

# Automated Object Extraction for Medical Image Retrieval Using the Insight Toolkit (ITK)

Henning Müller, Joris Heuberger, Adrien Depeursinge, Antoine Geissbühler

University and Hospitals of Geneva, Medical Informatics, Geneva, Switzerland  
henning.mueller@sim.hcuge.ch

**Abstract.** Visual information retrieval is an emerging domain in the medical field as it has been in computer vision for more than ten years. It has the potential to help better managing the rising amount of visual medical data. One of the most frequent application fields for content-based medical image retrieval (CBIR) is diagnostic aid. By submitting an image showing a certain pathology to a CBIR system, the medical expert can easily find similar cases. A major problem for is the background surrounding the object in many medical images. System parameters of the imaging modalities are stored around the images in text as well as patient name or a logo of the institution. With such noisy input data, image retrieval often rather finds images where the object appears in the same area and is surrounded by similar structures. Whereas in specialised application domains, segmentation can focus the research on a particular area, PACS-like (Picture Archiving and Communication System) databases containing a large variety of images need a more general approach. This article describes an algorithm to extract the important object of the image to reduce the amount of data to be analysed for CBIR and focuses analysis to the important object. Most current solutions index the entire image without making a difference between object and background when using varied PACS-like databases or radiology teaching files. Our requirement is to have a fully automatic algorithm for object extraction. Medical images have the advantage to normally have one particular object more or less in the centre of the image. The database used for evaluating this task is taken from a radiology teaching file called *casimage* and the retrieval component is an open source retrieval engine called *medGIFT*.

## 1 Introduction

Content-based visual information or image retrieval (CBIR) has been an extremely active research area in the computer vision and image processing domains [1, 2]. A large number of systems has been developed, mostly research prototypes [3] but also commercial systems such as IBM's QBIC [4] or Virage [5]. The main reason for the development of these systems is the fact that an ever-growing amount of visual data is produced in many fields, for example with the availability of digital consumer cameras at low prices, but also with the possibility to make the visual data accessible on the Internet. The data commonly

analysed includes photographs, graphics, and videos. One typical application of image retrieval is trademark images [6].

The medical field is no exception to this, and a rising amount of visual data is being produced [7]. The radiology department of the University Hospitals of Geneva alone produces currently more than 25,000 images per day, mostly in multi-slice tomographic series. The importance of retrieval of medical images was identified early [8–10] and a large number of projects has started to index various kinds of medical images [11]. Not all of the projects are analysing the visual image content, some simply use the accompanying textual information for retrieval [12]. This is often called CBIR but should rather be called context-based retrieval as the text describes the context, in which the image was taken [13] rather than its content. Very few projects are currently used in clinical routine. Most projects have been developed as research prototypes but without a direct link to a need in a clinical application [14, 15]. An example for a system that was used as a prototype in a clinical setting is *Assert* that showed a significant improvement of correct diagnosis when using the system [16, 17], especially among less experienced radiologists. Another active medical image retrieval project is *IRMA*<sup>1</sup> [18, 19], that concentrates on image classification and the segmentation of medical images for retrieval. An annotated database was developed in this project and a multi-axial code for image annotation [20].

*medGIFT*<sup>2</sup>, our CBIR tool, extracts mainly local and global features such as textures (based on Gabor filter responses) and grey level descriptors for the visual similarity retrieval from our teaching file. It is based on the GNU Image Finding Tool (*GIFT*<sup>3</sup>) [21] and includes modifications to use a slightly different feature space. One of the identified problems of retrieval is that a large part of the images does not contain any important information for retrieval but rather noise. This noise can be in the form of black areas around the principal object, but more often in the form of text and logos that occur around a large number of objects in the images. In this case, the area where the main object appears and the kind of noise in the image has a significant influence on the retrieval that is sometimes hindering a good performance. Our goal was to develop a completely automatic algorithm to reduce this background noise and extract the important object in the medical images of our database *casimage*<sup>4</sup>, a teaching file containing almost 9000 extremely varied images from several modalities (CT, MRI, PET, ...) as well as photographs and even powerpoint slides. To our knowledge no such algorithm for the extraction of objects from a large variation of medical images exists as of yet. Another goal was to have only an extremely small number of images with too much information being removed from the images, so a rather conservative approach was taken. Some roughly related algorithms have already been used in the analysis of images and also videos to identify text regions in

---

<sup>1</sup> <http://www.irma-project.org/>

<sup>2</sup> <http://www.sim.hcuge.ch/medgift/>

<sup>3</sup> <http://www.gnu.org/software/gift/>

<sup>4</sup> <http://www.casimage.com/>

the visual data [22, 23] but we do not only have to deal with text but with a large variety of structures that need to be removed.

