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

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

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.

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

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

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
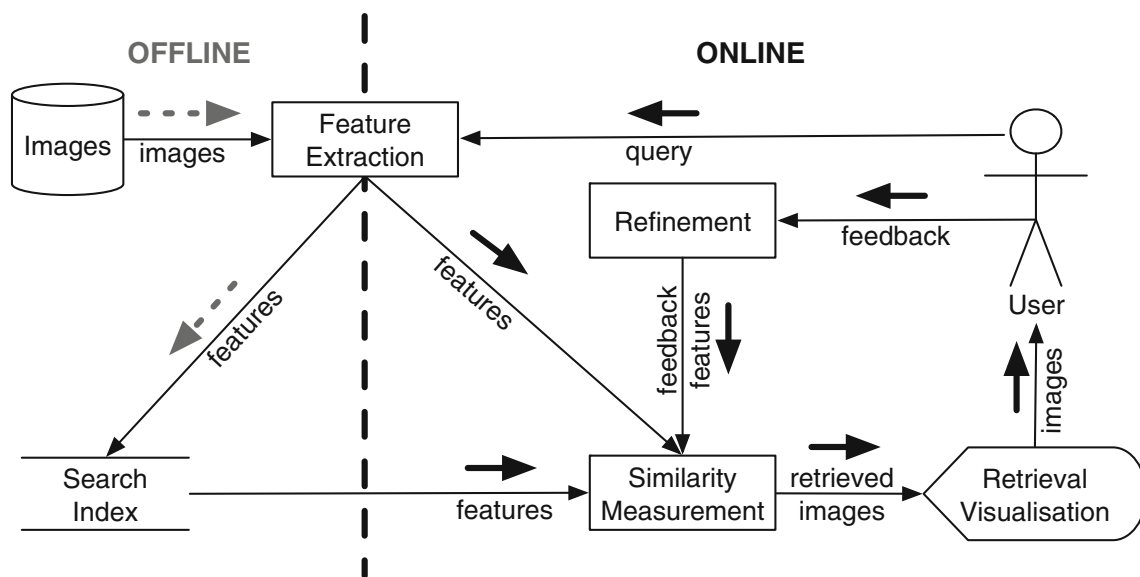

Fig. 1 A generic CBIR framework. The *dashed arrows* show the offline creation of the feature index from the image repository. The *solid arrows* show the online query process. The *dashed line* divides the offline and online processes. Note that feature extraction participates in both the offline and 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
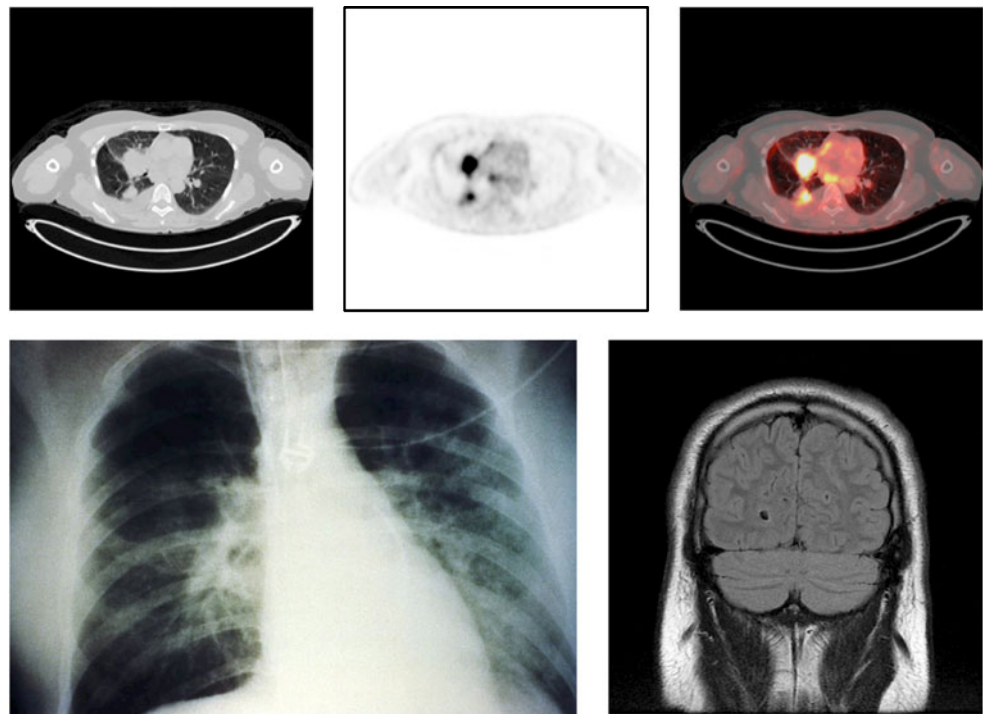
and viewpoints and provides a generic framework [36] that allows the derivation of flexible implementations that are optimized for specific applications.

