THESIS TITLE

CONTENT-BASED INDEXING OF LOW RESOLUTION DOCUMENTS

DANIAL BIN MD NOR

A thesis submitted in
fulfilment of the requirement for the award of the
Doctor of Philosophy

Faculty of Electrical and Electronic Engineering
Tun Hussein Onn University of Malaysia

SEPTEMBER 2016

DEDICATION

My dearest mother and father,
Hjh Puteh bte Ahmad and Hj Md Nor bin Majid

and

My beloved wife,
Noor Aini bte Idris

and

My sweet children,
Irfan, Imran, Idlan and Ilman

May Allah protect and bless us all.

ACKNOWLEDGEMENTS

It is with great pleasure to have this chance to sincerely appreciate the following individuals:

- Prof. Dr. Jean Ogier from ULR, my supervisors for his assistance and knowledge in providing me with the necessary guidance through the difficulties faced throughout the lifetime of the thesis. In addition, he proof read my thesis's draft and offered some understanding on how I can improve it further when it come to the final copy of my thesis. Moreover, Prof. Dr Muhammad Zarar Muhd Jenu (from UTHM) who is my local supervisor that always encouraged me whenever I struggled, his generous support always came at just the right time.
- Ministry of Higher Education Malaysia and Universiti Tun Hussein Onn Malaysia (UTHM) for awarded me the scholarship.
- My Joint PhD programme coordinator, Prof. Dr. Patrice Bousier. The management of UTHM who initiated the MoA between UTHM and ULR, Prof. Dr. Hashim Saim and Prof. Dr. Idrus Masirin. My dean Dr Afandi Ahmad. To my colleagues Prof. Ir. Dr. Zainal Abidin Ahmad, Dr. Rosli Omar, Prof. Dr. Talib Bon and Assoc. Prof. Dr.Mohd. Najib Mohd. Salleh. Discussions with them at different stages of the thesis were very rewarding.
- My wife Noor Aini Idris and my sons. Without Noor Aini's support and constant motivation, nothing would be possible.

Abstract

In any multimedia presentation, the trend for attendees taking pictures of slides that interest them during the presentation using capturing devices is gaining popularity. To enhance the image usefulness, the images captured could be linked to image or video database. The database can be used for the purpose of file archiving, teaching and learning, research and knowledge management, which concern image search. However, the above-mentioned devices include cameras or mobiles phones have low resolution resulted from poor lighting and noise. Content-Based Image Retrieval (CBIR) is considered among the most interesting and promising fields as far as image search is concerned. Image search is related with finding images that are similar for the known query image found in a given image database. This thesis concerns with the methods used for the purpose of identifying documents that are captured using image capturing devices. In addition, the thesis also concerns with a technique that can be used to retrieve images from an indexed image database. Both concerns above apply digital image processing technique. To build an indexed structure for fast and high quality content-based retrieval of an image, some existing representative signatures and the key indexes used have been revised. The retrieval performance is very much relying on how the indexing is done. The retrieval approaches that are currently in existence including making use of shape, colour and texture features. Putting into consideration these features relative to individual databases, the majority of retrievals approaches have poor results on low resolution documents, consuming a lot of time and in the some cases, for the given query image, irrelevant images are obtained. The proposed identification and indexing method in the thesis uses a Visual Signature (VS). VS consists of the captures slides textual layout's graphical information, shape's moment and spatial distribution of colour. This approach, which is signature-based are considered for fast and efficient matching to fulfil the needs of real-time applications. The approach also has the capability to overcome the problem low resolution document such as noisy image, the environment's varying lighting conditions and complex backgrounds. We present hierarchy indexing techniques, whose foundation are tree and clustering. K-means clustering are used for visual features like colour since their spatial distribution give a good image's global information. Tree indexing for extracted layout and shape features are structured hierarchically and Euclidean distance is used to get similarity image for CBIR. The assessment of the proposed indexing scheme is conducted based on recall and precision, a standard CBIR retrieval performance evaluation. We develop CBIR system and conduct various retrieval experiments with the fundamental aim of comparing the accuracy during image retrieval. A new algorithm that can be used with integrated visual signatures, especially in late fusion query was introduced. The algorithm has the capability of reducing any shortcoming associated with normalisation in initial fusion technique. Slides from conferences, lectures and meetings presentation are used for comparing the proposed technique's performances with that of the existing approaches with the help of real data. This finding of the thesis presents exciting possibilities as the CBIR systems is able to produce high quality result even for a query, which uses low resolution documents. In the future, the utilization of multimodal signatures, relevance feedback and artificial intelligence technique are recommended to be used in CBIR system to further enhance the performance.

TABLE OF CONTENTS

List of Figures

List of Tables

Nomenclature

AMORE	:	A World Wide Web Image Retrieval Engine
CBIR	:	Content-Based Image Retrieval
CBVIR	:	Content-Based Visual Information Retrieval
CIE	:	International Commission on Illumination
CST	:	Colour, Shape and Texture
DB	:	Database
ER	:	Extended Rectangle
ERBIR	:	Efficient Region Based Image Retrieval
GBIR	:	Global feature Based Image Retrieval
GLCM	:	Gray-Level Co-Occurrence Matrix
GRBIR	:	Global and Region Feature Based Image Retrieval
GUI	:	Graphical User Interface
GCH	:	Global Colour Histogram
HSV	:	Hue-Saturation-Value
HVS	:	HumanVisual System
IRM	:	Integrated Region Matching
LCH	:	Local Colour Histogram
OCR	:	Optical Character Recognition
PSNR	:	Peak Signal to Noise Ratio
PWT	:	Pyramid-structured Wavelet Transform
QBIC	:	Query by Image Content
RBIR	:	Region Based Image Retrieval
RLSA	:	Run-length Smoothing Algorithm
RGB	:	Red-Green-Blue
SIFT	:	Scale-Invariant Feature Transform
SOM	:	Self Organising Map
SURF	:	Speeded Up Robust Features

LIST OF APPENDICES

CHAPTER 1

Introduction

"The greatest enemy of knowledge is not ignorance, it is the illusion of knowledge."
– *Stephen Hawking*

Content :

Abstract

Chapter 1 presents the overall research picture. In particular, it details the current problems related to low resolution document and some proposals to address them. In addition, the chapter also highlights the outcomes that have been achieved throughout the research.

1.1 Research Motivation

Recently, focus is being given to Content-Based Image Retrieval (CBIR) due to the remarkable growth in terms of the number and size of video collections and digital image found on web. Consequently, it has made it necessary for developers to develop powerful tools that can be used for retrieving this imagery that is unconstrained. In addition, CBIR has also become the key technology in relation to improving the interface between computers and users.

The majority of conventional image databases are text annotated. Because of this, image retrieval is based on keyword searching. Text-annotated images are easy and simple to manipulate. Nevertheless, this image retrieval method has two main problems. The first problem is that, creating keywords for a large number of images takes a lot of time. In addition, the keywords are not unique and inherently subjective. Because of these disadvantages, automatic indexing and retrieval, which is based on image content has become more desirable for developing large-volume image retrieval applications.

Research on multimedia systems and CBIR has been given tremendous importance over the last decades. This is because multimedia databases are concerned with image, video, audio and text data that offers large amount of information. This has influence our lifestyle positively. The bottleneck as far as the access of multimedia databases is concerned in CBIR. This is attributed to the fact that a significant difference does exist between human capacity perception and that of a computer.

The investigation of document analysis, recognition, retrieval and integration with other media is motivated by the cutting-edge applications associated with multimedia. Today, a good number of documents along with visual and audio information are captured, archived, delivered and managed in digital forms. The numerous uses of such documents and digital media files provokes many new challenges in multimodal mobile information acquisition and large multimedia database management. Some of the most prominent areas are:

a) automatic meeting/lecture segmentation and indexing: creates a structured, searchable view of archives of the multimodal meeting/lecture content (Behera, Lalanne, & Ingold, 2007)(Chiu, Foote, Girgensohn, & Boreczky, 2000);

b) library digitization: digitises journals, magazines, newspapers and various videos using advanced imaging for digital libraries (DLs) (Akram & Dar, 2010)(Deselaers, Keysers, & Ney, 2007);
c) mobile translation: extracts and translates textual content as well as visual signs for tourist use. For example, English documents could be translated to French, German or Asian languages using a handheld translator. Text could be translated to speech for the visually impaired, or vice-versa for the hearing impaired (Zhang & Yang, 2012)(Tsai, 2012);
d) automatic indexing of broadcasts: indexes broadcast content for content-based retrieval (Munesawang & Ling, 2005)(Poullot, Cnam, & Crucianu, 2009).
e) digital media asset management: archives digital media files for efficient media management (Shao, 2008)(Chan, Lin, Ho, Fann, & Wang, 2012) (Chang et al., 2012); and
f) document cataloguing: document database catalogues on the basis of content relevance (Bhute, 2012); (Matusiak, 2006).

The fundamental techniques that address the above-mentioned challenges are content-based multimedia annotations and retrieval. The annotation should be structured in order to speed up real-time processing using computers. One fundamental aspect of annotation is that the process should be automated and the content relevant, which would make browsing and retrieval comfortable for users. In order to obtain this, content-based indexing and retrieval tasks require that descriptive features, which are significant to the subject materials like video, audio, documents should be extracted. This research therefore also addresses the use of documents for browsing and retrieval of other related multimedia information.

