Original citation:
Ng, C. U. and Martin, Graham R. (2002) Content-description interfaces for medical imaging. Department of Computer Science. (Department of Computer Science Research Report). (Unpublished) CS-RR-391

Content-Description Interfaces for Medical Imaging

Chee Un Ng and Graham R. Martin
Department of Computer Science,
University of Warwick
Coventry CV4 7AL

August 29, 2002

Abstract

This report details work which has been done in the second year of the research in content-based information retrieval (CBIR) for medical images. A hierarchical data structure suitable for medical CBIR, which is a combination of XML, MPEG-7 and DICOM (where possible) has been proposed, and a prototype testbed is presented. The emphasis of the prototype system is the attribute graph, which enables the visualisation of the search results / dataset as a whole.

Contents

List of Figures

1 Introduction

In the last report[1], the potential of content-based image retrieval and content description standards, particularly MPEG-7[2], were discussed in detail. The implementation of the Curvature Scale Space (CSS)[3], an algorithm which is used to extract the contour shape of an image, was also discussed in the last report. CSS is a standard visual descriptor, which is part of the evolving MPEG-7 standard. Emphasis is placed on the development of a user friendly interface, as it will be the only contact point which the clinician has with the system.

During the past year, work has been done on the investigation of other relevant descriptors and the development of a hierarchical content description which uses a combination of XML and MPEG-7 description schemes. A query interface has also been created.

Several approaches to the problem of content-based image management have been proposed, however few are specifically designed for medical images. Medical images differ from other conventional images as they usually are obtained by special equipment, and most of them are a representation instead of a real photographic image. In this sense, medical images deserve special attention, and to develop a CBIR system specifically for this domain is highly desirable, as there are millions of medical images created every day.

Similar systems which have been developed include I^2Cnet[4], developed at Foundation for Research and Technology - Hellas (FORTH), which is a network of content-based similarity search engines which supporting attribute and text searches, in conjunction with the geometric and texture properties of selected regions. I^2Cnet is a further development of I^2C[5], which is a stand-alone search and retrieval system. Another system known as PICSearch[6] is a non-medical system, which is primarily designed to be a modular environment for content-based pictorial retrieval algorithms. There are also several popular CBIR systems, which were already discussed in the last report.

In the current research work, a hierarchical data structure has been proposed and will be discussed in the following section. Introduction to the standard which the data structure utilises will also be discussed. The current prototype testbed will be presented, with an emphasis on the search and retrieval system and also the relationship between the search results and the proposed attribute graph.

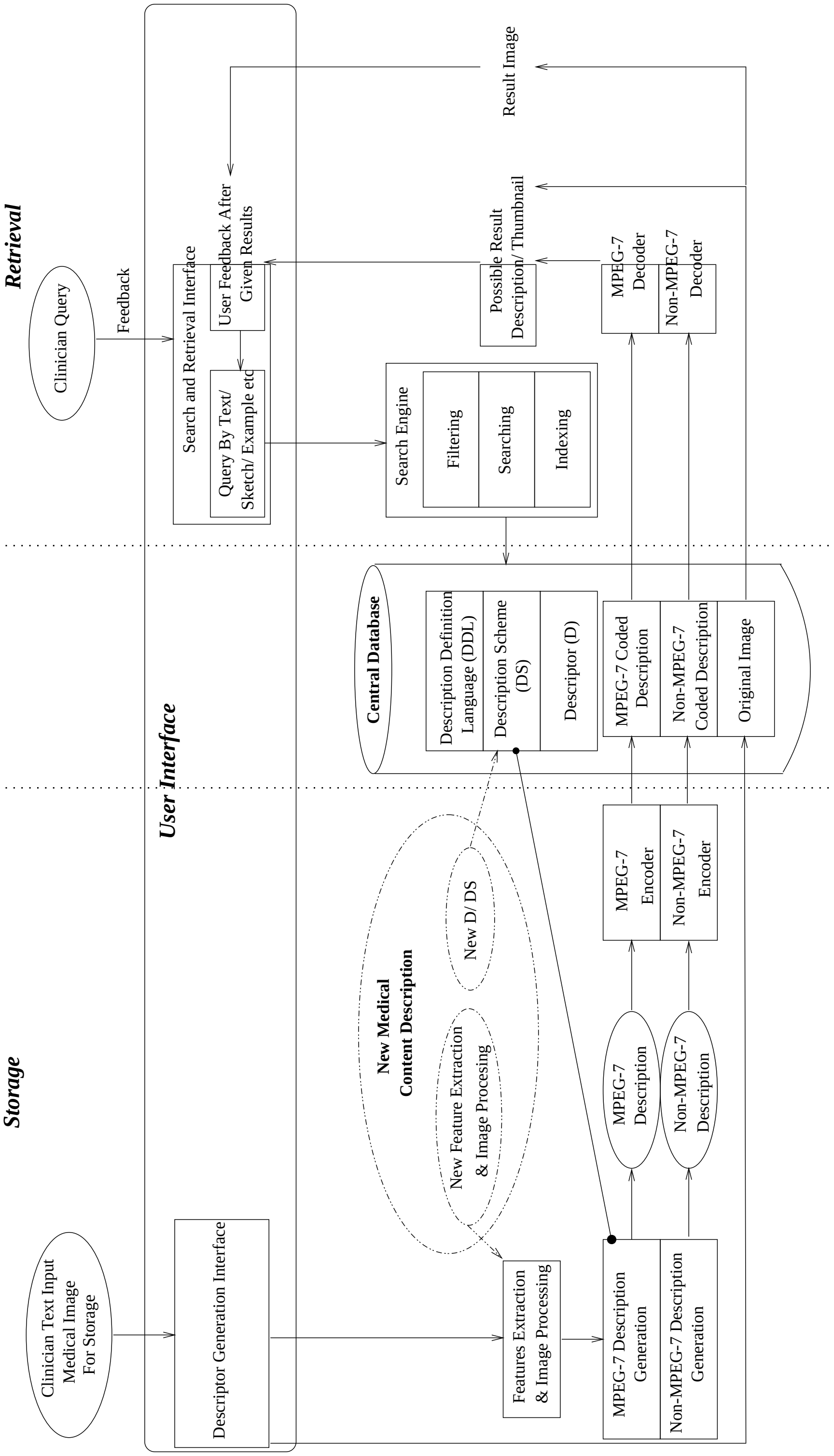


Figure 1: Overview of the proposed system

2 Data Structure

For a successful and easy to use search and retrieval system to perform well and be easy to maintain, it is essential to have a well designed data structure. Most CBIR systems in existence define a new data structure in which to store the relevant information, and little effort has been placed on the creation of a generally acceptable and universally readable data structure. This is seen as a very important step in our approach, as an appropriately designed, easily extensible data structure can aid widespread acceptance.