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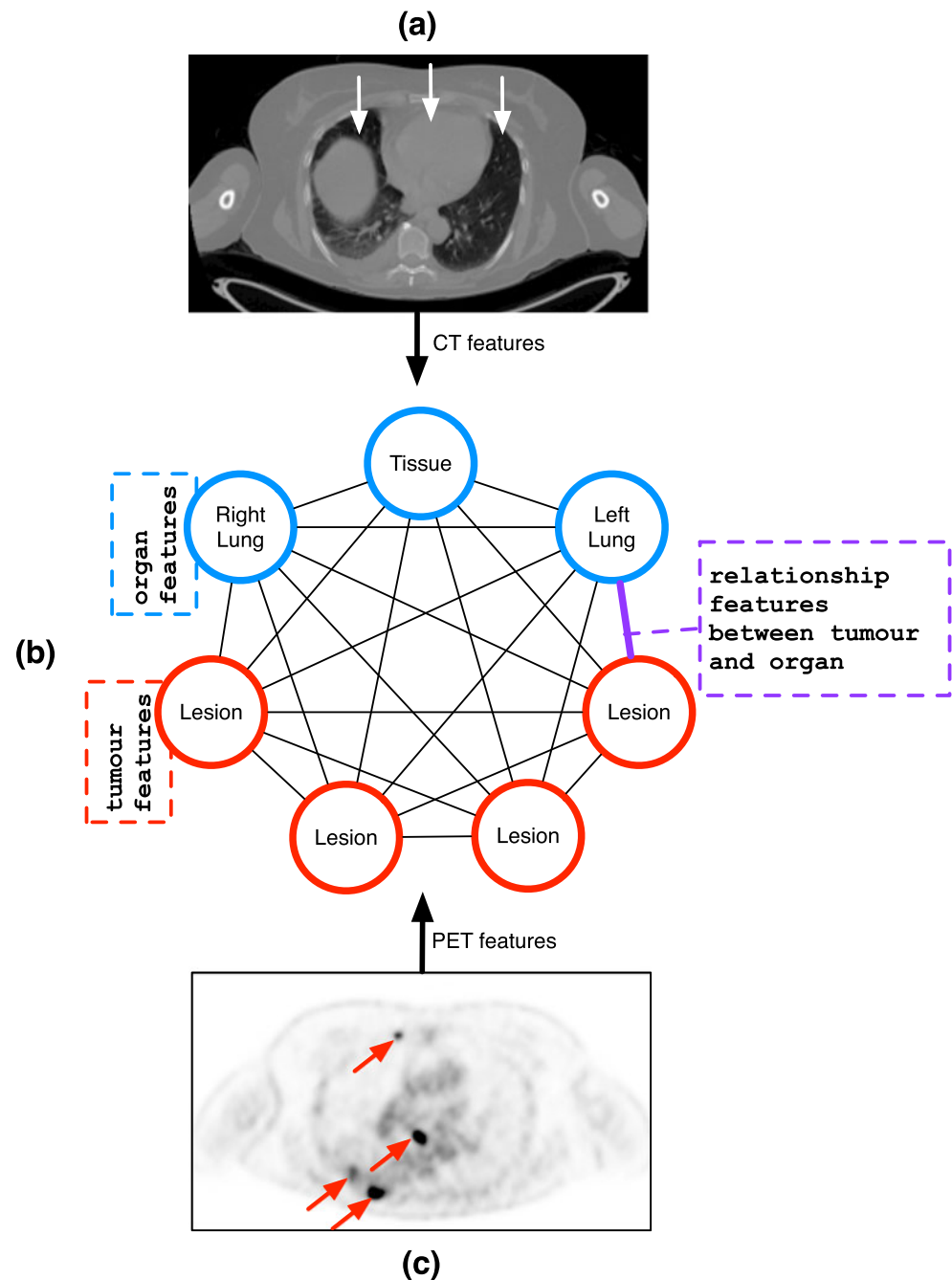
method of showing the images to the user is a critical aspect of CBIR systems.

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].

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.

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 intervertebral 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

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: MIRA-CLE 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