## 2 Insight Toolkit

*ITK*<sup>5</sup> is an open-source (OS) software system for medical image segmentation and registration. It has been developed since 1999 on initiative of the US National Library of Medicine (NLM) of the National Institutes of Health (NIH). As an OS project, *ITK* is used, debugged, maintained, and upgraded by several developers around the world. It is composed of a large collection of functions and algorithms designed specifically for medical images, and particularly for registration and segmentation tasks. The entire library is implemented in C++ and can be used on most platforms such as Linux, Mac OS and Windows. The decision to use *ITK* was taken because of the quantity of simple manipulation tools and filters it offers and the amount of medical segmentation research done based on it [24]. This allows us to concentrate on integrating tools rather than reprogramming and reinventing them. Many medical images are stored in DICOM [25], a complex standard, and *ITK* offers to open these images as well as other common standards such as JPG and GIF, which constitute much of our medical teaching file. *ITK* is the standard open source environment for medical image processing at the moment.

## 3 Methods Applied

### 3.1 Functions Used for Background Removal

The algorithm employed assumes that the object to extract has gray levels highly different from the background. Basically, an edge detection method is used to find these fast transitions. Several other steps are needed to handle a maximum of image types and remove some very specific problems that we identified. Much of the fine tuning was performed based on results on a small subset of images. Steps for the object extraction are:

- Removal of specific structures (University logo, typical large structures such as a grey square at the bottom right of images);
- Smoothing;
- Edge detection;
- Removal of small structures;
- Cropping;

The *casimage* collection that we use [26] presents two main image parts that interfere with the planned method. Several images contain the logo of the University hospitals in the upper left corner. Another problem is caused by a gray level square in the lower right corner (Figure 3). These two structures are too

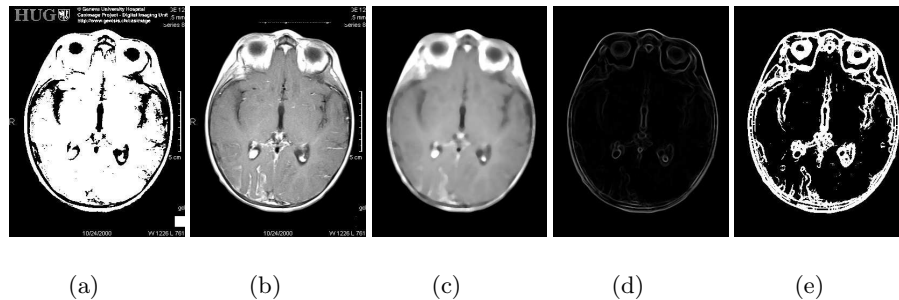
---

<sup>5</sup> <http://www.itk.org/>



**Fig. 1.** Logo of the University Hospitals of Geneva

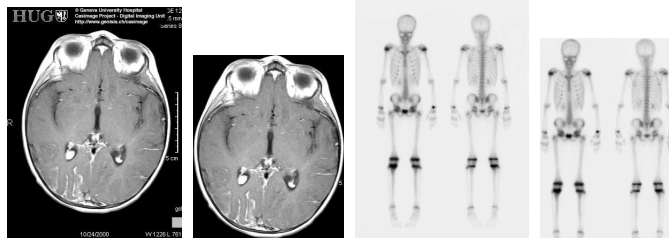
large to be removed during the foreseen noise cleaning step that well removes text. For this reason, the logo and the square have to be removed first. The hospital logo can be seen in Figure 1. It always appears in the upper left area in roughly the same size, so we cut 90 pixels horizontally and 30 pixels vertically for further analysis containing mainly the large characters of the logo. Part of the remaining text next to the logo is removed automatically in the following steps. To detect whether there is a logo in the upper left corner two criteria were defined. First, the number of white pixels in this area has to be in a certain range (350–400 pixels) and an erosion with a structuring element of size 1 has to eliminate all white pixels. The second criterion is based on a specific aspect of the logo: it is composed of fine horizontal lines. Thus, the erosion with a one pixel radius structuring element is highly destructive. If the two criteria are positive, the logo is removed by filling the region in black.