The data structure proposed is a combination of existing standards. To the author's best knowledge, little research has been done to utilise current meta-data standards for medical image search and retrieval purposes. The US National Library of Medicine uses XML for its PubMed record database (available at http://www.ncbi.nlm.nih.gov/PubMed/). The Lane Medical Library (available at http://lane.stanford.edu/index.html) converted over 200,000+ bibliographic and authority records into XML. All the previously mentioned systems are database systems rather than CBIR systems. Our approach is to propose a data structure which utilises both XML and MPEG-7. Each have their own advantage, but none are capable of encapsulating all the required information effectively. MPEG-7 is suitable for describing low level image content, such as colour structure, textual etc, but its descriptors are limited although extensible; XML can describe anything general, but no standardisation is available to describe medical images. By using MPEG-7 to describe low level image content (available in current MPEG-7 version) and XML to describe that content not included in MPEG-7, all information can be successfully stored. Noting that in MPEG-7, XML schema is used as the Description Definition Language (DDL). Lastly, if the image is available in DICOM format, its meta-data, including UIDs, will be integrated as part of the data structure.

The goal of the proposed data structure is to contain useful information about the image's content. Traditional CBIR systems mostly utilise low level content, which are proprietary data structures for convenience. The proposed data structure aims to accommodate both high and low level content. With hyperlinks in the data structure, users will be able to retrieve details remotely.

The following subsection will briefly discuss image formats, the utilised standards, and the proposed combination.

2.1 Image/Transfer Format

2.1.1 JPEG and Other Conventional Image Format

JPEG[7] is a popular image format, the acronym standing for the Joint Photographic Experts Group. With 24-bit pixel representation, it is especially suitable for photographic reproduction. However, it does not offer facilities to store medical metadata, and there is no UID to uniquely identify each image. Most other popular image formats are also not useful for storing medical or medical equipment information.

Even so, due to their widespread use, they are frequently used for the storage and traansmission of medical images. Sakusabea at al [8] discuss the combining DICOM with Internet technology and indicate that although the source image may be DICOM, a conversion to a JPEG image at the user's browser may ensure the user is able to view the image. Hence, it is important to incorporate and associated popular image formats and meta-data files in the data structure and system proposed.

2.1.2 DICOM

The Digital Imaging and Communications in Medicine (DICOM) standard[9] was created by the National Electrical Manufacturers Association (NEMA) in order to aid the distribution and viewing of various medical images. Popular medical images such as CT scans, MRIs, ultrasound etc. are can all be represented in the DICOM standard. Although most DICOM images are in black and white, DICOM is able to store colour images as well. In the current version 3.0, the protocol is capable of managing almost any type of medical image and its related information, even a complex medical report.

A DICOM file contains both a header and the image data itself. The header stores information about the patient's name, the type of scan, image dimensions, etc. Every day several thousands of medical images are produced by DICOM compatible machines. The risk of losing essential demographic data such as the name of the patient, type of examination and hospital is extremely great if using conventional image formats. With the DICOM standard, each image is associated with specific information including unique identifiers. Hence, each image is autonomous and unique, its origin and patient data can be identified.

In DICOM, several UID (" Unique IDentifiers ") are automatically generated by DICOM modalities and this is mandatory information in each

DICOM file or during transfer. Two identical UIDs indicating different information cannot exist. This is necessary not only because of medical and legal reasons, but also to ensure data integrity in the medical DBMS.

Examples of typical DICOM header information are included below (in incomplete form): (Several numbers are replaced by x to protect patient identity)

Meta Element Group Length (0002,0000) 1 UL [212]

File Meta Information Version (0002,0001) 1 OB [2 * 1 bytes at offset 156 (0x9c) in CT-MONO2-16-ort using bflags 0x2901 (has —little endian—str size 2—explicit hint—)]

Media Storage SOP Class UID (0002,0002) 1 UI [x.x.xxx.10008.5.1.4.x.x.x]

Media Storage SOP Instance UID (0002,0003) 1 UI [x.x.xxx.1113619. ...]

Transfer Syntax UID (0002,0010) 1 UI [x.x.xxx.10008.1.2]

Implementation Class UID (0002,0012) 1 UI [x.x.xxx.9999999.6.29.xx.x.x]

Implementation Version Name (0002,0013) 1 SH [v1.0 Windows]

Source Applica... Entity Title (0002,0016) 1 AE [DIGITAL JACKET]

Identifying Group Length (0008,0000) 1 UL [416]

Specific Character Set (0008,0005) 1-n CS [ISO IR 100]

Image Type (0008,0008) 1-n CS [ORIGINAL PRIMARY AXIAL]

SOP Class UID (0008,0016) 1 UI [x.x.xxx.10008.5.1.4.1.x.x]

SOP Instance UID (0008,0018) 1 UI [x.x.xxx.113619.2.30.1.xxxx ...]

DICOM is the most common standard for storing medical image data, which makes it very important that the proposed system can work with the standard. The proposed system also makes use of various in-built features of DICOM. One main feature is the use of UIDs as mentioned in the above paragraph, which are unique and cannot be duplicated.

