on visual information search behavior and requirements of radiologists," *Methods of Information in Medicine* **51**(6), pp. 539–548, 2012.
14. D. G. Lowe, "Object recognition from local scale invariant features," in *Proceedings of the International Conference of Computer Vision*, (Corfu, Greece), Sept. 1999.
15. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision* **60**(2), pp. 91–110, 2004.
16. G. Mori, "Guiding model search using segmentation," in *Proceedings of the Tenth IEEE International Conference on Computer Vision - Volume 2, ICCV '05*, pp. 1417–1423, IEEE Computer Society, (Washington, DC, USA), 2005.
17. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**, pp. 2274 –2282, nov. 2012.
18. A. Burner, R. Donner, M. Mayerhoefer, M. Holzer, F. Kainberger, and G. Langs, "Texture bags: Anomaly retrieval in medical images based on local 3D–texture similarity," in *Medical Content-based Retrieval for Clinical Decision Support*, H. Greenspan, H. Müller, and T. Syeda-Mahmood, eds., *MCBR-CDS 2011* **7075**, pp. 116–127, Lecture Notes in Computer Sciences (LNCS), September 2012.
19. Y. Chen and J. Z. Wang, "A region–based fuzzy feature matching approach to content–based image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**, pp. 1252–1267, Sept. 2002.
20. P. Soille, *Morphological Image Analysis: Principles and Applications*, ch. 6. Springer, second edition ed., 2003.
21. D. Hähn, R. N., G. P.E., and P. R., "Slice:Drop," in *IEEE VisWeek, SciVis Poster Session*, 2012.