1.2 Problem Definition

Resolution is the assessment of pixels in a display, often expressed in measurements of height and width. Consequently, image resolution denotes the number of a picture's pixels that can fit within every inch of a document when printed. The Low

Resolution Documents (LRD) normally referring to noise in image display due to quality of camera and light exposure. Hence it creates many issues, for example to Optical Colour Recognition (OCR) since it is almost impossible to carry out recognition independently from character segmentation. As the development and network of multimedia inventions become more popular, the conventional information retrieval techniques do not satisfy users. The most recent decade witnessed considerable interest in studies on CBIR. Consequently, this endeavour paved the way for numerous novel systems and techniques, and a great interest for related fields to support these systems (Datta, Jia, & James, 2005). Similarly, there was a great expansion of the digital imagery, bringing about an explosion in the quantity of image data that require organization.

Currently, the indexing of CBIR is a source of fast and exact retrieval. This thesis considers the layout and graphic of an image along with its colour to improve efficiency. Colour, layout and shape features are combined in our retrieval system to retrieve images that are similar for the given query image obtained from the database. However, the fact that the device identifies notes using colour alone does not guarantee good results. It should be noted that the device can only be used for identifying Canadian and USA dollar notes. This thesis considers the layout and graphic of an image along with its colour to improve efficiency. Colour, layout and shape features are combined in our retrieval system to compute images that are similar for the given query image that was obtained from the database.

Trends to take pictures of the images of slides that interest attendees during a presentation is growing significantly. The images that the attendees capture will be of much use for knowledge management, academics and research only if they can be linked to the presentation videos that were recorded during a presentation. Apart from the problems related to indexing, there are other main issues that were necessary to address in order to ensure better feature representation, these issues include noise, feature extraction of poor resolution, the capture environment's varying lighting conditions and complex backgrounds problems. Yet another way is to provide feature vectors for low-resolution images, which come from textual information.

The objectives of this research are:

1. to examine and to analyse different CBIR techniques for finding the best low-level features tools that can be used with low-resolution CBIR;
2. to enhance the extraction of layout and spatial feature through making use of new techniques, which utilise a simple matching process that can be indexed and searched;
3. to create a fusion technique that is multi-technique in nature that makes use of simple techniques in order to ensure enhanced performance as compared to that of the techniques that are already in existence;
4. to present a proposal on an indexing scheme that can be used for low-level visual signatures of low-resolution images;
5. to introduce a new algorithm for large-scale similarity searching for fast retrieval and scalability; and
6. to evaluate the CBIR system performance using precision and recall techniques.

1.3 Scope

This research is confined to CBIR system using three features of an image, which include layout, shape and colour in order to improve the efficiency of CBIR systems. The image features, which are described in Chapters 4 and 5, are computed on natural databases that are mixed in nature and are implemented using the layout, shape and colour feature approach. Through making use of the three image features, we perform image retrieval using the databases in existence for the specified query image.

1.4 Thesis Outline

This particular thesis has six distinct chapters as follows,

a) Chapter 1: Introduction

Chapter 1 offers the research's overall picture. It presents the solutions that are currently in existence to the identified problem and offers some proposals on how they can be improved. In addition, it generally highlights the things that were achieved in the process of conducting the project.

b) Chapter 2: State of The Art: Content-based Image Indexing

Chapter 2 gives an overview of image indexing based on content and its existing issues. In addition, it presents a comprehensive review of the problems generally experienced and offers some understanding into the problems whose solutions were found. It offers a comprehensive coverage as far as all the solutions that have been researched for a given problem are concerned. Moreover, it presents the reasons why a specific solution was selected over the others.

c) Chapter 3: Low-Resolution Document Identification and Segmentation

Chapter 3 tells a comprehensive discussion on the implementation of the solutions that were suggested. Because this was primarily a software-based project, a good number of the implementations were also software based.

d) Chapter 4: Extraction of Low-Level Features

Chapter 4 details the implementation of the process to extract low-level features including enhancement of existing techniques.

e) Chapter 5: Proposed Scheme for Content-based Indexing

Chapter 5 details the proposed scheme for content-based indexing that was executed for the system. It surveys the existing methods of bridging the semantic gap using advanced search settings and iterative searching. It also reviews the existing proposals for content-based indexing and introduces a new scheme. Performance evaluation was the most vital part of the system. The main evaluation was performed on the precision and recall of the system.

f) Chapter 6: Conclusion and Future Research

Chapter 6 summarizes the entire project and offers a comprehensive discussion of the objectives that were realized in the project as well as those that were not realized. In addition, for future implementations and further improvements of content-based indexing of low-resolution documents were suggested.

g) Appendix

The appendix includes the table and information for running the project.

CHAPTER 2

State of The Art: Content-based Image Indexing

"If you think you are worth what you know, you are very wrong. Your knowledge today does not have much value beyond a couple of years. Your value is what you can learn and how easily you can adapt to the changes this profession brings so often."
– Jose M. Aguilar

Content:

Abstract:

Chapter 2 gives an overview of content-based image indexing and its current issues. Generally, it also defines the problems that have occurred and gives some understanding to the problems that have already been solved. On the other hand, this

chapter presents a comprehensive coverage on literature review regarding the solutions that have been addressed to solve a specific problem. This chapter also explains why a given solution has been selected over other solutions.

2.1 Introduction

In the literature, the term Content-Based Image Retrieval (CBIR) was used by Kato (1992) when describing his experiments on automatic retrieval of images stored in a database by use of colour and shape features. The use of CBIR differs from any classical information retrieval in the sense that image repositories are normally unstructured. This is because digital images consist of arrays of pixel intensities that have no inherent meaning (Eakins & Graham, 1999). Fundamentally, image databases differ from text databases, since the raw material for text databases are application words (stored in form of ASCII character strings) that are already logically structured (Santini & Jain, 1997). CBIR, which is also referred to as Content-based Visual Information Retrieval (CBVIR) and Query by Image Content (QBIC) is an (computer vision application) that is primarily used to offer solutions towards problem associate with image retrieval. In particular, the problem relates to searching for digital images found in large databases. ‘Content-based’ simply means that during a search, the actual image contents are analysed. In this context, the term ‘content’ is used to refer to the shapes (Kekre et al., 2011), colours (Wang & Wu, 2008); (Kekre, Mukherjee, Kakaiya, & Singh, 2010); textures (Shi Dong Cheng, Xu Lan, 2007); (Tamer Mehyar, 2012), (Zhang Dengsheng & Lu, 2004), or any information (Danzhou, Hua, & Vu, 2008), (Singh, 2012) which can be obtained from a given digital image itself. Without examining image content, it would be necessary for searches to depend on metadata such as keywords and captions, which can expensive or laborious to produce.

In CBIR systems, the visual contents the database’s images are extracted and described by multi-dimensional feature vectors. The feature vectors of the database’s images form a feature database. One widely-used model for a CBIR system is the vector space model as shown in Figure 2.1. In image retrieval vector space model, both the user queries and database images are mapped into points that are found in the vector (feature) space. The point's proximity in the vector space is considered to be an indication of similarity in the image.

Most CBIR systems either fit to this model as a whole or employ significant parts of it. The idea behind the vector space model is to represent each image as a point or more generally, a collection of points, in a vector space. When a distance metric is defined for the vector space, the similarity scoring problem reduces to a geometric problem, e.g. finding points closest to the point that is the representation of a given example image. When the retrieval problem is reduced into vector space formulation, it becomes largely similar to the general CBIR problem, and conceptually same techniques can be applied as for example in text retrieval (Moens, 2001)(Sembok, 2005). It is just the challenge to perform the reduction properly that makes image retrieval different from general information retrieval.

The vector space model is, once again, a simplification of the CBIR problem. Still, there seems to be a large enough body of issues to be examined even within this framework. The following sections will mostly be based on the vector space model, but sometimes issues not fitting to the model will be also discussed.