There are four mandatory UID's for each DICOM image. They are:

1. SOP Class UID
2. Study Authority UID
3. Series Authority UID
4. SOP Instance UID or Image UID

The SOP Class UID is used to identify the type of service for which the image is intended, whereas the Study Authority UID uniquely identifies a whole examination, time and place. The Series Authority UID identifies a series of images within the examination. Lastly, the SOP Instance UID identifies the image associated with the file. The UIDs have periods separating the numbers (as shown in the example), which makes that look like they have a particular meaning, however, they do not. The UID exists solely to give a unique identity and it is not designed to carry any information about the item itself.

All four of these UIDs will form part of the meta-data file in XML format. This ensures that a one to one relationship is established between the DICOM image and its meta-data file.

Although DICOM assists with the transmission and storage of medical image data, it lacks any form of hierarchical data structure. The proposed work will complement DICOM with a hierarchical structure to enable effective retrieval and storage.

2.2 Content Description Standard

2.2.1 XML

XML[10] is subset of the Standard Generalised Mark-up Language (SGML) defined in ISO standard 8879:1986. Its creation is to make it easy to interchange structured documents over the Internet. XML also defines how Internet Uniform Resource Locators (URLs) can be used to identify parts of XML data streams. However, XML was not designed to be a standardised way of coding text, and certainly it is not a replacement of Hyper-Text Mark-up Language (HTML), which is also another subset of SGML. XML is designed to be a formal language which is used to pass information about the component parts of a document to another system. Furthermore XML is designed to be flexible enough to describe any logical text structure.

In the same manner, XML could be used to describe medical information as a text structure. High level contents such as patient name and address can easily be stored in XML format. However, due to the complexity of low level details, such as colour structure, they are usually stored in proprietary binary format, but with the recent standardisation of MPEG-7, low level details such as colour structure etc. can also be stored in XML format.

2.2.2 MPEG-7

MPEG-7 [2] is an ISO/IEC standard developed by the Moving Picture Experts Group (ISO/IEC JTC1/SC29 WG11). In the Overview of the MPEG-7 Standard, it states that MPEG-7 "... aims to create a standard for describing the multimedia content data that will support some degree of interpretation of the information's meaning, which can be accessed by or passed onto a device or a computer code.". MPEG-7 is not aimed at any one application, but aims to support as broad a range of applications as possible, including medical applications. However, as MPEG-7 is lead mainly by the entertainment industry, the descriptors and description schemes (DS) are very focussed on film and broadcasting systems. For medical applications, usage is restricted.

In MPEG-7 it is anticipated that new descriptors and DS will be required. In this case, a new DS might need to be created to suit medical applications. However it is pointless to solely develop a new DS which only a single system understands, as this contradicts our objective of having a globally understood hierarchical data structure.

For a more comprehensive definition and introduction to MPEG-7 terminologies and technologies, please refer to the MPEG-7 overview document[2] or the first year report of this research work[1].

MPEG-7 is used in the work due to its ability to define common low level details in binary and XML (Description Definition Language of MPEG-7) form and is expected to gain widespread acceptance. However, as MPEG-7 is still in its infancy, the selection of low level algorithms to represent the features is limited. This limitation is expected to lessen as MPEG-7 matures. The current approach is the maximise the utilisation of the existing algorithms to avoid duplication of effort and with any implementation of our own algorithm, the resulting data should be XML compliant.

2.2.3 eXperimental Model (XM)

In MPEG-7 development, evaluation of the constituant algorithms has been done in the XM (eXperimental Model). Various descriptors, description schemes [11] and their respective algorithms for feature extraction are available. During our work, an enormous amount of time has been spent in compiling and investigating various algorithms in the XM. The descriptors and description schemes which are the most interesting to the research work are Visual descriptors [12].

During our evaluation, it was found that the parameters and designs of numerous descriptors and description schemes were not exactly the same as stated in the official documents. Most occurred due to the speed of developing and testing the standard. For example, in the MPEG-7 documents, the Curvature Scale Space (CSS) descriptor algorithm is designed to process multiple passes on the contour (obtain peak information), and calculate the circularity and eccentricity of the contour. However, only the calculation of circularity and eccentricity (calcualtion of the Global Curvature Vector) is done in XM in order to speed up the algorithm for testing, but compromising the accuracy of the result. Peak information and thus the Prototype Curvature Vector is not determined.

2.3 Integration of Technologies

A combination of XML, MPEG-7 and DICOM header information will be effective and comprehensive. The design of an appropriate data structure will enable ease of retrieval from a conventional relational or object-oriented database, as all the information is in textual form. Low level details such as the histogram of an image will also be represented in textual form by DDL. This enables information retrieval (IR) [13] methods, which include relevance feedback mechanisms, to be applied. Search and retrieval in the IR domain is a well researched area, with many effective methods available. With DICOM unique UIDs, the identity of the images can be uniquely determined. In the case where the DICOM image transfer format isn't used, a special UID will be given in place of DICOM UID. In [14], a method of creating a UID similar to the DICOM UID is given, and this can be used when no DICOM header is available.

With MPEG-7, various low level details can be easily stored as text and shared between systems which understand the common standard. XML will accomodate both MPEG-7 and non-MPEG-7 information. The benefits of using XML are its evolving popularity and that its parser can be designed to skip over information unknown to the parser, and only extract the information required.

2.3.1 DICOM Header

In DICOM, devices that create new objects (patient records, image series, image instances) are required to generate UIDs using a method based on the OSI object identification defined by ISO 8824.

The DICOM standard also defines a way to uniquely identify objects by using Unique Identifiers (UIDs) which guarantee uniqueness across multiple countries, sites, vendors and equipment. In the proposed system, although all the DICOM header information will be retained in the header, four of the UIDs, should be also stored in the database and other related files, as it can easily help identify the correct image, even when the position of the files has changed.

2.3.2 MPEG-7 Descriptor

In the proposed data structure, any kind of MPEG-7 descriptors and description schemes can be included, even future standard descriptors and description schemes [15]. For demonstration and evaluation purpose, colour structure, texture browsing and contour shape descriptors are included in the current implementation.

