International Journal of Medical Informatics (2004) **73**, 1—23

International Journal of Medical Informatics

www.intl.elsevierhealth.com/journals/ijmi

REVIEW

A review of content-based image retrieval systems in medical applications—clinical benefits and future directions

Henning Müller*, **Nicolas Michoux**, **David Bandon**, **Antoine Geissbuhler**

Service of Medical Informatics, University Hospital of Geneva, Rue Micheli-du-Crest 24, 1211 Geneva 14, Switzerland

Received 22 July 2003; accepted 13 November 2003

KEYWORDS
Medical image retrieval;
Content-based search;
Visual information retrieval;
PACS

Summary Content-based visual information retrieval (CBVIR) or content-based image retrieval (CBIR) has been one on the most vivid research areas in the field of computer vision over the last 10 years. The availability of large and steadily growing amounts of visual and multimedia data, and the development of the Internet underline the need to create thematic access methods that offer more than simple text-based queries or requests based on matching exact database fields. Many programs and tools have been developed to formulate and execute queries based on the visual or audio content and to help browsing large multimedia repositories. Still, no general breakthrough has been achieved with respect to large varied databases with documents of differing sorts and with varying characteristics. Answers to many questions with respect to speed, semantic descriptors or objective image interpretations are still unanswered.

In the medical field, images, and especially digital images, are produced in ever-increasing quantities and used for diagnostics and therapy. The Radiology Department of the University Hospital of Geneva alone produced more than 12,000 images a day in 2002. The cardiology is currently the second largest producer of digital images, especially with videos of cardiac catheterization ($\sim$1800 exams per year containing almost 2000 images each). The total amount of cardiologic image data produced in the Geneva University Hospital was around 1 TB in 2002. Endoscopic videos can equally produce enormous amounts of data.

With digital imaging and communications in medicine (DICOM), a standard for image communication has been set and patient information can be stored with the actual image(s), although still a few problems prevail with respect to the standardization. In several articles, content-based access to medical images for supporting clinical decision-making has been proposed that would ease the management of clinical data and scenarios for the integration of content-based access methods into picture archiving and communication systems (PACS) have been created.

This article gives an overview of available literature in the field of content-based access to medical image data and on the technologies used in the field. Section 1 gives an introduction into generic content-based image retrieval and the technologies used. Section 2 explains the propositions for the use of image retrieval in medical practice and the various approaches. Example systems and application areas are

*Corresponding author. Tel.: +41-22-372-6175; fax: +41-22-372-8680.
E-mail address: henning.mueller@sim.hcuge.ch (H. Müller).

1386-5056/$ — see front matter
doi:10.1016/j.ijmedinf.2003.11.024

described. Section 3 describes the techniques used in the implemented systems, their datasets and evaluations. Section 4 identifies possible clinical benefits of image retrieval systems in clinical practice as well as in research and education. New research directions are being defined that can prove to be useful.
This article also identifies explanations to some of the outlined problems in the field as it looks like many propositions for systems are made from the medical domain and research prototypes are developed in computer science departments using medical datasets. Still, there are very few systems that seem to be used in clinical practice. It needs to be stated as well that the goal is not, in general, to replace text-based retrieval methods as they exist at the moment but to complement them with visual search tools.

1. Introduction to image retrieval

This section gives an introduction to content-based image retrieval systems (CBIRSs) and the technologies used in them. Image retrieval has been an extremely active research area over the last 10 years, but first review articles on access methods in image databases appeared already in the early 1980s [1]. The following review articles from various years explain the state-of-the-art of the corresponding years and contain references to a large number of systems and descriptions of the technologies implemented. Enser [2] gives an extensive description of image archives, various indexing methods and common searching tasks, using mostly text-based searches on annotated images. In [3], an overview of the research domain in 1997 is given and in [4], the past, present and future of image retrieval is highlighted. In [5] an almost exhaustive overview of published systems is given and an evaluation of a subset of the systems is attempted [6]. Unfortunately, the evaluation is very limited and only for very few systems. The most complete overview of technologies to date is given by Smeulders et al. [7]. This article describes common problems such as the semantic gap or the sensory gap and gives links to a large number of articles describing the various techniques used in the domain. For an even deeper introduction into the domain, several theses and books are available [8—11].

The only article reviewing several medical retrieval systems so far, is to our knowledge [12]. It explains using one paragraph per topic a number of medical image retrieval systems. No systematic comparison of the techniques employed and the data/evaluation used has been attempted.

This review paper in contrast is the first review that concentrates on image retrieval in the medical domain and that does a systematic overview of techniques used, visual features employed, images indexed and medical departments involved. It also offers future perspectives for image retrieval in the medical domain and will be a good starting point for research projects on medical image retrieval as useful techniques for certain sorts of images can be isolated and past errors can be avoided.

1.1. Content-based image retrieval systems

Although early systems existed already in the beginning of the 1980s [13], the majority would recall systems such as IBM's Query By Image Content[1] (QBIC) as the start of content-based image retrieval [14,15]. The commercial QBIC system is definitely the most well-known system. Another commercial system for image [16] and video [17] retrieval is Virage[2] that has well known commercial customers such as CNN.

Most of the available systems are, however from academia. It would be hard to name or compare them all but some well-known examples include Candid [18], Photobook[3] [19] and Netra [20] that all use simple color and texture characteristics to describe the image content. Using higher level information, such as segmented parts of the image for queries, was introduced by the Blobworld[4] system [21,22]. PicHunter [23] on the other hand is an image browser that helps the user to find an exact image in the database by showing to the user images on screen that maximize the information gain in each feedback step. A system that is available free of charge is the GNU Image Finding Tool (GIFT[5]) [24,25]. Some systems are available as demonstration versions on the web such as Viper,[6] WIPE[7] or Compass.[8]

[1] http://wwwqbic.almaden.ibm.com/.
[2] http://www.virage.com/.
[3] http://www-white.media.mit.edu/vismod/demos/facerec/basic.html.
[4] http://elib.cs.berkeley.edu/photos/blobworld/.
[5] http://www.gnu.org/software/gift/.
[6] http://viper.unige.ch/demo/php/demo.php.
[7] http://wang.ist.psu.edu/IMAGE/.
[8] http://compass.itc.it/demos.html.

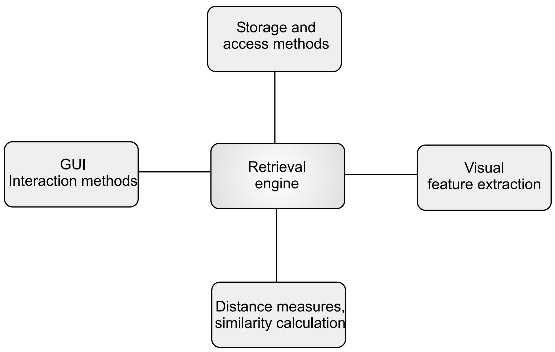

Fig. 1 The principal components of all content-based image retrieval systems.

Most of these systems have a very similar architecture for browsing and archiving/indexing images comprising tools for the extraction of visual features, for the storage and efficient retrieval of these features, for distance measurements or similarity calculation and a type of graphical user interface (GUI). This general system setup is shown in Fig. 1. All shown components are described in more detail further on.

1.2. Visual features used

Visual features were classified in [5] into *primitive* features such as color or shape, *logical* features such as identity of objects shown and *abstract* features such as significance of scenes depicted. Still, all currently available systems only use primitive features unless manual annotation is coupled with the visual features as in [26]. Even systems using segments and local features such as Blobworld [21,22] are still far away from identifying objects reliably. No system offers interpretation of images or even medium level concepts as they can easily be captured with text. This loss of information from an image to a representation by features is called the *semantic gap* [7]. The situation is surely not satisfactory and the semantic gap definitely accounts for part of the rejection to use image retrieval applications, but the technology can still be valuable when advantages and problems are understood by the users. The more a retrieval application is specialized for a certain, limited domain, the smaller the gap can be made by using domain knowledge. Another gap is the *sensory* gap that describes the loss between the actual structure and the representation in a (digital) image.

1.2.1. Color

In stock photography (large, varied databases for being used by artists, advertisers and journalists), color has been the most effective feature and almost all systems employ colors. Although most of the images are in the red, green, blue (RGB) color space, this space is only rarely used for indexing and querying as it does not correspond well to the human color perception. It only seems reasonable to be used for images taken under exactly the same conditions each time such as trademark images. Other spaces such as hue, saturation, value (HSV) [24,27,28] or the CIE Lab [15] and Luv [29] spaces are much better with respect to human perception and are more frequently used. This means that differences in the color space are similar to the differences between colors that humans perceive. Much effort has also been spent on creating color spaces that are optimal with respect to lighting conditions or that are invariant to shades and other influences such as viewing position [30,31]. This allows to identify colors even under varying conditions but on the other hand information about the absolute colors is lost. In specialized fields, namely in the medical domain, absolute color or grey level features are often of very limited expressive power unless exact reference points exist as it is the case for computed tomography images.

1.2.2. Texture

Partly due to the imprecise understanding and definition of what exactly visual texture actually is, texture measures have an even larger variety than color measures. Some of the most common measures for capturing the texture of images are wavelets [32,33] and Gabor filters [24,34,35] where the Gabor filters do seem to perform better and correspond well to the properties of the human visual cortex for edge detection [36,37]. These texture measures try to capture the characteristics of the image or image parts with respect to changes in certain directions and the scale of the changes. This is most useful for regions or images with homogeneous texture. Again, invariances with respect to rotations of the image, shifts or scale changes can be included into the feature space but information on the texture can get lost in this process [38].

Other popular texture descriptors contain features derived from co-occurrence matrices [39—41], features based on the factors of the Fourier transform [38] and the so-called Wold features [42].

1.2.3. Local and global features

Both, color and texture features can be used on a global image level or on a local level on parts of the image. The easiest way to use regional features is to use blocks of fixed size and location, so-called *partitioning* of the image [7,24] for local feature extraction. These blocks do not take into account any semantics of the image itself. When allowing the user to choose image regions (regions of interest (ROI)) [43], to delineate objects in the image [44] or when segmenting the image into areas with similar properties [45], the locally extracted features contain more information about the image objects or underlying structures.

1.2.4. Segmentation and shape features

Fully automated segmentation of images into objects itself is an unsolved problem. Even in fairly specialized domains, fully automated segmentation causes many problems and is often not easy to realize. In image retrieval, several systems attempt to perform an automatic segmentation of the images in the collection for feature extraction [21,46]. To have an effective segmentation of images using varied image databases the segmentation process has to be done based on the color and texture properties of the image regions [45].

Much has also been written on medical image segmentation with respect to browsing image repositories [47,48]. After segmentation, the resulting segments can be described by shape features that commonly exist, including those with invariances with respect to shifts, rotations and scaling [49,50].

1.2.5. Semantics?

All these visual features, and even features derived from segmented regions, are still fairly low-level compared to high level concepts that are contained in the images. They do not necessarily correspond to objects in the images or the semantic concepts or structures that a user is interested in. Several articles speak of semantic or cognitive image retrieval [51—54] but in the end this has not yet been realized with visual features alone. It often comes down to connecting visual low-level features with textual high level features which has already been proposed in [55] as early as 1996. The annotation of image collections for retrieval or for the combination with visual features for retrieval is another very active research area [26,56]. Many problems such as the subjectiveness of annotations need to be addressed even when working with restricted vocabularies. The users' annotations do not only vary between persons, they are also varying in time for the same person and they depend strongly on the users' actual search tasks. However, in the medical domain, good annotated atlases of medical images do exist that contain objective knowledge, for example based on the images of the visible human.[9] The definition of visual similarity or relevance with respect to visual similarity are also philosophical questions that have been discussed for a long time [57].

1.3. Comparison techniques used

Basically all systems use the assumption of equivalence of an image and its representation in feature space. These systems often use measurement systems such as the easily understandable Euclidean vector space model [15,58] for measuring distances between a query image (represented by its features) and possible results representing all images as feature vectors in an *n*-dimensional vector space. This is done, although metrics have been shown to not correspond well to human visual perception (Tversky [59]). Several other distance measures do exist for the vector space model such as the city-block distance, the Mahalanobis distance [15] or a simple histogram intersection [60]. Still, the use of high-dimensional feature spaces has shown to cause problems and great care needs to be taken with the choice of distance measurement to be chosen in order to retrieve meaningful results [61,62]. These problems with a similarity definition in high-dimensional feature spaces is also known as the *curse of dimensionality* and has also been discussed in the domain of medical imaging [63].

Another approach is a *probabilistic* framework to measure the probability that an image is relevant [64]. A relationship between probabilistic image retrieval and vector-space distance measures is given in [65]. This paper concludes that the vector space distance measurements described in the literature correspond, in principal, to probabilistic retrieval under certain assumptions of the feature distributions. Another probabilistic retrieval form is the use of support vector machines (SVMs) [66] for a classification of images into classes for relevant and non-relevant.

Various systems use methods that are well known from the text retrieval field and apply them to visual features where the visual features have to correspond roughly to words in text [24,67,68]. This is based on the two principles.

- A feature frequent in an image describes this image well.
- A feature frequent in the collection is a weak indicator to distinguish images from each other.

Several weighting schemes for text retrieval that have also been used in image retrieval are described in [69]. A formal definition of vector-space, probabilistic and boolean models for information retrieval is attempted in [70]. A general overview

[9] http://www.nlm.nih.gov/research/visible/visible/human.html.

of pattern recognition methods and various comparison techniques is given in a very good review article [187]. This article describes the feature extraction, selection, feature space reduction techniques that are equally important in the image retrieval domain.

1.4. Storage and access methods

Although most systems do not talk in detail about the underlying storage and access methods [23,52] this is extremely important for interactive systems to keep response times at bay. Common storage methods used are relational databases [15,71], inverted files [24], self-made structures or simply to keep the entire index in the main memory which will inevitably cause problems when using large databases.

These methods often need to use dimension reduction techniques or pruning methods [72] to allow for an efficient and quick access to the data. Some indexing techniques such as the KD-trees are described in [73]. Principal component analysis (PCA) for feature space reduction is used in [74]. This technique is also called Karhunen-Loeve transform (KLT) [75]. Another feature space reduction technique is the independent component analysis (ICA) described in [76,187]. [187] also explains a variety of other techniques such as for feature selection.

1.5. Other important techniques

There is a large number of other important techniques to improve the performance of retrieval systems. One of the most prominent techniques is *relevance feedback* that is well known from text retrieval [77]. This technique has proven to be important for image retrieval as well [78—80] because often unexpected or unwanted images show up in the result of a similarity query. The active selection of relevant and unrelevant images by the user represents an interactive method for controlling the pertinence of the results adequately. Often, the performance of a retrieval system with feedback is regarded as being even more important than without as only with feedback the users subjectivity can seriously be taken into account. An overview of interaction techniques in image retrieval is given in [81].

Other techniques from the artificial intelligence community are also used for image retrieval such as long-term learning from user behavior based on data mining in usage log files [82] using the well-known market basket analysis.

Some interesting and innovative *user interfaces* are described in [83,84]. This includes a three-dimensional representation of the similarity space as well as the *El Niño* system, where the user moves images together into clusters that (s)he thinks are similar.

The *correlation across various media* (text, image, video, audio) should also not be forgotten if these are available. Whenever additional information is available such as annotations of the images, it should be used for the retrieval.

2. Use of image retrieval in medical applications

The number of digitally produced medical images is rising strongly. In the radiology department of the University Hospital of Geneva (HUG) alone, the number of images produced per day in 2002 was 12,000, and it is still rising. Videos and images produced in cardiology are equally multiplying and endoscopic videos promise to be another very large data source that are planned to be integrated into the PACS. The management and the access to these large image repositories become increasingly complex. Most accesses to these systems are based on the patient identification or study characteristics (modality, study description) [85] as it is also defined in the DICOM standard [86].

Imaging systems and image archives have often been described as an important economic and clinical factor in the hospital environment [87—89]. Several methods from the computer vision and image processing fields have already been proposed for the use in medicine more than ten years ago [90,91]. Several radiological teaching files exist [92,93] and radiology reports have also been proposed in a multimedia form in [94]. Web-interfaces to medical image databases are described in [95].

Medical images have often been used for retrieval systems and the medical domain is often cited as one of the principal application domains for content-based access technologies [7,18,96—98] in terms of potential impact. Still, there has rarely been an evaluation of the performance and the description of the clinical use of systems is even rarer.

Two exceptions seem to be the Assert[10] system on the classification of high resolution CTs of the lung [40,99] and the IRMA[11] system for the classification of images into anatomical areas, modalities and view points [100].

Content-based retrieval has also been proposed several times from the medical community for the

[10] http://rvl2.ecn.purdue.edu/~cbirdev/www/CBIRmain.html.

[11] http://irma-project.org/.

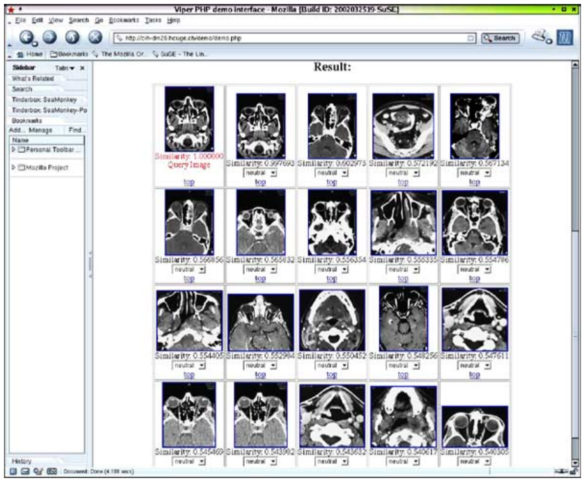

Fig. 2 A screenshot of a typical image retrieval system showing retrieved images similar to an example in a web browser interface.

inclusion into various applications [101—103], often without any implementation. Still, for a real medical application of content-based retrieval methods and the integration of these tools into medical practice a very close cooperation between the two fields is necessary for a longer period of time and not simply an exchange of data or a list of the necessary functionality.

An interface of a typical content-based retrieval system is shown in Fig. 2. The interface shows the images retrieved with their similarity score to an example image. The user can then mark images as relevant, non-relevant or leave them as neutral, change the parameters for retrieval and start a new query for refinement.

2.1. The need for content-based medical image retrieval

There are several reasons why there is a need for additional, alternative image retrieval methods apart from the steadily growing rate of image production. It is important to explain these needs and to discuss possible technical and methodological improvements and the resulting clinical benefits.

The goals of medical information systems have often been defined to deliver the needed information at the right time, the right place to the right persons in order to improve the quality and efficiency of care processes [104]. Such a goal will most likely need more than a query by patient name, series ID or study ID for images. For the clinical decision-making process it can be beneficial or even important to find other images of the same modality, the same anatomic region of the same disease. Although part of this information is normally contained in the DICOM headers and many imaging devices are DICOM-compliant at this time, there are still some problems. DICOM headers have proven to contain a fairly high rate of errors, for example for the field anatomical region, error rates of 16% have been reported [105]. This can hinder the correct retrieval of all wanted images.

Clinical decision support techniques such as case-based reasoning [106] or evidence-based medicine [107,108] can even produce a stronger need to retrieve images that can be valuable for supporting certain diagnoses. It could even be imagined to have image-based reasoning (IBR) as a new discipline for diagnostic aid. Decision support systems in radiology [109] and computer-aided diagnostics for radiological practice as demonstrated at the Radiological Society of North America (RSNA) [110] are on the rise and create a need for powerful data and meta-data management and retrieval.

The general clinical benefit of imaging systems has also already been demonstrated in [111]. In [112] an initiative is described to identify important tasks for medical imaging based on their possible clinical benefits. It needs to be stated that the purely visual image queries as they are executed in the computer vision domain will most likely not be able to ever replace text-based methods as there will always be queries for all images of a certain patient, but they have the potential to be a very good complement to text-based search based on their characteristics. Still, the problems and advantages of the technology have to be stressed to obtain acceptance and use of visual and text-based access methods up to their full potential. A scenario for hybrid, textual and visual queries is proposed in the CBIR2 system [113].

Besides diagnostics, teaching and research especially are expected to improve through the use of visual access methods as visually interesting images can be chosen and can actually be found in the existing large repositories. The inclusion of visual features into medical studies is another interesting point for several medical research domains. Visual features do not only allow the retrieval of cases with patients having similar diagnoses but also cases with visual similarity but different diagnoses. In teaching, it can help lecturers as well as students to browse educational image repositories and visually inspect the results found. This can be the case for navigating in image atlases.[12] It can also be used to cross-correlate visual and textual features of the images.

[12] http://www.loni.ucla.edu/MAP/index.html.

2.2. The use in PACS and other medical databases

There is a large number of propositions for the use of content-based image retrieval methods in the medical domain in general [101—103]. Other articles describe the use of image retrieval with an image management framework [114—119], sometimes without stating what has actually been implemented and what is still in the status of ideas. Also the integration into PACS systems [85,120—123] or other medical image databases [92,124—126] has been proposed often, but implementation details are generally rare.

Most of the general articles such as [101] state that the medical domain is very specialized so that general systems cannot be used. This is true but it is the case for all specialized domains such as trademark retrieval or face recognition, and specialized solutions need to be found. The more specialized the features of a system are the smaller the range of application and compromises for each specific application area needs to be found. Domain knowledge needs to be integrated into specialized query engines.

Another proposition of what is needed for an efficient use in the medical domain is given in [102], including some implementation details. Clinically relevant indexing and selective retrieval of biomedical images is explained in [103]. Some examples are given but no implementation details. It is proposed to change the DICOM headers which is in principal not allowed according to the standard for the storage of DICOM images, but would, however, be allowed in DICOM structured reporting. Most of these articles ask for semantic retrieval based on images that are segmented automatically into objects and where diagnoses can be derived easily from the objects' visual features. This is still a dream, as it has been in the computer vision domain for general segmentation methods for a while. Steps into the direction of solutions have to be taken using machine learning techniques and by including specific domain knowledge. Implementations of image retrieval systems are a step-by-step process and first systems will definitely not meet all the high requirements that are asked for.