**Fig. 2.** The steps of the removal process: (a) thresholding for logo and grey square removal, (b) logo and gray square removed, (c) median filtering for smoothing and removal of small structures, (d) edge detection (e) thresholding.

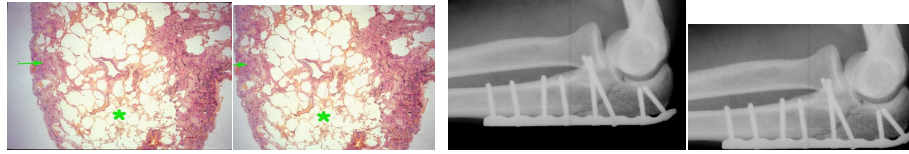
Gray squares (see Figure 3, first image, bottom right) also appear in many images. To remove them, the lower right area is thresholded to select only very light pixels. The resulting binary image is eroded then dilated to eliminate small objects. If a square object is remaining, it is the gray square. This binary image is used as a mask to eliminate the square from the original image.

Then, a median filter of size 4 is applied to smooth the image and already remove many small structures such as part of the text on the black background. Examples for the results of the various treatment steps can be seen in Figure 2.



**Fig. 3.** Result of the algorithm on a CT scan and a scintigraphy image.

The edge detection filter has as a consequence a weaker response and only the main structures will remain. A `GradientMagnitudeImageFilter` edge detection filter from the itk package is used in this step to detect the structures in the image. The aim of this structure removal step is to remove structures not being part of the main object. This can be annotations (patient name, system parameters, ...) but also simple frames around the image or a ruler. A binary image is produced by thresholding the result of the edge detection (threshold 5, 5-255 are mapped to 255 the rest to 0). Remaining small objects are removed by measuring the size of connected components. The size for removing objects depends on the image size itself. We use a simple cutoff for images having more than  $1000 \times 1000$  pixels where we remove objects up to a size of 300 pixels and smaller images where we remove objects of a maximum size of 50 pixels. Some small structures can be part of the main object, so the image is dilated (filter size 5) and then eroded (filter size 4) before the removal (a closing operation). This leads to a merging of neighbouring structures, which keeps slightly fragmented structures together. Unfortunately this also leads to connecting some text parts where the characters are fairly close and large. The bounding box of the result is finally computed, and the output image is created with the part of the original image contained in the box. The parameter settings of the filters and for removing connected components were obtained by systematic testing and trials with several “harder” images. Figure 4 contains some more results of the object extraction process, where a small part of the images was removed. On a Pentium IV computer with 2.8 GHz and 1 GB of RAM, the entire extraction process takes slightly more than 1 second making it feasible for a collection of 9000 images on a simple desktop computer.



**Fig. 4.** Results on a colour image and a radiograph with small enhancements.

### 3.2 Encountered Problems

Due to the variety of image types and acquisition systems, our algorithm can not handle all specific problems. In particular, the text part in images can be too large or too dense to be considered as noise (see Figure 5). Another problem can occur in CT scans, when the patient support under the patient appears on the image and is considered as part of the main object (Figure 6).



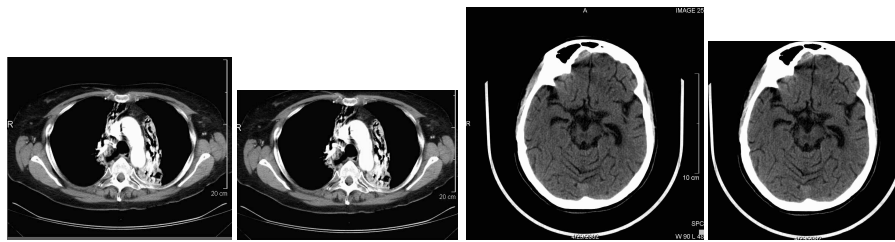
**Fig. 5.** Images where the text is not recognised and as a consequence not removed.