2.3.3 Proposed Hierarchical Structure

The proposed hierarchical data structure is in XML form, which incorporates MPEG-7 descriptors and other information. Other data will include either DICOM or non-DICOM UIDs, related patient information, other image related information, and links to other related data and images. The design of such a structure requires that it be easily readable, acceptable and portable.It is modular as only a module of the structure need to be exchanged when a different class of medical image is given. In our implementation, XML Schema is used to define and regulate the possible data content and format, although Data Type Definition (DTD) [16] or other can also be used.

Figure 2 shows the proposed data structure comprising a combined a combination of DICOM header and a XML compliant meta-data file type (CDIMI file). The CDIMI file is split into two different types of information. Image Class Independent information is non-changeable regardless of the image class. Image Class Dependent information will however change depending on the different class of images. For example, the information of a Dermoscopic image class will be different from the information of a CT scan.

An example of a data structure for a dermoscopic image is shown below (Several numbers are replaced by x to protect patient identity):

```
<CDIMI id=Derma01”>
        <ImageClassIndependent>
                <NonChangeableImageInfo>
                        <UIDs>
```

```
                <SOPClass> x.x.xxx.10008.5.1.4.1.x.x
                <Study> xxx.xxx.2.19960619.xxxxxx
                <Series> xxx.xxx.2.19960619.xxxxxx.x
                <Instance> xxx.xxx.2.19960619.xxxxx.x.xxx
            < /UIDs>
            <General>
                ...
        < /General>
        < /NonChangeableImageInfo>
        <PatientInfo>
            ...
        < /PatientInfo>
    < /ImageClassIndependent>
    <ImageClassDependent>
        <ChangeableImageInfo>
            <StillRegion id=”Derma01”>
                <SegmentDecomposition gap=”true”
                overlap=”false” decompositionType=”spatial”>
                    ...
                < /SegmentDecomposition>
            < /StillRegion>
        < /ChangeableImageInfo>
        <MedicalInfo>
            <Race> African < /Race>
            <ABCD>
                ...
            < /ABCD>
            ...
        < /MedicalInfo>
    < /ImageClassDependent>
< /CDIMI>
```

Under Image Class Independent, there are two branches. The first branch is Non-changeable Image Info. Within it, UIDs from DICOM can be found; the other data is general image information which will be image independent. An example is the Image Class (essential for the system to know what is in the Image Class Dependent), Colour Format, Links to Next in Series (if DICOM), Links to this patient's images index etc. The other branch is Patient Info, which stores information about the patient, doctor and the hospital. It is currently implemented using an MPEG-7 compliant format where feasible. This creates a situation in which some information is not MPEG-7 compliant, especially at the higher level, and yet the lower branches are MPEG-7 compliant.This compatibility problem is currently under investigation and once a suitable method has been found, it is expected that the whole Image Class Independent will have the same standard format.

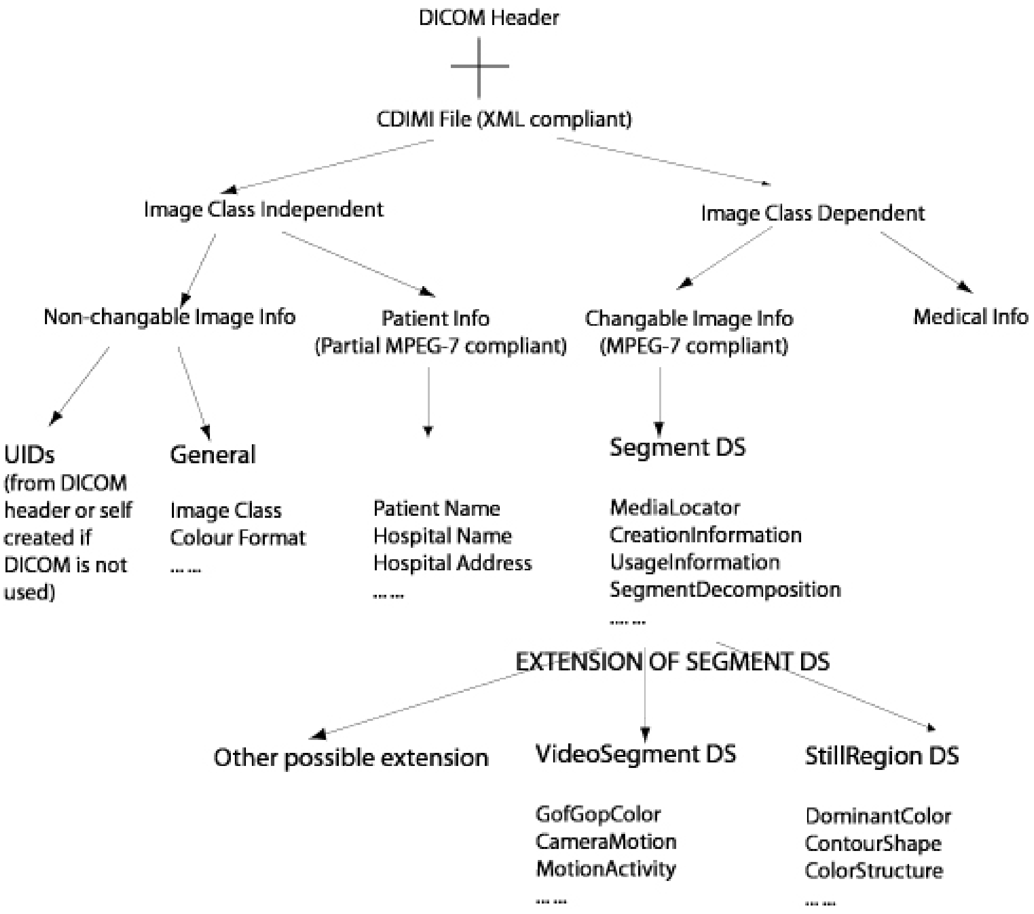


Figure 2: Overview of the proposed data structure

The Image Class Dependent section is also split into two branches, Changeable Image Info (versus Non-changeable Image Info under Image Class Independent), and Medical Info. In Changeable Image Info, all data is MPEG-7 compliant. Segment DS is a description Scheme (DS) found in the Multimedia Description Schemes (MDS) under the MPEG-7 Descriptors and Description Schemes(DS). It is an abstract type of which its specialised DS can be StillRegion DS, VideoSegment DS, MovingRegion DS, AudioVisualSegment DS etc. Segment DS aims to describe the content of a segment, as a video, image or audio.