Several frameworks for distributed image management solutions have been developed such as I^2Cnet [98,115]. When reading articles on these frameworks it is often not clear what had and had not been implemented. Image retrieval based on visual features is often proposed but unfortunately nothing is said about the visual features used or the performance obtained. [117] describes a telemedicine and image management framework

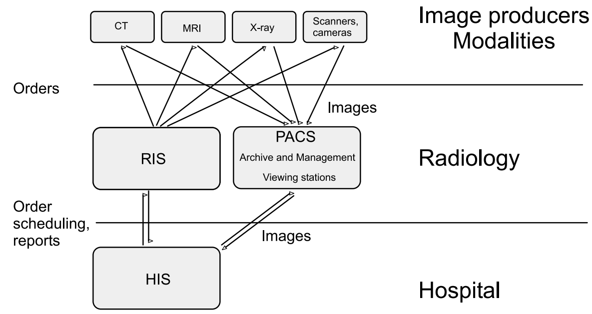

Fig. 3 The basic position of a PACS within the information system environment in a hospital.

and [114] is another very early article on the architecture of a distributed multimedia database. [127] describes an active index for medical image data management, and in [116] a newer image management environment is described. In [118,119], two frameworks for image management and retrieval are described focusing on technical aspects and stating application areas. One of the few frameworks with at least a partial implementation is the image retrieval in medical applications (IRMA) framework [100,128] that allows for a relatively robust classification of incoming images into anatomical regions, modality and the taken orientation. This project also developed a classification code for medical images based on four axes (modality, body orientations, body region, biological system) to uniquely classify medical images and allow to test and measure the performance of classification [129].

The use of content-based techniques has been proposed several times in a PACS environment. PACS are the main software components to store and access the large amount of visual data used in medical departments. Often, several-layer architectures exist for quick short-term access and slow long-term storage. More information on PACS can be found in [130]. A web-based PACS architecture is proposed in [131]. The general schema of a PACS system within the hospital is shown in Fig. 3. The Integrating the Healthcare Enterprise (IHE)[13] standard is aiming at data integration in healthcare including all the systems described in Fig. 3.

An indexing of the entire PACS causes problems with respect to the sheer amount of data that needs to be processed to efficiently allow access by content to all the images. This issue of the amount of data that needs to be indexed is not discussed in any of the articles. [122] proposes to use content-based image retrieval techniques in a PACS system as a search method but no implementation

[13] http://www.rsna.org/IHE/index.shtml.

details are given. In [120] an integration into the PACS is described that uses the text attached to the images as content. More on this IDEM project can be found at[14] [123] proposes an extension to the database management system for integrating content-based queries based on simple visual features into PACS systems. A classification of systems is given in [121] proposing an integration into the PACS, but no implementation details are stated in the text. A coupling of a PACS and an image classification system is given in [85]. Here, it is possible to search for certain anatomic regions, modalities or views of an image. A simple interface for coupling the PACS and the image retrieval system is stated as well. The identification is based on the DICOM unique identifier (UIDs) of the images. Still, there is lack of publications describing the integration of image retrieval into the workflow in a medical institution and visual knowledge management in a learning institution has not been the subject of publications either. Besides the use directly within a PACS system or very general image database environment, content-based image retrieval has also been used or proposed in a couple of specialized collections. In [92], CBIR is proposed in the context of a case database containing images and attached case descriptions. [124] describes the use in a medical reference database and [132] the use within a teaching file assistant. An object-oriented approach to store and access medical databases is given in [126]. But it remains unclear what kind of visual features are supposed to be used. In [133] an online pathology atlas uses the search-by-similarity paradigm.

Decision support systems are another application of content-based medical image retrieval [134]. In [135] access-control models for content-based retrieval are discussed. It can be seen that the number and sort of applications is large and diverse, and the techniques used or proposed for an implementation contain a variety almost as large as for general image retrieval.

2.3. The use in various medical departments

The same variety that exists with respect to proposed applications exists also with respect to the medical departments where the use of content-based access methods has been implemented or proposed. Obviously, most applications are centered around images produced in radiology departments, but there are also several other departments where CBIRSs have been implemented.

[14] http://www.hbroussais.fr/Broussais/InforMed/IDEM/InterrogerBase.html.

A categorization of images from various departments has been described in [54,100]. A classification of *dermatologic* images is explained in [75,136,137]. *Cytological* specimens have already been described very early (in 1986, [138]) and also later on [139] whereas the search for 3D cellular structures followed later on [96].

Pathology images have often been proposed for content-based access [43,140] as the color and texture properties can relatively easy be identified. The tasks of a pathologist when searching for reference cases also supports the use of an image retrieval system instead of only reference books. The use with tuberculosis smears is described in [141]. An application with *histopathologic* images is described in [142] and histologic images are analyzed in [134,143,144]. Within *cardiology*, CBIR has been used to discover stenosis images [97]. MRIs of the heart have been used in [145].

Within the *radiology* department, mammographies are one of the most frequent application areas with respect to classification and content-based search [146—149]. The negative psychological effects of removing tissue for false positive patients have been described of one of the principal goals to be reduced. Ultrasound images of the breast are used in [41]. Varied ultrasound images are used in [150].

Another active area is the classification of high resolution computed tomography (HRCT) scans of the lung as done by the Assert project [151,152]. A study about the diagnostic quality with and without using the system showed a significant improvement of the diagnostic quality with using a retrieval system for finding similar cases [99]. A less sophisticated project also using HRCT lung images is described in [125,132]. A justification of use in this area is the hard decision-making task and the strong dependence of the diagnoses from texture properties. Descriptions of HRCT lung images, their visual features and their pathologies are given in [153,154]. The use of thorax radiographies is proposed in [110]. This will be an even harder task as several layers are superposed and many factors other than the pathology can influence the visual content strongly.

Many other articles use medical images to demonstrate their algorithms but a clinical evaluation of their use has rarely been done. In [53,54,155], magnetic resonance images (MRIs) of the brain are used to demonstrate the image search algorithms but the articles do not talk about any medical integration. [115,156] also use MRIs of the head for testing their algorithms. CT brain scans to classify lesions are used in [157]. The search for medical tumors by their shape properties (after segmentation) have

Table 1 Various image types and the systems that are using these images

Images used	Names of the systems
HRCTs of the lung	ASSERT
Functional PET	FICBDS
Spine X-rays	CBIR2, MIRS
Pathologic images	IDEM, I-Browse, PathFinder, PathMaster
CTs of the head	MIMS
Mammographies	APKS
Images from biology	BioImage, BIRN
Dermatology	MELDOQ, MEDS
Breast cancer biopsies	BASS
Varied images	I^2C, IRMA, KMed, COBRA, MedGIFT, ImageEngine

been described in [147]. Functional photon emission tomography (PET) images for retrieval are used in [158]. Spine X-rays are used in [113,159].

Table 1 shows an overview of several image types and the systems that are used to retrieve these images.

2.4. The use in fields close to medicine

There is a number of fields close to the medical domain where the use of content-based access methods to visual data have been proposed as well or are already implemented. In the USA, a biomedical research network is about to be set up, and the sharing of visual data and their management include the use of similarity queries [160]. Multidimensional biological images from various devices are handled in the BioImage project [161]. In [162] drug tablets are retrieved by their visual similarity which is mainly for the identification of ecstacy tablets. Another pharmaceutical use is described in [163] where powders are retrieved based on visual properties.

3. Techniques used in medical image retrieval

This section describes the various techniques that are currently-used or that have been proposed for the use in medical image retrieval applications. Many of the techniques are similar to those used for general content-based retrieval but also techniques that have not yet been used in medical applications are identified. A special focus is put on the data sets that are used to evaluate the image retrieval systems and on the measurements used for evaluation. Unfortunately, the performance evaluation of systems is currently strongly neglected.

Machine learning in medical applications also gets increasingly more important and it is essential to research the various possibilities. Specialized workshops exist for this area [164].

3.1. Features used

This section describes the (visual) features that are used in the various applications. The section text is added to discuss whether this should be named content-based retrieval or rather not. As the formulation of similarity queries without text can be quite a problem, another subsection is added to describe the various possibilities to formulate queries without text.

3.1.1. Query formulation

The query formulation with using exclusively visual features can be a big problem. Most systems in CBIR use the query by example(s) (QBE) paradigm which needs an appropriate starting image for querying. This problem of a sometimes missing starting image is known as the *page zero problem*.

If text is attached to the images, which is normally the case in medical applications, then the text can be used as a starting point and once visually relevant images have been found, further queries can be entirely visual [115] to find visually similar cases not able to be found by text or to sort the found cases by their visual similarity. In the medical decision-making process, there are often images produced and available for the current case. The starting point does thus not need to be further defined but the images of the case can be used directly [121]. In connection with the segmentation of the images the user can also restrict the query to a certain region of interest (ROI) in the image [121], which can lead to much more specific queries than if using an image in its entirety.

The use of human sketches has already been proposed in generic image retrieval [33,165] and it has also been proposed for the use in medical applications [113,115,121,166]. Considering the difficulty in exact drawing and the need for some artistic skills and time, this method will only be applicable for a very small subset of queries, such as tumor shapes or spine X-rays, where outlines are possible directly in the image. For general image retrieval, sketches are too time-consuming and the retrieved results often not exact enough.

3.1.2. Text

Many systems propose to use text from the patient record [120] or studies [121] to search by content. Others define a context-free grammar [97], a standardized vocabulary for image description [142] or an image definition language [126] for the querying of images in image repositories. [167,168] uses text from radiology reports to transform it into concepts in the UMLS metathesaurus to then retrieve the images. The use of text for queries is undeniable efficient but the question is whether this can really be called content-based queries as the text does not necessarily define the image content. It rather puts the images into the context they have been taken in, so it should maybe called context-based queries as defined in [67]. The combination of textual with visual features or content and context of the images does have the most potential to lead to good results [113]. One can also be used to control the quality of the other or to obtain a better recall of the retrieval results.

Besides the free text that is frequently used for retrieval, medical patient records also contain very valuable *structured information* such as age, sex and profession of the patient. This information is just as important as free text to put the images into a context.

3.1.3. Visual features

Unfortunately, most articles that propose content-based queries do not explain in detail which visual features have been used or are planned to be used. Sometimes, only a very vague description such as general texture and color or grey level features are given as in [54,127,169].

Basically all systems that do give details use color and grey level features, mostly in the form of a histogram [134,143,150,151]. Local and global grey level features are used in [170]. [100,128] use statistical distributions of grey levels for the classification of images and [122] proposes a brightness histogram. As many of the images in the medical domain do not contain colors or are taken under controlled conditions, the color properties are not at all in the center of research and the same holds for invariants to lighting conditions. This can change when using photographs such as in dermatology. Pathologic images will need to be normalized in some way as different staining methods can produce different colors [171]. Within radiology, the normalization of grey levels between different modalities or even for the same modality can cause problems when there is no exact reference point as is for the density of the CT, for example. [172] illustrates the dependency of intensity values of the brain from the used modalities.

As color and grey level features are of less importance in medical images than in stock photography, the texture and shape features gain in importance. Basically all of the standard techniques for texture characterization are used from edge detection using Canny operators [141] to Sobel descriptors [151]. [113,139,151] also use Fourier descriptors to characterize shapes, [113,123,139] use invariant moments and [113] also scale-space filtering. Features derived from co-occurrence matrices are also frequently used [96,115,150,151], as well as responses of Gabor filters [134,143,170], wavelets [140,150] and Markov texture characteristics [139]. In mammography, denseness is used for finding small nodules [148]. It would be interesting to have a comparison of several texture descriptors. Many of them model the same information and will most likely deliver very similar results.

In connection with segmentation, the *shape* of the segments can be used as a powerful feature. Again, often the exact nature of the shape features is not described [115] which makes it impossible to define what exactly had been used. In [145] no segmentation has been done for the acquisition of shape features but computer-assisted outlining. The segmentation of pathologic images is described in [140]. In [156] even shape descriptors for 3D structures using modal modeling are described. Most common shape descriptors are Fourier descriptors [43,132,141] that easily allow to obtain invariant descriptions. The pattern spectrum is proposed in [147] and morphological features are used in [147].

Using segments in the images also allows to use *spatial relationships* as visual descriptors of the images. This is often proposed [114,116,121,169,173] but rarely any detail is given on how to obtain the objects/segments in the images, which does not permit to judge whether an implementation is possible. Another article not taking into account the problems of automatic segmentation is [116].

The use of Eigenimages for the retrieval of medical images in analogy to Eigenfaces for face recognition is proposed in [74,124]. These features can be used for classification when a number of images for each class exist. Still, the features are purely statistical and it is hard to actually explain the similarity of two images based on these features which can more easily be done for a histogram intersection, for example.

In [121], signatures of the manually segmented objects of the images are proposed to reduce the list of resulting images. It is hard to say whether these features can count as visual features as they are not extracted automatically but based on semiautomatic segmentations and marking of the segments.

Tissue time—activity curve (TTAC) curves for the retrieval of PET images are used in [158]. These are not really image features but rather 1D temporal signals that are compared. However, the results seem to be good.

Similar to general CBIR, *semantic features* are proposed for visual similarity queries with medical images [143,144]. But again, it comes down to simple textual labels attached to the images and a mapping between the text and the low-level features. A project for automatically attaching semantic labels to images or regions is described in [134] and in ProjetImage.[15]

3.2. Comparison methods and feature space reductions

Most systems do not give many details on the distance measurements or comparison methods used which most likely implies an Euclidian vector space model using either a simple Euclidean distance (L2) or something close such as city-block distance or L1. To efficiently work with these distances even in large databases, the dimensionality is often reduced. This can be done with methods such as *principal component analysis* (PCA) [74,124] or *minimum description length* (MDL) [151] that try to reduce the dimensionality while staying as discriminative as possible. In principle, redundant information is removed but this can also remove small but important changes from the feature space. Techniques such as *KD-trees* [145] and R-trees [173] are also used in medicine for efficient access to such a large feature spaces.

On the other hand, statistical methods are used for the comparison of features that can be trained with existing data and that can then be used on new, incoming cases. These can be *neural networks* for the classification of mammography images [63,148] or on images extremely reduced in size (18×12 pixels) in [166]. Other statistical approaches use *Bayesian networks* [157] or *Hidden Markov Models* (HMMs) [96]. In [174], an associative computing approach is proposed for retrieval assuming that a query is performed with a local part of the images.

A receiver operating characteristic (ROC) curve for the comparison of methods is used in [136]. This is well known in the medical domain and easily interpretable.

3.3. Image databases used for evaluation

The data used for demonstrating the capabilities of the visual access methods are extremely varied in size and quality. From 15 PET studies in [158] to more than 25,000 images in [92] is the spectrum of the articles analyzed for this review.

Often, the images are pre-processed into sometimes fairly small blocks (18×12, [166], 64×64, [134] and 256×256, [170]) before the visual features are extracted. In some cases, even a reduction to 32×32 pixels has proven not to influence the quality of the results compared with using the original size [100]. Some systems use pre-processing to remove artifacts from the image or to improve image quality such as the removal of hairs from dermatologic images [75].

Unfortunately, most of the larger databases such as [124] containing 10,000 MRI images, [116,173] containing 13,500 CT and MRI images and [147] using 1,000 tumor shapes only use *simulated* images. Although these simulated images are easy and cheap to obtain, their use for any qualitative assessments is more than questionable. On the other hand, only [121] uses a very large database containing 22,000 images of a PACS but without any further assessment of image categories and qualities and without an evaluation. [159] uses 17,000 spinal X-ray images as the basis of their research [92] proposes even more images, but here as well, no content-based access mechanisms are implemented as of yet.

An interesting approach to obtain a large database is taken in [54], where 2000 images from freely available medical image databases on the web are taken. A database containing more than 8000 varied medical images is available free of charge from the casimage webpage or can be ordered from the author of this article.[16] [123] uses a varied set of 4247 medical images.

The other, often specialized image collections for content-based retrieval are unfortunately sometimes too small for delivering any statistically significant measurements: [158] uses 15 PET studies, [149] 41 biopsy slides, [157] 48 brain CTs, [167] 50 varied images with radiology reports, [141] 65 smears for tuberculosis identification, [145] 85 MRI images, [74] 100 axial brain images, [75] 100 dermatologic images, [43] 261 cell images, [41] 263 ultrasound breast images, [132] 266 CT images and [96] 300 cell images. 312 HRCTs of the lung are used in [151,152], 345 liver disorders in [175], 404 biopsy proven mammography masses in [148] and 749 dermatological images in [136].

Almost as interesting as the image database itself is the question of how to choose query topics and then how to assess relevance for the query topics. The subject of relevance alone can fill several books [176,177]. This is relatively easy for simulated images as there is a model plus some added

[15] http://perso-iti.enst-bretagne.fr/~brunet/Boulot/ProjetImage/ProjetImage.html.

[16] http://www.casimage.com/.

noise and the noise level basically determines the measured quality of retrieval. Simulated images are consequently only usable for showing efficiency of an algorithm using large image repositories. Nothing can really be said about retrieval quality when using simulated images.

For the future, it is extremely important that image databases are made available free of charge and/or copyright for the comparison and verification of algorithms. Only such reference databases allow to compare systems and to have a reference for the evaluation that is done based on the same images. Some medical image collections are freely available on the Internet.[17] [18] [19] [20] An important effort is underway by the European Federation of Medical Informatics (EFMI) in a working group on medical image processing[21] to generate reference databases and identify important medical imaging tasks [112].

3.4. System evaluations

Already in the general image retrieval domain it is difficult to compare any two retrieval systems. For medical image retrieval systems, the evaluation issue is almost non-existent in most of the papers [54,102,114,115,118—120,126,127,174]. Still, there are several articles on the evaluation of imaging systems in medicine [111] or on general evaluation of clinical systems and the problems with it [178].

Those systems that do perform evaluation often only use screenshots of example results to queries [121—124,145,149,169]. A single example result does not reveal a great deal about the real performance of the system and is not objective as the best possible query result can be chosen arbitrarily by the authors. This problematic in retrieval system evaluation is described in detail in [179]. Most other system evaluations show measures with a limited power for comparison. In [151], the precision of the four highest ranked images is used which does not reveal much about the number of actually relevant items and gives very limited information about the system. [74] measures the number of times a differently scaled or rotated image retrieves the original which is also not very close to medical image retrieval reality.

[17] http://marathon.csee.usf.edu/Mammography/Database.html.
[18] http://cir.ncc.go.jp/pub/gmain.html.
[19] http://www.meduniv.lviv.ua/links/index_multimedia.html.
[20] http://wwwihm.nlm.nih.gov/.
[21] http://www.efmi-wg-mip.net/.

In medical statistics commonly used measurements are sensitivity and specificity defined as follows:

$$\text{sensitivity} = \frac{\text{positive items classified as pos.}}{\text{all positive items}} \tag{1}$$

$$\text{specificity} = \frac{\text{negative items classified as neg.}}{\text{all negative items}} \tag{2}$$

Systems that use sensitivity and specificity include [41,136,141]. These values can also be presented in the form of a ROC curve which contains much more information and is done in [136,157]. As many of the presented systems use classifications of images, accuracy is very often used to evaluate the system [96,100,125,141,143,146]. This can be defined as follows:

$$\text{accuracy} = \frac{\text{items classified correctly}}{\text{all items classified}} \tag{3}$$

Still, it has to be kept in mind that content-based retrieval systems are not mainly being employed for classification of the images but for finding similar images or cases. This is often more helpful as the practitioner must still judge the retrieved cases and the reasons for retrieving the images are often clearer whereas classification results are sometimes hard to detail and need to be explained.

Only rarely are measurements used that are common to the domains of information retrieval [180] or content-based image retrieval [179] such as precision and recall defined as follows:

$$\text{precision} = \frac{\text{number of relevant items retrieved}}{\text{number of items retrieved}} \tag{4}$$

$$\text{recall} = \frac{\text{number of relevant items retrieved}}{\text{number of relevant items}} \tag{5}$$

In [140], for example, the precision after 50 images are retrieved is measured to describe the system performance [123] mentions precision and recall for the evaluation but then, does not use it. [116] uses the precision at five different cutoff points. These data are incomplete and hard to interpret as little is known about the number of relevant images and thus on the difficulty of the query task. Much better is the use of a precision vs. recall graph that puts the two values on the axis of a graph as in [147].

Another rarely mentioned evaluation parameter is the speed of the system which is very important for an interactive system. In [123] it is only mentioned that the speed is reduced from hours to minutes for a set of 4000 images which is completely insufficient for an interactive system where response times should be around one second at a maximum.

This list with few in depth evaluations shows that evaluation is very often neglected in medical image retrieval. It is extremely important and crucial for the success of this technology. Measurement parameters need to show the usefulness of an application and the possible impact that an application of the method can have.

Such an evaluation does not only contain the validation of a technology which is commonly evaluated with measures such as specificity and sensitivity but also the inclusion of human factors into the process such as usability issues and acceptance of the technology [178], which can be obtained through real user tests. Finally, it will be interesting to evaluate the clinical impact of the application when it is used in real clinical practice. Are these technologies able to reduce the length of stay of patients or do they manage to reduce the use of human resources for the patient care?

Studies on clinical effects of image retrieval technologies might still be a distance away but there are several necessities that can be done at the moment such as the definition of standard databases that are freely available, the definition of query topics for these databases including the creation of a ''gold standard'' or ground truth for these topics. This can, in the long run, make way for real clinical studies once the general retrieval performance is proven.