## 4 Results

### 4.1 Extraction Results

To evaluate the accuracy of the extraction algorithm, a subset of 500 medical images from the 9000-image collection was randomly extracted from the *casim-age* database and the algorithm was executed on these images. Then, each result was rated with respect to extraction quality. To simplify this task, a visual PHP interface was built that presents each extracted object next to its corresponding original image. It allows the validator to classify the result into one of these four quality classes:

1. Class 1: The object is extracted as wanted.



**Fig. 6.** In these CTs, the structure under the body is too important to be discarded.

2. Class 2: The image is fine but no work was needed.
3. Class 3: The result contains the object, but some background remains.
4. Class 4: Parts of the object are lost.

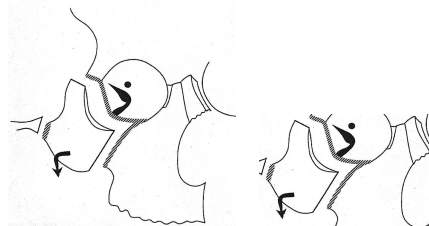
The 500 results were classified into these four categories by one validator familiar with the database. An optimal result was achieved for 389 images (204 in class 1 and 185 in class 2). For 105 images, parts of the text or background could not be removed, with most of the images having at least part of the background removed, so quality is at least better than before, although not perfectly satisfying. Six images have too much of the image being removed in error.

Table 2.1  
Classification topographique des principaux travaux et zones de recherche:

Classe	Indicateurs positifs	Indicateurs négatifs	Indicateurs positifs	Indicateurs négatifs
1	Précision de la détection	Précision de la détection	Précision de la détection	Précision de la détection
2	Précision de la détection	Précision de la détection	Précision de la détection	Précision de la détection
3	Précision de la détection	Précision de la détection	Précision de la détection	Précision de la détection
4	Précision de la détection	Précision de la détection	Précision de la détection	Précision de la détection

Table 2.2  
Classification topographique des principaux travaux et zones de recherche:

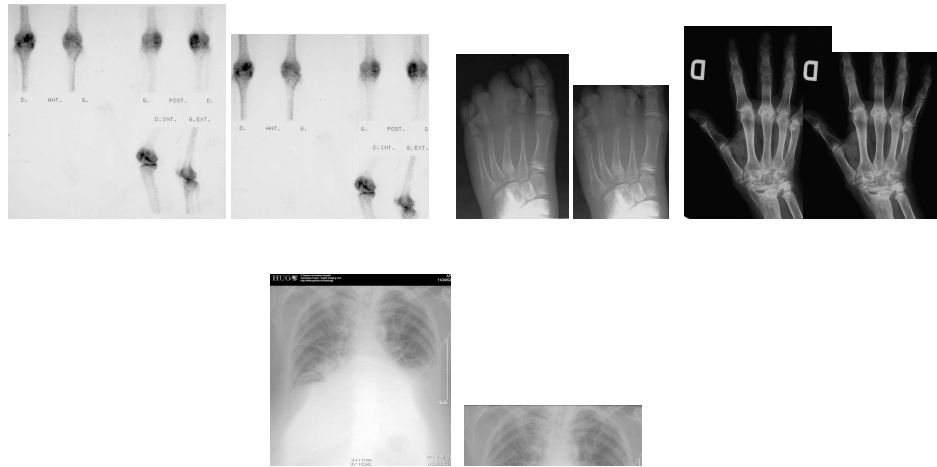
Classe	Indicateurs positifs	Indicateurs négatifs	Indicateurs positifs	Indicateurs négatifs
1	Précision de la détection	Précision de la détection	Précision de la détection	Précision de la détection
2	Précision de la détection	Précision de la détection	Précision de la détection	Précision de la détection
3	Précision de la détection	Précision de la détection	Précision de la détection	Précision de la détection
4	Précision de la détection	Précision de la détection	Précision de la détection	Précision de la détection



**Fig. 7.** Structures on these two images are too fine and taken for noise.

Of these, two are not medical images but drawings and text (Figure 7). Part of the fine structures was removed in these images. The quantity of lost information is negligible for three of the other four images (Figure 8). Only one image has a serious loss due to the object extraction but even this image contains all diagnostically relevant image information. It can clearly be seen that the algorithm has some problems with very slow changes in the contrast of the main object is not marked enough for edge detection.

For our goal of CBIR, it is important to eliminate or reduce the amount of useless information but not lose any important parts of the images for retrieval.