For example, if a dermoscopic image is given, the Segment DS will be StillRegion DS. Under the StillRegion DS, the infected skin area will have its own segment decomposition and segment descriptors. The descriptors will be Contour Shape, Dominant Colour, Texture Browsing etc. All the descriptors aim to describe the content of the infected skin area only. Non-visual descriptors, could also be included. For example, FreeTextAnnotation allows the creator to fill in any comment for the segment (infected skin area). However, this is not really recommended, as the design of this data structure is to store other text based, or high level information in Medical Info category.

An example of a Segment DS for a dermoscopic image is included below:

```
<StillRegion id=”Derma12”>
        <SegmentDecomposition gap=”true” overlap=”false”
        decompositionType=”spatial”>
                <Segmentxsi:type=”StillRegionType
                id=”InfectedArea”>
                        <ContourShape numberOfPeaks=”59”>
                                <GlobalCurvatureVector>
                                        7 1
                                < /GlobalCurvatureVector>
                                <PrototypeCurvatureVector>
                                        3 1
                                < /PrototypeCurvatureVector>
                                <HighestPeak>
                                        8
                                < /HighestPeak>
                                <Peak>
                                        <xpeak>17< /xpeak>
                                        <ypeak>7< /ypeak>
                                < /Peak>
                                ... ...
                        < /ContourShape>
                        <ColorStructure colorQuant=”1”>
                                <Values>
                                        24 2 12 6 22 26 ...
                                < /Values>
```

```
            < /ColorStructure>
            <TextureBrowsing>
                <Regularity>irregular
                    <Direction>
                       150 degree
                    < /Direction>
                    <Scale>
                       medium
                    < /Scale>
                < /Regularity>
            < /TextureBrowsing>
            ... ...
        < /Segment>
    < /SegmentDecomposition>
< /StillRegion>
```

Although most of the work represented will be based on still images, the wide choice of Segment DS means that the data structure can also be implemented even if the data is a section of medical audio etc.

The next category in the Image Class Dependent section is the Medical Info. It is used to store other medical information, especially high level content information related to the diagnosis of the disease. It should be sufficiently modular that with different classes of image, the required content will change accordingly. For example, the Medical Info for a dermoscopic image class may be as below:

```
<MedicalInfo>
    <Race> African < /Race>
    <ABCD>
        <A> 2 < /A>
        <B> 0 < /B>
        <C> 2 < /C>
        <D> 2 < /D>
    < /ABCD>
    <ABCDweight>
        <A> 1.3 < /A>
        <B> 0.1 < /B>
        <C> 0.5 < /C>
        <D> 0.5 < /D>
    < /ABCDweight>
    <Score> 4.6 < /Score>
    <Diagnose> malanocytic naevus < /Diagnose>
< /MedicalInfo>
```

3 System and Interface

A prototype system, provisionally named CDIMI (Content Description Interface for Medical Imaging) has been designed and built to demonstrate how the proposed data structure, together with search and retrieval functions, can be effective in finding the appropriate information. The system will be useful for several purposes. 1. training of junior clinicians to recognise symptoms by comparing similar images; 2. Searching for similar images to the model images which the clinician has (for example, evidence-based medicine); 3. assisting clinician with diagnostic procedures by generating prompts and comments as a result of comparision of image features. The diagnostic support environment could conceivably lessen the clinician's liability during any case of conflict with the patient or the patient's family.

3.1 Overall System Functionality

The system is composed of several components:

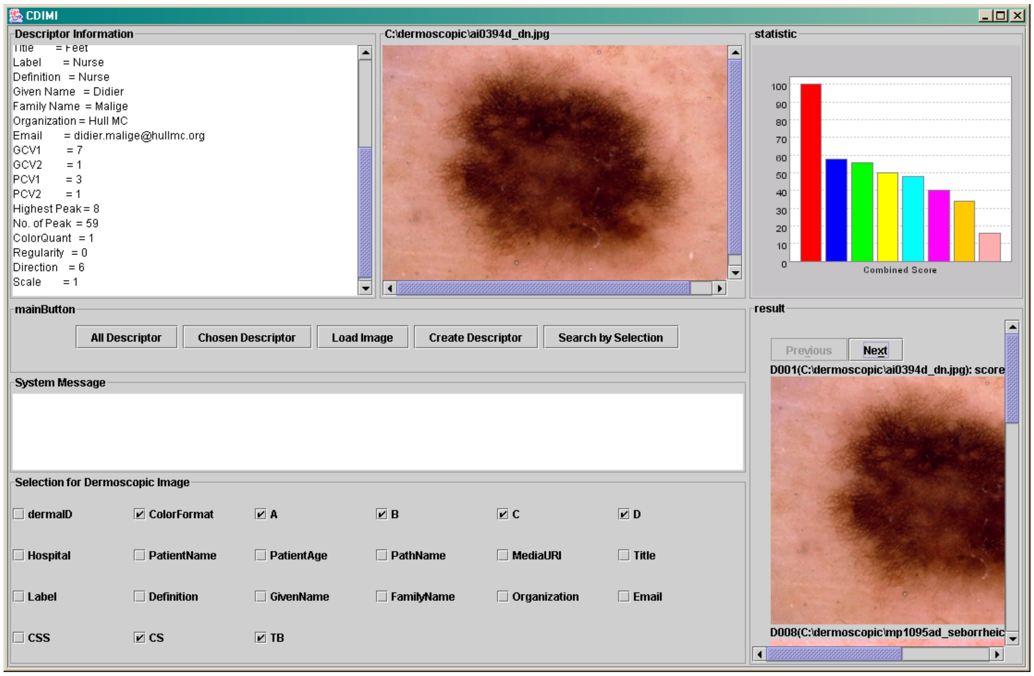
1. Creation (of Data Structure and database)
2. Search and Retrieval
3. Attribute Graph
4. Database

The creation, or data construction component is designed to build the relevant data structure for a given image.