3.5. Techniques not yet used in the medical field

The preceding subsections already showed the large variability in techniques that are used for the retrieval of images. Still, several very successful techniques from the image retrieval domain have not been used for medical images as of yet. The entire discussion on *relevance feedback* that first improved the performance of text retrieval systems and then, 30 years later, of image retrieval systems has not at all been discussed for the medical domain. A few articles mention it but without any details on use and performance. Often the argument for omitting relevance feedback is that medical doctors do not have the time to look at cases and judge them. If the systems are interactive (response times below 1 s, [181]) this should not be a reason as an expert can mark a few images as positive and negative relevance feedback within less than a minute and the improved quality will more than compensate for a minute lost. Also the prospect of long-term learning from this marking of images should motivate people to use it. Long-term learning has shown to be an extremely effective tool for system improvements.

Another domain not discussed at all for medical images are the user interfaces. Sometimes web-based interfaces are proposed [170,182] but no comparison of interfaces is reported and no real usability studies have been published to the authors knowledge so far. As there are several creative solutions in image retrieval it will be interesting to study the effects of interfaces, ergonomics and usability issues on the acceptance and use of the technology in clinical practice.

Performance comparisons for different feature sets have also never been performed and are important to identify well-performing visual features and the applications that they can successfully be used for. This would help a great deal to start new projects in the domain and also to optimize existing systems.

4. Potential clinical benefits and future research

This section gives an overview of the potential application areas of medical image retrieval systems by the image content and the potential clinical benefits of it. Some propositions for future research are made that can influence the research outcome of content-based retrieval methods in the medical domain.

4.1. Application fields in medicine and clinical benefits

Already in Section 2.3 it has been shown that content-based retrieval methods are used in a large variety of applications and departments. This section gives a more ordered view on what in medicine image retrieval can be used for and what the effects can be if proper applications are developed.

Three large domains can instantly be stated for the use of content-based access methods: *teaching, research* and *diagnostics*. Other very important fields are the automatic *annotation/codification* of images and the classification of medical images.

First to benefit will most likely be the domain of *teaching*. Here, lecturers can use large image repositories to search for interesting cases to present to the students. These cases can be chosen not only based on diagnosis or anatomical region but also visually similar cases with different diagnoses can be presented which can augment the educational quality. Indeed, in multiplying the routes to access the right data, cross-correlation approaches between media and various data can be eased. On the other hand, anonymized image archives can be made available for medical students for educational purposes. Content-based techniques allow browsing databases and then comparisons of diagnoses of visually similar cases. Especially for Internet-based teaching, this can offer new possibilities. As most of the systems are based on Internet technologies this does not cause any implementation problems.

Research can also benefit from visual retrieval methods. Researchers have more options for the choice of cases to include into research and studies by allowing text-based and visual access. It can also be imagined that by including visual features directly into medical studies, new correlations between the visual nature of a case and its diagnosis or textual description could be found. Visual data can also be mined to find changes or interesting patterns which can lead to the discovery of new knowledge by combining the various knowledge sources.

Finally, *diagnostics* will be the hardest but most important application for image retrieval. To be used as a diagnostic aid, the algorithms need to prove their performance and they need to be accepted by the clinicians as a useful tool. This also implies an integration of the systems into daily clinical practice which will not be an easy task. It is often hard to change the methods that people are used to, confidence needs to be won. For domains such as evidence-based medicine or case-based reasoning it is essential to supply relevant, similar cases for comparison. Such retrieval will need special visual features that model the visual detection of an MD using as much domain knowledge as possible. Images are normally taken for a very specific reason and this needs to be modeled.

There are two principal ideas for supporting the clinical decision-making process. The first one is to supply the medical doctor with cases that offer a *similar visual appearance*. This can supply a second opinion for the MD and (s)he can perform the reasoning based on the various cases that are supplied by the system and the data that is available on the current patient. Another idea is the creation of databases containing normal (non-pathologic) cases and compare the distance of a new case with the existing cases doing thus *dissimilarity retrieval* as opposed to similarity retrieval (distance to normality). This is even more natural compared to the normal workflow in medicine where the first requirement is to find out whether the case is pathologic or not. A tumor or fracture are such differences from normal cases, for example. A dissimilarity could be combined with highlighting regions in the image where the strongest dissimilarity occurred. Such a technique can help to find cases that might otherwise be missed. A combination of the two approaches is also possible where firstly, the requirement is whether the image contains abnormalities and if it does, a query to find similar cases is done with another image database containing the pathologic cases.

High quality *annotation/codification* is a problem not only in radiology but also in other medical departments. Good annotation and codification takes time and experience that is unfortunately sometimes not available in medical routine. Much research is done on natural language processing techniques to extract diagnoses from the patient record [183] and many tools exist to ease the coding, for example for the American College of Radiology (ACR) codes[22] in radiology.[23] When large databases of correctly coded images are available, image retrieval systems can be used for semi-automatic coding by retrieving visually similar cases and proposing the codes of the images from the database. Studies will need to prove the quality of the coding but time can be saved even when a medical doctor only has to control the codes that the system is proposing. Retrieval methods can also be used as simple tools to have a quality control on the DICOM headers. The combination of textual and visual attributes definitely promises the best results.

In principle, all image-producing departments can profit from content-based technologies but there are some departments and some sorts of images that seem to stand out as textures and colors do play an important role for the diagnostics. Color and texture features are normally easy to index with current retrieval systems.

This includes *Pathology* where microscopic images are analyzed and the clinical decision-making depends on the color changes and textures within the images. Many books with example images for typical or hard cases exist and it is relatively easy to provide these books in a digital form and search for them not only based on text or hierarchies but also based on the visual content. Care needs to be taken with respect to different staining methods. Images need to be normalized with respect to that [171].

Hematology already contains a large number of tools to automatically count blood cells but an interesting application would be the classification of abnormal white blood cells and the comparison of diagnoses between a new case and cases with similar abnormalities stored in the databases.

Dermatology already has classification applications for potential melanoma cases that work fairly

[22] http://www.acr.org/.
[23] http://www.casimage.com/ACR.html.

well. Content-based access can help to make understand the decision of an expert system to the practitioner.

Within the *Radiology* department there are a number of possible applications that can deliver good results. For HRCTs of the lung, computer-based tools have already been proven to help in the diagnostics process and diagnostics in this case are fairly difficult. Three-dimensional retrieval can also help to retrieve tumor forms and to classify observed tumors. As a tool for the use in PACS systems, a large number of people can profit from the methods to retrieve similar cases for a number of applications, often without realizing that the results come from a content-based retrieval engine.

4.2. Future research

When thinking about future research directions it becomes apparent that the goal needs to be a real *clinical integration* of the systems. This implies a number of changes in the ways that research is done at the moment. It will become more important to design applications in a way that they can be integrated easier into existing systems through open communication *interfaces*, for example based on extensible markup language (XML) as a description language of the data or Hyper-Text Transport Protocol (HTTP) as a transport protocol for the data [184]. Such a use of standard Internet technologies can help for the integration of retrieval methods into other applications. Such access methods are necessary to make the systems accessible to a larger group of people and applications and to gain experience that goes far beyond a validation of retrieval results. This can not only be seen as engineering but as research as the practical use of the integrated methods needs to be researched.

The integration into PACS is an essential step for the clinical use of retrieval systems. PACS solutions currently allow search by patient and study characteristics and are mainly a storage place for images. A project to allow further search methods in medical image databases based on a standard communication interface is the Medical Image Resource Center (MIRC).[24] Here, search by several characteristics, including free-text, is allowed based on a standard platform. The future of PACS or medical image storage systems might be in a separate architecture with a *storage component* just as PACS systems currently are and an automatic *indexing system* where important characteristics from the images and the linked case information are stored to allow for retrieval methods based on structured information, free text and the visual image content.

Of course, *evaluation* of the retrieval quality is an extremely important topic as well. Research will need to focus on the development of open test databases and query topics plus defined gold standards for the images to be retrieved. Retrieval systems need to be compared to identify good techniques. This can advance the field much more than any single technique developed so far.

But evaluation also needs to go one step further and prepare *field studies* on the use and the influence of retrieval techniques on the diagnostic process. So far, only one study on the impact of image retrieval system on the diagnostics of HRCT images of the lung has been published and shows a significant improvement in diagnostic quality even for senior radiologists [99]. Practitioners need to give their opinion on the usability and applicability of the technologies and acceptance needs to be gained before they can be used in daily practice. Such communication with the system users can also improve the interface and retrieval quality significantly when good feedback is delivered.

User interaction and relevance feedback are two other techniques that need to be integrated more into retrieval systems as this can help to lead to much better results. Image retrieval needs to be interactive and all the interaction needs to be exploited for delivering the best possible results.

Multimedia data mining will also be made possible once features of good quality are available to describe the images. This will help to find new relationships among images and certain diseases or it will simply improve the retrieval quality of medical image search engines.

Although first applications will most likely be on large image archives for teaching and research, a *specialization* of the retrieval systems for promising domains such as dermatology or pathology will be necessary to include as much domain knowledge as possible into the retrieval. This will be necessary for decision-support systems such as systems for case-based reasoning. Such a specialization can be done in the easiest way with a modular retrieval system based on components where feature sets can be exchanged easily and modules for new retrieval techniques or efficient storage methods can be integrated easily. Fig. 4 shows such a component-based architecture where system parts can be changed and optimized easily. Easy plug-in mechanisms for the different components need to be defined.

Besides the use of images, system developments also need to put a focus on *higher-dimensional data*. Already tomographic images contain three

[24] http://mirc.rsna.org/.

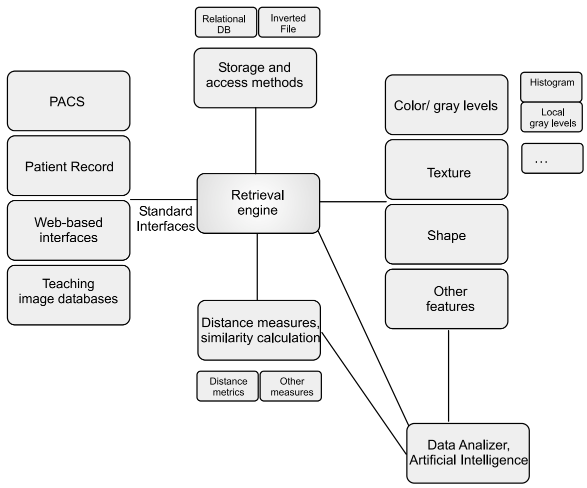

Fig. 4 A modular schema for retrieval system development.

dimensions as do video sequences of endoscopy or ultrasound. Tools for retrieval of videos for example by motion parameters do exist for general videos [185,186] but to our knowledge do not exist specialized for the medical domain. Fast scanners also allow for the registration of 4D-data streams such as tomographic images taken over time. Combinations of modalities such as PET/CT scanners or the use of image fusion techniques also create multi-dimensional data that needs to be analyzed and retrieved. Omitting these high-dimensional informations will result in a significant lack of knowledge.

5. Conclusion

The large number of research publications in the field of content-based medical image retrieval especially in recent years shows that it is very active and that it is starting to get more attention. This will hopefully advance the field as new tools and technologies will be developed and performance will increase. Content-based visual information retrieval definitely has a large potential in the medical domain. The amount of visual data produced in medical departments shows the importance of developing new and alternative access methods to complement text. Content-based methods can be used on a large variety of images and in a wide area of applications. Still, much work needs to be done to produce running applications and not only research prototypes. When looking at most current systems, it becomes clear that few to none of them are actually in routine use.

An important factor is to build prototypes that are integrated with a hospital-wide communication structure and that use open standards, so data can be exchanged with other applications. It needs to become easy to integrate these new functionalities into other existing applications such as a hospital information system (HIS)/radiology information system (RIS)/PACS or other medical image management or viewing software. In this way, it will become much easier to have prototypes running for a sample of users and to get feedback on the clinical use of systems. To get acceptance, it is important to be integrated into the current applications and with interfaces that the users are familiar with. To win acceptance from the users it is also important to show the performance of the systems and to optimize the performance of systems for certain specialized tasks or people.

The development of open toolboxes is another important factor for successful applications. Not only do interfaces for the communication with other applications need to be developed, also within the application it is important to stay modular, so parts and pieces can be exchanged easily. This will help to reduce the number of applications developed and will make it possible to spend more time on the important tasks of integration and development of new methods and system optimizations.

It is clear that new tools and methods are needed to manage the increasing amount of visual information that is produced in medical institutions. Content-based access methods have an enormous potential when used in the correct way. It is now the time to create medical applications and use this potential for clinical decision-making, research and teaching.

6. Summary

This article gives an overview of the currently available literature on content-based image retrieval in the medical domain. It evaluates after a few years of developments the need for image retrieval and presents concrete scenarios for promising future research directions.

The necessity for additional, alternative access methods to the currently-used, text-based methods in medical information retrieval is detailed. This need is mainly due to the large amount of visual data produced and the unused information that these data contain, which could be used for diagnostics, teaching and research. The systems described in the literature and published propositions for image retrieval in medicine are critically reviewed and sorted by medical departments, image categories and technologies used. A short overview of nonmedical image retrieval is given as well.

The lack of evaluations of the retrieval quality of systems becomes apparent along with the unavailability of large image databases free of charge with defined query topics and gold standards. However, some databases are available, from the National Institutes of Health (NIH), for example. Ideas for creating such image databases and evaluation methods are proposed. Also, several research directions for improving the retrieval quality based on the experiences from other closely related research fields are given in the paper. Possible clinical benefits from the use of content-based access methods are described as well as promising fields of applications.

Acknowledgements

The authors would like to thank the reviewers for their comments that helped to improve the quality of this paper.

References

[1] S.-K. Chang, T. Kunii, Pictorial data-base applications, IEEE Comput. 14 (11) (1981) 13—21.

[2] P.G.B. Enser, Pictorial information retrieval, J. Document. 51 (2) (1995) 126—170.

[3] A. Gupta, R. Jain, Visual information retrieval, Commun. ACM 40 (5) (1997) 70—79.

[4] Y. Rui, T.S. Huang, S.-F. Chang, Image retrieval: past, present and future, in: M. Liao (Ed.), Proceedings of the International Symposium on Multimedia Information Processing, Taipei, Taiwan, 1997.

[5] J.P. Eakins, M.E. Graham, Content-based image retrieval, Tech. Rep. JTAP-039, JISC Technology Application Program, Newcastle upon Tyne, 2000.

[6] C.C. Venters, M. Cooper, Content-based image retrieval, Tech. Rep. JTAP-054, JISC Technology Application Program, 2000.

[7] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, Content-based image retrieval at the end of the early years, IEEE Trans. Pattern Anal. Machine Intel. 22 (12) (2000) 1349—1380.

[8] H. Müller, User interaction and performance evaluation in content-based visual information retrieval, Ph.D. thesis, Computer Vision and Multimedia Laboratory, University of Geneva, Geneva, Switzerland, June 2002.

[9] J.R. Smith, Integrated spacial and feature image systems: retrieval, compression and analysis, Ph.D. thesis, Graduate School of Arts and Sciences, Columbia University, 2960 Broadway, New York, NY, USA, 1997.

[10] A. del Bimbo, Visual Information Retrieval, Academic Press, New York, 1999.

[11] S.M. Rahman, Design & Management of Multimedia Information Systems: Opportunities & Challenges, Idea Group Publishing, London, 2001.

[12] L.H.Y. Tang, R. Hanka, H.H.S. Ip, A review of intelligent content-based indexing and browsing of medical images, Health Informatics J. 5 (1999) 40—49.

[13] N.-S. Chang, K.-S. Fu, Query-by-pictorial-example, IEEE Trans. Software Eng. SE 6 (6) (1980) 519—524.

[14] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, Query by image and video content: the QBIC system, IEEE Comput. 28 (9) (1995) 23—32.

[15] W. Niblack, R. Barber, W. Equitz, M.D. Flickner, E.H. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, G. Taubin, QBIC project: querying images by content, using color, texture, and shape, in: W. Niblack (Ed.), Proceedings of the of SPIE Conference on Storage and Retrieval for Image and Video Databases, vol. 1908, 1993, pp. 173—187.

[16] J.R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, C.-F. Shu, The Virage image search engine: an open framework for image management, in: I.K. Sethi, R.C. Jain (Eds.), Proceedings of the SPIE Conference on Storage & Retrieval for Image and Video Databases IV, vol. 2670, San Jose, CA, USA, 1996, pp. 76—87.

[17] A. Hampapur, A. Gupta, B. Horowitz, C.-F. Shu, C. Fuller, J. Bach, M. Gorkani, R. Jain, Virage video engine, in: I.K. Sethi, R.C. Jain (Eds.), Proceedings of the of SPIE Conference on Storage and Retrieval for Image and Video Databases V, vol. 3022, 1997, pp. 352—360.

[18] P.M. Kelly, M. Cannon, D.R. Hush, Query by image example: the Candid approach, in: W. Niblack, R.C. Jain (Eds.), Proceedings of the of SPIE Conference on Storage and Retrieval for Image and Video Databases III, vol. 2420, 1995, pp. 238—248.

[19] A. Pentland, R.W. Picard, S. Sclaroff, Photobook: tools for content-based manipulation of image databases, Int. J. Comput. Vis. 18 (3) (1996) 233—254.

[20] W.Y. Ma, Y. Deng, B.S. Manjunath, Tools for texture- and color-based search of images, in: B.E. Rogowitz, T.N. Pappas (Eds.), Proceedings of the of SPIE Conference on Human Vision and Electronic Imaging II, vol. 3016, San Jose, CA, 1997, pp. 496—507.

[21] C. Carson, M. Thomas, S. Belongie, J.M. Hellerstein, J. Malik, Blobworld: a system for region-based image indexing and retrieval, in: D.P. Huijsmans, A.W.M. Smeulders (Eds.), Proceedings of the Third International Conference On Visual Information Systems (VISUAL'99), no. 1614 in Lecture Notes in Computer Science, Springer-Verlag, Amsterdam, The Netherlands, 1999, pp. 509—516.

[22] S. Belongie, C. Carson, H. Greenspan, J. Malik, Color- and texture-based image segmentation using EM and its application to content-based image retrieval, in: Proceedings of the International Conference on Computer Vision (ICCV'98), Bombay, India, 1998, pp. 675—682.

[23] I.J. Cox, M.L. Miller, S.M. Omohundro, P.N. Yianilos, Target testing and the PicHunter Bayesian multimedia retrieval system, in: Advances in Digital Libraries (ADL'96), Library of Congress, Washington, D.C., 1996, pp. 66—75.

[24] D.M. Squire, W. Müller, H. Müller, T. Pun, Content-based query of image databases: in-spirations from text retrieval, Pattern Recognition Letters, vol. 21, 2000, pp. 1193—1198 (Selected papers from the 11th Scandinavian Conference on Image Analysis, SCIA'99).

[25] D.M. Squire, H. Müller, W. Müller, S. Marchand-Maillet, T. Pun, Design and evaluation of a content-based image retrieval system, in: Design & Management of Multimedia Information Systems: Opportunities & Challenges, IDEA Group Publishing, pp. 125—151 (Chapter 7), 2001.

[26] T. Pfund, S. Marchand-Maillet, Dynamic multimedia annotation tool, in: G. Beretta, R. Schettini (Eds.), Proceedings of the SPIE Photonics West Conference on Internet Imaging III, vol. 4672, San Jose, CA, USA, 2002, pp. 216—224.

