TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY
Julkaisu 792 • Publication 792

Esin Guldogan

Improving Content-Based Image Indexing and Retrieval Performance

Tampere 2008

Tampereen teknillinen yliopisto. Julkaisu 792
Tampere University of Technology. Publication 792

Esin Guldogan

Improving Content-Based Image Indexing and Retrieval Performance

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB103, at Tampere University of Technology, on the 9th of January 2009, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2008

ISBN 978-952-15-2100-3 (printed)
ISBN 978-952-15-2238-3 (PDF)
ISSN 1459-2045

Abstract

Latest image technology improvements along with the Internet growth have led to a huge amount of digital multimedia during the recent decades. Various methods, algorithms and systems have been proposed addressing image storage and management problems. Such studies revealed the indexing and retrieval concepts, which have further evolved to Content-Based Image Retrieval (CBIR). CBIR systems often analyze image content via the so-called low-level features for indexing and retrieval, such as color, texture and shape. In order to achieve significantly higher semantic performance, recent systems seek to combine low-level with high-level features that contain perceptual information for human. However, such combinations increase the feature extraction processing time and the memory requirements as well as the retrieval complexity. Performance improvements of indexing and retrieval play an important role for providing advanced CBIR services on every hardware platform.

In this thesis, we propose novel techniques for improving the overall performance of CBIR. We define general CBIR challenges as memory and disk space requirements, computational complexity, semantic retrieval performance, and usability. Bringing generic and feasible solutions to these challenges is the main contribution of this thesis.

A novel system for feature selection is introduced for enhancing semantic image retrieval results, decreasing retrieval process complexity, and improving the overall system usability for end-users of CBIR systems. Three feature selection criteria and a decision method construct the proposed feature selection system. A majority voting based method is adapted for efficient selection of features and feature combinations. The performance of the proposed criteria is assessed over a large image database and a number of features, and compared against other techniques from the literature. Experiments show that the proposed feature selection system improves semantic performance results in image retrieval systems.

We introduce a novel Transform-Based Layered Query (TLQ) Scheme designed for efficient handling of visual media retrieval, which mainly aims at decreasing processing time and run-time memory consumption without degrading semantic retrieval results. The proposed scheme is based on abstract layers in indexing and retrieval phases, where each indexing layer of TLQ corresponds to a retrieval layer. The layers are independent from the underlying indexing and retrieval methods, and mainly constructed using multimedia and

feature data transformations for reducing data dimensions. A two-layer TLQ scheme is implemented and integrated into MUVIS content-based multimedia indexing and retrieval framework.

A new feature dimension reduction method referred to as Mapping by Adaptive Threshold (MAT) is also proposed as a solution for memory requirements and computational complexity of retrieval processes. Theoretical and practical advantages of TLQ over existing methods are validated experimentally on image databases using the MAT method for feature data. Experimental studies also show that the proposed MAT method is a fast feature transformation for successfully reducing the dimension of feature data without degrading semantic retrieval performance significantly. We also studied the effects of image downscaling techniques on semantic retrieval performance via dedicated experiments in order to utilize the downscaling methods in TLQ scheme. The evaluation results show that image downscaling does not have significant impact on color and moderately affects texture-based retrieval in general, while it degrades edge-based retrieval performance significantly.

In order to accomplish the primary objective of the thesis, a novel study on system profiles and adaptation of parameters for CBIR applications is presented. The main aim of the study is to improve the overall CBIR system performance in different hardware platforms having different technical capabilities and conditions. We define CBIR system profiles in terms of hardware and system platform properties and propose CBIR parameters for each defined system profile. The performances of the proposed parameters for each system profile are assessed over a large set of experiments. Experimental studies show that the proposed parameters for each system profile improve semantic performance, while reducing computational complexity and storage requirement.

Preface

The research presented in this thesis has been carried out at the Department of Signal Processing, Tampere University of Technology, Finland as a part of the MUVIS research project.

First of all, I would like to express my deepest gratitude to my supervisor, Prof. Moncef Gabbouj for his guidance and support throughout my studies.

I owe special thanks to the reviewers of the thesis, Prof. Vladimir Lukin from Kharkov National Aerospace University in Ukraine, and Dr. Faouzi Alaya Cheikh from Gjøvik University College in Norway, for their constructive feedback and helpful comments.

I would like to thank all my current and former colleagues in MUVIS team, especially Dr. Serkan Kiranyaz, Dr. Mari Partio, and Murat Birinci, whom I have had the honor to work with. I am also grateful to the Department of Signal Processing for the friendly working environment and its well-equipped facilities.

My special thanks are due to Virve Larmila, Elina Orava, Pirkko Ruotsalainen, Kirsi Järnström, and Ulla Siltaloppi for their friendly support and kind assistance in practical issues.

Warmest thanks go to all my friends in Finland for their spiritual support and friendship within all these years. I also would like to thank my parents, my grandparents and family-in-law for their support throughout my life.