The search and retrieval system is considered the most important part of the overall system, as without it, the system is rendered useless. A query image must be first obtained by the system. The clinician will then select the required search criterias, which can be as simple as the age of a patient, or a descriptor, such as texture browsing. With the search criteria and query image, the search is performed and the retrieved result returned to the interface.

An attribute graph accompanies the result, which gives the clinician guidelines for further refined searching.

Finally, there is a database where all data is stored for fast search and retrieval.



Title = Feet
Label = Nurse
Definition = Nurse
Given Name = Didier
Family Name = Malige
Organization = Hull MC
Email = didier.malige@hullmc.org
GCV1 = 7
GCV2 = 1
PCV1 = 3
PCV2 = 1
Highest Peak = 8
No. of Peak = 59
ColorQuant = 1
Regularity = 0
Direction = 6
Scale = 1

☐ dermalD ☑ ColorFormat ☑ A ☑ B ☑ C ☑ D
☐ Hospital ☐ PatientName ☐ PatientAge ☐ PathName ☐ MediaURI ☐ Title
☐ Label ☐ Definition ☐ GivenName ☐ FamilyName ☐ Organization ☐ Email
☐ CSS ☑ CS ☑ TB

D001(C:\dermoscopic\ai0394d_dn.jpg): score
D008(C:\dermoscopic\mp1095ad_seborrheic

Figure 3: User interface, showing attribute graph of combined score

3.2 Data Construction Sub-System

The Creation component is designed to construct the relevant data structure for a given image. With the image, the system can generate or import the relevant information, which can be low level details from MPEG-7 descriptor, or high level information such as a patient address. The output of the creation component is a meta-data file which consists of all relevant information, as mentioned in the section 2. Although the system allows the meta-data file to be stored in different locations from the image, the meta-data file is usually stored in the same location (directory). A copy of similar information is also stored in a database for ease of searching and retrieval.

3.3 Search and Retrieval Sub-System

The Search and Retrieval System works with the database. The first step is the calculation of the weight of different criteria and the overall ranking points. The second step is to send the queries to the database, and utilise the database to return the result in the order of ranking.

First, let $\mathbf{x}$ represent the feature vector of a random image in the database

$$\mathbf{x} = [x_1 x_2 \ldots x_n] \tag{3.1}$$

where n is the number of criteria selected.

The similarity measure is then calculated as

$$S = \sum_{i=1}^{n} \frac{w_i |x_i^{(r)} - x_i^{(q)}|}{k_i} \tag{3.2}$$

where w_i is the weighting for the ith criterion; $x_i^{(r)}$ is the output/feature value of a sample image in the database under the ith criterion and $x_i^{(q)}$ is the feature value of the query image under the ith criterion; k_i is the normalisation factor for the ith criterion, which is calculated as the value range of the distance of sample images to the query image. The normalisation is performed by obtaining the minimum and maximum values of (r), the samples in the database, and linearly normalising (r) (0 to 1) between the minimum and maximum values.

The algorithm can be summarised as follows. If more than 1 criterion is selected, use algorithm A, else, use algorithm B.

A. More than 1 criterion selected.

A1. Retrieve the selected criterion and its respective weighting.

A2. For each individual criterion, compare the values from the database to the values of the query image.

A3. For each individual criterion, get the highest and lowest absolute difference in values, and let them be H and L respectively.

A4. Assign 100 points to L, and 0 point to H, proportionally score those which are in the middle of L and H.

For example, if the query value is 100, and in the database there exist values of $A = 0$, $B = 250$, $C = 300$, the difference x will be 100, 150 and 200 respectively. In this case, $H = 200$ and $L = 100$. One more value that needs to be evaluated is B, which has a difference of 150. The score for B can be calculated as

$$B = 100/(H - L) * (x_b - L) \tag{3.3}$$

The result is 50 points.

A5. Since all of the values have been converted into a common unit (points), calculation of the actual score can be performed.

For example, assume 3 criteria are given, with weights w be $w_A = 0.1$, $w_B = 0.4$, $w_C = 0.5$. Assume also that a sample image in the database was awarded the following points p where $p_A = 90$, $p_B = 20$, $p_C = 45$, the final result should be: $w_A * p_A + w_B * p_B + w_C * p_C = 39.5$

A6. Rank the results according to the final points awarded, with the highest points as the closest match.

B. Only one criteria selected.

B1. Retrieve the value of the selected criterion from the database, compare the value from the database to the value of the query image.

B2. Rank the results according to the difference in values. The smaller the difference, the higher the ranking.

B3. No points will be calculated, as the original values will be displayed in the search result column in the interface.

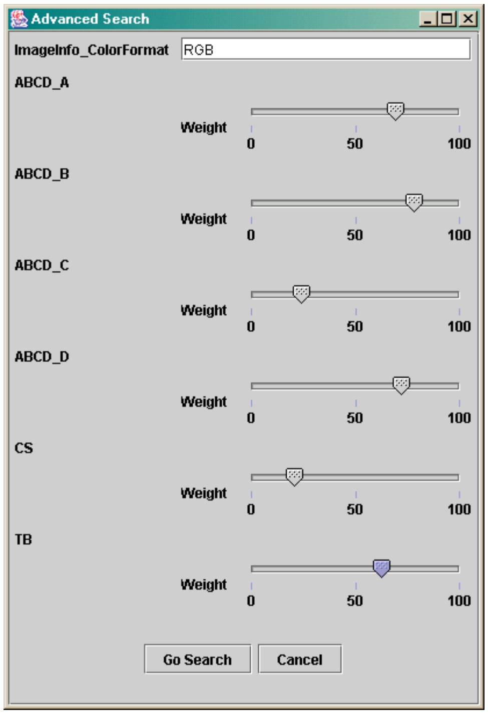
The above method is suitable for a small sample set as it requires calculation of the distance of sample images to the query image in real time. When the sample set is sufficiently large as in Law of Large Number (LLN)[17], which means the set of sample images would generously represent the image population, the normalization factor for each feature dimension, i.e. each criterion, would normally be pre-calculated as the difference between the maxima and the minima of image values for that feature dimension.