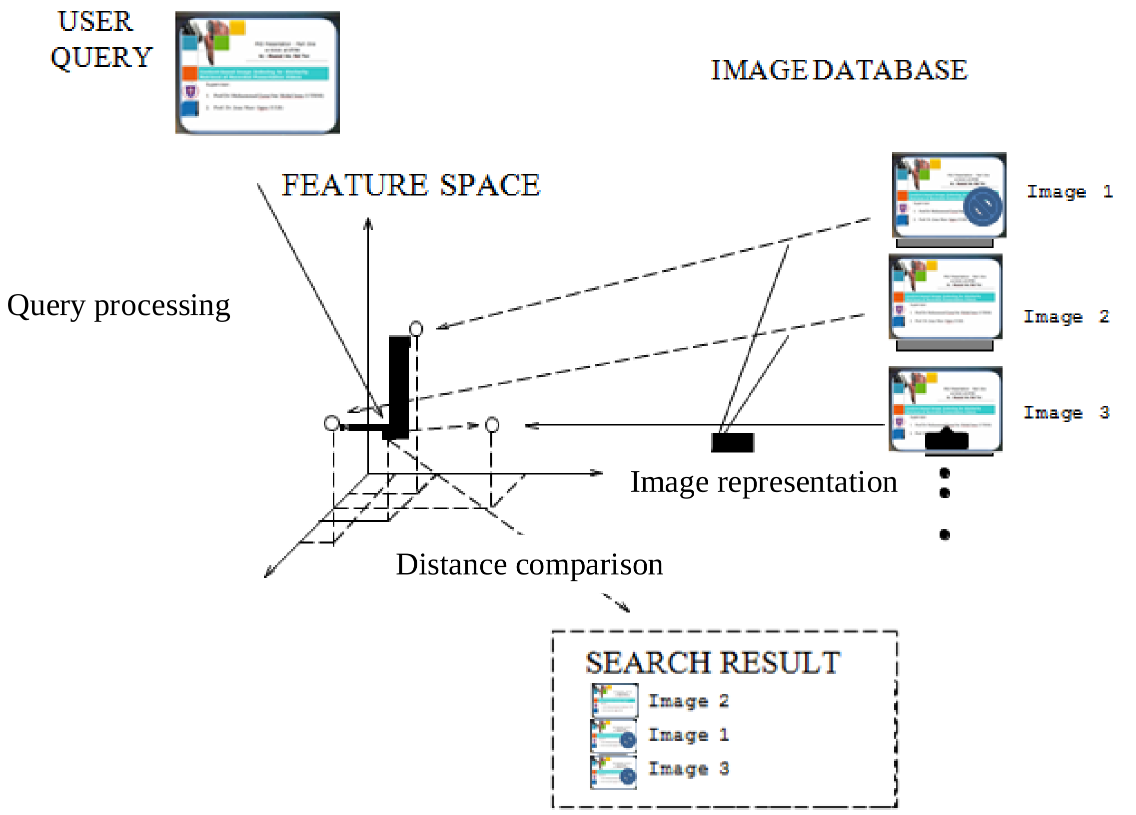

Figure 2.1: Model for a CBIR system is the vector space model

In order to be able to retrieve the required images, the users have to present the system with necessary sample figures (sketched) or images. The image retrieval system would then alter these samples into its feature vector's internal representation. The distances or similarities that exist between the images found in the database and the sketch's or query example's feature vectors are then calculated and the retrieval process is carried out with the help of an indexing scheme. In particular, the indexing scheme plays an essential role of providing an efficient way through the image on the database can be searched.

CBIR is gaining popularity simply because there is a rising need to search image databases, which are ever-growing size. Since precision and speed are imperative, it is important for us to come up with a system that would be used for image retrieval in effective and efficient manner. The arrival of multimedia technology and the fact that image and video collections are expanding rapidly on the Internet have attached research efforts significantly with an aim of providing tools that would be effective for retrieving and managing of visual data. It is vital to bear in mind that the foundation of CBIR is the existence of an image content representation scheme. Thus, one can make use of CBIR in order for images to be retrieved from the database only by using visual cues. The descriptors of an image may be visual features like, shape, colour, texture or any other type of feature that describes the context of the image context. This study aims at designing and implementing an effective and reliable framework as far as image retrieval techniques are concerned for documents with low-resolution, using various visual features like shape and colour. When a query image is made available to the CBIR system, the image features are usually obtained from it and then compared to those that are found in the database using similarity measure as the foundation.

2.2 Overview of CBIR

The amount of digital images produced by educational, industrial, medical, scientific, and other applications that are made available to users increased significantly in the early 1990s because of advances in the new digital image sensor and the technologies. The challenges that were being experienced by text-based image retrieval started to become more serious. This made the fast expanding visual information efficiency management to be an urgent problem, which made it necessary for CBIR techniques to be established. In the year 1992, the National Science Foundation (NSF) of the US organised for a workshop (Jain, 1993) that focused on systems used for visual information management in order to identify new directions as far as image database management systems are re concerned. It was recognized widely that a more intuitive and efficient way of representing and indexing visual information would be based on properties, which are inherent in nature in the images themselves. Researchers from the human-computer interface, the communities of computer vision, information retrieval and database management were attracted to this particular field. Ever since, CBIR research has developed significantly. Since the year 1997, the volume of research publications relative to organisation, indexing, database management, user query and interaction and the techniques of visual information extraction, has increased enormously. On the same note, various commercial and academic retrieval systems have been developed by hospitals, universities, companies and government organisations.

The need for CBIR is to retrieve more appropriate images, together with many features that would ensure a better image retrieval accuracy. Image Conference and Labs of the Evaluation Forum (CLEF), International Conference on Document Analysis and Recognition (ICDAR) and Moving Picture Experts Group (MPEG) are among the organisations actively promoting CBIR through journals and conferences.

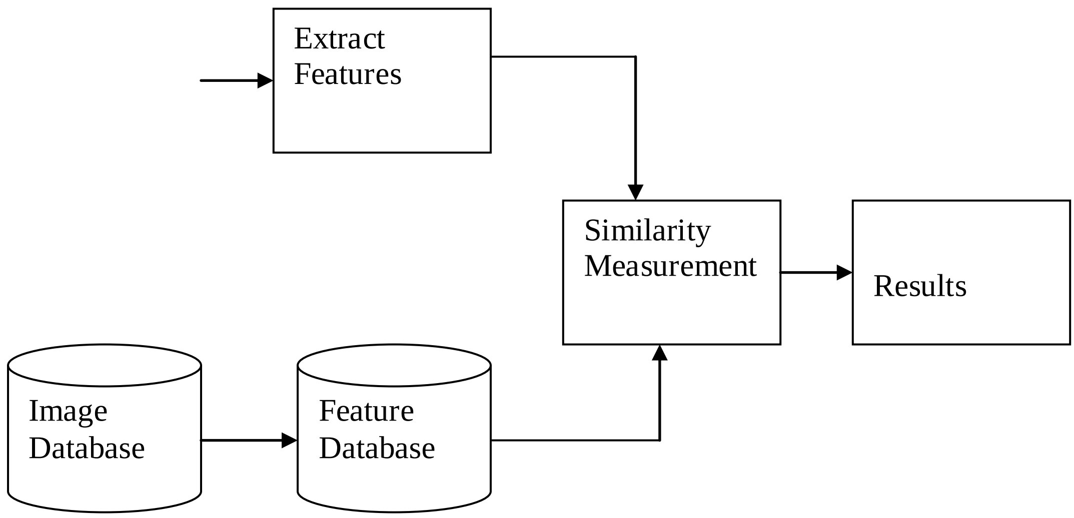

Figure 2.2a: General CBIR System.

‘Content-based’ in CBIR infers that during a search, the search uses image’s contents, instead of depending on human-inputted metadata like keywords or captions. In particular context, the term ‘Content’ is used to refer to texture, shapes and colours. Searches have depended on that the user provides, without the capacity to examine video content (Patel, 2012)(Chang et al., 2008). The similarities that exist between the visual features representation and the visual measurement are two vital concerns in CBIR. With a query image provided that has a single multiple object found in its this work’s mission would be to retrieve the same type of images found on the database based on the query features of the image. Figure 2.2 depicts (a) the CBIR system and (b) the multimedia archival and retrieval architecture.