Finally, I am deeply grateful to my dear husband Olcay. His love, support and endless encouragement helped me to finish this thesis. My deepest gratitude goes to my beloved daughter Deniz, for being so wonderful and being in my life.

Tampere, December 2008 — Esin Guldogan

Contents

List of Publications

This thesis monograph is written on the basis of the following publications:

[Publication 1] Esin Guldogan and Moncef Gabbouj, "Feature Selection for Content-based Image Retrieval", Springer Journal on Signal, Image and Video Processing, Vol. 2, No. 3, September 2008, pp. 241-250.

[Publication 2] Esin Guldogan and Moncef Gabbouj, "System Profiles in Content-Based Image Indexing and Retrieval", Springer Journal on Signal, Image and Video Processing, submitted.

[Publication 3] Esin Guldogan, Olcay Guldogan and Moncef Gabbouj, "Visual Media Retrieval Using Transform-Based Layered Query Scheme", Proceedings of IEEE International Conference on Image Processing, ICIP 2005, Genoa, Italy, September 2005, pp. 521-524.

[Publication 4] Esin Guldogan and Moncef Gabbouj, "Transform-Based Layered Query System for Image Indexing and Retrieval", Proceedings of IEEE International Workshop on Nonlinear Signal and Image Processing, NSIP 2005, Sapporo, Japan, 18-20 May 2005, pp. 381-384.

[Publication 5] Esin Guldogan and Moncef Gabbouj, "Mapping By Adaptive Threshold Method for Dimension Reduction of Content-Based Indexing and Retrieval Features", Proceedings of European Signal Processing Conference, EUSIPCO 2005, Antalya, Turkey, September 2005.

[Publication 6] Esin Guldogan, Olcay Guldogan and Moncef Gabbouj, "Efficient Image Retrieval with JPEG 2000 and DWT-Based Downscaling", Proceedings of the International Workshop on Spectral Methods and Multirate Signal Processing, SMMSP 2006, Florence, Italy, September 2-3, 2006, pp. 199-203.

[Publication 7] Esin Guldogan, Olcay Guldogan and Moncef Gabbouj, "DCT-Based Downscaling Effects On Color And Texture-Based Image Retrieval", Proceedings of the EWIMT 2004, IEE European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology, London, UK, November 2004, pp. 79-86.

[Publication 8] Esin Guldogan and Moncef Gabbouj, "Unsupervised Elimination of Media Items in Content-Based Image Retrieval", Proceedings of the IEEE International Conference on Signal Processing and Communications, ICSPC 2007, Dubai United Arab Emirates (UAE), 24–27 November 2007, pp. 5-8.

[Publication 9] Esin Guldogan, Olcay Guldogan, Serkan Kiranyaz, Kerem Caglar and Moncef Gabbouj, "Compression Effects on Color and Texture Based Multimedia Indexing and Retrieval", Proceedings of the IEEE International Conference on Image Processing, ICIP 2003, Barcelona, Spain, September 2003, pp. 9-12.

[Publication 10] Olcay Guldogan, Esin Guldogan, Serkan Kiranyaz, Kerem Caglar and Moncef Gabbouj, "Dynamic Integration of Explicit Feature Extraction Algorithms into MUVIS Framework", Proceedings of the FINSIG 2003, Finnish Signal Processing Symposium, Tampere, Finland, May. 2003, pp. 120-123.

[Publication 11] Serkan Kiranyaz, Kerem Caglar, Esin Guldogan, Olcay Guldogan and Moncef Gabbouj, "MUVIS: A Content Based Multimedia Indexing and Retrieval Framework", Proceedings of the ISSPA 2003 International Symposium on Signal Processing and its Applications, July 2003, pp. 1-8.

[Publication 12] Serkan Kiranyaz, Kerem Caglar, Esin Guldogan, Olcay Guldogan and Moncef Gabbouj, "MUVIS: A Content Based Multimedia Indexing and Retrieval Framework", Proceedings of the CBMI 2003, International Workshop on Content-Based Multimedia Indexing, Rennes, France, September 2003, pp. 405-412.

List of Acronyms

AAC	Advanced Audio Codec
ANMRR	Average Normalized Modified Retrieval Rank
API	Application Programming Interface
BSP	Baseline System Profile
CBIR	Content-Based Image Retrieval
DCT	Discrete Cosine Transform
DLL	Dynamic Link Library
DSP	Distributed System Profile
DWT	Discrete Wavelet Transform
FSRL	Feature Selection Ranking List
GUI	Graphical User Interface
ICR	Intra-Cluster Relation
JPEG	Joint Photographic Experts Group
KMCC	K-Means with Connectivity Constraint
LSP	Limited System Profile
MAT	Mapping by Adaptive Threshold
MFCC	Mel-Frequency Cepstral Coefficients
MI	Mutual Information
MPEG	Moving Pictures Experts Group
PC	Personal Computer
PCA	Principal Component Analysis
PPMC	Pearson's Product Moment Correlation
PSP	Powerful System Profile
PV	Performance Value
QBE	Query By Example
QBS	Query By Sketch
SBD	Shot Boundary Detection
SPFL	Semantic Performance Feature List
TLQ	Transform-Based Layered Query
UI	User Interface

List of Tables

List of Figures

Chapter 1

Introduction

Advances in hardware and network technologies have provided extensive generation, storage, and transmission capabilities for digital audio/visual information. Rapid increase in the use of digital visual information brought the storage, handling and accessibility problems. Consequently, organization of visual information by means of image and video databases became inevitable for efficient accessibility. Such management and accessing processes have revealed the need for *Indexing and Retrieval* concept.