**Fig. 8.** The four other images where too much was removed.

The results have to be evaluated in terms of improvement and deterioration. Classes 1 and 3 constitute an improvement of the image for content-based image retrieval (about 60% of the images), and class 4 which contains around 1% of the images is a deterioration of the images. For about 39% of the images, the object extraction was not necessary.

An analysis of the number of removed pixels per image on the entire 9000-image dataset shows that 3000 images have more than 10% of the pixels removed, 1600 images more than 20% and 500 images even more than 50% of the pixels. This shows the large amount of data that could be removed in a simple object extraction algorithm.

#### 4.2 A simple improvement

To improve the results of our system on the images of class 3, we ran our algorithm a second time on all the already segmented images. We regarded this step as a necessary trial because some images contained a manually introduced border plus the actual image background. The majority of the images stayed unchanged but results were much better for 11 images among the 500 images observed. On the images of the class 1 and 2, no change appeared, but 2 images of class 4 present worse results. These two images are the text fragment and the thoracic radiograph shown beforehand. An analysis of the second step of segmentation shows that in 90% of the images less than 10% of the pixels were removed in this second step. Still, it also shows that more parts of the background could be removed by simply applying the same algorithm twice.



### 4.3 Retrieval results

The entire *casimage* database in its original form as well as after a simple segmentation and after running the segmentation twice were indexed using *medGIFT*. *medGIFT* first scales the images to  $256 \times 256$  pixels and then indexes them with the following feature groups:

- a global colour and gray level histogram;
- local gray level information by partitioning the image successively four times into four subregions and taking the mode colour of each region as feature;
- a global histogram of Gabor filter responses (4 directions, 3 scales and 10 strengths for quantisation);
- local Gabor filter responses within the entire images in blocks of size 16.

To compare the features, a frequency-based weighting similar to the text retrieval tf/idf weighting is used (see [21] for more details):

$$feature\ weight_j = \frac{1}{N} \sum_{i=1}^N (tf_{ij} \cdot R_i) \cdot \log^2 \left( \frac{1}{cf_j} \right), \quad (1)$$

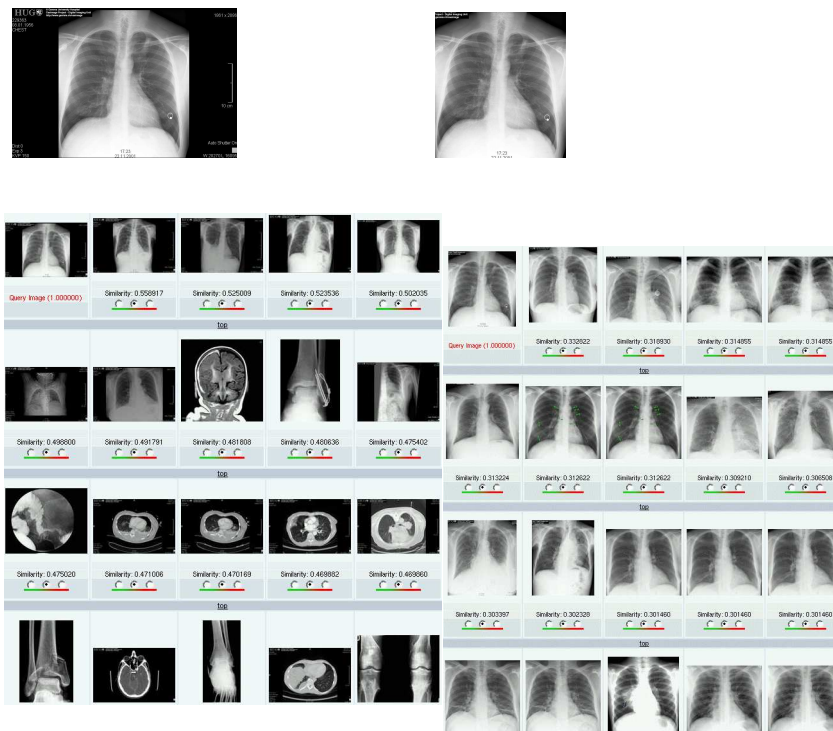
where  $tf$  is the *term frequency* of a feature,  $cf$  the *collection frequency*,  $j$  a *feature number*,  $q$  corresponds to a query with  $i = 1..N$  input images, and  $R_i$  is the *relevance* of an input image  $i$  within the range  $[-1; 1]$ .