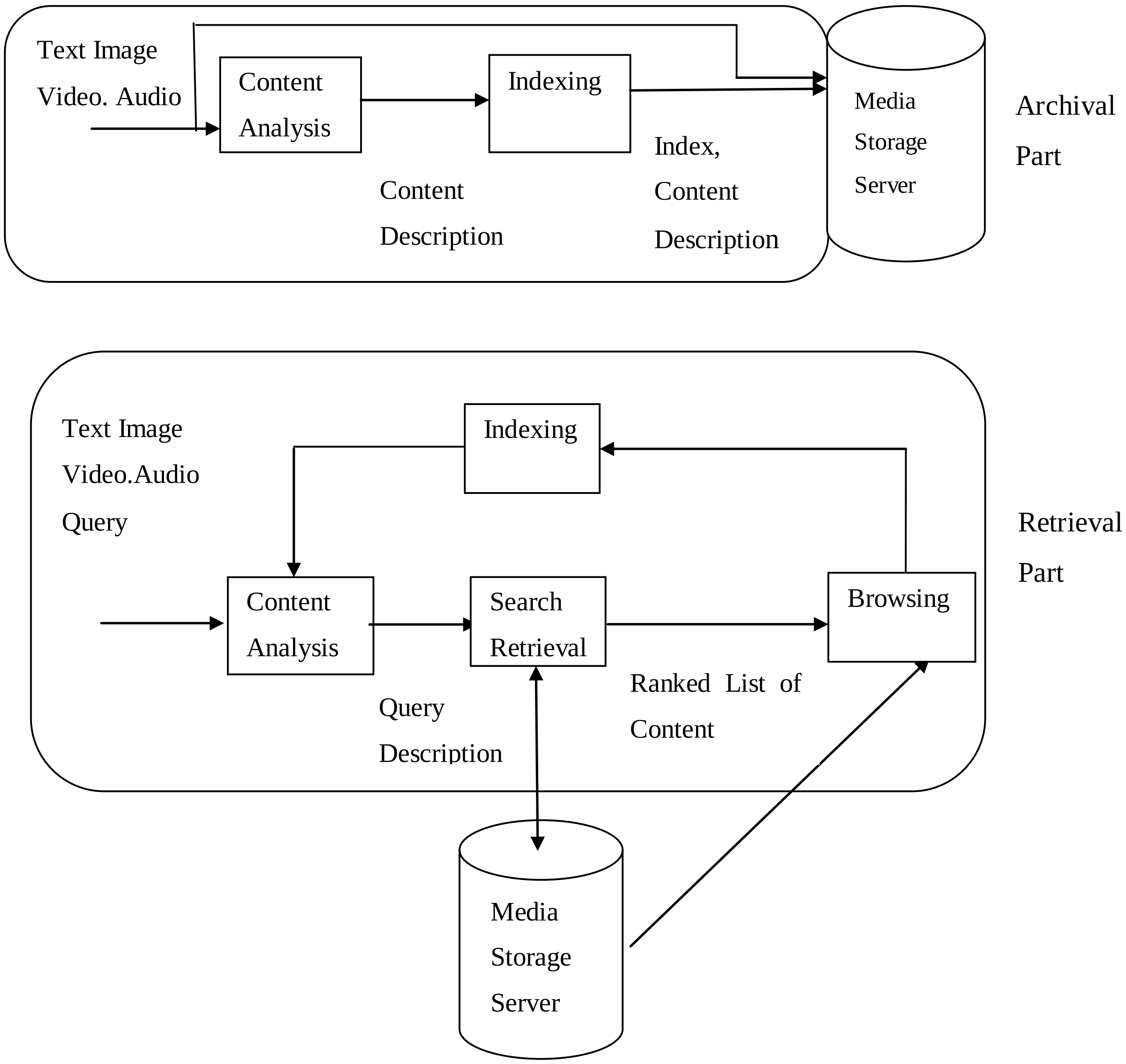

Figure 2.2b: Multimedia archival and retrieval architecture (Abdel-Mottaleb & Manor, 1998)

2.3 Related Work

On the issue of organizing images, humans have conventionally outperformed machines for many tasks. However, people made ambitious attempts over the past decade to make machines learn to comprehend, annotate, and index images depicting various concepts, with much progress. Likewise, because of the rising demand for accurate and fast information retrieval access, the literature proposes many CBIR systems (R. Zhang & Zhang, 2003b). In spite of the fact that CBIR is most preferred in numerous applications, it experiences a few issues including matching the images, extracting features from the images, which capture the semantic and perceptual as well as semantic meanings and image segmentation in a database that has a query image that is based on the features that have been extracted. The issue of extensive scale image databases that have effective searching and indexing, is still unsolved for CBIR (Balafar, 2012); (Datta et al., 2008).

Existing CBIR systems used for general purposes can be basically classified into three categories. They are categorized based on the approach used to extract the features of an image: histogram, the region-based search and colour layout (Behera et al., 2005)(Huiskes et al., 2005). These approaches have a common characteristic: the image objects are often represented as vectors of *d* numeric features and access them via the feature vectors and similarity measure. The feature vector dimensions of typical vector based descriptors are quite large. The high dimensionality of the feature vectors leads to high computational complexity in distance calculation for similarity retrieval, and inefficiency in indexing and search. However, colour histograms result in large feature vectors that suffer from high index and retrieval cost (Pabboju & Reddy, 2009). The number of bins in a typical colour histogram ranges from few tens to a few hundreds. To make the CBIR truly scalable to large size image databases, efficient multidimensional indexing techniques need to be explored (Lai, Visani, Boucher, & Ogier, 2012).

Recently, various methods have been suggested to provide solutions to these problems. The techniques are basically categorised into three classes (1) the multidimensional indexing approach, (2) the filter-based approach, and (3) the dimensionality reduction (DR) approach (Snoek & Worring, 2005).

Presentations recording technologies are currently becoming very much available. The trend as far as attendees to capturing pictures of slides that interest then during a presentation is also gaining popularity. The images that the attendees capture will become more useful in knowledge management, research and academics if it can be linked to the presentation videos recorder during a presentation. Relative to this, there exists a system (Lai et al., 2012)(Behera et al 2008)(Erol, Hull, & Lee, 2002) that makes it possible for users to submit their presentation slide images that have been taken using consumer devices in the form of queries to various presentation videos that have been recorded, and retrieve the presenter's audiovisual recording that talks about those presentation slides. Other issues apart from indexing issues also do exist, which include noisy, feature extraction of poor resolution, capture environment's varying lighting conditions and complex backgrounds. These issues also need to be addressed in order to ensure better representation of image features.

In a good number of areas of the government facilities, hospitals, academia and commerce, various digital images are being established. The majority of these collections are the product of diagrams, prints, paintings, drawings and digitising existing collections of analogue photographs. In most cases, the only way that was used for searching these collections was through browsing the digital image database or using keyword indexing. However, digital imaging database creates a way for content-based searching.

In content-based searching, an image content is provided and similar images whose content matches with the provided content are retrieved. There are various ways of specifying the images to retrieve from a given database that contain images. One way is to browse database of one image after another until you find the image that you want to retrieve. Specifying the image based on image features like colour histogram or keywords extracted from the image is another way of specifying the image to be retrieved. In addition, providing a sketch or an image where features are of the same kind have to be extracted can be also used for specifying the image to retrieve.

Various classes of features do exist that are used for the purpose of specifying queries. These features include, texture, spatial layout, faces, colour and shape. It is very easy to obtain colour features directly from the intensities of the pixel, e.g. colour histogram over a fixed subimage, over a segmented or over the whole image.

Despite the fact that it is not possible to trace a precise definition of texture, the notion of texture refers to the existence of a spatial pattern in general, which has some homogeneity properties. In specific, the homogeneity cannot exist because of the existence of one colour in the regions, but needs the interaction of different colours. A universal definition of what shape is either does not exist. Impressions of shape can be presented using intensity or colour patterns or colour by colour or intensity patterns, or texture, from which a geometrical representation can be derived.

2.4 CBIR Systems

Several CBIR systems have been suggested in recent years by researchers and the community for CBIR:

a) Chabot by University of California, Berkeley, USA (Ogle & Stonebraker, 1995)
b) Query by Image Content (QBIC) by IBM Almaden Research Center, San Jose, CA, USA (Flickher et al., 1995)
c) Photobook by Vision and Modeling Group, MIT Media Laboratory, Cambridge, MA, USA (Pentland & Picard, 1996)
d) WebSEEk by Image and Advanced Television Lab, Columbia University, New York (Chang, 1997)
e) Image Rover by Department of Computer Science, Boston University, MA (Sclaroff, Taycher, & Cascia, 1997)
VIR Image Engine by VirageInc (Gupta, 1996)
f) WebSeer by Department of Computer Science, University of Chicago, Illinois, USA (Swain & Frankel, 1996).
g) VisualSEEk by Image and Advanced Television Lab, Columbia University, NY, USA (Smith, 1997)
h) NETRA by Department of Electrical and Computer Engineering, University of California, Santa Barbara, USA (Manjunath & Ma, 1997)

Chabot (Ogle & Stonebraker, 1995)(Antani, Kasturi, & Jain, 2002) is considered to be among the early systems. The system aimed at utilizing both image analysis and text based descriptions in the process of retrieving images from a large volume of photographs belonging to the California Department of Water Resources. The utilized an existing database's text description of the collection, adding other kinds of textual information for querying like the picture location, the perspective of the photo and shooting date. For every image a colour histogram that has only a total 20 bins is usually computed. For the purpose of querying, the user is usually offered a list of search criteria like photographer, film format, location, shooting date, colours, Keywords and perspective). The images and their associated data are normally stored in the database known as POSTGRES, created at the University of California, Berkley.

Image Rover (Sclaroff et al., 1997)(Chadha & Carolina, 2012), which is one of the image search engines used with the WWW. For querying purposes, the features used include text orientation and colour. Image colour histograms are normally computed in the CIE Luv colour space and every histogram quantises the colour space into 64 bins. The calculation of texture direction distribution using is normally done using steerable pyramids. At each level of the pyramid, texture strength and direction for every pixel is calculated resulting in an orientation histogram, quantised to 16 bins. The user can begin a query session through specifying a collection of keywords that are related to the preferred images. From the collection of images displayed, the user can identify and mark more than images that are the same to what she/he is searching for. Apart from choosing relevant images, the user can choose not to select one or more query images before actually reiterating a new query. It should be noted that there is no limitation to the number of iterations required in providing the relevant feedback, nor in the quantity of the example images. For search algorithm and indexing purposes, an optimised kD tree is normally used in this structure, which makes use of approximation factor. The user can choose through the search accuracy button one of the three values for this approximation factor.

The NETRA system (Manjunath & Ma, 1997) is made up of 2,500 images from the Corel photo gallery. The images were arranged in 25 images categories, with each image category made up of 100 images. A user can select any of them and use it as the query image. It is imperative to note that all the database's images have been segmented into regions that are homogeneous in nature. Of those regions, the features extracted include texture, spatial location, shape and colour. A user can click on any of the regions and choose one image attribute from the four image features/attribute. Rather than making use of an image example, it is also possible for one to directly specify the spatial location and colour. Based on training a collection of images, the RGB colour space is usually quantised, and represented using a colour codebook of 256 colours, the centroids of the quantisation cells. On the other side, the representation of texture is done using a feature vector that has the standard deviation and the normalised mean of a collection of Gabor wavelet transforms of the image. The foundation of the shape feature is curvature function of the contour centroid distance function, complex coordinate function.