3.4 Search Result and Attribute Graph

After the search has been performed, the next step is to present the result in an easily understood and informative way. All results are displayed in a "Closest match" first order, ensuring that the user will easily see the nearest match. A total of five closest matches will be displayed on the first page, with an option to view the next five.

To make the result as informative as possible, we have to incorporate the attribute graph concept in the system. An attribute graph is a scaled down version of Attribute Explorer[18].

An attribute graph belongs to the category of information visualisation. It is different from scientific visualisation, which usually involves very large data sets and three-dimensional representations. Some examples include stress analysis in building materials and medical scanning. Instead, the emphasis of information visualisation is on meaningfulness. Literal representation is of less importance. Typical examples include bar charts and pie charts. Here,



ImageInfo_ColorFormat RGB

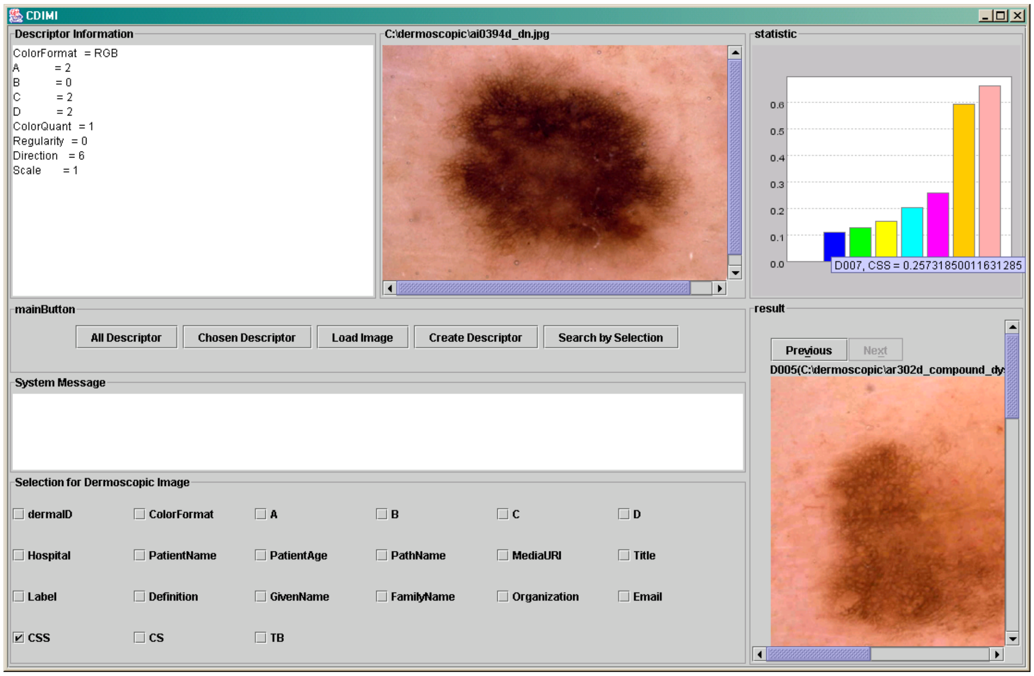
Figure 4: The Search box

our attribute graph uses bar charts to represent the meaning of the search result, and it provides an indication of how well the search is performed, instead of giving a statistical result.

In cases where there is only one criterion involved, the attribute graph will display a bar chart in which each bar represents the actual values of all the sample images in the database. The first bar will be the sample images with the closest match, as in the search result, and the last bar will represent subsequent matches.

In cases where there are more than one criterion involved, the attribute graph will display a bar chart in which each bar represents the awarded score for each sample image in the database. The first bar will represent the sample image with the highest score, which is the closest match, and the last bar will represent the sample image with the lowest score.

Initially, after the search result has been generated, the bar chart will display each result. This is to provide the clinician with a visualisation of the overall result, and consequently an indication of the quality of the search operation. While with only few result images shown without the attribute graph, the clinician will have no idea of what the score of the next page will



ColorFormat = RGB
A = 2
B = 0
C = 2
D = 2
ColorQuant = 1
Regularity = 0
Direction = 6
Scale = 1

☐ dermaID ☐ ColorFormat ☐ A ☐ B ☐ C ☐ D
☐ Hospital ☐ PatientName ☐ PatientAge ☐ PathName ☐ MediaURI ☐ Title
☐ Label ☐ Definition ☐ GivenName ☐ FamilyName ☐ Organization ☐ Email
☑ CSS ☐ CS ☐ TB

Figure 5: The interface, throwing the attribute graph with individual values instead of overall score

be, let alone understand the overall situation.

The graph can be enhanced (zoom) and individual bar (sample image) information can be displayed. The clinician can then select the interested bar and the relevant data is provided.

In short, the attribute graph can be used to assist the clinician in preparing the next search criteria. It focuses on giving an abstract understanding of the whole dataset. With a few searches done, and if the user focuses on the difference in the graph, abstract meaning of the whole dataset from the graph can always be determined. It is also useful in that the graph can assist in the statistical understanding of images in the database.

3.5 Database and File Format

Currently, the database system is implemented using MySQL, a popular and powerful relational DBMS. The database function is a back end system, which enables faster query and retrieval. Query and retrieval is also possible through the original meta-data file, however, the speed of retrieval is extremely slow when compared with retrieval from the database.