[27] C. Carson, S. Belongie, H. Greenspan, J. Malik, Region-based image querying, in: Proceedings of the 1997

IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97), IEEE Computer Society, San Juan, Puerto Rico, 1997, pp. 42—51.
[28] J.R. Smith, S.-F. Chang, Visualseek: a fully automated content-based image query system, in: Proceedings of the Fourth ACM International Multimedia Conference and Exhibition, Boston, MA, USA, 1996.
[29] S. Sclaroff, L. Taycher, M. La Cascia, ImageRover: a content-based browser for the world wide web, in: IEEE Workshop on Content-Based Access of Image and Video Libraries, San Juan, Puerto Rico, 1997, pp. 2—9.
[30] T. Gevers, A.W.M. Smeulders, A comparative study of several color models for color image invariants retrieval, in: Proceedings of the First International Workshop ID-MMS'96, Amsterdam, The Netherlands, 1996, pp. 17—26.
[31] J.-M. Geusebroek, R. van den Boogaard, A.W.M. Smeulders, H. Geerts, Color invariance, IEEE Trans. Pattern Anal. Machine Intel. 23 (12) (2001) 1338—1350.
[32] M. Ortega, Y. Rui, K. Chakrabarti, K. Porkaew, S. Mehrotra, T.S. Huang, Supporting ranked boolean similarity queries in MARS, IEEE Trans. Knowledge Data Eng. 10 (6) (1998) 905—925.
[33] J. Ze Wang, G. Wiederhold, O. Firschein, S. Xin Wei, Wavelet-based image indexing techniques with partial sketch retrieval capability, in: Proceedings of the Fourth Forum on Research and Technology Advances in Digital Libraries, Washington D.C., 1997, pp. 13—24.
[34] W. Ma, B. Manjunath, Texture features and learning similarity, in: Proceedings of the 1996 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'96), San Francisco, California, 1996, pp. 425—430.
[35] S. Santini, R. Jain, Gabor space and the development of preattentive similarity, in: Proceedings of the 13th International Conference on Pattern Recognition (ICPR'96), IEEE, Vienna, Austria, 1996, pp. 40—44.
[36] J.G. Daugman, An information theoretic view of analog representation in striate cortex, Computat. Neurosci. 2 (1990) 9—18.
[37] J.G. Daugman, High confidence visual recognition of persons by a test of statistical independence, IEEE Trans. Pattern Anal. Machine Intel. 15 (11) (1993) 1148—1161.
[38] R. Milanese, M. Cherbuliez, A rotation, translation and scale-invariant approach to content-based image retrieval, J. Visual Commun. Image Represent. 10 (1999) 186—196.
[39] J.S. Weszka, C.R. Dyer, A. Rosenfeld, A comparative study of texture measures for terrain classification, IEEE Trans. Sys. Man Cybernetics 6 (4) (1976) 269—285.
[40] C.-R. Shyu, C.E. Brodley, A.C. Kak, A. Kosaka, A.M. Aisen, L.S. Broderick, ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases, Comput. Vis. Image Understand. 75 (1—2) (1999) 111—132.
[41] W.-J. Kuo, R.-F. Chang, C.C. Lee, W.K. Moon, D.-R. Chen, Retrieval technique for the diagnosis of solid breast tumors on sonogram, Ultrasound Med. Biol. 28 (7) (2002) 903—909.
[42] C.-S. Lu, P.-C. Chung, Wold features for unsupervised texture segmentation, in: Proceedings of the 14th International Conference on Pattern Recognition (ICPR'98), IEEE, Brisbane, Australia, 1998, pp. 1689—1693.
[43] D. Comaniciu, P. Meer, D. Foran, A. Medl, Bimodal system for interactive indexing and 18 retrieval of pathology images, in: Proceedings of the Fourth IEEE Workshop on Applications of Computer Vision (WACV'98), Princeton, NJ, USA, 1998, pp. 76—81.
[44] S.T. Perry, P.H. Lewis, A novel image viewer providing fast object delineation for content based retrieval and navigation, in: I.K. Sethi, R.C. Jain (Eds.), Proceedings of the of SPIE Conference on Storage and Retrieval for Image and Video Databases VI, vol. 3312, 1997.
[45] A. Winter, C. Nastar, Differential feature distribution maps for image segmentation and region queries in image databases, in: IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL'99), Fort Collins, Colorado, USA, 1999, pp. 9—17.
[46] L. Lucchese, S.K. Mitra, Unsupervised segmentation of color images based on κ-means clustering in the chromaticity plane, in: IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL'99), Fort Collins, CO, USA, 1999, pp. 74—78.
[47] S. Ghebreab, Strings and necklaces—on learning and browsing medical image segmentations, Ph.D. thesis, Faculty of computer science, University of Amsterdam, Amsterdam, The Netherlands, July 2002.
[48] R.J. Lapeer, A.C. Tan, R. Aldridge, A combined approach to 3D medical image segmentation using marker-based watersheds and active contours: The active watershed method, in: T. Dohi, R. Kikin (Eds.), International Conference on Medical Image Computing and Computer-Assisted Interventions (MICCAI 2002), no. 2488 in Lecture Notes in Computer Science, Springer-Verlag, Tokyo, Japan, 2002, pp. 596—603.
[49] S. Loncaric, A survey of shape analysis techniques, Pattern Recog. 31 (8) (1998) 983—1001.
[50] R.C. Veltkamp, M. Hagedoorn, State-of-the-art in shape matching, in: Principles of Visual Information Retrieval, Springer, Heidelberg, 2000, pp. 87—119.
[51] D. Dori, Cognitive image retrieval, in: A. Sanfeliu, J.J. Villanueva, M. Vanrell, R. Alcézar, J.-O. Eklundh, Y. Aloimonos (Eds.), Proceedings of the 15th International Conference on Pattern Recognition (ICPR 2000), IEEE, Barcelona, Spain, 2000, pp. 42—45.
[52] J.Z. Wang, J. Li, G. Wiederhold, Simplicity: semantics-sensitive integrated matching for picture libraries, IEEE Trans. Pattern Anal. Machine Intel. 23 (9) (2001) 1—17.
[53] Y. Liu, A. Lazar, W.E. Rothfus, M. Buzoiano, T. Kanade, Classification-driven feature space reduction for semantic-based neuroimage retrieval, in: Proceedings of the International Symposium on Information Retrieval and Exploration in Large Medical Image Collections (VISIM 2001), Utrecht, The Netherlands, 2001.
[54] A. Mojsilovis, J. Gomes, Semantic based image categorization, browsing and retrieval in medical image databases, in: IEEE International Conference on Image Processing (ICIP'2000), Rochester, NY, USA, 2000.
[55] T.P. Minka, R.W. Picard, Interactive learning using a society of models, in: Proceedings of the 1996 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'96), San Francisco, California, 1996, pp. 447—452.
[56] C. Jörgensen, Retrieving the unretrievable in electronic imaging systems: emotions, themes and stories, in: B. Rogowitz, T.N. Pappas (Eds.), Proceedings of the of SPIE Photonics West Conference on Human Vision and Electronic Imaging IV, vol. 3644, San Jose, California, USA, 1999.
[57] K. Koffka, Principles of Gestalt Psychology, Lund Humphries, London, 1935.
[58] A.K. Jain, A. Vailaya, Image retrieval using color and shape, Pattern Recog. 29 (8) (1996) 1233—1244.
[59] A. Tversky, Features of similarity, Psychol. Rev. 84 (4) (1977) 327—352.

[60] M.J. Swain, D.H. Ballard, Color indexing, Int. J. Comput. Vis. 7 (1) (1991) 11—32.
[61] C.C. Aggarwal, A. Hinneburg, D.A. Keim, On the surprising behavior of distance metrics in high dimensional space, in: Proceedings of the International Conference on Database Theory (ICDT 2001), no. 1973 in Lecture Notes in Computer Science, Springer-Verlag, London, UK, 2001.
[62] A. Hinneburg, C.C. Aggarwal, D.A. Keim, What is the nearest neighbor in high-dimensional spaces?, in: Proceedings of 26th International Conference on Very Large Databases (VLDB 2000), Cairo, Egypt, 2000, pp. 506—516.
[63] R. Hanka, T.P. Harte, Curse of dimensionality: classifying large multi-dimensional images with neural networks, in: Proceedings of the European Workshop on Computer-Intensive Methods in Control and Signal Processing (CIMCSP 1996), Prague, Czech Republic, 1996.
[64] N. Vasconcelos, A. Lippman, A probabilistic architecture for content-based image retrieval, in: Proceedings of the 2000 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2000), IEEE Computer Society, Hilton Head Island, South Carolina, USA, 2000, pp. 216—221.
[65] N. Vasconcelos, A. Lippmann, A unifying view of image similarity, in: A. Sanfeliu, J.J. Villanueva, M. Vanrell, R. Alcézar, J.-O. Eklundh, Y. Aloimonos (Eds.), Proceedings of the 15th International Conference on Pattern Recognition (ICPR 2000), IEEE, Barcelona, Spain, 2000, pp. 1—4.
[66] K.-S. Goh, E. Chang, K.-T. Cheng, Support vector machine pairwise classifiers with error reduction for image classification, in: Proceedings of the ACM Multimedia Workshop on Multimedia Information Retrieval (ACM MIR 2001), The Association for Computing Machinery, Ottawa, Canada, 2001, pp. 32—37.
[67] T. Westerveld, Image retrieval: content versus context, in: Recherche d'Informations Assistée par Ordinateur (RIAO'2000) Computer-Assisted Information Retrieval, vol. 1, Paris, France, 2000, pp. 276—284.
[68] L. Zhu, C. Tang, A. Rao, A. Zhang, Using thesaurus to model keyblock-based image retrieval, in: Proceedings of the second International Conference on Multimedia and Exposition (ICME'2001), IEEE Computer Society, IEEE Computer Society, Tokyo, Japan, 2001, pp. 237—240.
[69] G. Salton, C. Buckley, Term weighting approaches in automatic text retrieval, Inform. Proces. Manage. 24 (5) (1988) 513—523.
[70] S. Dominich, Formal foundation of information retrieval, in: Proceedings of the Workshop on Mathematical/Formal Methods in Information Retrieval at the International ACM SIGIR Conference on Research and Development in Information Retrieval, Athens, Greece, 2000.
[71] A.P. de Vries, M.G.L.M. van Doorn, H.M. Blanken, P.M.G. Apers, The Mirror MMDBMS architecture, in: Proceedings of 24th International Conference on Very Large Databases (VLDB'98), New York, NY, USA, 1998, pp. 758—761.
[72] H. Müller, D.M. Squire, W. Müller, T. Pun, Efficient access methods for content-based image retrieval with inverted files, in: S. Panchanathan, S.-F. Chang, C.-C.J. Kuo (Eds.), Proceedings of the SPIE Conference on Multimedia Storage and Archiving Systems IV (VV02), vol. 3846, Boston, MA, USA, 1999, pp. 461—472.
[73] D.A. White, R. Jain, Algorithms and strategies for similarity retrieval, Tech. Rep. VCL-96-101, Visual Computing Laboratory, University of California, San Diego, 9500 Gilman Drive, Mail Code 0407, La Jolla, CA 92093-0407, July 1996.
[74] U. Sinha, H. Kangarloo, Principal component analysis for content-based image retrieval, RadioGraphics 22 (5) (2002) 1271—1289.
[75] P. Schmidt-Saugeon, J. Guillod, J.-P. Thiran, Towards a computer-aided diagnosis system for pigmented skin lesions, Comput. Med. Imag. Graphics 27 (2003) 65—78.
[76] T. Kämpfe, T.W. Nattkemper, H. Ritter, Combining independent component analysis and self-organizing maps for cell image classification, in: B. Radig, S. Florczyk (Eds.), Pattern Recognition: DAGM Symposium 2001—Lecture Notes in Computer Science 2191, Munich, Germanny, 2001, pp. 262—268.
[77] J.J. Rocchio, Relevance feedback in information retrieval, in: The SMART Retrieval System, Experiments in Automatic Document Processing, Prentice Hall, Englewood Cliffs, NJ, USA, 1971, pp. 313—323.
[78] H. Müller, W. Müller, D.M. Squire, S. Marchand-Maillet, T. Pun, Strategies for positive and negative relevance feedback in image retrieval, in: A. Sanfeliu, J.J. Villanueva, M. Vanrell, R. Alcézar, J.-O. Eklundh, Y. Aloimonos (Eds.), Proceedings of the 15th International Conference on Pattern Recognition (ICPR 2000), IEEE, Barcelona, Spain, 2000, pp. 1043—1046.
[79] Y. Rui, T.S. Huang, M. Ortega, S. Mehrotra, Relevance feedback: a power tool for interactive content-based image retrieval, IEEE Trans.Circuits Sys. Video Technol. 8 (5) (1998) 644—655 (special issue on segmentation, description, and retrieval of video content)..
[80] Y. Rui, T.S. Huang, S. Mehrotra, Relevance feedback techniques in interactive content-based image retrieval, in: I.K. Sethi, R.C. Jain (Eds.), Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases VI, vol. 3312, 1997, pp. 25—36.
[81] M. Worring, A.W.M. Smeulders, S. Santini, Interaction in content-based image retrieval: An evaluation of the state of the art, in: R. Laurini (Ed.), Proceedings of the Fourth International Conference On Visual Information Systems (VISUAL'2000), no. 1929 in Lecture Notes in Computer Science, Springer-Verlag, Lyon, France, 2000, pp. 26—36.
[82] H. Müller, D.M. Squire, T. Pun, Learning from user behavior in image retrieval: application of the market basket analysis, Int. J. Comput. Vis. (special issue on content-based image retrieval) vol. 56 (1—2), pp. 65—77, 2004.
[83] M. Nakazato, T.S. Huang, 3D MARS: Immersive virtual reality for content-based image retrieval, in: Proceedings of the Second International Conference on Multimedia and Exposition (ICME'2001), IEEE Computer Society, IEEE Computer Society, Tokyo, Japan, 2001, pp. 45—48.
[84] S. Santini, R. Jain, Direct manipulation of image databases, Tech. rep., Department of Computer Science, University of California San Diego, San Diego, CA, November 1998.
[85] T.M. Lehmann, M.O. Güld, C. Thies, B. Fischer, M. Keysers, D. Kohnen, H. Schubert, B.B. Wein, Content-based image retrieval in medical applications for picture archiving and communication systems, in: Proceedings of the SPIE Conference on Medical Imaging, vol. 5033, San Diego, CA, USA, 2003.
[86] B. Revet, DICOM Cook Book for Implementations in Madalities, Philips Medical Systems, Eindhoven, Netherlands, 1997.
[87] R.A. Greenes, J.F. Brinkley, Imaging systems, in: Medical Informatics: Computer Applications in Healthcare, second ed., Springer, New York, 2000, pp. 485—538 (Chapter 14).
[88] C. Kulikowski, E. Ammenwerth, A. Bohne, K. Ganser, R. Haux, P. Knaup, C. Maier, A. Michel, R. Singer, A.C. Wolff, Medical imaging informatics and medical informatics: opportunities and constraints, Methods Inform. Med. 41 (2002) 183—189.

[89] M.W. Vannier, E.V. Staab, L.C. Clarke, Medical image archives—present and future, in: H.U. Lemke, M.W. Vannier, K. Inamura, A.G. Farman, J.H.C. Reiber (Eds.), Proceedings of the International Conference on Computer-Assited Radiology and Surgery (CARS 2002), Paris, France, 2002.

[90] A.P. Sarvazyan, F.L. Lizzi, P.N.T. Wells, A new philosophy of medical imaging, Med. Hypotheses 36 (1991) 327—335.

[91] T. Pun, G. Gerig, O. Ratib, Image analysis and computer vision in medicine, Comput. Med. Imag. Graphics 18 (2) (1994) 85—96.

[92] A. Rosset, O. Ratib, A. Geissbuhler, J.-P. Vallée, Integration of a multimedia teaching and reference database in a PACS environment, RadioGraphics 22 (6) (2002) 1567—1577.

[93] E. Binet, J.H. Trueblood, K.J. Macura, R.T. Macura, B.D. Morstad, R.V. Finkbeiner, Computer-based radiology information system: from floppy disk to CD-ROM, RadioGraphics 15 (1995) 1203—1214.

[94] K. Maloney, C.T. Hamlet, The clinical display of radiologic information as an interactive multimedia report, J. Digital Imag. 12 (2) (1999) 119—121.

[95] T. Frankewitsch, U. Prokosch, Navigation in medical internet image databases, Med. Informatics 26 (1) (2001) 1—15.

[96] S. Beretti, A. Del Bimbo, P. Pala, content-based retrieval of 3D cellular structures, in: Proceedings of the Second International Conference on Multimedia and Exposition (ICME'2001), IEEE Computer Society, IEEE Computer Society, Tokyo, Japan, 2001, pp. 1096—1099.

[97] M.R. Ogiela, R. Tadeusiewicz, Semantic-oriented syntactic algorithms for content recognition and understanding of images in medical databases, in: Proceedings of the second International Conference on Multimedia and Exposition (ICME'2001), IEEE Computer Society, IEEE Computer Society, Tokyo, Japan, 2001, pp. 621—624.

[98] S.C. Orphanoudakis, C.E. Chronaki, S. Kostomanolakis, I^2Cnet: a system for the indexing. storage and retrieval of medical images by content, Med. Informatics 19 (2) (1994) 109—122.

[99] A.M. Aisen, L.S. Broderick, H. Winer-Muram, C.E. Brodley, A.C. Kak, C. Pavlopoulou, J. Dy, C.-R. Shyu, A. Marchiori, Automated storage and retrieval of thin-section CT images to assist diagnosis: system description and preliminary assessment, Radiology 228 (2003) 265—270.

[100] D. Keysers, J. Dahmen, H. Ney, B.B. Wein, T.M. Lehmann, A statistical framework for model-based image retrieval in medical applications, J. Electronic Imag. 12 (1) (2003) 59—68.

[101] H.D. Tagare, C. Jaffe, J. Duncan, Medical image databases: a content-based retrieval approach, J. Am. Med. Informatics Assoc. 4 (3) (1997) 184—198.

[102] H.J. Lowe, I. Antipov, W. Hersh, C. Arnott Smith, 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, pp. 882—886.

[103] W.D. Bidgood, B. Bray, N. Brown, A.R. Mori, K.A. Spackman, A. Golichowsky, R.H. Jones, L. Korman, B. Dove, L. Hildebrand, M. Berg, Image acquisition context: procedure description attributes for clinically relevant indexing and selective retrieval of biomedical images, J. Am. Med. Informatics Assoc. 6 (1) (1999) 61—75.

[104] A. Winter, R. Haux, A three-level graph-based model for the management of hospital information systems, Methods Information Med. 34 (1995) 378—396.

[105] M.O. Güld, M. Kohnen, D. Keysers, H. Schubert, B.B. Wein, J. Bredno, T.M. Lehmann, Quality of DICOM header information for image categorization, in: Proceedings of the International Symposium on Medical Imaging, vol. 4685, San Diego, CA, USA, 2002, pp. 280—287.

[106] C. LeBozec, M.-C. Jaulent, E. Zapletal, 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, 1998.

[107] A.A.T. Bui, R.K. Taira, J.D.N. Dionision, D.R. Aberle, S. El-Saden, H. Kangarloo, Evidence-based radiology, Acad. Radiol. 9 (6) (2002) 662—669.

[108] J.-P. Boissel, M. Cucherat, E. Amsallem, P. Nony, M. Fardeheb, W. Manzi, M.C. Haugh, Getting evidence to prescribers and patients or how to make EBM a reality, in: Proceedings of the Medical Informatics Europe Conference (MIE 2003), St. Malo, France, 2003.

[109] C.E. Kahn, Artificial intelligence in radiology: decision support systems, RadioGraphics 14 (1994) 849—861.

[110] H. Abe, H. MacMahon, R. Engelmann, Q. Li, J. Shiraishi, S. Katsuragawa, M. Aoyama, T. Ishida, K. Ashizawa, C.E. Metz, K. Doi, Computer-aided diagnosis in chest radiography: Results of large-scale observer tests at the 1996-2001 RSNA scientific assemblies, RadioGraphics 23 (1) (2003) 255—265.

[111] B. Kaplan, H.P. Lundsgaarde, Toward an evaluation of an integrated clinical imaging system: Identifying clinical benefits, Methods Inform. Med. 35 (1996) 221—229.

[112] A. Horsch, R. Thurmayr, How to identify and assess tasks and challenges of medical image processing, in: Proceedings of the Medical Informatics Europe Conference (MIE 2003), St. Malo, France, 2003.

[113] S. Antani, L.R. Long, G.R. Thoma, A biomedical information system for combined content-based retrieval of spine X-ray images and associated text information, in: Proceedings of the Third Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2002), Ahamdabad, India, 2002.

[114] W.W. Chu, A.F. Cárdenas, R.K. Taira, KMED: A knowledge-based multimedia distributed database system, Inform. Sys. 19 (4) (1994) 33—54.

[115] S.C. Orphanoudakis, C.E. Chronaki, D. Vamvaka, I^2Cnet: content-based similarity search in geographically distributed repositories of medical images, Comput. Med. Imag. Graphics 20 (4) (1996) 193—207.

[116] E.G.M. Petrakis, Content-based retrieval of medical images, Int. J. Comput. Res. 11 (2) (2002) 171—182.

[117] M. Tsiknakis, D. Katehakis, C. Orphanoudakis, Stelios, intelligent image management in a distributed PACS and telemedicine environment, IEEE Commun. Magazine 34 (7) (1996) 36—45.

[118] H.J. Lowe, B.G. Buchanan, G.F. Cooper, J.K. Vreis, Building a medical multimedia database system to integrate clinical information: an application of high-performance computing and communications technology, Bull. Med. Library Assoc. 83 (1) (1995) 57—64.

[119] S.T. Wong, H.K. Huang, Networked multimedia for medical imaging, IEEE Multimedia Magazine April—June (1997) 24—35.

[120] C. Le Bozec, E. Zapletal, M.-C. Jaulent, D. Heudes, P. Degoulet, 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, pp. 477—481.