The Photobook system (Pentland & Picard, 1996) is simply a collection of interactive tools that are used searching and browsing digital images. The main idea behind these tools is semantics that are responsible for preserving images. The main reason for preserving the images is in order to see into it that an image is reduced to a relatively smaller collection of perceptually important coefficients. Photobook is made up of three main sub-books, which include texture, shape and appearance photobook that are responsible for extracting texture, shape and face respectively when making use of then system, a user can create a query that queries an image based on the corresponding feature of every sub-book. On the same note, a user can query an image using a combination of various mechanisms that have text-based descriptions. It would be noted that other CBIR systems enable users to query an image based on global shape, colour, texture, semantic content spatial relationship and colour layout of the regions of the image. Despite the fact that the systems offer various image querying features, none of them has the capabilities to combine qualitative spatial relation, colour region, global colour, shape, and colour sensation features .

Query By Image Content (Flickher et al., 1995) was the first commercial system to be developed by Flickner. With this system, the colour features computed include a

256 dimensional RBG and the Munsell colour space histogram, the 3D average colour vector of the whole image in RBG or an object. The texture features made use of in this system are simply the modified versions of contrast, directional and coarseness features that were proposed by Tamura (H. Tamura, S. Mori, T. Yamawaki, 1978) . The shape features are made up of a collection of algebraic moment invariants, major axis orientation, circularity and shape area. The system makes it possible for users to query an image based on an existing image example, texture patterns and sketches constructed by the user. For texture patterns, a user may select textures or colours from a sampler. An image's desired colour percentage is normally adjusted through moving the sliders. The system was among the first CBIR systems that made use of multidimensional indexing to in order to see into it that the system's speed performance is enhanced.

VIR Image Engine system (Gupta, 1996) is basically extensible framework as far as building CBIR systems is concerned. It offers a collection of general primitives like shape, texture, local and global colour. Beside these general primitives, it is possible to create primitives that are domain specific in nature, in the process of developing the system. In the process of defining the primitives that are domain specific in nature, the developer simply supplies functions that are used for computing the feature data of the primitives, from the raw image. The system offers a collection of Graphical User Interface (GUI) tools that are needed when developing a user interface. These tools include facilities to support various common image file formats, inclusion of keywords, weight adjustment for the purpose or re-querying, image query and insertion. It is vital to note that these GUI facilities were integrated into the CBIR system Sybase. The Illustra's Visual Intelligence system and the AltaVista Photofinder are two Virage technology applications.

VisualSEEk CBIR system (Smith, 1997) has the capabilities to support spatial location and colour features. During the process of populating the database, every image is normally decomposed automatically into regions that have equal dominant colours. The spatial and feature properties are retained for the preceding queries in every region. It should be noted that a query is responsible for finding those images that have the same organizations of the same regions. In order to initiate a query, a user has to sketch various regions and then position and dimension them the existing grid. The user then selects a colour for every region. In addition, in the process of initiating a query, it is possible for a user to indicate the spatial relationships that exist between the regions, and/or the boundaries for size and location. After the system has returned the best matches' thumbnail images, the user would then be allowed to search by examples with the help of the returned images. It should be beared in mind that R-tree is used for the purpose of indexing the maximum bounding rectangles.

The Excalibur Visual Retrieval Ware SDK (Software Developers Kit) (Excalibur Technologies, 1997) is one of the open CBIR systems. Excalibur Corp is responsible for distributing the system's development environment. The SDK is toolkit that is usually used by a developer, which offers a collection of development tools that are used for building content-based image software. The development tools found in SDK includes reference documentation, sources code, sample programs, Tk/Tcl interpreter, Java, C++, C and three API's. The toolkit is made up of various examples, whose primary purposes is to demonstrate what a developer can be able to build with the help of the SDK. Among other sample programs, one of the sample programs found in the SDK is the CST (colour, shape, texture) demo. This particular sample program makes it possible for queries by example that based on curvature and contrast in the image, texture attributes that measure the roughness and flow in the image, relative orientation and HSV (hue-saturation-value) colour histograms to be used. During the initial step, the user begins by defining the desired visual similarity through specifying the relative importance of the above attributes of the image. The user then selects one displayed images as an image query. The system has been made use of in the Image Surfer system that was used by search engines such as the Infoseek and Yahoo!.

WebSEEk (Smith, 1997) is a CBIR system that was developed at Image and Advanced Television Lab, Columbia University, New York. The system makes colour and text-based queries with the help of videos and a catalogue of images obtained from the Internet. In order to initiate a query, a user has to enter a topic or select one from the available catalogue. The results that would be obtained from the query may be used for organizing the result list through decreasing colour similarity to the selected item or for a colour query in the whole catalogue. When using the system, it is possible for a user to select negative and positive examples from the query's result in order to be able to reformulate the query.

Finally, WebSeer (Swain & Frankel, 1996) as one of the CBIR systems is a system developed at the Department of Computer Science, University of Chicago, Illinois (Yang, 2004). When making use of this system, the images obtained from the Internet are usually submitted to various colour tests in order for the photographs to be separated from the drawings. Some simple tests measure the ratio of the image dimensions, the fraction of pixels which have one of the N most frequent colours for a given threshold N, the number of different colours in the image, the fraction of pixels with transition value greater than a threshold T and the fraction of pixels with saturation level greater than a threshold T. for a test that is more elaborate, the test first creates the graphics' average colour histogram, Hg, and one for photographs, Hp, using two large images sets. The user then gives some keywords that describe the desired images' content, and specifies (optional) some image characteristics like file size, dimensions. The keywords also describe what the user is looking for drawings or photographs. In cases where a user is looking for people, the user will have to indicate the number of faces and the size of the portrait. In order to classify an image as graphics or photograph, the colour tests are usually combined with the help of various decision trees that have been constructed with the help of a training set of images that are hand-classified.

2.5 Content-based Indexing: Previous Works

This section presents a literature review of most of the significant works that address the issues related to meeting/lecture analysis and retrieval.

2.5.1 Meeting and Lectures Analysis and Retrieval

Several research groups have studied the problem of analysing and indexing the multimodal information captured during meetings presentation/lectures (e.g. video, voice, whiteboard), which would allow searching and browsing at a later time point once the recording is completed. Analysis of the recorded multimodal meetings presentation/lectures materials takes place either during or after recording. The goal of the analysis for the captured multimedia streams is to add the structural information to the captured unstructured data. The structural information provides indices to access the captured multimodal information in a random manner. The robustness of user access to the captured data is highly dependent on the granularity and meaningfulness of the indexing during the analysis.

Most of the multimodal meeting/lecture capturing, indexing, annotation and retrieval systems can be conceptually described by the framework illustrated in Figure 2.3. The existing research projects, which capture and analyse meetings presentation/lectures, can be classified into application-specific groups and are based on the target areas. Furthermore, the analysis and indexing approach of each application is carried out using the captured media streams, which could be further categorised as sensor-specific indexing.

2.5.2 Classification Based on Application

There are two categories into which the existing research projects for the analysis and recording of meetings can be classified based on their application. These include industrial and academic classification.

This research examines the exclusive use of the retrieved visual content in the academic and industrial arena. In this setting, the information recording and analysis during meetings and lectures presentations is done on a visual display.

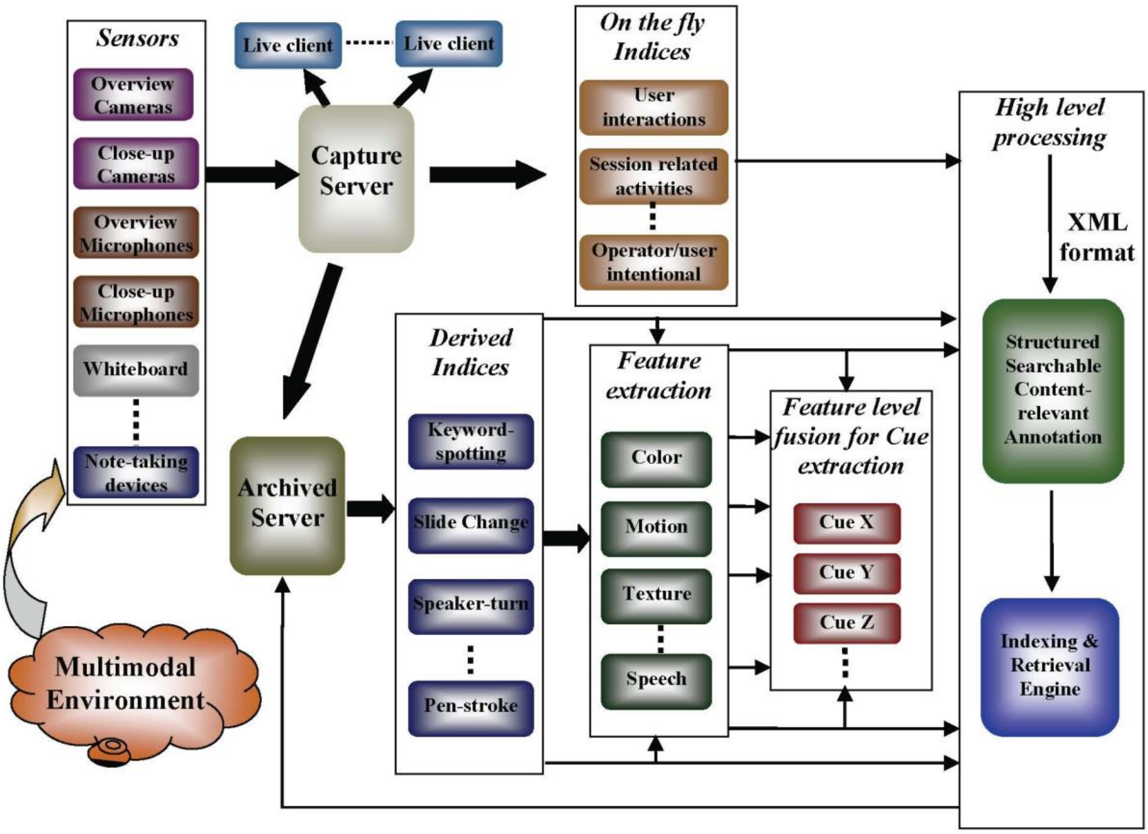

Figure 2.3: A simple model of a multimodal meeting/lecture capturing, indexing and annotation system. (R. Zhang & Zhang, 2003a)