Subjective impressions when using the system show an important improvement in retrieval quality. A few queries deliver worse results in a first step, but much better results once feedback is applied. Figure 9 shows a retrieval result without the use of feedback using a single image as query, and Figure 10 shows a result after one step of feedback for the same image once on the segmented and once on the non-segmented database.

Besides the subjective evaluation of the retrieval results, we also used the query topics and relevance judgements that were created in the *ImageCLEF*<sup>6</sup> competition [27, 28] for retrieval and compared them with the *medGIFT* base system. This competition created 26 query topics that contain one example image per topic, only, and no text. The lead measure for the competition is the mean average precision (MAP) that is used in most text retrieval benchmarking events such as TREC (Text REtrieval Conference) and CLEF (Cross Language Evaluation Forum). This measure is averaged over all 26 query topics. The *medGIFT* system has a MAP of 0.3757 and was among the three best visual systems in the competition. The results for the segmented database are slightly surprising as they do not appear to be better but rather slightly worse (MAP 0.3562). We also tried out two more configurations of grey levels and Gabor filters. Using 8 grey levels instead of 4 and 8 directions for the Gabor filters leads to even worth results (MAP 0.3350). When using 8 grey levels but the same Gabor filters, the results are even slightly worth than the latter (MAP 0.3322). The second run of the segmentation lead to very similar results (MAP 0.3515). This surprising

---

<sup>6</sup> <http://ir.shef.ac.uk/imageclef/>



**Fig. 9.** Comparison of retrieval results without user feedback.

result can partly be explained with several effects that are due to the way the groundtruth is being produced after the CLEF submissions. As only part of the database (a ground truth pool) is controlled for the ground truthing, systems with very differing techniques have a disadvantage [29], even more so if not included into the pool before the ground truthing [30], which was the case as this technique did not participate at the actual competition. We discovered that some of the images found to be relevant with the algorithm do not appear in the relevance set as in the competition no other system retrieved these images at a high enough position to be included into the pool.

Another problem is the loss of shape information when cutting off directly next to the object. Leaving a few background pixels around the object might improve retrieval as artifacts of the Gabor filters are reduced and more shape information is available in this case.

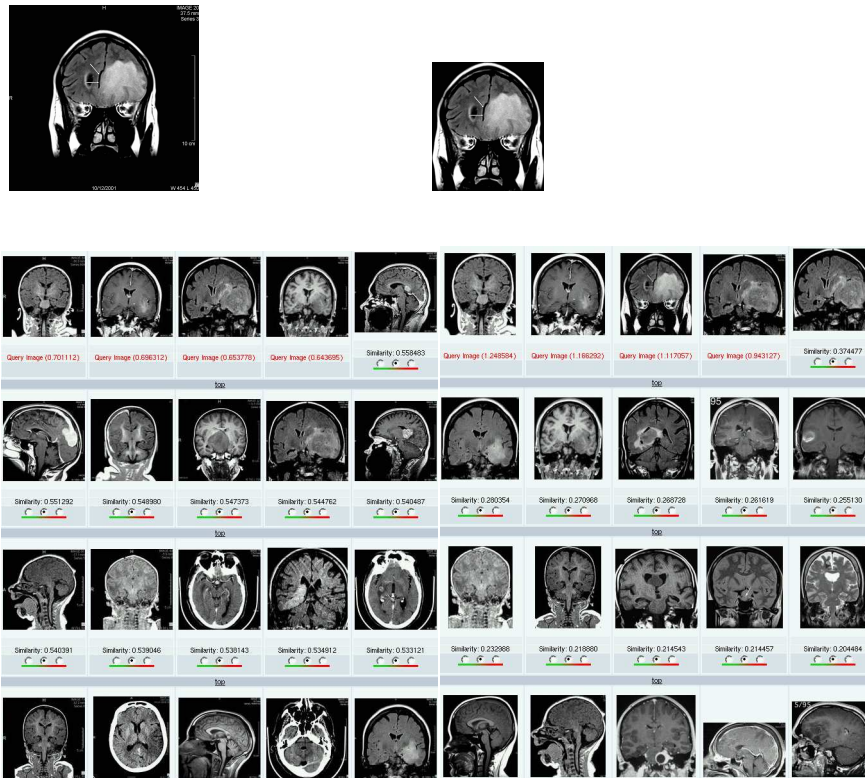


Fig. 10. Comparison of retrieval results with one step of user feedback.