[121] E. El-Kwae, H. Xu, M.R. Kabuka, Content-based retrieval in picture archiving and communication systems, J. Digital Imag. 13 (2) (2000) 70—81.
[122] H. Qi, W.E. Snyder, Content-based image retrieval in PACS, J. Digital Imag. 12 (2) (1999) 81—83.
[123] J.M. Bueno, F. Chino, A.J.M. Traina, C.J. Traina, P.M. Azevedo-Marques, How to add content-based image retrieval capacity into a PACS, in: Proceedings of the IEEE Symposium on Computer-Based Medical Systems (CBMS 2002), Maribor, Slovenia, 2002, pp. 321—326.
[124] G. Bucci, S. Cagnoni, R. De Domicinis, Integrating content-based retrieval in a medical image reference database, Comput. Med. Imag. Graphics 20 (4) (1996) 231—241.
[125] C.-T. Liu, P.-L. Tai, A.Y.-J. Chen, C.-H. Peng, J.-S. Wang, A content-based medical teaching file assistant for CT lung image retrieval, in: Proceedings of the IEEE International Conference on Electronics, Circuits, Systems (ICECS 2000), Jouneih-Kaslik, Lebanon, 2000.
[126] C. Traina, Jr., Agma J.M. Traina, R.R. dos Santos, E.Y. Senzako, A support system for content-based medical image retrieval in object oriented databases, J. Med. Sys. 21(6) (1997) 339—352.
[127] S.-K. Chang, Active index for content-based medical image retrieval, Comput. Med. Imag. Graphics 20 (4) (1996) 219—229.
[128] M.O. Güld, B.B. Wein, D. Keysers, C. Thies, M. Kohnen, H. Schubert, T.M. Lehmann, A distributed architecture for content-based image retrieval in medical applications, in: Proceedings of the International Conference on Enterprise Information Systems (ICEIS 2001), Setúbal, Portugal, 2001, pp. 299—314.
[129] T.M. Lehmann, H. Schubert, D. Keysers, M. Kohnen, B.B. Wein, The irma code for unique classification of medical images, in: Proceedings of the SPIE Conference on Medical Imaging, vol. 5033, San Diego, CA, USA, 2003.
[130] H.U. Lemke, PACS developments in Europe, Comput. Med. Imag. Graphics 27 (2002) 111—120.
[131] J. Zhang, J. Sun, J.N. Stahl, PACS and web-based image distribution and display, Comput. Med. Imag. Graphics 27 (2/3) (2003) 197—206.
[132] C.-T. Liu, P.-L. Tai, A.Y.-J. Chen, C.-H. Peng, T. Lee, J.-S. Wang, 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), IEEE Computer Society, IEEE Computer Society, Tokyo, Japan, 2001, pp. 241—244.
[133] L. Zheng, A.W. Wetzel, Y. Yagi, M.J. Becich, A graphical user interface for content-based image retrieval engine that allows remote server access through the Internet, in: Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA), Nashville, TN, USA, 1998.
[134] L.H. Tang, R. Hanka, R. Lan, H.H.S. Ip, Automatic semantic labelling of medical images for content-based retrieval, in: Proceedings of the International Conference on Artificial Intelligence, Expert Systems and Applications (EXPERSYS 1998), Virginia Beach, VA, USA, 1998, pp. 77—82.
[135] S. Tzelepi, D.K. Koukopoulos, G. Pangalos, A exible content and context-based access control model for multimedia medical image database systems, in: Proceedings of the Ninth ACM International Conference on Multimedia (ACM MM 2001), The Association for Computing Machinery, Ottawa, Canada, 2001.
[136] R. Pompl, W. Bunk, A. Horsch, W. Stolz, W. Abmayr, W. Brauer, A. Glässl, G. Morfill, MELDOQ: Ein System zur Unterstützung der Früherkennung des malignen Melanoms durch digitale Bildverarbeitung, in: Proceedings of the Workshop Bildverarbeitung für die Medizin, Munich, Germany, 2000.
[137] A. Sbober, C. Eccher, E. Blanzieri, P. Bauer, M. Cristifolini, G. Zumiani, S. Forti, A multiple classifier system for early melanoma diagnosis, Artificial Intel. Med. 27 (2003) 29—44.
[138] F. Meyer, Automatic screening of cytological specimens, Comput. Vis. Graphics Image Proces. 35 (1986) 356—369.
[139] M.E. Mattie, L. Staib, E. Stratmann, H.D. Tagare, J. Duncan, P.L. Miller, PathMaster: Content-based cell image retrieval using automated feature extraction, J. Am. Med Informatics Assoc. 7 (2000) 404—415.
[140] Pathfinder: Region-based searching of Pathology Images using IRM.
[141] K. Veropoulos, C. Campbell, G. Learnmonth, Image processing and neural computing used in the diagnosis of tuberculosis, in: Proceedings of the Colloquium on Intelligent Methods in Healthcare and Medical Applications (IMHMA), York, UK, 1998.
[142] M.-C. Jaulent, C. Le Bozec, Y. Cao, E. Zapletal, P. Degoulet, A property concept frame representation for exible image content retrieval in histopathology databases, in: Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA), Los Angeles, CA, USA, 2000.
[143] L.H. Tang, R. Hanka, H.H.S. Ip, R. Lam, Extraction of semantic features of histological images for content-based retrieval of images, in: Proceedings of the IEEE Symposium on Computer-Based Medical Systems (CBMS 2000), Houston, TX, USA, 2000.
[144] H.Y. Tang, Lilian, R. Hanka, H.H.S. Ip, K.K.T. Cheung, R. Lam, Semantic query processing and annotation generation for content-based retrieval of histological images, in: International Symposium on Medical Imaging, vol. 3976, San Diego, CA, USA, 2000.
[145] G.P. Robinson, H.D. Targare, J.S. Duncan, C.C. Jaffe, Medical image collection indexing: shape-based retrieval using KD-trees, Comput. Vis. Graphics Image Proces. 20 (4) (1996) 209—217.
[146] A.S. Constantinidis, M.C. Fairhurst, A.F.R. Rahman, A new multi-expert decision combination algorithm and its application to the detection of circumscribed masses in digital mammograms, Pattern Recog. 34 (2001) 1527—1537.
[147] P. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, Z. Protopapas, Fast and effective retrieval of medical tumor shapes, IEEE Trans. Knowledge Data Eng. 10 (6) (1998) 889—904.
[148] S. Baeg, N. Kehtarnavaz, Classification of breast mass abnormalities using denseness and architectural distorsion, Electronic Lett. Comput. Vis. Image Anal. 1 (1) (2002) 1—20.
[149] F. Schnorrenberg, C.S. Pattichis, C.N. Schizas, K. Kyriacou, Content-based retrieval of breast cancer biopsy slides, Technol. Health Care 8 (2000) 291—297.
[150] D.-M. Kwak, B.-S. Kim, O.-K. Yoon, C.-H. Park, J.-U. Won, K.-H. Park, Content-based ultrasound image retrieval using a coarse to fine approach, Annals New York Acad. Sci. 980 (2002) 212—224.
[151] C. Brodley, A. Kak, C. Shyu, J. Dy, L. Broderick, A.M. Aisen, Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision, in: Proceedings of the 10th National Conference on Artificial Intelligence, Orlando, FL, USA, 1999, pp. 760—767.

[152] C.-R. Shyu, A. Kak, C. Brodley, L.S. Broderick, Testing for human perceptual categories in a physician-in-the-loop CBIR system for medical imagery, in: Proceedings of the IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL'99), Fort Collins, CO, USA, 1999, pp. 102—108.
[153] C. Schaefer-Prokop, M. Prokop, D. Fleischmann, C. Herold, High-resolution CT of diffuse interstitial lung disease: key findings in common disorders, Europe. Radiol. 11 (2001) 373—392.
[154] D.M. Hansell, High-resolution CT of diffuse lung disease, Radiol. Clin. North Am. 39 (6) (2001) 1091—1113.
[155] C.Y. Han, H. Chen, L. He, W.G. Wee, A web-based distributed image processing system, in: S. Santini, R. Schettini (Eds.), Proceedings of the SPIE Photonics West Conference on Internet Imaging IV, vol. 5018, San Jose, CA, USA, 2003, pp. 111—122.
[156] S. Sclaroff, A.P. Pentland, On modal modeling for medical images: Underconstrained shape description and data compression, in: Proceedings of the IEEE Workshop on Biomedical Image Analysis (BIA'1994), Seattle, WA, USA, 1994, pp. 70—79.
[157] Y. Liu, F. Dellaert, Classification-driven medical image retrieval, in: Proceedings of the ARPA Image Understanding Workshop, 1997.
[158] W. Cai, D.D. Feng, R. Fulton, Content-based retrieval of dynamic PET functional images, IEEE Trans. Information Technol. Biomed. 4 (2) (2000) 152—158.
[159] L.R. Long, G.R. Thoma, L.E. Berman, A prototype client/server application for biomedical text/image retrieval on the internet, in: I.K. Sethi, R.C. Jain (Eds.), Proceedings of the Conference on Storage and Retrieval for Image and Video Databases VI, vol. 3312, 1997, pp. 362—372.
[160] S. Santini, A. Gupta, The role of internet images in the biomedical informatics research network, in: S. Santini, R. Schettini (Eds.), Proceedings of the Conference on Internet Imaging IV, vol. 5018, San Jose, CA, USA, 2003, pp. 99—110.
[161] J.M. Carazo, H.K. Stelter, The BioImage database project: organizing multidemensional biological images in an object-relational database, J. Struct. Biol. 125 (1999) 97—102.
[162] Z. Geradts, H. Hardy, A. Poortmann, J. Bijhold, Evaluation of contents based image retrieval methods for a database of logos on drug tablets, in: E.M. Carapezza (Ed.), Proceedings of the Conference on Technologies for Law Enforcement, vol. 4232, Boston, MA, USA, 2000.
[163] N. Laitinen, O. Antikainen, J.-P. Mannermaa, J. Yliruusi, Content-based image retrieval: a new promising technique in powder technology, Pharma. Dev. Technol. 5 (2) (2000) 171—179.
[164] G.D. Magoulas, A. Prentza, Machine learning in medical applications, in: G. Paliouras, V. Karkaletsis, C.D. Spyrpoulos (Eds.), Machine Learning and its Applications, Lecture Notes in Computer Science, Springer-Verlag, Berlin, 2001, pp. 300—307.
[165] M.J. Egenhofer, Spatial-query-by-sketch, in: Proceedings of the IEEE Symposium on Visual Languages (VL 1996), Boulder, CO, USA, 1996, pp. 6—67.
[166] T. Ikeda, M. Hagiwara, Content-based image retrieval system using neural networks, Int. J. Neural Sys. 10 (5) (2000) 417—424.
[167] W. Hersh, M. Mailhot, C. Arnott-Smith, H. Lowe, Selective automated indexing of findings and diagnoses in radiology reports, J. Biomed. Informatics 34 (2001) 262—273.
[168] M.M. Wagner, G.F. Cooper, Evaluation of a meta-1-based automatic indexing methods for medical documents, Comput. Biomed Res. 25 (1992) 336—350.
[169] R. Chbeir, Y. Amghar, A. Flory, MIMS: A prototype for medical image retrieval, in: Recherche d'Informations Assistée par Ordinateur (RIAO'2000) Computer-Assisted Information Retrieval, vol. 1, Paris, France, 2000.
[170] H. Müller, A. Rosset, J.-P. Vallée, A. Geissbuhler, 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, pp. 480—485.
[171] T. Würinger, J. Stockhausen, D. Meyer-Ebrecht, A. Böcking, Automatic coregistration, segmentation and classification for multimodal cytopathology, in: Proceedings of the Medical Informatics Europe Conference (MIE 2003), St. Malo, France, 2003.
[172] A.W. Toga, P. Thompson, An introduction to brain warping, in: Brain Warping, Academic Press, New York, 1998, pp. 1—28 (Chapter 1).
[173] E.G.M. Patrakis, C. Faloutsos, Similarity searching in medical image databases, IEEE Trans. Knowledge Data Eng. 9 (3) (1997) 435—447.
[174] J.I. Khan, D.Y.Y. Yun, Holographic image archive, Comput. Med. Imag. Graphics 20 (4) (1996) 243—257.
[175] L. Yu, Kaia nd Ji, X. Zhang, Kernel nearest-neighbor algorithm, Neural Proces. Lett. 15(2) (2002) 147—156.
[176] T. Saracevis, Relevance: a review of and a framework for the thinking on the notion in information science, J. Am. Soc. Inform. Sci., November—December (1975) 321—343.
[177] L. Schamber, M.B. Eisenberg, M.S. Nilan, A re-examination of relevance: toward a dynamic, situational definition, Inform. Proces. Manage. 26 (6) (1990) 755—775.
[178] T. Bürkle, E. Ammenwerth, H.-U. Prokosch, J. Dudeck, Evaluation of clinical information systems. what can be evaluated and what cannot, J. Evaluat. Clin. Practice 7 (4) (2001) 373—385.
[179] H. Müller, W. Müller, D.M. Squire, S. Marchand-Maillet, T. Pun, Performance evaluation in content-based image retrieval: Overview and proposals, Pattern Recog. Lett. 22 (5) (2001) 593—601.
[180] G. Salton, The evaluation of computer-based information retrieval systems, in: Proceedings of the 1965 Congress International Federation for Documentation (IFD 1965), Spartan Books Washington, Washington DC, USA, 1965, pp. 125—133.
[181] J. Nielsen, Usability Engineering, Academic Press, Boston, MA, 1993.
[182] T.M. Lehmann, B.B. Wein, H. Greenspan, Integration of content-based image retrieval to picture archiving and communication systems, in: Proceedings of the Medical Informatics Europe Conference (MIE 2003), St. Malo, France, 2003.
[183] P. Franz, A. Zaiss, U. Schulz, Stefana nd Hahn, R. Klar, Automated coding of diagnoses—three methods compared, in: Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA), Los Angeles, CA, USA, 2000.
[184] A. Geissbuhler, C. Lovis, A. Lamb, S. Spahni, Experience with an XML/HTTP-based federative approach to develop a hospital-wide clinical information system, in: R. Rogers, R. Haux, V. Patel (Eds.), Proceedings of the International Medical Informatics Conference (Medinfo 2001), London, UK, 2001, pp. 735—739.
[185] J.R. Smith, VideoZoom: A spatial-temporal video browser for the internet, in: C.-C.J. Kuo, S.-F. Chang, S. Panchanathan (Eds.), Proceedings of the Conference on Mul-

timedia Storage and Archiving Systems III (VV02), vol. 3527, Boston, MA, USA, 1998, pp. 212—222.

[186] D. Schonfeld, D. Lelescu, VORTEX: Video retrieval and tracking from compressed multimedia databases—template matching from MPEG2 video compression standard, in: C.-C.J. Kuo, S.-F. Chang, S. Panchanathan (Eds.), Proceedings of the Conference on Multimedia Storage and Archiving Systems III (VV02), vol. 3527, Boston, MA, USA, 1998, pp. 233—244.

[187] Anil K. Jain, Robert P.W. Dvin, Jianchang Mao, Statistical Pattern Recognition: A Review, IEEE Transaction on Pattern Analysis and Machin Intelligence, vol. 22 (1), pp. 4—37, 2000.

J Digit Imaging (2013) 26:1025–1039
DOI 10.1007/s10278-013-9619-2

Content-Based Medical Image Retrieval: A Survey of Applications to Multidimensional and Multimodality Data

Ashnil Kumar · Jinman Kim · Weidong Cai · Michael Fulham · Dagan Feng

Published online: 12 July 2013

Abstract Medical imaging is fundamental to modern healthcare, and its widespread use has resulted in the creation of image databases, as well as picture archiving and communication systems. These repositories now contain images from a diverse range of modalities, multidimensional (three-dimensional or time-varying) images, as well as co-aligned multimodality images. These image collections offer the opportunity for evidence-based diagnosis, teaching, and research; for these applications, there is a requirement for appropriate methods to search the collections for images that have characteristics similar to the case(s) of interest. Content-based image retrieval (CBIR) is an image search technique that complements the conventional text-based retrieval of images by using visual features, such as color, texture, and shape, as search criteria. Medical CBIR is an established field of study that is beginning to realize promise when applied to multidimensional and multimodality medical data. In this paper, we present a review of state-of-the-art medical CBIR approaches in five main categories: two-dimensional image retrieval, retrieval of images with three or more dimensions, the use of nonimage data to enhance the retrieval, multimodality image retrieval, and retrieval from diverse datasets. We use these categories as a framework for discussing the state of the art, focusing on the characteristics and modalities of the information used during medical image retrieval.

A. Kumar (✉) · J. Kim · W. Cai · M. Fulham · D. Feng
Biomedical and Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, University of Sydney, Building J12, Sydney, NSW 2006, Australia
e-mail: ashnil@it.usyd.edu.au

M. Fulham
Department of Molecular Imaging, Royal Prince Alfred Hospital, Sydney, Australia

M. Fulham
Sydney Medical School, University of Sydney, Sydney, Australia

D. Feng
Med-X Research Institute, Shanghai Jiao Tong University, Shanghai, China

Keywords Content-based image retrieval · Medical images · Multimodality data · Multidimensional data · Review

Introduction

Imaging is a fundamental component of modern medicine and is used widely for diagnosis [1], treatment planning [2], and assessing response to treatment [3]. The question of image similarity has important applications in the medical domain because diagnostic decision-making has traditionally involved using evidence from a patient's data (image and nonimage) coupled with the physician's prior experiences of similar cases [4]. A recent study [5] has shown that clinical staff selected these similar cases primarily based upon visual properties. It has been suggested that the reliance on imaging for various clinical workflows means that access to relevant stored data will allow for more informed and effective treatment [6].

Digitization and the development of picture archiving and communication systems (PACS) [7] have enabled the storage of medical images in large digital repositories, which can be accessed by clinical staff over a network. PACS allows physicians to consider a patient's image history by allowing them to find all images related to a particular patient

Large PACS repositories also provide new opportunities for image-based diagnosis, teaching, and research based on interpatient comparisons [8–11]. This requires searching the repository for images that have similar characteristics to the image of the patient under consideration. However, the search capabilities provided by PACS are based on textual keywords, including patient name, identifiers, and image device. Text descriptions limit the search capabilities of PACS and mean that users must read through clinical reports or already know the keywords of the images to be retrieved [12, 13]. While a text-based PACS search is useful when clinical staff already know the identifiers and characteristics of the images they wish to find, the search is limited for interpatient comparative studies because it does not consider the visual properties of the images in the repository. Further, the massive volume of imaging data stored in modern clinical environments means that PACS image retrieval is not viable on the basis of manually assigned labels, e.g., clinical keywords and annotated regions. An example of the problem is given by the volume of images acquired by the Radiology Department at the University Hospital of Geneva [10].

Modern hospitals acquire a diverse ranging of imaging data. Higher-resolution devices allow physicians to detect small lesions, such as small tumors and fractures [14]. Other devices produce multidimensional images (three or more dimensions) that provide additional three-dimensional (3D) spatial or temporal information. It is also common to use different imaging modalities to provide complementary information about a particular patient. The first multimodality imaging technique to be routinely used in clinical environments was combined positron emission tomography and computed tomography (PET-CT), which enables improved cancer diagnosis, localization, and staging compared to its single modality counterparts [15]. Image search using existing PACS techniques is unfeasible due to the high amount of information encoded by these modern medical images; manual annotation is impractical, not to mention uneconomical. Furthermore, manual annotation is a subjective task with a high dependence on the skill, training, experience, and alertness of the expert performing the annotation [16].

Content-based image retrieval (CBIR) is an image search technique that does not rely upon manually assigned annotations. Instead, CBIR uses quantifiable (objectively calculated) features as the search criteria [16]. These features can be automatically or semiautomatically extracted directly from the images, thereby eliminating uneconomical and subjective manual labeling. In this paper, we review CBIR developments that have enabled medical image access for clinical applications. There are detailed, previous reviews in this field [8, 9, 17–19] but they have mainly catalogued the different methods (image features and algorithms) that were applied for medical CBIR. Our review takes a different approach. We describe CBIR methods based on clinical imaging data that are modern, multidimensional, and acquired from multimodality devices.

Our approach is as follows. We have surveyed different applications and approaches to medical CBIR and classified these into five groups: (1) two-dimensional (2D) image retrieval, (2) retrieval of images with three or more dimensions, (3) the use of nonimage data to enhance the retrieval, (4) retrieval from diverse datasets, and (5) the retrieval of multiple images (patient cases and multimodality images). We use these groups as a framework for discussing the state of the art, focusing on the characteristics and modalities of the information used during medical image retrieval.

An Overview of Content-Based Image Retrieval

CBIR is an image search technique designed to find images that are most similar to a given query. It complements text-based retrieval by using quantifiable and objective image features as the search criteria [16]. Essentially, CBIR measures the similarity of two images based on the similarity of the properties of their visual components, which can include the color, texture, shape, and spatial arrangement of regions of interest (ROIs). The nonreliance of CBIR on labels makes it ideal for large repositories where it is not feasible to manually assign keywords and other annotations. The objective features used by CBIR mean that it is also possible to show *what* images are similar and to explain *why* they are similar in an objective, nonqualitative manner. The *what* is essentially the set of retrieved images; the *why* is the difference in specific image features between the query and the retrieved results.

The major challenges for CBIR include the application-specific definition of similarity (based on users' criterion), extraction of image features that are relevant to this definition of similarity, and organizing these features into indices for fast retrieval from large repositories [16, 20–22]. The choice of features is a critical task when designing a CBIR system because it is closely related to the definition of similarity. Features fall into several categories. General purpose features can be extracted from almost all images but are not necessarily appropriate for all applications, e.g., color is inappropriate for grayscale ultrasound images. Application-specific features are tuned to a particular problem and describe characteristics unique to a particular problem domain; they are semantic features intended to encode a specific meaning [16]. Global features capture the overall characteristics of an image but fail to identify important visual characteristics if these characteristics occur in only a relatively small part of an image. Local features describe the characteristics of a small set of pixels (possibly even one pixel), i.e., they represent the details. In recent years, there has been a shift towards using local features largely driven by the belief that most images are too complex to be described in a general manner; however, the combination of local and global features remains an area of investigation for practical computer vision applications [22].

An underlying assumption of most CBIR systems is that the chosen image features used are sufficient to describe the image accurately. The choice of image features must, therefore, be made to minimize two major limitations: the *sensory gap* and the *semantic gap* [16]. The sensory gap is the difference between the object in the world and the features derived from the image. It arises when an image is noisy, has low illumination, or includes objects that are partially occluded by other objects. The sensory gap is further compounded when 2D images of physical 3D objects are considered; some information is lost as the choice of viewpoint means an object may occlude part of itself. The semantic gap is the conflict between the intent of the user and the images retrieved by the algorithm. It occurs because CBIR systems are unable to interpret images; they do not understand the "meaning" in the images in the same way that a human does. Retrieval is performed on the basis of image features not image interpretations.

The similarity of image features can be measured in a number of ways. When the features are represented as a vector, distance metrics such as the Euclidean distance can be used. The notion of elastic deformation can be used to define similarity when subtle geometric differences between images are important. Graph matching enables the comparison of images based upon a combination of image features and the arrangement of objects in the images (or the relationships between them). Finally, statistical classifiers can be trained to categorize the query image into known classes. Classifier-based approaches constitute an attempt to overcome the semantic gap through training a similarity measure on known labeled data. A detailed discussion of various similarity measures can be found in [19].

The large volume of modern image repositories and high feature dimensionality of images has also contributed to challenges in efficient real-time retrieval. In many cases, it is no longer viable to compare a query to every element of the dataset. Efficient indexing schemes are necessary to store and partition the dataset so the data can be accessed and traversed quickly, without needing to visit or process irrelevant data. Alternatively, the search space can be pruned by using only a subset of the features or applying weights to features [22]. The large datasets also mean that exact search paradigms, which look for images in the dataset that exactly satisfy all query criterion, may no longer be viable. This has led to the rise of approximate search schemes, which rank the images in the dataset according to how well they satisfy each search criterion [16]. Perhaps the most well-known approximate scheme is k-nearest neighbor search, which retrieves the k most similar (highly ranked) images as measured by distance from the query in the feature space.