2.5.2.1 Academic Applications

A cursory look at the academic application that has now transformed into e-learning platform indicates that many schools and teachers have adopted the use of computers (soft-copy material) for lectures as a new way of teaching in classrooms (Baptista-nunes, Court, & Mcpherson, 2002)(Brotherton & Abowd, 2004), There is modern e-library emerging in the academic setting, with examples such as, Georgia Technical that has adopted E-classes and projector presenter in University of Washington classes among others. Students capture the audio-visual lesson content flow from the instructor based on their course. It is this data from the subject and lecture obtained and stored in the repository, and in case need to refer arises, students go through captured audio and visual presentation slides material. One of the notable aspects of

REFERENCES

A. Kapuskar, et al. (2007). *Room With A Rear View: Meeting Capture in a Multimedia Conference Room*. IEEE Multimedia 7(4).

Abdel-Mottaleb, M., & Manor, B. (1998). *A Scalable Algorithm For Image Retrieval Bycolor*. International Conference on Image Processing, 119–122.

Abowdl, D., Atkeson, G., Koope, R., & Long, S. (1996). *Teaching and Learning as Multimedia Authoring : The Classroom 2000 Project*. In In Proc. of ACM Multimedia.

Aghajari, G., Shanbehzadeh, J., & Sarrafzadeh, A. (2010). *A Text Localization Algorithm in Color Image via New Projection Profile*. International MultiConference of Engineers and Computer Scientists, *II*, 17–20.

Akram, S., & Dar, M. (2010). *Document Image Processing - A Review. 10*(5), 35–40.

Amir, A., Srinivasan, S., & Efrat, A. (2003). Search the Audio , Browse the Video — A Generic, 209–222.

Antani, S., Kasturi, R., & Jain, R. (2002). *A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video*. Pattern Recognition, *35*(4), 945–965.

Anthimopoulos, M., Gatos, B., & Pratikakis, I. (2008). The Eighth IAPR Workshop on Document Analysis Systems *A Hybrid System for Text Detection in Video Frames*. In IAPR on Document Analysis Systems (pp. 286–292).

Babu, R. S., Vengattaraman, T., Baskaran, R., & Dhavachelvan, P. (2010). *Intelligent Layout Based Retrieval from Document Images*. In International Conference on Information Science an Application (pp. 2–5).

Bakken, T. (2007). *An evaluation of the SIFT algorithm for CBIR*. Telenor R&I Research Note, (0809–1021).

Balafar, M. A. (2012). *Application Of Content Based Image Retrieval In Diagnosis*

Brain Disease. International Journal on Technical and Physical Problems of Engineering, (December), 133–138.

Baptista-nunes, M., Court, R., & Mcpherson, M. (2002). *No Lectures On-Campus: Can eLearning Provide a Better Learning Experience? Action Research into eLearning*. International Conference on Advanced Learning Technologies, 442–447.

Beckmann, N., & Seeger, B. (2009). *A Revised R * -tree in Comparison with Related Index Structures*. Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data, 799–811.

Behera, A., D. Lalanne, et al. (2005). *Combining color and layout features for the identification of low-resolution documents*. Int'l Journal of Signal Processing.

Behera, A., Lalanne, D., & Ingold, R., Lalanne, D., & Ingold, R. (2005). *Enhancement of layout-based identification of low-resolution documents using geometrical color distribution*. Eighth International Conference on Document Analysis and Recognition (ICDAR'05), *2*(1), 468–472 Vol. 1.

Behera, A., Lalanne, D., & Ingold, R. (2008). *An automatic document-based indexing system for meeting retrieval*. Multimed Tools Appl.

Behera, A. (2006). *A Visual Signature-based Identification Method of Low-resolution Document Images and its Exploitation to Automate Indexing of Multimodal Recordings*. Orissa , India.

Behera, A., Lalanne, D., & Ingold, R. (2004a). *Looking at projected documents: event detection & document identification*. Multimedia and Expo, 2004. …, *2*(1), 2127–2130.

Behera, A., Lalanne, D., & Ingold, R. (2004b). *Visual signature based identification of Low-resolution document images*. Proceedings of the 2004 ACM Symposium on Document Engineering - DocEng '04, 178.

Behera, A., Lalanne, D., & Ingold, R. (2005). *Combining Color and Layout Features for the Identification of Low-resolution Documents*. International Journal of Signal Processing, 2(1), ISSN: 1304-4478: 7-14.

Behera, A., Lalanne, D., & Ingold, R. (2007). *DocMIR: An automatic document-based indexing system for meeting retrieval*. Multimedia Tools and Applications (Vol. 37).

Behera, a., Lalanne, D., & Ingold, R. (2005). *Enhancement of layout-based* *identification of low-resolution documents using geometrical color distribution*. Eighth International Conference on Document Analysis and Recognition (ICDAR'05), 468–472 Vol. 1.

Beusekom, J. Van, Keysers, D., Shafait, F., & Breuel, T. M. (2006). *Distance measures for layout-based document image retrieval*. Second International Conference on Document Image Analysis for Libraries (DIAL'06), 232–242.

Bhute, A. & Meshram, B. (2013). *Content-based Image Indexing and Retrieval.* International Journal of Graphics & Image Processing, 3(4).

Bhute, A. N. (2012). *Multimedia Indexing and Retrieval Techniques : A Review*. International Journal of Computer Applications (0975-8887), 58(3), 35–42.

Brotherton, J. A., & Abowd, G. D. (2004). *Lessons Learned From eClass: Assessing Automated Capture and Access in the Classroom*. ACM Transactions on Computer-Human Interaction, 11(2), 121–155.

Brunner, G. (2007). *The Structure Features for Content Based Image Retrieval and Classification Problems*. S.L: S.N. Feature-Based Identification of Low-Resolution Documents. In Proc. of ACM Symposium on Document Engineering, 20–22.

Chadha, A., & Carolina, N. (2012). *Comparative Study and Optimization of Feature-Extraction Techniques for Content based Image Retrieval.* International Journal of Computer Applications (0975-8887), *52*(20), 35–42.

Chang, R., Lin, S.-Y., Ho, J.-M., Fann, C.-W., & Wang, Y.-C. (2012). *A novel content based image retrieval system using K-means/KNN with feature extraction*. Computer Science and Information Systems, 9(4), 1645–1661.

Chang, S.-F., Smith, J. R., Beigi, M., & Benitez, A. (1997). *Visual information retrieval from large distributed online repositories.* Communications of the ACM, 40(12), 63–71.

Chang, Y., Lee, D. J., Hong, Y., & Archibald, J. (2008). *Unsupervised video shot detection using clustering ensemble with a color global scale-invariant feature transform descriptor*. Eurasip Journal on Image and Video Processing, 2008.

Chatterjee, K., & Chen, S. (2006). *Affinity Hybrid Tree: An Indexing Technique for Content-Based Image Retrieval in Multimedia Databases*. Eighth IEEE International Symposium on Multimedia (ISM'06), 47–54.

Chen, Y.-L., & Wu, B.-F. (2009). *A multi-plane approach for text segmentation of* *complex document images*. Journal of Pattern Recognition, 42(7), 1419–1444.

Chiu, P., Foote, J., Girgensohn, A., & Boreczky, J. (2000). *Automatically Linking Multimedia Meeting Documents by Image Matching.* In In Proc. of ACM Hypertext (pp. 244–245).

Chiu, P., & Kapuskar, A. (2000). *Room with a rear view: Meeting capture in a multimedia conference room.* In IEEE Multimedia (Vol. 7, pp. 48–54).

Choraś R.S. (2015). *Image processing and communications challenges*. Springer Science and Business Media.

Cooper, M. (2013). *Presentation video retrieval using automatically recovered slide and spoken text*, Retrived Mac 2014 http://lucene.apache.org

Cutler, R., Rui, Y., Gupta, A., & Cadiz, J. (2002). *Distributed meetings: A meeting capture and broadcasting system*. Proceedings of ACM Multimedia, 503–512.

Danzhou, L., Hua, K., & Vu, K. (2008). *Fast Query Point Movement Techniques for Large CBIR Systems*, Journal IEEE Transactions on Knowledge and Data Engineering 1–14.

Deselaers, T., Keysers, D., & Ney, H. (2007). *Features for Image Retrieval : An Experimental Comparison*. Information Retrieval. Springer.