## 5 Conclusions

In image retrieval systems for specialised medical domains, image segmentation can focus the search very precisely. Retrieval in broad, PACS-like databases needs different algorithms to extract the most important parts of the image for indexing and retrieval. We present such an algorithm that works completely automatic and quickly, and as a consequence can be applied to very large databases such as teaching files or even entire PACS systems. Some of the recognised problems might be particular for our setting but they will appear in a similar fashion in other teaching files. The particular problems will need to be detected for any other collection.

Our solution is optimised to have very few images where too much is cut off as this could prevent images from being retrieved. This fact leads to a larger number of images where part of the background remains. We need to work on removing these missed parts as well while keeping the number of images with

too much being removed low. One idea is to not only use the properties of text for removal but to really recognise text boxes entirely and remove them from the images. Maybe, together with OCR (optical character recognition) it might be possible to even use the obtained textual data to improve retrieval quality.

We also plan to participate in the 2005 *ImageCLEF* competition so our system is taken into account for the ground truthing and results become better comparable with the other techniques used. We also need to find out whether we cut off too much around the objects for retrieval and we should rather leave a few pixels around the objects so our shape detectors work better on the object form. Data reduction for general medical image retrieval is necessary and our algorithm is one step towards a more intelligent indexing of general medical image databases removing part of the noise surrounding the objects in the images.

## 6 Acknowledgements

Part of this research was supported by the Swiss National Science Foundation with grant 632-066041.

## References

1. Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** No 12 (2000) 1349–1380
2. Rui, Y., Huang, T.S., Chang, S.F.: Image retrieval: Past, present and future. In Liao, M., ed.: *Proceedings of the International Symposium on Multimedia Information Processing*, Taipei, Taiwan (1997)
3. Carson, C., Thomas, M., Belongie, S., Hellerstein, J.M., Malik, J.: Blobworld: A system for region-based image indexing and retrieval. In Huijismans, D.P., Smeulders, A.W.M., eds.: *Third International Conference on Visual Information Systems (VISUAL'99)*. Number 1614 in *Lecture Notes in Computer Science*, Amsterdam, The Netherlands, Springer (1999) 509–516
4. Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, P.: Query by Image and Video Content: The QBIC system. *IEEE Computer* **28** (1995) 23–32
5. Bach, J.R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R., Shu, C.F.: The Virage image search engine: An open framework for image management. In Sethi, I.K., Jain, R.C., eds.: *Storage & Retrieval for Image and Video Databases IV*. Volume 2670 of *IS&T/SPIE Proceedings*, San Jose, CA, USA (1996) 76–87
6. Graham, M.E., Eakins, J.P.: Artisan: A prototype retrieval system for trademark images. *Vine* **107** (1998) 73–80
7. Müller, H., Rosset, A., Vallée, J.P., Geissbuhler, A.: Integrating content-based visual access methods into a medical case database. In: *Proceedings of the Medical Informatics Europe Conference (MIE 2003)*, St. Malo, France (2003)
8. Tagare, H.D., Jaffe, C., Duncan, J.: Medical image databases: A content-based retrieval approach. *Journal of the American Medical Informatics Association* **4** (1997) 184–198