It is possible that some images retrieved by approximate search paradigms will fail to meet the expectations of the users. Precision and recall are two quality measures defined to calculate the accuracy of an approximate search paradigm. Precision refers to the proportion of retrieved images that are relevant, i.e., the proportion of all retrieved images that the user was expecting. Recall is the proportion of all relevant images that were retrieved, i.e., the proportion of similar images in the dataset that were actually retrieved. The ideal case would be a retrieval system that achieves 100 % precision and 100 % recall. The reality is that most current algorithms fail to find all similar images, and many of the retrieved images contain dissimilar images (false positives).

Figure 1 shows a generic CBIR framework, which can be adapted for specific applications. The dashed arrows indicate the offline process that constructs the search index, while the solid arrows indicate the online query process. The dashed line divides the offline and online processes. During the offline processes, features are extracted from each of the images from the dataset. These features are then indexed for

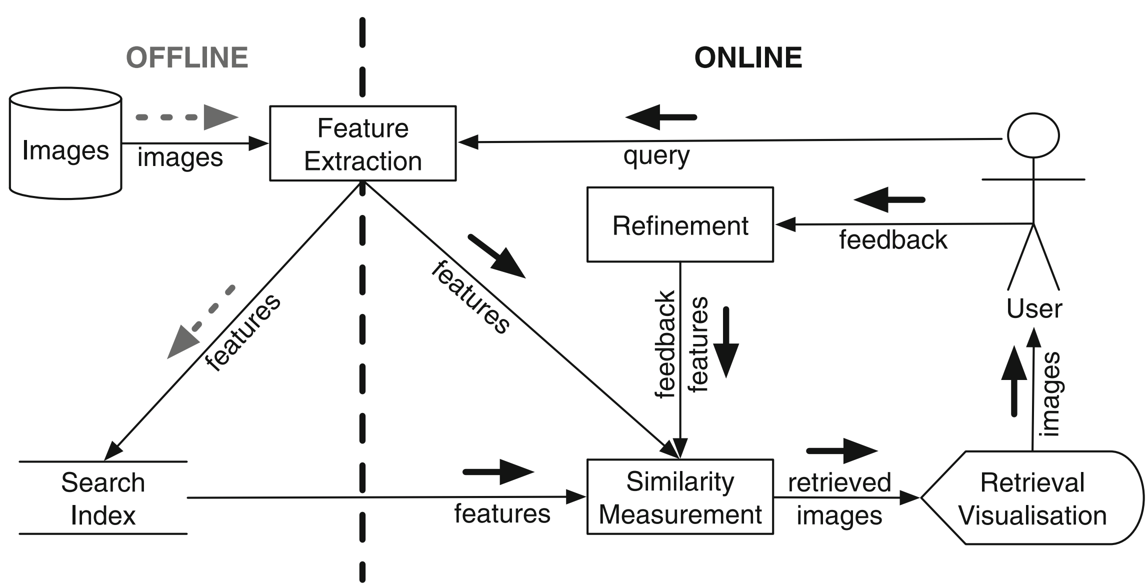

Fig. 1 A generic CBIR framework. The *dashed arrows* show the offline creation of the feature index from the image repository. The *solid arrows* show the online query process. The *dashed line* divides the offline and online processes. Note that feature extraction participates in both the offline and online processes

searching. During online processing, the same feature extraction process is performed on the query image. The query image's features are then compared to the features of indexed images using a defined similarity measurement algorithm. The measurements can then be used to rank the images in order of similarity or can be used to classify the images as "similar" or "not similar." This ranking is then displayed to the user. In many cases, the user can provide feedback in the form of weights or similarity indication to further refine the search results. The feedback and retrieval process is repeated until the user is satisfied with the retrieved results. The papers [16] and [20–22] in the reference list provide detailed overviews of general CBIR frameworks and components.

Early examples of CBIR use include IBM's *Query By Image Content* (QBIC)[1] system [23], which was used to search for famous artworks; others include the Virage framework [24] and Photobook [25]. More recently, Google Search by Image[2] used the points, colors, lines, and textures in images uploaded by users to find similar images [26]. These recent developments mean that CBIR is a technology that is available to the masses.

In recent years, a paradigm shift has changed the focus of CBIR research towards application-oriented, domain-specific technologies that would have greater impact on daily life [22]. Due to advances in acquisition technologies, ongoing CBIR research has moved towards images with more dimensions, with an aim towards increasing image understanding. Modern medical imaging is one such domain, where the retrieval of multidimensional and multimodal images from repositories of diverse data has potential applications in diagnosis, training, and research [8]. The content of medical images is complex: there is a high variability in the detail of anatomical structures across patients; misalignment of structures can occur in volumetric and multimodality images; some imaging modalities suffer from low signal-to-noise ratios; and occlusion of structures is a common occurrence. In addition, there can be large variability among patients with the same health condition [27]. It is essential that the characteristics of particular medical images are taken into account when designing CBIR systems for them. The following section presents a summary of the state of the art in medical CBIR.

Content-Based Image Retrieval in Medicine

PACS and other hospital information systems store a large variety of information, ranging from patient demographics and clinical measurements (age, weight, and blood pressure) to free text reports, test results, and images. The image types include 2D modalities, such as images of cell pathologies and plain X-rays, and volumetric images including CT, PET, and magnetic resonance (MR). Recent advances have introduced multimodality devices, e.g., PET-CT [28, 29] and PET-MR [30] scanners, which are capable of acquiring two co-aligned modalities during the same imaging session. Figure 2 shows a subset of the different types of medical images.

Several studies have already reported on the potential clinical benefits of CBIR in clinical applications. The ASSERT CBIR system used for high-resolution CT (HRCT) lung images [31] showed an improvement in the accuracy of the diagnosis made by physicians [32]. Another study for liver CT concluded that CBIR could provide real-time decision support [33]. CBIR was also shown to have benefits when used as part of a radiology teaching system [34].

In the following section, we begin our review by presenting a summary of CBIR research for 2D medical images and examine how these technologies have evolved and been applied to images with higher dimensions, e.g., volumetric CT scans, and images with a temporal dimension, e.g., dynamic PET. The integration of image with nonimage data will then be presented. We will also examine how studies have dealt with the challenge of retrieving images from datasets containing images from a diverse range of modalities. Finally, we will discuss how multiple images from different modalities have enhanced medical CBIR capabilities. Table 1 provides a brief summary of the studies that we will examine in this review and the types of data used during retrieval. Readers should refer to the relevant article for further details, e.g., figures showing the retrieval outcomes.

2D Image Retrieval

The majority of CBIR research on 2D medical images has focused on radiographic images, such as plain X-rays and mammograms. Our focus in this section is on techniques that mainly use traditional features, e.g., shape and texture. These techniques are representative of how standard techniques in nonmedical CBIR [16] have been adapted to the medical domain.

The Image Retrieval in Medical Applications (IRMA)[3] project has been a sustained effort in the CBIR of X-ray images for medical diagnosis systems. The IRMA approach is divided into seven interdependent steps [35]: (1) categorization based on global features, (2) registration using geometry and contrast, (3) local feature extraction, (4) category-dependent and query-dependent feature selection, (5) multiscale indexing, (6) identification of semantic knowledge, and (7) retrieval on the basis of the previous steps. The IRMA method classifies images into anatomical areas, modalities,

[1] QBIC: http://wwwqbic.almaden.ibm.com/.
[2] Click the camera icon in the search bar on http://images.google.com/.
[3] IRMA Homepage (English): http://www.irma-project.org/index_en.php.

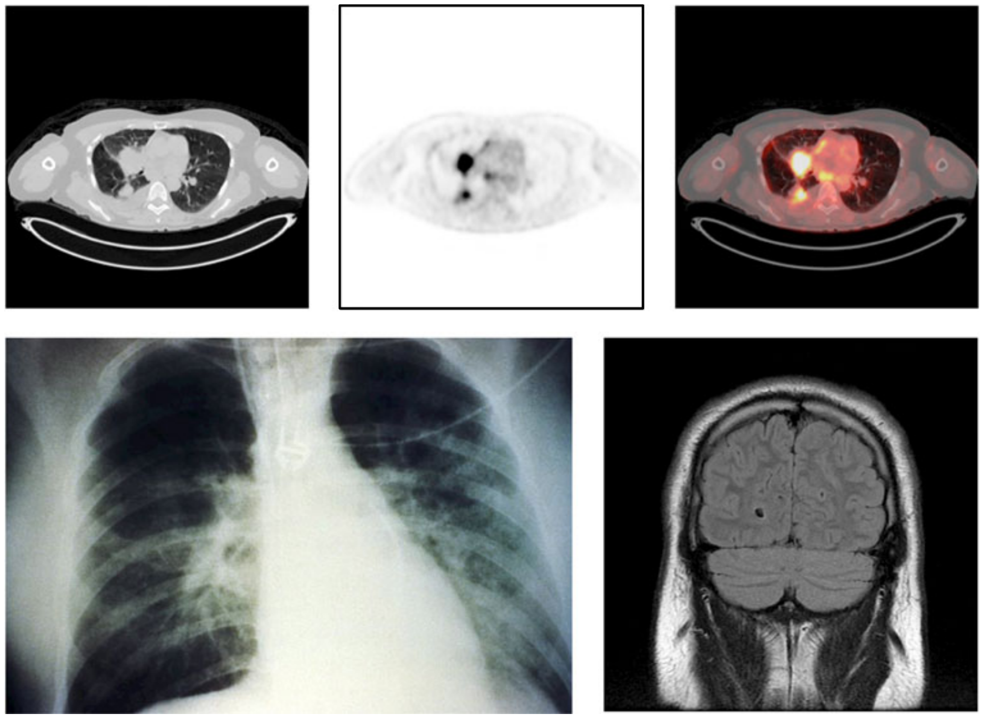

Fig. 2 A subset of the medical images available in many hospitals. *Clockwise from the top left*, they are axial CT slice, axial PET slice, axial fused PET-CT slice, coronal MR slice, and chest X-ray

and viewpoints and provides a generic framework [36] that allows the derivation of flexible implementations that are optimized for specific applications.

Other approaches for radiograph retrieval have tried to group features into semantically meaningful patterns. In one such study [37], multiscale statistical features were extracted from images by a 2D discrete wavelet transform. These features were then clustered into small patterns; images were represented as complex patterns consisting of sets of these smaller patterns. Experimental results revealed that the method had significantly higher precision and recall compared to two conventional approaches: local and global gray-level histograms.

A number of papers [38–44] have described investigations into every component of CBIR for spine X-ray retrieval, including feature extraction [39, 40, 43], indexing [44], similarity measurement [41, 44], and visualization and refinement [42]. The initial methods of matching whole vertebrae shapes [39, 40] had a major drawback: in 2D X-rays, regions of the vertebrae that were not of pathologic interest could obscure differences between critical regions. Xu et al. [41] proposed partial shape matching as a way to deal with occlusion when comparing incomplete or distorted shapes. An application-specific feature, the nine-point landmark model used by radiologists and bone morphometrists in marking pathologies, was localized to improve the computational performance of their algorithm for partial shape matching. In experiments, their method achieved a precision >85 %. While the users could apply weights to angles, lengths, and the cost to merge points on the model, it was difficult to determine the effect these weights had on the retrieval results, i.e., there was no feedback in regards to what each weight did to the shape.

This was resolved in a later study by Hsu et al. [42]; a web-based spine X-ray retrieval system allowed a user to alter the appearance of a shape and to assign weights to points on the shape to emphasize their importance. The integration of relevance feedback further improved the performance of the algorithm. Originally, 68 % of the retrieved images were relevant (what the user expected); three iterations of feedback increased this by a further 22 %. Assigning weights to parts of the shape allowed the user to specify *why* the images were similar. Furthermore, the web-based shape retrieval algorithm was shown to also work with uterine cervix images; the system was able to distinguish between three tissue types with an accuracy of 64 % [45, 46].

Table 1 Studies divided by data types

Type of data	Studies
2D images	Radiographs: [35–37]; spine X-rays: [38–44]; cervicographs: [45, 46]; mammograms: [47–49], [50, 51][a]; retinopathy: [49], [50, 51][a]
3D+ images	CT: [31, 32, 52], [33][a]; MRI: [53–55]; dynamic PET: [56, 57][a]; PET-CT: [58–69]
Nonimage data	Text: [56, 57, 70–76][b], [77, 78]; annotation or ontology: [33, 79, 80][b]; others: [50, 51][b]
Multiple images	ImageCLEF: [81–85]; pathology: [86]; general [87, 88]; PET-CT: [58–69]

[a] Also used nonimage data

[b] Also used image data

The spine retrieval framework was further enhanced with the introduction of several domain-specific features: the geometric and spatial relationships between adjacent vertebrae [43]. Combining these features with a voting consensus algorithm improved retrieval accuracy by about 8 %. To improve the speed of the retrieval, Qian et al. [44] indexed the images by embedding the shapes in a Euclidean space. This index resulted in a significantly faster retrieval time of 0.29 s compared to 319.42 s. In addition, the embedded Euclidean distance measure was a very good approximation of the Procrustes distance used previously; the first 5 retrieved images were identical in both cases over 100 queries.

Korn et al. [47] proposed a tumor shape retrieval algorithm for mammography images. In particular, the study introduced application-specific features to model the "jaggedness" of the periphery of tumors; tumors were represented by a pattern spectrum consisting of shape characteristics with high discriminatory power, such as shape smoothness and area in different scales. This was done to differentiate benign and malignant masses, which are more likely to have higher fractal dimensions. Experiments on a simulated dataset revealed that the proposed application-specific approach achieved 80 % precision at 100 % recall. Their use of pruning to reduce the search space resulted in computational performance that was up to 27 times better than sequential scans of the entire dataset.

Yang et al. [48] used a boosting framework to learn a distance metric that preserved both semantic and visual similarity during medical image retrieval. Initially, sets of binary features for data representation were learned from a labeled training set. To preserve visual similarity, sets of visual pairs (pairs of similar images) were used alongside the binary features for training the distance function. The proposed approach had higher retrieval accuracy than other retrieval methods on mammograms and comparable accuracy to the best approach on the X-ray images from the medical dataset of the Cross Language Evaluation Forum's imaging track (ImageCLEF)[4]. By learning dataset-specific features and distance functions, the retrieval framework performed more consistently than other state-of-the-art approaches across different datasets.

3D+ Image Retrieval

In recent years, many retrieval algorithms have been adapted for use in 3D medical image retrieval. A common approach is to transform a 3D image retrieval algorithm into a different problem. One such example is to select *key slices* from the volume to reduce a complex 3D retrieval to a 2D image retrieval problem. Other techniques involve representing 3D features in domains where the dimensionality of the image is not a factor, e.g., graph representations. This section described how such techniques have been adapted for images with more than two dimensions.

The most well-known example of 3D image retrieval is perhaps the ASSERT system [31], which retrieved volumetric HRCT images on the basis of key slices selected from the volumes. The system retrieved images with the same type of lung pathology (e.g., emphysema, cysts, metastases, etc.), preferably within the same lung lobe as the query. During the query process, a physician would mark a pathology-bearing region in the HRCT lung slice; gray-level texture features, as well as other statistics, were then extracted from these regions. Relational information about the lung lobes was also captured. In experiments, the ASSERT system achieved a retrieval precision of 76.3 % when matching the type of disease; this dropped to 47.3 % when the lobular location of the pathology was also considered. During clinical evaluation [32], physicians used the ASSERT system to retrieve and display four diagnosed cases that were similar to an unknown case; this was shown to improve the accuracy of their diagnosis.

An improvement to the ASSERT system involved a two-stage unsupervised feature selection method to "customize" the query [52]. During the first stage, the features that best discriminated different classes of images were used to classify the query into the most appropriate pathology class. In the second stage, the features that best discriminated between images within a class were used to identify the "subclass" of the query, i.e., to find the most similar images within the class. The customized query approach had an effective retrieval precision of 73.2 % compared to 38.9 % using a single vector of all the features. The study showed that finding images on the basis of class was suboptimal; there was a need to also find the most similar images within a particular class.

Local structure information in ROIs was used for the retrieval of brain MR slices [53]. Two feature sets for the representation of structural information were compared. The first, local binary patterns (LBPs), treated every local ROI equally. The other, Kanade–Lucas–Tomasi (KLT) feature points, gave greater emphasis to the more salient regions. The results revealed interesting insights about the trade-offs inherent in structure-based retrieval. LBPs were very dominant when spatial information was included, and its accuracy was consistently higher than its rivals in experiments involving pathological cases or other anomalies. The experiments also showed that accuracy was degraded when KLT points were not matched.

Petrakis [54] proposed a graph-based methodology for retrieving MR images. Each image was represented by an attributed graph; vertices represented ROIs, while edges represented relationships between ROIs. Their results showed that a similarity measure based on the concept of graph edit distance achieved the best retrieval precision, at the cost of

[4] ImageCLEF Homepage: http://www.imageclef.org/.

computational efficiency. Alajlan et al. [55] proposed a tree representation that achieved improved computational performance by only indexing relationships between ROIs that were included (completely surrounded) within other ROIs.

Dynamic PET images consist of a sequence of PET image frames acquired over time. Cai et al. [56] proposed a CBIR system that utilized the temporal features in these images. They exploited the activity of pixels or voxels across different time frames by basing their retrieval on the similarity of tissue time–activity curves (TTACs) [89]. Cai et al. [56] allowed three query input methods: textual attributes, definition of a query TTAC, and a combination of these features. Kim et al. [57] extended this retrieval to four dimensions (three spatial and one temporal) by registering 3D brain images to an anatomical atlas and defining the structures to search using the atlas' labels.

Retrieval Enhancement Using Nonimage Data

The majority of image search in clinical environments is performed using nonimage data. The wealth of nonimage information stored in hospitals (clinical reports and patient demographics) means that these data could enhance the image retrieval process. In this section, we focus on studies that present the use of nonimage data to add semantic information to image features as a means of reducing the semantic gap.

Text information is a common complement to image features in general [90], as well as medical CBIR research. Several examples of studies including nonimage data have been described [56, 57]. Textual information has also been used to complement several studies that were part of the ImageCLEF medical challenge or used the same data [70–76].

An initial approach to using text as the input query mechanism for image data together was presented by Chu et al. [77]. The spatial properties of ROIs and the relationships between them were indexed in a conceptual model consisting of two layers. The first layer abstracted individual objects from images, while the second layer modeled hierarchical, spatial, temporal, and evolutionary relations. The relationships represented the users' conceptual and semantic understanding of organs and diseases. Users constructed text queries using an SQL-like language; each query specified ROI properties, e.g., organ size, as well as relationships between ROIs. This retrieval approach was expanded in [78] with the introduction of a visual method for query construction and by the inclusion of a hierarchy for grouping related image features.

Rahman et al. [75] presented a technique that used the correlation between text and visual components to expand the query. Their comparison of text, visual, and combined approaches revealed that the text retrieval had a higher mean average precision than the purely visual method, while the combined method outperformed both text and visual features alone. This outcome was also visible in a comparison of different retrieval algorithms in [76] but could be explained by the nature of the dataset that was used. The medical images in the ImageCLEF dataset were highly annotated and this made text-based retrieval inherently easier than purely visual approaches.

A comparison of text, images, and combined text and image features was conducted by Névéol et al. [79], using a dataset that was not as well annotated. The text features were extracted from the caption of the images in the document, as well as paragraphs referring to those images. The experiments consisted of an indexing task that produced a single IRMA annotation for an image and a retrieval task that matched images to a query. The results showed that image analysis was better than text for both indexing and retrieval, though there were a few circumstances where indexing performed better with text data. The results also revealed that caption text provided more suitable information than the paragraph text. While combined image and text data seemed beneficial for indexing, the retrieval accuracy was not significantly higher than that of using images alone.

A preliminary clinical study [33] evaluated different features for the retrieval of liver lesions in CT images. In particular, the study compared texture, boundary features, and semantic descriptors. Twenty-six unique descriptors, from a set of 161 terms from the RadLex terminology [80], were manually assigned by trained radiologists to the 30 lesions in the dataset; each lesion was given between 8 and 11 descriptors. The semantic descriptors were a feature that explained *why* images were clinically similar. The similarity of a pair of lesions was defined as the inverse of a weighted sum of differences of their respective feature vectors. Evaluation identified that the semantic descriptors outperformed the other features in precision and recall. However, the highest accuracy was obtained when a combination of all the features was used for retrieval.

Quellec et al. [50] used unsupervised classification to index heterogeneous information (in the form of wavelets [49] and semantic text data) on decision trees. A committee was used to ensure that individual attributes (either text or image features) were not weighted too highly. A boosting algorithm was applied to reduce the tendency of decision trees to be biased towards larger classes. The proposed algorithm achieved an average precision at five retrieved items of about 79 % on a retinopathy dataset and of about 87 % on a mammography dataset. Without boosting, the results were lower, with 74 % for retinopathy and 84 % for mammography. The approach was robust to missing data, with a precision of about 60 % for the retinopathy data when <40 % of the attributes were available in the query images.

Similarly, in [51], wavelets were fused with contextual semantic data for case retrieval. A Bayesian network was used to estimate the probability of unknown variables, i.e., missing features. Information from all features was then used to estimate a correspondence between a query case and a reference case in the dataset, again using the conditional probabilities of a Bayesian network. An uncertainty component modeled the confidence of this correspondence. The highest precision was achieved when using all features, though the Bayesian method alone outperformed Bayesian plus confidence information on a mammography dataset. On the retinopathy dataset, the highest precision was achieved by the Bayesian plus confidence component.

Retrieval from Diverse Datasets

The diverse nature of medical imaging means that CBIR capabilities must have the capacity to differentiate between modalities when searching for images. This problem has been taken up by the medical image retrieval challenge at ImageCLEF. Participants submit retrieval algorithms that are evaluated on a large diverse medical image repository [91]. Overviews of submissions to the ImageCLEF medical imaging task can be found in [81–83]. A major focus of the works included is modality classification or annotation of regions, allowing effective retrieval on a subset of the diverse repository.