Doermann, D., H. Li, et al. (1997). *The detection of duplicates in document image databases*. *I*n Proc. of ICDAR, (August).

Eakins, J., & Graham, M. (1999). *Content-based Image Retrieval*. In A report to the JISC technology applications programme. Institute for Image Data Research, University of Northumbria.

Eglin, V., & Bres, S. (2003). *Document page similarity based on layout visual saliency: application to query by example and document classification*. In Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings. Retrieved May 2014 from http://doi.org/10.1109/ICDAR.2003.1227849

Erol, B., Hull, J. J., & Lee, D. (2002). *Linking Multimedia Presentations with their Symbolic Source Documents : Algorithm and Applications*. In In Proc. of ACM Multimedia (pp. 498–507).

Excalibur Technologies. (1997). *Visual Retrieval Ware*. SDK.

Flickher, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Lee, D. (1995). *Query by image and video content: The QBIC system*. In IEEE

Computer, 38, 23–31.

Fu, L., Meng, Y., Xia, Y., & Yu, H. (2010). *Web Content Extraction based on Webpage Layout Analysis*. International Conference on Information Technology and Computer Science, 40–43.

Garcia, C., & Apostolidis, X. (2000). *Text detection and segmentation in complex color images*. In: Proc. IEEE Int, 2(1), 6–9.

Geetha, P., & Narayanan, V. (2008). *A Survey of Content-Based Video Retrieval*. Journal of Computer Science, 4(6), 474–486.

Geyer, W., & H. Richter et al. (2001). *A team collaboration space supporting capture and access of virtual meetings*. In Proc. of ACM SIGGROUP Conf. on Supporting Group Work, 188–196.

Golub, G. H., & Strebel, R. (1994). *Least-Squares Fitting of Circles and Ellipses*. BIT, 34, 558–578.

Gonzalez, R. C. ., & Woods, R. E. (2004). *Digital Image Processing*. Pentice Hall.

Gupta, A. et al. (1996). *The Virage image search engine: an open framework for image management*. In In Storage and Retrieval for Image and Video Databases IV, Proc SPIE 2670 (pp. 76–87).

H. Tamura, S. Mori, T. Yamawaki. (1978). *Textural Features Corresponding to Visual Perception*. IEEE Transactions. On Systems, Man and Cybern, *Vol. 8*, 460–472.

Hai-tao, W. (2008). Image Retrieval Based on Bit-plane Distribution Entropy, *1*, 532–535. Retrieved April 2013, http://doi.org/10.1109/CSSE.2008.270

Hanif, S. M., Prevost, L., Syst, I., & Robotique, D. (2009). *Text Detection and Localization in Complex Scene Images using Constrained AdaBoost Algorithm*. International Conference on Document Analysis and Recognition.

He, D. (2010). *Three New Methods For Color And Texture Based Image Matching In Content-Based Image Retrieval*. Dalhousie University: Ph.D. Thesis.

Hirano, T., Okano, Y., Okada, Y., & Yoda, F. (2007). *Text and Layout Information Extraction from Document Files of Various Formats Based on the Analysis of Page Description Language*, ICDAR.

Hsiao, M., Huang, Y., Tsai, T., & Chiang, T. (2010). *An Efficient and Flexible Matching Strategy for Content-based Image Retrieval*, Life Science Journal 7(1).

Hu, X., Gao, L., Lin, X., Tang, Z., Lin, X., & Baker, J. (2013). *WikiMirs: A* *mathematical information retrieval system for wikipedia*. Proceedings Of The 13Th ACM/IEEE-CS Joint Conference On Digital Libraries - JCDL '13.

Hu.M.K. (1962). *Visual pattern recognition by moment invariants*. IRE Transactions on Information Theory.

Hu, J. H. J., Kashi, R., & Wilfong, G. (1999). *Document image layout comparison and classification.* Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR '99 (Cat. No.PR00318), (September 1999), 285–288.

Hu, W., Member, S., Xie, N., Li, L., Zeng, X., & Maybank, S. (2011). *A Survey on Visual Content-Based Video Indexing and Retrieval, 41*(6), 797–819.

Hull, & J.J. (1994). *Document image matching and retrieval with multiple distortion-invariant descriptors*. In In Proc. of Document Analysis System (DAS) (pp. 383–400).

Hunter, & J. and S. Little. (2001). *Building and indexing a distributed multimedia presentation archive using SMIL*. In In Proc. of ECDL.

IDIAP (2009). *Smart Meeting Room*.

Jain, R. (1993). *NSF*. Workshop on Visual Information Management Systems Editor : Ramesh Jain.

K. Hadjar, M. Rigamonti, D. L. and R. I. (2004a). Xed: *A new tool for eXtracting hidden structures from Electronic Documents*. In Proc. of Int'l Workshop on Document Image Analysis for Libraries (DIAL), (pp. 212–224).

K. Hadjar, M. Rigamonti, D. L. and R. I. (2004b). *Xed: A new tool for eXtracting hidden structures from Electronic Documents*. In In Proc. of Int'l Workshop on Document Image Analysis for Libraries (DIAL) (pp. 212–224).

Kato, T. (1992). *Database architecture for content-based image retrieval.* In Proc. of SPIE: Image Storage and Retrieval Systems1662.

Kekre, H. B., Mukherjee, P., Kakaiya, M., & Singh, S. (2010). *Image Retrieval with Shape Features Extracted using Morphological Operators with BTC, 12*(3), 2–6.

Kekre, H. B., Thepade, S., Kakaiya, M., Mukherjee, P., Singh, S., & Wadhwa, S. (2011). *Image Retrieval using Shape Texture Content as Row Mean of Transformed Columns of Morphological Edge Images.* International Journal of Computer Science and Information Technologies, 2(2), 641–645.

Kise, K., Tsujino, M., & Matsumoto, K. (2005). *Retrieval of Relevant Parts of*

Document Images Based on 2D Density Distributions of Characters. IEEJ Transactions on Electronics, Information and Systems, 125(1), 113–119.

Kokare, M. B., & A, B. L. D. E. (2010). *Document Image Retrieval : An Overview*. Life Science Journal 1(7), 114–119.

Kumar, V. V., Rao, N. G., Rao, A. L. N., & Krishna, V. V. (2009). *IHBM : Integrated Histogram Bin Matching For Similarity Measures of Color Image Retrieval.* International Journal of Signal Processing, Image Processing and Pattern Recognition, 2(3), 109–120.

Lai, H. P., Visani, M., Boucher, A., & Ogier, J.-M. (2012). *An experimental comparison of clustering methods for content-based indexing of large image databases*. Pattern Analysis and Applications, 15(4), 345–366.

Lalanne, D., Lisowska, A., Bruno, E., & Flynn, M. (2005). *The IM2 multimodal meeting browser family* . In Interactive Multimodal Information Management Tech. Report, IDIAP, Margtiny, Switzerland. (pp. 7–14).

Lalanne, D., Sire, S., Ingold, R., Behera, A., Mekhaldi, D., & Rotz, D. (2003a). *A research agenda for assessing the utility of document annotations in multimedia databases of meeting recordings.* 3rd International Workshop on Multimedia Data and Document Engineering, in Conjunction with VLDB, *2*(1), 47–55. Lalanne, D., Sire, S., Ingold, R., Behera, A., Mekhaldi, D., & Rotz, D. Von. (2003b). *A research agenda for assessing the utility of document annotations in multimedia databases of meeting recordings* . Proceeding oh the 29th VLDB Conference.

Liew, Guo Min. (2008). *Slide Image Retrieval : A Preliminary Study.* JCDL 08, June 16-20(ACM 1-58113-000-0/00/0008).

Lin, X., Gao, L., Hu, X., Tang, Z., Xiao, Y., & Liu, X. (2014). *A Mathematics Retrieval System for Formulae in Layout Presentations*. Proceedings Of The 37Th International ACM SIGIR Conference On Research & Development In Information Retrieval - SIGIR ’14.

Liu, G.-H., & Yang, J.-Y. (2013). *Content-based image retrieval using color difference histogram*. Pattern Recognition, 46(1), 188–198.

Liu, Z., Xue, Y., Shen, L., & Zhang, Z. (2010). *Nonparametric saliency detection using kernel density estimation*. International Conference on Image Processing, 253–256.

Liwicki, M., & Bunke, H. (2008). IAM-OnDB - *An On-Line English Sentence Database Acquired From Handwritten Text On A Whiteboard*. In Proc. of ICDAR, 2(2), 956–961.

Ma, D., & Agam, G. (2013). *A super resolution framework for low resolution document image OCR*. 8658(1), 86580P. Retrieved May 2012 http://doi.org/10.1117/12.2008354

Manjunath, B. S., & Ma, W. Y. (1997). *Netra : A Toolbox for Navigating Large Image Database*. In Proceedings of IEEE International Conference on Image Processing, 568–571.

Mar, P., Oria, V., & Rukoz, M. (2005). *Multi-level index for global and partial content-based image retrieval*. In 21st International Conference on Data Engineering.

Marinai, S., Marino, E., Soda, G., & Marta, V. S. (2006). *Tree clustering for layout-based document image retrieval.* In Document Image Analysis for Libraries.

Martin, T., A.Boucher, & et al. (2006). *Multimodal Interactions For Multimedia Content Analysis.* In ICTACS.