The actual image itself is not stored in the database, however a link is provided. When the tuple (image information) in the database is selected, the relevant image can be obtained by using the provided link. Although images which require values of arbitrary length can be stored in the database, it is not recommended because of flexibility and speed considerations.

4 Conclusions and Future Work

Content-based image retrieval (CBIR) for medical imaging us of interest for future advanced medical care systems. As more hospitals and health authorities are implementing Picture Archiving and Communications Systems (PACS), it almost becomes essential that a flexible and powerful CBIR system is included.

This report describes work to date, which includes the investigation of various MPEG-7 and XML components, and how to utilise the flexibility of the DICOM protocol in the research. A hierarchical data structure has been proposed, which is a combination of XML, MPEG-7 and possibly DICOM. The research focuses on utilising available technology, rather than to propose another standard. The Information Technology community is already concerned with many different and confusing technologies.

A prototype system has been developed. Emphasis has been placed on the ease of use interface, as it is the communicating point for the clinician and the underlying system. The search and retrieval system has been explained, as well as the relationship between the results and the attribute graph. The attribute graph is useful for providing a visual impression of the overall results and the underlying datasets.

Currently, work is continuing to implement a new version of the prototype, which will probably update the data structure to a more user friendly and comprehensive format. The new version of the prototype should include the following concepts:

1. Object Oriented and Modular design. This is absent in the current version which is insufficiently modular.

2. Demonstrate the capability of natural language processing (NLP) in a user friendly medical CBIR system. It is strongly believed that with NLP, the acceptance rate for such a system will be increased, as clinicians are not prepared to spend learning about the system or using a Structural Query Language (SQL) in order to communicate effectively

with the system. With NLP, the gap between the system and human can be significantly reduced.

3. Further extend the capability of the attribute graph. Ideas on possible improvement include making the graph more interactive and dynamic. Currently, only the distribution of the result is shown and there is no dynamic feedback mechanism from either the user or the system. Ideally, the attribute graph should present information dynamically, and respond to user interactions with immediate feedback to finely adjust various criteria.

4. Provide an immediate feedback mechanism to create a dynamic and adaptive system. The system should be actively interacting with the user. Relevance feedback mechanisms from the information retrieval domain can be applied, which enables the search to be more relevant each time feedback is provided.

5. Implementation of a Search Editor and Refinement system. The Search Editor will act as a Image editor, but for legal reasons, disallows the image to be saved as an original image. Such systems will enable the user to remove unnecessary features or refine the search image. For example, dermoscopic images will give better results without hair interfering with the border contour. The user can also select a region of interest, where only the content from that region is used for searching.

References

[1] C.U. Ng and G.R.Martin, "Content-description interfaces for medical imaging," Tech. Rep., University of Warwick, 2001.

[2] MPEG Requirements Group, *MPEG-7 Overview (version 3.0), ISO/IEC JTC1/SC29/WG11-N3445*, Geneva, May/June 2000.

[3] F. Mokhtarian, S. Abbasi, and J. Kitter, "Efficient and robust retrieval by shape content through curvature scale space," in *Image DataBases and Multi-Media Search*, pp. 51–58. World Scientific Publishing, Singapore, 1997.

[4] S. C. Orphanoudakis, C. Chronaki, and D. Vamvaka, "I^2Cnet: Content-based similarity search in geographically distributed repositories of medical images," *Computerized Medical Imaging and Graphics*, vol. 20, no. 4, pp. 193–207, 1996.

[5] S. C. Orphanoudakis, C. Chronaki, and S. Kostomanolakis, "I^2C: A system for the indexing, storage, and retrieval of medical images by content," *Medical Informatics*, vol. 19, no. 2, pp. 109–122, 1994.

[6] K. Lemstrom, J. Korte, P. Kuusi, P. Kyheroinen, and P. Paivakumpu, "PICSearch - a platform for image content-based searching algorithms," in *The Sixth Int. Conf. in Central Europe on Coputer Graphics and Visualization*, 1998, pp. 222–229.

[7] ITU-T Rec. T.81 —, *Digital Compression and Coding of Continuous-Tone Still Images*, Geneva, 1994.

[8] Takaya Sakusabea, Baatar Shirchinb, and Michio Kimura, "Proposal of a new internet standard for dicom: Dicom-qr url," in *Proceedings of the Ninth World Congress on Medical Informatics*, 1998, pp. 438–441.

[9] National Electrical Manufacturers Association, *Digital Imaging and Communications in Medicine (DICOM) PS 3.1-2001*, Virginia, USA, 2001.

[10] Tim Bray, Jean Paoli, C. M. Sperberg-McQueen, and Eve Maler, *Extensible Markup Language (XML) 1.0 (Second Edition)*, World Wide Web Consortium, MIT, Cambridge, USA, October 2000.

[11] Koiti Hasida and John R. Smith, *MPEG-7 Multimedia Description Schemes XM (Version 13.0)*, Fairfax USA, May 2002.

[12] Sylvie Jeannin, Leszek Cieplinski, Jens Rainer Ohm, and Munchurl Kim, *MPEG-7 Visual part of eXperimentation Model Version 7.0*, Beijing, July 2000.

[13] G. Salton and M. J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill Book Company, 1983.

[14] Mallinckrodt Institute of Radiology Stephen M. Moore, Electronic Radiology Laboratory, *Programmers Guide to the UID Facility*, August 1998.

[15] L. Cieplinski, J.-R. Ohm W. Kim, M. Pickering, and A. Yamada, *MPEG-7 Visual part of eXperimentation Model Version 14.0*, Fairfax USA, May 2002.

[16] The Organization for the Advancement of Structured Information Standards, *The DocBook Document Type*, Billerica, MA, USA, July 2002.

[17] Henry Stark and John W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, Prentice Hall, 1986.

[18] L. Tweedie, R. Spence, D. Williams, and R. Bhogal, “The attribute explorer,” in *Video Proceedings of CHI’94 and Conference Companion Proceedings*, Boston, USA, 1994, pp. 435–436.