In 2006, Liu et al. [84] proposed two methods for solving this retrieval challenge. The first method used global features such as the average gray levels in blocks, the mean and variance of wavelet coefficients in blocks, spatial geometric properties (area, contour, centroid, etc.) of binary ROIs, color histograms, and band correlograms. The second method divided the image into patches and used clusters of high dimensional patterns within these patches as features. Using multiclass support vector machines (SVMs), they were able to achieve a mean average precision of about 68 % when using visual features.

Tian et al. [92] used a feature set consisting of LBPs and the MPEG-7 edge histogram to compare the effect of dimensionality reduction using principle component analysis (PCA); the classification was performed using multiclass SVMs. The accuracy of the dimensionally reduced feature set (80.5 % at 68 features) was not very different from the accuracy using all features (83.5 % at 602 features). The highest accuracy was achieved by the feature set falling between these two extremes (83.8 % at 330 features).

Rahman et al. [85] proposed a method for the automatic categorization of images by modality and prefiltering of the search space. The authors reduced the semantic gap by associating low-level global image features with high-level semantic categories using supervised and unsupervised learning via multiclass SVMs and fuzzy *c*-means clustering. The retrieval efficiency was increased by using PCA to reduce the feature dimension, while the learned categorization and filtering reduced the search space. Experiments on the ImageCLEF medical dataset showed that prefiltering resulted in higher precision and recall than executing queries on the entire dataset.

In a similar approach, the associations between features in MPEG-7 format and anatomical concepts in the University of Washington Digital Anatomist reference ontology were used to annotate new, unlabeled images [87]. The most similar images, based upon feature distance, were retrieved from the dataset on the basis of feature similarity. The semantic annotation for the unlabeled image was derived from the annotations of the similar images. Experiments on the Visible Human dataset [93] demonstrated that their retrieval and annotation framework achieved an accuracy of about 93.5 %.

Retrieval of Multiple Images and Modalities

The storage of patient histories in PACS and the emergence of multimodality imaging devices have introduced challenges for the retrieval of multiple related images. The most important challenge is using complementary information from different images to perform the retrieval. The works described in this section address this challenge by grouping images by the information they provide or by using relationships between features from different images.

A recent study [86] proposed the use of multiple query images to augment the retrieval process. These images were of the same modality: microscopic images of cells. Texture and color features were used in a two-tier retrieval approach. In the first tier, SVMs were used to classify the major disease type (similar to the approach used by [52]). The second tier was further subdivided into two levels: the first level found the most similar images, while the second tier ranked individual slides using a nearest neighbor approach for slide-level similarity. The slide-level similarity was weighted according to the distribution of the disease subtypes appearing on the slide and the frequency of that subtype across the entire dataset. The method achieved a classification accuracy of 93 and 86 % on two separate disease types.

Zhou et al. [88] presented a case-based retrieval algorithm for images with fractures. The algorithm combined multi-image queries consisting of data from different imaging modalities to search a repository of diverse images. The cases in the repository included X-ray, CT, MR, angiography, and scintigraphy images. The cases were represented by a bag of visual keywords and a local scale-invariant feature transform [94] descriptor. Retrieval was achieved by calculating the similarity of every image in the query case with every image in the dataset to find the set of most similar images (for a particular image in the query case). The list of all similar images was then reduced to a list of unique cases in the dataset. Three feature selection strategies were evaluated, and it was demonstrated that feature selection based on case offered the best performance and stability.

The studies described earlier in this section operated on multiple images or multiple modalities but were not designed to retrieve multimodality images that were acquired on a combined scanner. Devices such as the PET-CT and PET-MR scanners produce co-aligned images from two different modalities. The co-alignment of the different modalities offers opportunities for searches based on complementary features in the different modalities and spatial relationships between regions in either modality.

While clinical utilization of co-aligned PET-CT has grown rapidly [95, 96], few studies have investigated PET-CT CBIR [58–69]. Kim et al. [58] presented a PET-CT retrieval framework that enabled a user to search for images with tumors (extracted from PET) that were contained within a particular lung (extracted from CT) using overlapping pixels. The study introduced the capability to search for tumors by their location or size. Song et al. [59] presented a PET-CT retrieval method using Gabor texture features from CT lung fields and the SUV normalized PET image. Experiments showed that the method had higher precision than approaches that used traditional histograms and Haralick texture features. A scheme for matching tumors and abnormal lymph nodes by pairwise mapping across images was presented in [62]. A weight learning approach using regression for feature selection was presented in [64]. While the algorithms were restricted to thoracic images, they showed promise for adaptation to whole body images.

Kumar et al. [65] proposed a graph-based approach to PET-CT image retrieval by indexing PET-CT features on attributed relational graphs [97]; graph vertices represented organs extracted from CT and tumors extracted from PET. The graph-based methodology exploited the co-alignment of the two modalities to extract spatial relationship features [54] between tumors and organs; these were represented as graph edges. This allowed their graph representation to model tumor localization information, relative to a patient's anatomy. Retrieval was achieved by using graph matching to compare the query graph to graphs of images in the dataset. The approach was extended to volumetric ROIs instead of key slices, thereby enabling retrieval based upon 3D spatial features [66]. They also demonstrated that constraining tumors to the nearest anatomical structures by pruning the graph improved the retrieval process on simulated images [67]. Furthermore, they exploited their graph-based retrieval algorithm to explain *why* the retrieved images were similar to the query by designing user interfaces that enabled the interpretation of the retrieved 2D PET-CT key slices [68] and 3D PET-CT volumes [69].

Figure 3 shows the PET-CT graph representation proposed by Kumar et al. [65, 66]. Each graph vertex represents an anatomical structure or a tumor. The graph vertices are essentially feature vectors that characterize the properties of the regions they represent. The graph edges represent relationships between regions. Of particular interest are the intermodality relationships between tumor and organs. The representation can be expanded with the addition of new vertex and edge attributes to represent more image features and with the addition of extra vertices and edges to represent more complex images.

Summary and Future Directions

A number of approaches in the literature have been validated for different image modalities and clinical applications (breast cancer, spinal conditions, etc.). The multiplicity of 2D CBIR research has led to many 2D approaches being applied to images with higher dimensions, e.g., the representation of volumetric images through the use of key slices.

The ImageCLEF medical retrieval task has encouraged research into retrieval from diverse datasets. The CBIR technologies developed as part of the task are well positioned to tackle the challenges in clinical environments where a variety of image modalities are acquired. In particular, the ImageCLEF task has led to the development of methodologies for classifying image modalities based on features. In past years, most of the images in the ImageCLEF medical dataset were inherently 2D or 2D constructions of multidimensional data. The dataset is expanding to include volumetric, dynamic, and multimodality images to inspire further research into the retrieval of such data.

The use of nonimage features to complement image features has been widely investigated because all patients have some associated textual data, such as clinical reports and measurements. It has been demonstrated that combining visual features together with text data improves the accuracy of the search, but further research is necessary to make the contribution of this combination statistically significant [79].

In this review, we have presented the evolution of CBIR towards the retrieval of multidimensional and multimodality images. While great progress has been made, there are still several challenges to be solved. In the following subsections, we detail specific areas for future research that should be pursued to improve CBIR capabilities for multidimensional and multimodality medical image retrieval from repositories containing a diverse collection of data.

Visualization and User Interfaces

There has been limited investigation into visualization methods for CBIR systems, with most studies focusing on improving retrieval accuracy and speed. However, image retrieval tasks are often carried out for a particular purpose. In medicine, these purposes can include image-based reasoning, image-based training, or research. As such, an effective

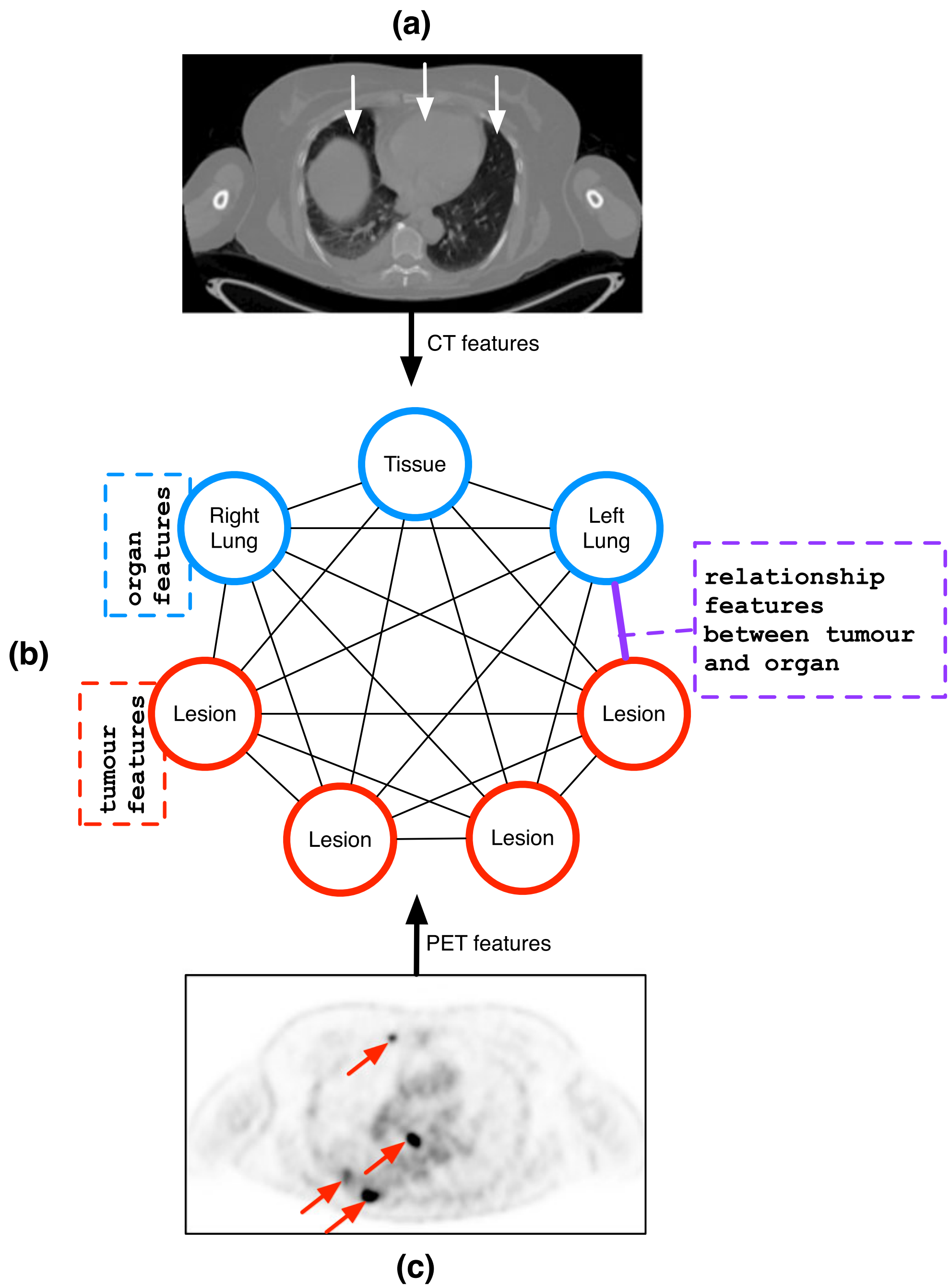

Fig. 3 The graph representation used by Kumar et al. [64, 65] for PET-CT retrieval. **a**, **c** The CT and PET images acquired by the scanner, respectively; **b** the graph representing the relationships between the ROIs, including intermodality relationships between PET tumors and CT organs

method of showing the images to the user is a critical aspect of CBIR systems.

Existing research works that address these problems are often 2D or key slice CBIR systems, such as [98] for nonmedical images. The introduction of multidimensional and multimodality data introduces new visualization challenges. CBIR systems need to have the capacity to display multiple volumes or time series (one for each retrieved image), as well as fusion information in the case of multimodality images. The systems need to optimize hardware use, especially when volume rendering is being used. In addition, Tory and Moller [99] presented a number of human factors that also need to be considered to enable the interpretation of visualized data by users. The visualization should exploit the retrieval process to demonstrate *why* the retrieved images are relevant.

The development of effective user interfaces is an area of increasing interest, especially if the CBIR systems are to be trialed in clinical environments. User interface guidelines for search applications should be followed to ensure that users are able to easily integrate the CBIR system into their clinical workflow [100]. Context-aware multimodal search interfaces, such as [101], should be pursued to give users the flexibility to overcome the sensory and semantic gaps.

Feature Selection

The curse of dimensionality has always been an issue for medical CBIR algorithms and remains relevant as algorithms are developed for modern medical images. Feature extraction and selection algorithms will need to form a core component of retrieval technologies to ensure that indexing and retrieval can be performed in an efficient manner. Methods that extract multidimensional local features from every pixel are no longer feasible for volume and types of images routinely acquired in modern hospitals.

Furthermore, the increasing clinical utilization of multimodality images offers the opportunity to derive complementary information from different modalities, the fusion of which will provide extra multidimensional features that may not be available from a single image type. Future studies should make full use of these features by defining similarity in terms of features from both modalities. In addition, useful indexing features can potentially be extracted from the relationships between ROIs in different modalities. Feature selection algorithms will need to examine the balance between features from individual modalities, as well as relationship features between modalities.

Multidimensional Image Processing

Multidimensional images are now acquired as a routine part of clinical workflows. However, despite the prevalence of volumetric images (CT, PET, MR, etc.) and time-varying images (4D CT, dynamic PET, and MR), some medical CBIR algorithms adopt key slices to represent the entire set of multidimensional image data. While this has proven effective in some scenarios, it is highly dependent on the selection of appropriate key slices; manual selection is subjective. In applications where key slices are still viable, subjective selection can be avoided by using a selection algorithm trained by unsupervised learning, as in [102]. In other cases, the use of key slices may not be possible as it may sacrifice spatial information, such as clinically relevant information (a fracture, multiple tumors, etc.) that is spread across multiple sites and slices. Multiple key slices, as in [63, 102], become less viable in cases where the disease potentially spreads throughout the body, e.g., cancer. As such, it is important that future medical CBIR studies do not rely on key slices and are optimized to operate directly on the rich multidimensional image data acquired in modern hospitals.

The direct use of multidimensional images will require the integration of image processing techniques (compression, segmentation, registration, etc.) that are designed for such images. The trend towards using local features in generic CBIR [22] indicates that the development of accurate segmentation algorithms will become critical for the development of ROI-based CBIR solutions. The efficiency of some existing algorithms will also need to be optimized for real-time operation. As an example, a recent adaptive local multi-atlas segmentation algorithm [103] requires about 30 min to segment the heart from chest CT scans with a mean accuracy of about 87 %; such processing times are not feasible for rapid data access.

Registration will be important for the retrieval of multimodality images. In particular, registration will be necessary for the extraction of relational features, segmentation tumors given anatomical priors, and fused visualization. Fortunately, hybrid multimodality PET-CT and PET-MR scanners inherently provide co-alignment information that can be used for these purposes.

Standardized Datasets for Evaluation

Most medical CBIR research is evaluated on private datasets that are collected for specific studies or purposes, e.g., retrieval of lung cancer images. These datasets are described in the studies where they are used. Such datasets have the advantage of enabling CBIR that is optimized for particular clinical applications or objectives. It also has the potential to improve outcomes by reducing the number of variables that the algorithm must consider, e.g., by having fixed image acquisition protocols, devices, resolutions, etc. Researchers can thus solve a specific problem before generalizing their algorithms for a wider array of circumstances.

However, the use of private datasets makes it difficult to compare different CBIR algorithms across different studies. To alleviate this problem, there has been a push for the creation and use of large and varied publicly available datasets with standardized gold standards or ground truth. We list several such datasets in this section.

The ImageCLEF medical image dataset [91] contained over 66,000 images between 2005 and 2007. The collection was derived from numerous sources and contained radiology, pathology, endoscopic, and nuclear medicine images. In 2013, the ImageCLEF medical image task[5] contained over 300,000 images including MR CT, PET, ultrasound, and combined modalities in one image.

The PEIR Digital Library [104][6] is a public access pathology image database for medical education. Text descriptions have been added to the images in this collection as its original purpose was for the creation of teaching materials. These text descriptions can form the ground truth from which retrieval algorithms can be evaluated.

The National Health and Nutrition Examination Surveys (NHANES)[7] were a family of surveys conducted over 30 years to monitor a number of health trends in the USA [105]. The dataset includes spine X-ray images (as used in [41]), as well

[5] ImageCLEF medical image task: http://www.imageclef.org/2013/medical.
[6] PEIR Digital Library: http://peir.path.uab.edu/.
[7] NHANES: http://www.cdc.gov/nchs/nhanes.htm.

as hand and knee X-rays. However, only a part of this dataset is publicly available.

The Cancer Imaging Archive (TCIA) [106][8] is a set of several image collections, each of which was built for a particular purpose, such as the Lung Imaging Database Consortium (LIDC) [107] of chest CT and X-rays. The images in the TCIA collection include various different image modalities, numerous subjects, and various forms of supporting data.

To enable retrieval on large collections, the VISCERAL project [108] is a new initiative where a major aim is to provide 10 TB of medical image data for research and validation. In particular, the project intends to hold challenges that exploit the knowledge stored in repositories for the development of diagnostic tools. The VISCERAL dataset will contain two annotation standards: a *gold corpus* annotated by domain experts and a *silver corpus* annotated by deriving a consensus among research systems developed by challenge participants.

Clinical Adoption

There is a dearth of clinical examples of CBIR utility despite many years of CBIR research. This is partially due to the focus of most medical CBIR research: solving technical challenges (optimizing feature selection, similarity measurement) as opposed to fulfilling a clinical goal. In addition, the majority of CBIR research is evaluated purely in nonclinical environments; collaboration between physicians and computer scientists is generally limited to sharing data [10]. Clinical evaluation of CBIR will allow the examination of the benefits and drawbacks of current algorithms and will enable greater clinical relevance in future CBIR investigations.

The use of medical literature to guide CBIR design is another avenue that requires investigation. Disease staging and classification schemes in cancer [109, 110] provide contextual information that can be used to optimize medical CBIR systems based on the guidelines used by physicians. Furthermore, the integration of medical terminology in ontologies such as RadLex [80] and the Unified Medical Language System [111] by learning correspondences between image features and text labels should also be investigated for the case of multidimensional images.

Closer communication is needed with clinical staff to ensure that medical CBIR research has outcomes that are relevant to healthcare. Clinical staff should be involved in the design of CBIR systems; medical specialists should be consulted especially if a domain-specific paradigm [22] is being adapted. An example of such research is given by Depeursinge et al. [112], who implemented three clinical workflows to assist students, radiologists, and physicians in the diagnosis of interstitial lung disease using a hybrid detection-CBIR diagnosis system. The implementation of CBIR research as integral components of the clinical workflow, as opposed to stand-alone applications, will facilitate its adoption in routine clinical practice [113].

[8] TCIA: http://cancerimagingarchive.net/.

Conclusions

In this review, we examined how state-of-the-art medical CBIR studies have been applied in the retrieval of 2D images, images with multiple dimensions, and multimodality images from repositories containing a diverse collection of medical data. We also examined the manner in which nonimage data were used to complement visual features during the retrieval process.

Even though methods have evolved from 2D image retrieval to multidimensional and multimodality image retrieval, there still remain several challenges to face. In particular, these challenges relate to retrieval visualization and interpretation, feature selection from multiple modalities, efficient image processing, and making retrieval algorithms and systems that are relevant for clinical applications. Further investigations in these areas should be pursued to produce CBIR frameworks that are practical, usable, and most importantly, have a positive impact on healthcare.

Acknowledgments We are grateful to our collaborators at the Royal Prince Alfred Hospital, Sydney, Australia for their direct and indirect contributions to this work. This work was supported in part by ARC grants.