Matusiak, K. K. (2006). *Towards user-centered indexing in digital image collections*, Retrieved May 2012 http://doi.org/10.1108/10650750610706998

Mejdoub, M., Fonteles, L., BenAmar, C., & Antonini, M. (2009). *Embedded lattices tree: An efficient indexing scheme for content based retrieval on image databases.* Journal of Visual Communication and Image Representation, 20(2), 145–156.

Moens, M. F. (2001). *Innovative techniques for legal text retrieval*. Artificial Intelligence and Law, 9(1), 29–57.

Moghaddam, B. (2000). *Regions-of-Interest and Spatial Layout for Content-Based Image Retrieval.* Mitsubishi Electric Research Laboratory.

Mostefai, S., & Jeyasekhar, S. (2012). *Improved CBIR through Combined Indexing Technique*. International Journal of Computer Theory and Engineering, 4(2), 217–219.

Mukhopadhyay, & S. and B. Smith. (1999). *Passive capture and structuring of lectures.* In Proc. of ACM Multimedia, (Oktober-November), 477–487.

Munesawang, P., & Ling, G. (2005). *Adaptive Video Indexing and Automatic/Semi-Automatic Relevance Feedback.* IEEE Transactions On Circuits And Systems

For Video Technology, 15(8), 1032–1046.

Mustaffa, M. R., Ahmad, F., Rahmat, R. W. O. K., & Mahmod, R. (2008). *Content-based image retrieval based on color-spatial features*. Malaysian Journal of Computer Science, 21(1), 1–12.

Nguyen, N. Van, Ogier, J., & Charneau, F. (2012). *PEDIVHANDI : Multimodal Indexation and Retrieval System for Lecture Videos.* In ACCV'12 Proceedings of the 11th Asian conference on Computer Vision (pp. 382–393).

Ogle, V. E., & Stonebraker, M. (1995). *Chabot: retrieval from a relational database of images.* Computer, 28(9), 40–48.

Otsu, & N. (1996). *A threshold selection method from gray-level histograms.* IEEE Trans. on Systems, Man and Cybernetics, 9(1), 62–66.

Ozawa, N., & H. Takebe et al. (2004). *Slide identification for lecture movies by matching characters and images.* In Proc. of SPIE-Document Recognition and Retrieval XI5296.

Papaeti, A., & Grant, M. J. (2013). *Similarity Searching in Large Image Databases*. Torture : Quarterly Journal on Rehabilitation of Torture Victims and Prevention of Torture, 23(2), 1–3. Retrieved July 2012 from http://www.ncbi.nlm.nih.gov/pubmed/24482839

Pardhi, P. D., Paikrao, P. L., & Chaudhari, D. S. (2012). *Introduction to Query Techniques for Large CBIR Systems*. International Journal of Soft Computing and Engineering, (1), 181–184.

Patel, B. V. (2012). *Content Based Video Retrieval Systems*. International Journal of UbiComp, 3(2), 13–30.

Pentland, A., & Picard, R. (1996). *Photobook : Content-based manipulation of image databases* . Intl. Journal of Computer Vision, 18(3), 233–254.

Poullot, S., Cnam, C., & Crucianu, M. (2009). *Indexing Local Configurations of Features for Scalable Content-Based Video Copy Detection*. ACM 978-1-60558-756-1/09/10, 43–50.

Ramanathan, S., & Kankanhalli, M. (2009). *A robust framework for aligning lecture slides with video*. 2009 16th IEEE International Conference on Image Processing (ICIP), 249–252.

Rao, N. G. (2009). *Texture Based Image Indexing and Retrieval*. International Journal of Computer Science and Network Security, 9(5), 206–210.

Rui, Y., Gupta, A., Grudin, J., & He, L. (2004). *Automating lecture capture and broadcast: technology and videography*. ACM Multimedia Systems, 10, 3–15.

Saidane, Z., & Garcia, C. (2008). *An Automatic Method for Video Character Segmentation. An Automatic Method for Video Character Segmentation*, 557–566.

Schellenberg, T. (2011). *Layout-Based Substitution Tree Indexing and Retrieval for Mathematical Expressions*. Ph.D. Thesis.

Sclaroff, S., Taycher, L., & Cascia, M. La. (1997). *ImageRover : A Content-Based Image Browser for the World Wide Web 1 Introduction 2 Approach*. In Proc. IEEE Workshop on Content-based Access of Image and Video Libraries, (June), 2–9.

Sembok, T. M. T. (2005). *Word Stemming Algorithms and Retrieval Retrieval Systems*. In Proceeding of World Academy of Science, Engineering and Technology (pp. 1–3).

Shao, J. (2008). *Challenges and Techniques for Effective and Efficient Similarity Search in Large Video Databases*. VLDB Endowment, ACM, 1598–1603.

Shen, H. T., Zhou, X., & Zhou, A. (2007). *An adaptive and dynamic dimensionality reduction method for high-dimensional indexing*, 16, 219–234. Retrieved May 2012 from http://doi.org/10.1007/s00778-005-0167-3

Shi Dong Cheng, Xu Lan, H. L. Y. (2007). *Image retrieval using both color and texture features*. The Journal Of China Universities Of Posts And Telecommunications, 14(October).

Shin, C., & Doermann, D. (2006). *Document image retrieval based on layout structural similarity*. International Conference on Image Processing, Computer Vision, and Pattern Recognition, 2, 606–612.

Shobana, D., Sc, M. I. T., Phil, M. C. S., & Ph, D. (2012). *Layout Based Information Retrieval from Document Images*. IOSR Journal of Computer Engineering (IOSRJCE), 4(4), 31–35.

Singh, C., Walia, E., & Upneja, R. (2013). *Accurate calculation of Zernike moments*. Information Sciences. Journal of Information Sciences, 233, 255–275.

Singh, J. (2012). *Different Approaches of CBIR Techniques*. Life Science1(2), 76–78.

Smith, J. R. (1997). *VisualSEEK : A Fully Automated Content-Based Image Query System*. In Proceedings of the fourth ACM international conference on

Multimedia (pp. 87–98).

Steinmetz, & A. and M. Kienzle. (2001). *The e-Seminar lecture recording and distribution system*. In Proc. of SPIE MultiMedia Computing and Networking (MMCN)4312, 25–36.

Swain, M., & Frankel, C. (1996). *WebSeer : An image search engine for the world wide web* . Technical report tr-96-14, Department of Computer Science, University of Chicago.

Tamer Mehyar, J. O. A. (2012). *An Enhancement on Content-Based Image Retrieval using Color and Texture Features*. Journal of Emerging Trends in Computing and Information Sciences, *3*(4), 488–496.

Teague.M.R. (1979). *Image analysis via the general theory of moments*. Optical Society of America.

Tran, H., Lux, A., Nguyen, H. L. T., & Boucher, A. (n.d.). *A novel approach for text detection in images using structural features*. Springer.

Tran, L. V. (2003). *Efficient Image Retrieval with Statistical Color Descriptors*. Journal of Science And Technology, 2(810), 7–14.

Tsai, C.-F. (2012). *Bag-of-Words Representation in Image Annotation: A Review*. ISRN Artificial Intelligence, 2012, 1–19.

Vasconcelos, N. V. and M. (2012). *Minimum Probability of Error Image Retrieval : From Visual Features to Image Semantics*. Foundations and Trends® in Machine Learning, 265–389.

Veltkamp, R. C., Tanase, M., & Sent, D. (2000). *Chapter 5 Features In Content-Based Image Retrieval Systems : A Survey*. Department of Computing Science, Utrecht University, Utrecht, The Netherlands, (pp. 97–124).

Vicario E. (1998). *Image description and retrieval*. New York: Plenum Press.

Vorobyov, M. (2011). *Shape Classification Using Zernike Moments*. iCamp at University of California Irvine, (August), 1–22.

Wang, S., & Wu, Y. S. (2008). *Color Image Retrieval Using a Fuzzy Dominant Color Selection Clustering Algorithm*. Department of Computer Science, Chung Cheng Institute of Technology, National Defense University, *37*(1).

Wu, B. Y., & Chao, K. M. (2004). *Spanning trees and optimization problems*. *Chapman and Hall/CRC*. Retrieved June 2013 from http://scholar.google.com /scholar?hl=en&btnG=Search&q=intitle:Spanning+Trees+and+Optimization+P

roblems

Yuan, X., Yu, J., Qin, Z., & Wan, T. (2011). *A SIFT-LBP image retrieval model based on bag of features*. International Conference on Image Processing, 1061–1064. Retrieved June 2014 from http://icmll.buaa.edu.cn/members/jing.yu/YuanYuQinWan.pdf

Zhang, R., & Zhang, Z. (Mark). (2003a). *Addressing CBIR efficiency, effectiveness, and retrieval subjectivity simultaneously*. Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval - MIR '03, 71.

Zhang, R., & Zhang, Z. M. (2003b). *Addressing CBIR Efficiency , Effectiveness , and Retrieval Subjectivity Simultaneously*. ACM, MIR'03, November 7, 2003, Berkeley, California, USA

Zhang, S., & Yang, M. (2012). *Query Specific Fusion for Image Retrieval,* ECCV 2012, Part II, LNCS 7573, Springer-Verlag Berlin Heidelberg 660–673.

Zhang Dengsheng, & Lu, G. (2004). *Review of shape representation and description techniques*. Pattern Recognition, 37(1), 1–19.