9. Lowe, H.J., Antipov, I., Hersh, W., Arnott Smith, C.: Towards knowledge-based retrieval of medical images. The role of semantic indexing, image content representation and knowledge-based retrieval. In: Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA), Nashville, TN, USA (1998) 882–886
10. Tang, L.H.Y., Hanka, R., Ip, H.H.S.: A review of intelligent content-based indexing and browsing of medical images. *Health Informatics Journal* **5** (1999) 40–49
11. Müller, H., Michoux, N., Bandon, D., Geissbuhler, A.: A review of content-based image retrieval systems in medicine – clinical benefits and future directions. *International Journal of Medical Informatics* **73** (2004) 1–23
12. Le Bozec, C., Zapletal, E., Jaulent, M.C., Heudes, D., Degoulet, P.: Towards content-based image retrieval in HIS-integrated PACS. In: Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA), Los Angeles, CA, USA (2000) 477–481
13. Westerveld, T.: Image retrieval: Content versus context. In: Recherche d'Informations Assistée par Ordinateur (RIAO'2000) Computer-Assisted Information Retrieval. Volume 1., Paris, France (2000) 276–284
14. Orphanoudakis, S.C., Chronaki, C.E., Vamvaka, D.: *I<sup>2</sup>Cnet*: Content-based similarity search in geographically distributed repositories of medical images. *Computerized Medical Imaging and Graphics* **20** (1996) 193–207
15. Liu, C.T., Tai, P.L., Chen, A.Y.J., Peng, C.H., Lee, T., Wang, J.S.: A content-based CT lung retrieval system for assisting differential diagnosis images collection. In: Proceedings of the second International Conference on Multimedia and Exposition (ICME'2001), Tokyo, Japan, IEEE Computer Society, IEEE Computer Society (2001) 241–244
16. Aisen, A.M., Broderick, L.S., Winer-Muram, H., Brodley, C.E., Kak, A.C., Pavlopoulou, C., Dy, J., Shyu, C.R., Marchiori, A.: Automated storage and retrieval of thin-section CT images to assist diagnosis: System description and preliminary assessment. *Radiology* **228** (2003) 265–270
17. Shyu, C.R., Brodley, C.E., Kak, A.C., Kosaka, A., Aisen, A.M., Broderick, L.S.: ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases. *Computer Vision and Image Understanding* (special issue on content-based access for image and video libraries) **75** (1999) 111–132
18. Lehmann, T., Güld, M.O., Thies, C., Spitzer, K., Keysers, D., Ney, H., Kohnen, M., Schubert, H., Wein, B.B.: Content-based image retrieval in medical applications. *Methods of Information in Medicine* **43** (2004) 354–361
19. Güld, M.O., Kohnen, M., Keysers, D., Schubert, H., Wein, B.B., Bredno, J., Lehmann, T.M.: Quality of DICOM header information for image categorization. In: International Symposium on Medical Imaging. Volume 4685 of SPIE Proceedings., San Diego, CA, USA (2002) 280–287
20. Lehmann, T.M., Schubert, H., Keysers, D., Kohnen, M., Wein, B.B.: The IRMA code for unique classification of medical images. In: Medical Imaging. Volume 5033 of SPIE Proceedings., San Diego, California, USA (2003)
21. Squire, D.M., Müller, W., Müller, H., Pun, T.: Content-based query of image databases: inspirations from text retrieval. *Pattern Recognition Letters* (Selected Papers from The 11th Scandinavian Conference on Image Analysis SCIA '99) **21** (2000) 1193–1198 B.K. Ersboll, P. Johansen, Eds.
22. Chen, D., Odobez, J.M., Thiran, J.P.: A localization/verification scheme for finding text in images and video frames based on contrasts independent features and machine learning methods. *Signal Processing: Image Communication* **19** (2004) 205–217

23. Agnithotri, L., Dimitrova, N.: Text detection for video analysis. In: IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL'99), Fort Collins, Colorado, USA (1999) 109–113
24. Cates, J., Lefohn, A., Whitaker, R.: GIST: An interactive GPU-based level-set segmentation tool for 3d medical images. *Journal on Medical Image Analysis* **8** (2004 – to appear)
25. Revet, B.: DICOM Cook Book for Implementations in Modalities. Philips Medical Systems, Eindhoven, Netherlands (1997)
26. Rosset, A., Müller, H., Martins, M., Dfouni, N., Vallée, J.P., Ratib, O.: Casimage project – a digital teaching files authoring environment. *Journal of Thoracic Imaging* **19** (2004) 1–6
27. Clough, P., Sanderson, M., Müller, H.: A proposal for the CLEF cross language image retrieval track (imageCLEF) 2004. In: *The Challenge of Image and Video Retrieval (CIVR 2004)*, Dublin, Ireland, Springer LNCS (2004)
28. Müller, H., Rosset, A., Geissbuhler, A., Terrier, F.: A reference data set for the evaluation of medical image retrieval systems. *Computerized Medical Imaging and Graphics* (2004 (to appear))
29. Zobel, J.: How reliable are the results of large-scale information retrieval experiments? In Croft, W.B., Moffat, A., van Rijsbergen, C.J., Wilkinson, R., Zobel, J., eds.: *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Melbourne, Australia, ACM Press, New York (1998) 307–314
30. Sparck Jones, K., van Rijsbergen, C.: Report on the need for and provision of an ideal information retrieval test collection. *British Library Research and Development Report 5266*, Computer Laboratory, University of Cambridge (1975)