References

1. Doi K: Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. Comput Med Imaging Graph 31(4–5):198–211, 2007
2. Zaidi H, Vees H, Wissmeyer M: Molecular PET/CT imaging-guided radiation therapy treatment planning. Acad Radiol 16(9):1108–33, 2009
3. Marcus C, Ladam-Marcus V, Cucu C, Bouché O, Lucas L, Hoeffel C: Imaging techniques to evaluate the response to treatment in oncology: Current standards and perspectives. Crit Rev Oncol/Hematol 72(3):217–38, 2009
4. Holt A, Bichindaritz I, Schmidt R, Perner P: Medical applications in case-based reasoning. Knowl Eng Rev 20(03):289–92, 2005
5. Sedghi S, Sanderson M, Clough P: How do health care professionals select medical images they need? ASLIB Proc 64(4):437–56, 2012
6. Haux R: Health information systems—Past, present, future. Int J Med Inform 75(3–4):268–81, 2006
7. Huang HK. PACS and Imaging Informatics: Basic Principles and Applications. New York: Wiley, 2004
8. Müller H, Michoux N, Bandon D, Geissbuhler A: A review of content-based image retrieval systems in medical applications—Clinical benefits and future directions. Int J Med Inform 73(1):1–23, 2004

9. Müller H, Zhou X, Depeursinge A, Pitkanen M, Iavindrasana J, Geissbuhler A: Medical visual information retrieval: State of the art and challenges ahead. In: Proceedings of the IEEE International Conference on Multimedia and Expo, Beijing, 2007, pp 683–686
10. Müller H, Kalpathy-Cramer J, Caputo B, Syeda-Mahmood T, Wang F: Overview of the first workshop on medical content-based retrieval for clinical decision support at MICCAI 2009. In: Caputo B, Müller H, Syeda-Mahmood T, Duncan J, Wang F, Kalpathy-Cramer J Eds. Medical Content-Based Retrieval for Clinical Decision Support, Vol. 5853 of Lecture Notes in Computer Science. Berlin: Springer, 2010, pp 1–17
11. Huang HK: Utilization of medical imaging informatics and biometrics technologies in healthcare delivery. Int J Comput Assist Radiol Surg 3:27–39, 2008
12. Tagare HD, Jaffe CC, Duncan J: Medical image databases: A content-based retrieval approach. J Am Med Inform Assoc 4(3):184–98, 1997
13. Lehmann TM, Guld MO, Thies C, Fischer B, Keysers D, Kohnen M, et al: Content-based image retrieval in medical applications for picture archiving and communication systems. In: Huang HK, Ratib OM Eds. Proceedings of SPIE 5033, 2003, pp 109–117
14. Brown KR, Silver I, Musgrave J, Roberts A: The use of μCT technology to identify skull fracture in a case involving blunt force trauma. Forensic Sci Int 206(1–3):8–11, 2011
15. Blodgett TM, Meltzer CC, Townsend DW: PET/CT: Form and function. Radiology 242(2):360–85, 2007
16. Smeulders A, Worring M, Santini S, Gupta A, Jain R: Content-based image retrieval at the end of the early years. IEEE Trans Pattern Anal Mach Intell 22(12):1349–80, 2000
17. Cai TW, Kim J, Feng DD: Content-based medical image retrieval. In: Feng DD Ed. Biomedical Information Technology. Burlington: Academic Press, 2008, pp 83–113
18. Long LR, Antani S, Deserno TM, Thoma GR: Content-based image retrieval in medicine: Retrospective assessment, state of the art, and future directions. Int J Healthcare Inf Syst Inform 4(1):1–16, 2009
19. Akgül C, Rubin D, Napel S, Beaulieu C, Greenspan H, Acar B: Content-based image retrieval in radiology: Current status and future directions. J Digit Imaging 24:208–22, 2011
20. Lew MS, Sebe N, Djeraba C, Jain R: Content-based multimedia information retrieval: State of the art and challenges. ACM Trans Multimed Comput Commun Appl 2(1):1–19, 2006
21. Rui Y, Huang TS, Chang SF: Image retrieval: Current techniques, promising directions, and open issues. J Vis Commun Image Represent 10(1):39–62, 1999
22. Datta R, Joshi D, Li J, Wang JZ: Image retrieval: Ideas, influences, and trends of the new age. ACM Comput Surv 40(2):5:1–5:60, 2008
23. Flickner M, Sawhney H, Niblack W, Ashley J, Huang Q, Dom B, et al: Query by image and video content: The QBIC system. Computer 28(9):23–32, 1995
24. Bach JR, Fuller C, Gupta A, Hampapur A, Horowitz B, Humphrey R, et al: Virage image search engine: an open framework for image management. In: Sethi IK, Jain RC Eds. Proceedings of SPIE 2670, 1, 1996, pp 76–87
25. Pentland A, Picard RW, Sclaroff S: Photobook: Content-based manipulation of image databases. Int J Comput Vis 18:233–54, 1996
26. Chechik G, Sharma V, Shalit U, Bengio S: Large scale online learning of image similarity through ranking. J Mach Learn Res 11:1109–35, 2010
27. Duncan JS, Ayache N: Medical image analysis: Progress over two decades and the challenges ahead. IEEE Trans Pattern Anal Mach Intell 22(1):85–106, 2000
28. Townsend DW, Beyer T: A combined PET/CT scanner: The path to true image fusion. Br J Radiol 75(Supplement 9):S24–30, 2002
29. Townsend DW, Beyer T, Blodgett TM: PET/CT scanners: A hardware approach to image fusion. Semin Nucl Med 33(3):193–204, 2003
30. Judenhofer MS, Catana C, Swann BK, Siegel SB, Jung WI, Nutt RE, et al: PET/MR images acquired with a compact MR-compatible PET detector in a 7-T magnet. Radiology 244(3):807–14, 2007
31. Shyu CR, Brodley CE, Kak AC, Kosaka A, Aisen AM, Broderick LS: ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases. Comp Vision Image Underst 75(1–2):111–32, 1999
32. Aisen AM, Broderick LS, Winer-Muram H, Brodley CE, Kak AC, Pavlopoulou C, et al: Automated storage and retrieval of thin-section CT images to assist diagnosis: System description and preliminary assessment. Radiology 228(1):265–70, 2003
33. Napel SA, Beaulieu CF, Rodriguez C, Cui J, Xu J, Gupta A, et al: Automated retrieval of CT images of liver lesions on the basis of image similarity: Method and preliminary results. Radiology 256(1):243–52, 2010
34. Müller H, Rosset A, Garcia A, Vallée JP, Geissbuhler A: Benefits of content-based visual data access in radiology. Radiographics 25(3):849–58, 2005
35. Keysers D, Dahmen J, Ney H, Wein BB, Lehmann TM: Statistical framework for model-based image retrieval in medical applications. J Electron Imaging 12(1):59–68, 2003
36. Güld MO, Thies C, Fischer B, Lehmann TM: A generic concept for the implementation of medical image retrieval systems. Int J Med Inform 76(2–3):252–9, 2007
37. Iakovidis D, Pelekis N, Kotsifakos E, Kopanakis I, Karanikas H, Theodoridis Y: A pattern similarity scheme for medical image retrieval. IEEE Trans Inf Technol Biomed 13(4):442–50, 2009
38. Antani S, Lee D, Long LR, Thoma GR: Evaluation of shape similarity measurement methods for spine X-ray images. J Vis Commun Image Represent 15(3):285–302, 2004
39. Antani S, Long LR, Thoma GR, Lee DJ: Evaluation of shape indexing methods for content-based retrieval of X-ray images. In: Yeung MM, Lienhart RW, Li CS Eds. Proceedings of SPIE 5021, 2003, pp 405–416
40. Lee DJ, Antani S, Long LR: Similarity measurement using polygon curve representation and Fourier descriptors for shape-based vertebral image retrieval. In: Sonka M, Fitzpatrick JM Eds. Proceedings of SPIE 5032, 2003, pp 1283–1291
41. Xu X, Lee DJ, Antani S, Long L: A spine X-ray image retrieval system using partial shape matching. IEEE Trans Inf Technol Biomed 12(1):100–8, 2008
42. Hsu W, Antani S, Long LR, Neve L, Thoma GR: SPIRS: A web-based image retrieval system for large biomedical databases. Int J Med Inform 78(Supplement 1):S13–24, 2009
43. Lee DJ, Antani S, Chang Y, Gledhill K, Long LR, Christensen P: CBIR of spine X-ray images on inter-vertebral disc space and shape profiles using feature ranking and voting consensus. Data Knowl Eng 68(12):1359–69, 2009
44. Qian X, Tagare HD, Fulbright RK, Long R, Antani S: Optimal embedding for shape indexing in medical image databases. Med Image Anal 14(3):243–54, 2010
45. Xue Z, Antani S, Long LR, Jeronimo J, Thoma GR: Investigating CBIR techniques for cervicographic images. In: Proceedings of the Annual Symposium of American Medical Information Association, 2007, pp 826–830
46. Xue Z, Antani S, Long L, Thoma G: A system for searching uterine cervix images by visual attributes. In: IEEE International Symposium on Computer-Based Medical Systems, 2009, pp 1–5
47. Korn P, Sidiropoulos N, Faloutsos C, Siegel E, Protopapas Z: Fast and effective retrieval of medical tumor shapes. IEEE Trans Knowl Data Eng 10(6):889–904, 1998
48. Yang L, Jin R, Mummert L, Sukthankar R, Goode A, Zheng B, et al: A boosting framework for visuality-preserving distance metric learning and its application to medical image retrieval. IEEE Trans Pattern Anal Mach Intell 32(1):30–44, 2010

49. Quellec G, Lamard M, Cazuguel G, Cochener B, Roux C: Wavelet optimization for content-based image retrieval in medical databases. Med Image Anal 14(2):227–41, 2010
50. Quellec G, Lamard M, Bekri L, Cazuguel G, Roux C, Cochener B: Medical case retrieval from a committee of decision trees. IEEE Trans Inf Technol Biomed 14(5):1227–35, 2010
51. Quellec G, Lamard M, Cazuguel G, Roux C, Cochener B: Case retrieval in medical databases by fusing heterogeneous information. IEEE Trans Med Imaging 30(1):108–18, 2011
52. Dy JG, Brodley CE, Kak A, Broderick LS, Aisen AM: Unsupervised feature selection applied to content-based retrieval of lung images. IEEE Trans Pattern Anal Mach Intell 25(3):373–8, 2003
53. Unay D, Ekin A, Jasinschi R: Local structure-based region-of-interest retrieval in brain MR images. IEEE Trans Inf Technol Biomed 14(4):897–903, 2010
54. Petrakis EG: Design and evaluation of spatial similarity approaches for image retrieval. Image Vis Comput 20(1):59–76, 2002
55. Alajlan N, Kamel M, Freeman G: Geometry-based image retrieval in binary image databases. IEEE Trans Pattern Anal Mach Intell 30(6):1003–13, 2008
56. Cai W, Feng D, Fulton R: Content-based retrieval of dynamic pet functional images. IEEE Trans Inf Technol Biomed 4(2):152–8, 2000
57. Kim J, Cai W, Feng D, Wu H: A new way for multidimensional medical data management: Volume of interest (VOI)-based retrieval of medical images with visual and functional features. IEEE Trans Inf Technol Biomed 10(3):598–607, 2006
58. Kim J, Constantinescu L, Cai W, Feng DD: Content-based dual-modality biomedical data retrieval using co-aligned functional and anatomical features. In: Proceedings of the MICCAI Workshop on Content-Based Image Retrieval for Biomedical Image Archives: Achievements, Problems and Prospects, 2007, pp 45–52
59. Song Y, Cai W, Eberl S, Fulham M, Feng D: A content-based image retrieval framework for multi-modality lung images. In: IEEE International Symposium on Computer-Based Medical Systems, 2010, pp 285–290
60. Song Y, Cai W, Eberl S, Fulham M, Feng D: Structure-adaptive feature extraction and representation for multi-modality lung images retrieval. In: International Conference on Digital Image Computing: Techniques and Applications, 2010, pp 152–157
61. Song Y, Cai W, Eberl S, Fulham M, Feng D: Thoracic image case retrieval with spatial and contextual information. In: 2011 I.E. International Symposium on Biomedical Imaging: From Nano to Macro, 2011, pp 1885–1888
62. Song Y, Cai W, Eberl S, Fulham M, Feng D: Thoracic image matching with appearance and spatial distribution. In: International Conference of the IEEE Engineering in Medicine and Biology Society, 2011, pp 4469–4472
63. Song Y, Cai W, Feng D: Hierarchical spatial matching for medical image retrieval. In: Proceedings of the International ACM Multimedia Workshop on Medical Multimedia Analysis and Retrieval, 2011, pp 1–6
64. Cai W, Song Y, Feng DD: Regression and classification based distance metric learning for medical image retrieval. In: IEEE International Symposium on Biomedical Imaging, 2012, pp 1775–1778
65. Kumar A, Kim J, Cai W, Eberl S, Feng D: A graph-based approach to the retrieval of dual-modality biomedical images using spatial relationships. In: International Conference of the IEEE Engineering in Medicine and Biology Society, 2008, pp 390–393
66. Kumar A, Kim J, Wen L, Feng D: A graph-based approach to the retrieval of volumetric PET-CT lung images. In: Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2012, pp 5408–5411
67. Kumar A, Kim J, Fulham M, Feng D: Graph-based retrieval of multi-modality medical images: A comparison of representations using simulated images. In: IEEE International Symposium on Computer-Based Medical Systems, 2012, pp 1–6
68. Kumar A, Haraguchi D, Kim J, Wen L, Eberl S, Fulham M, et al: A query and visualisation interface for a PET-CT image retrieval system. Int J Comput Assist Radiol Surg 6(Supplement 1):69, 2011
69. Kumar A, Kim J, Bi L, Feng D: An image retrieval interface for volumetric multi-modal medical data: Application to PET-CT content-based image retrieval. Int J Comput Assist Radiol Surg 7(Supplement 1):475–7, 2012
70. Radhouani S, Lim J, Chevallet JP, Falquet G: Combining textual and visual ontologies to solve medical multimodal queries. In: IEEE International Conference on Multimedia and Expo, 2006, pp 1853–1856
71. Lacoste C, Lim JH, Chevallet JP, Le D: Medical-image retrieval based on knowledge-assisted text and image indexing. IEEE Trans Circ Syst Video Technol 17(7):889–900, 2007
72. Gobeill J, Müller H, Ruch P: Translation by text categorisation: Medical image retrieval in ImageCLEFmed 2006. In: Peters C, Clough P, Gey F, Karlgren J, Magnini B, Oard D, et al. Eds. Evaluation of Multilingual and Multi-modal Information Retrieval, Vol. 4730 of Lecture Notes in Computer Science, 2007, pp 706–710
73. Villena-Román J, Lana-Serrano S, González-Cristóbal J: MIRACLE at ImageCLEFmed 2007: Merging textual and visual strategies to improve medical image retrieval. In: Peters C, Jijkoun V, Mandl T, Müller H, Oard D, Peñas A, et al. Eds. Advances in Multilingual and Multimodal Information Retrieval, Vol. 5152 of Lecture Notes in Computer Science, 2008, pp 593–596
74. Caicedo JC, Moreno JG, Niño EA, González FA: Combining visual features and text data for medical image retrieval using latent semantic kernels. In: Proceedings of the International Conference on Multimedia Information Retrieval, ACM, 2010, pp 359–366
75. Rahman M, Antani S, Long R, Demner-Fushman D, Thoma G: Multi-modal query expansion based on local analysis for medical image retrieval. In: Caputo B, Müller H, Syeda-Mahmood T, Duncan J, Wang F, Kalpathy-Cramer J Eds. Medical Content-Based Retrieval for Clinical Decision Support, Vol. 5853 of Lecture Notes in Computer Science, 2010, pp 110–119
76. Müller H, Kalpathy-Cramer J, Charles E. Kahn J, Hersh W: Comparing the quality of accessing medical literature using content-based visual and textual information retrieval. In: Siddiqui KM, Liu BJ Eds. Proceedings of SPIE 7264, 2009, pp 726405:1–726405:11
77. Chu WW, Ieong IT, Taira RK: A semantic modeling approach for image retrieval by content. VLDB J—Int J Very Large Data Bases 3(4):445–77, 1994
78. Chu W, Hsu CC, Cardenas A, Taira R: Knowledge-based image retrieval with spatial and temporal constructs. IEEE Trans Knowl Data Eng 10(6):872–88, 1998
79. Névéol A, Deserno TM, Darmoni SJ, Güld MO, Aronson AR: Natural language processing versus content-based image analysis for medical document retrieval. J Am Soc Inf Sci Technol 60(1):123–34, 2009
80. Langlotz CP: RadLex: A new method for indexing online educational materials. Radiographics 26(6):1595–7, 2006
81. Müller H, Deselaers T, Deserno T, Kalpathy-Cramer J, Kim E, Hersh W: Overview of the ImageCLEFmed 2007 medical retrieval and medical annotation tasks. In: Peters C, Jijkoun V, Mandl T, Müller H, Oard D, Peñas A, et al. Eds. Advances in Multilingual and Multimodal Information Retrieval, Vol. 5152 of Lecture Notes in Computer Science, 2008, pp 472–491
82. Müller H, Kalpathy-Cramer J, Kahn C, Hatt W, Bedrick S, Hersh W: Overview of the ImageCLEFmed 2008 medical image retrieval task. In: Peters C, Deselaers T, Ferro N, Gonzalo J, Jones G, Kurimo M, et al. Eds. Evaluating Systems for Multilingual and

Multimodal Information Access, Vol. 5706 of Lecture Notes in Computer Science, 2009, pp 512–522
83. Müller H, Kalpathy-Cramer J, Eggel I, Bedrick S, Radhouani S, Bakke B, et al: Overview of the CLEF 2009 medical image retrieval track. In: Peters C, Caputo B, Gonzalo J, Jones G, Kalpathy-Cramer J, Müller H, et al. Eds. Multilingual Information Access Evaluation II. Multimedia Experiments, Vol. 6242 of Lecture Notes in Computer Science, 2010, pp 72–84
84. Liu J, Hu Y, Li M, Ma S, ying Ma W: Medical image annotation and retrieval using visual features. In: Evaluation of Multilingual and Multi-modal Information Retrieval, Vol. 4730 of Lecture Notes in Computer Science, 2007, pp 678–685
85. Rahman MM, Desai BC, Bhattacharya P: Medical image retrieval with probabilistic multi-class support vector machine classifiers and adaptive similarity fusion. Comput Med Imaging Graph 32(2):95–108, 2008
86. Akakin H, Gurcan M: Content-based microscopic image retrieval system for multi-image queries. IEEE Trans Inf Technol Biomed 16(4):758–69, 2012
87. Allampalli-Nagaraj G, Bichindaritz I: Automatic semantic indexing of medical images using a web ontology language for case-based image retrieval. Eng Appl Artif Intell 22(1):18–25, 2009
88. Zhou X, Stern R, Müller H: Case-based fracture image retrieval. Int J Comput Assist Radiol Surg 7:401–11, 2012
89. Huang SC, Phelps ME, Hoffman EJ, Sideris K, Selin CJ, Kuhl DE: Noninvasive determination of local cerebral metabolic rate of glucose in man. Am J Physiol—Endocrinol Metab 238(1):E69–82, 1980
90. Chang E, Goh K, Sychay G, Wu G: CBSA: Content-based soft annotation for multimodal image retrieval using Bayes point machines. IEEE Trans Circ Syst Video Technol 13(1):26–38, 2003
91. Hersh W, Müller H, Kalpathy-Cramer J: The ImageCLEFmed medical image retrieval task test collection. J Digit Imaging 22:648–55, 2009
92. Tian G, Fu H, Feng D: Automatic medical image categorization and annotation using LBP and MPEG-7 edge histograms. In: International Conference on Information Technology and Applications in Biomedicine, 2008, pp 51–53
93. Spitzer V, Ackerman MJ, Scherzinger AL, Whitlock D: The visible human male: A technical report. J Am Med Inform Assoc 3(2):118–30, 1996
94. Lowe DG: Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60:91–110, 2004
95. Czernin J, Dahlbom M, Ratib O, Schiepers C: Atlas of PET/CT Imaging in Oncology. Springer, Berlin, 2004
96. Goerres GW, von Schulthess GK, Steinert HC: Why most PET of lung and head-and-neck cancer will be PET/CT. J Nucl Med 45(Supplement 1):66S–71S, 2004
97. Fu KS: A step towards unification of syntactic and statistical pattern recognition. IEEE Trans Pattern Anal Mach Intell 8(3):398–404, 1986
98. Jing Y, Rowley H, Rosenberg C, Wang J, Zhao M, Covell M: Google image swirl, a large-scale content-based image browsing system. In: IEEE International Conference on Multimedia and Expo, 2010, p 267
99. Tory M, Moller T: Human factors in visualization research. IEEE Trans Vis Comput Graph 10(1):72–84, 2004
100. Wilson ML: Search user interface design. Synth Lect Inf Concepts Retr Serv 3(3):1–143, 2011
101. Etzold J, Brousseau A, Grimm P, Steiner T: Context-aware querying for multimodal search engines. In: Schoeffmann K, Merialdo B, Hauptmann A, Ngo CW, Andreopoulos Y, Breiteneder C Eds. Advances in Multimedia Modeling, Vol. 7131 of Lecture Notes in Computer Science. Berlin: Springer, 2012, pp 728–739
102. Ekin A, Jasinschi R, van der Grond J, Van Buchem M: Improving information quality of MR brain images by fully automatic and robust image analysis methods. J Soc Inf Disp 15(6):367–76, 2007
103. van Rikxoort EM, Isgum I, Arzhaeva Y, Staring M, Klein S, Viergever MA, et al: Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus. Medical Image Analysis 14(1):39–49, 2010
104. Jones KN, Woode DE, Panizzi K, Anderson PG: PEIR digital library: Online resources and authoring system. In: Proceedings of the American Medical Informatics Association Symposium, 2001, p 1075
105. Long LR, Antani SK, Thoma GR: Image informatics at a national research center. Comput Med Imaging Graph 29(2–3):171–93, 2005
106. The Cancer Imaging Archive. 2011. http://cancerimagingarchive.net/
107. Armato III, SG, McLennan G, McNitt-Gray MF, Meyer CR, Yankelevitz D, Aberle DR, et al: Lung image database consortium: Developing a resource for the medical imaging research community. Radiology 232(3):739–48, 2004
108. Langs G, Müller H, Menze BH, Hanbury A: VISCERAL: Towards large data in medical imaging—Challenges and directions. In: MICCAI Workshop on Medical Content-Based Retrieval for Clinical Decision Support 2012, Vol. 7723 of Springer LNCS, 2013, pp 92–98
109. Detterbeck FC, Boffa DJ, Tanoue LT: The new lung cancer staging system. Chest 136(1):260–71, 2009
110. Edge SB, Byrd DR, Compton CC, Frtiz AG, Greene FL, Trotti A Eds. AJCC Cancer Staging Manual. New York: Springer, 2010
111. Bodenreider O: The unified medical language system (UMLS): Integrating biomedical terminology. Nucleic Acids Res 32(Supplement 1):D267–70, 2004
112. Depeursinge A, Vargas A, Gaillard F, Platon A, Geissbuhler A, Poletti PA, et al: Case-based lung image categorization and retrieval for interstitial lung diseases: Clinical workflows. Int J Comput Assist Radiol Surg 7(1):97–110, 2012
113. Antani S, Xue Z, Long LR, Bennett D, Ward S, Thoma GR: Is there a need for biomedical CBIR systems in clinical practice? Outcomes from a usability study. In: Proceedings of SPIE 7967, 2011, pp 796708-1–796708-7