Studies on image indexing and retrieval started during 1970s with the basic annotation method, which relies on associated keywords. However, this method had limited performance on large image collections. Further studies brought the concept of *Content-Based Image Retrieval* (CBIR) in 1990s. CBIR is mainly based on the visual content of images such as color, texture and shape information. Several techniques have been proposed to extract content characteristics from the visual data automatically, and to use the extracted content information for retrieval purposes. Similar techniques have also been applied on digital video sequences, thus the scope of CBIR has been extended to C*ontent-Based Multimedia Indexing and Retrieval*.

CBIR applications became a part of a practical life and used in several commercial, governmental, and academic institutes such as libraries, TV broadcasting channels, governmental archives [3], [9], [16], [50], [85], [96], [113], [123]. Content-based image searching, browsing and retrieval applications are required for users from various domains such as remote sensing and surveillance. CBIR systems often analyze image content via the so-called low-level features for indexing and retrieval, such as color, texture and shape. Usually, such low-level descriptors have many limitations when dealing with broad image databases and cannot completely express the semantic concepts of the image from the user's perspective [109]. In order to achieve significantly higher semantic performance, recent systems seek to combine low-level features with high-level features that contain perceptual information for human [79], [132]. However, such combinations increase the feature extraction processing time and the memory requirements as well as the retrieval complexity. Besides, semantic performance of CBIR systems still needs improvement.

CBIR systems usually deal with large image collections with several low-level and high-level features, which directly influence indexing and retrieval complexity, and memory and disk space requirements. Due to high memory and processing power requirements, CBIR has not been widely applied on platforms having limited resources, such as mobile devices. However, multimedia capabilities of all computing devices are growing steadily. Recently multimedia became one of the key features of these devices for end-users. Hence, the necessity of multimedia services running on these platforms has arisen, where image indexing and retrieval is one of the biggest challenges. Improving the performance of indexing and retrieval processes plays an important role for providing successful CBIR services for every platform. In this thesis, studies are shaped with motivation of these facts. The main objectives of the studies are as follows:

- Improving the overall performance of CBIR systems, and
- Forming a scalable and adaptive CBIR framework for any type of users and platforms.

Improving CBIR applications involves improving the retrieval performance, efficiency, adaptability and scalability of the overall system. Therefore, this complicated task can be rather simplified by considering indexing and retrieval parts individually. In other words, offline and online processes can be managed separately for ease of handling the improvement process. First of all, the most common challenges in CBIR systems should be defined in order to improve the overall performance. Subsequently, generic solutions should be provided to the defined challenges, where generic solution refers to being independent from image file types and contents. In this thesis, we define four main categories of challenges in CBIR systems, which have significant impact on the overall performance:

- Computational complexity of indexing and retrieval process, which mainly represents the elapsed time for indexing and retrieval purposes,
- Memory and disk space requirements,
- Semantic retrieval performance, which shows the accuracy of the system, and
- Usability.

Although it seems that the most important issue is the semantic retrieval performance in CBIR systems, other categories have also considerable impact on overall system performance. For example, a user may not be satisfied with a slow CBIR system, if he/she has to wait an hour for satisfactory retrieval results, which is a subjective concept and will be defined later in this thesis with objective criteria.

In this thesis, we intend to address most of the aforementioned common challenges and propose generic, efficient and feasible solutions. The proposed techniques are spread among image indexing and retrieval functionalities. All the methods are implemented and integrated into MUVIS content-based multimedia indexing and retrieval framework [94].

1.1. OUTLINE OF THE THESIS

The thesis is organized as follows. In the introduction section, historical development of content-based image retrieval and motivations for the thesis are given. Chapter 2 represents the general overview about the MUVIS framework, in which all the proposed techniques are implemented and tested. Related subjects, such as distance measurement and semantic retrieval performance evaluation methods are also given in this chapter.

Chapter 3 discusses the mainline of the thesis, challenges in CBIR, and the relations of each chapter with the main subject of the thesis. This chapter gathers all studies into a general skeleton and defines the associations between them.

Chapter 4 presents a novel system for feature selection, which aims at enhancing semantic image retrieval results, decreasing retrieval process complexity, and improving the overall system usability for end-users of multimedia search engines. Three feature selection criteria and a decision method construct the feature selection system. Two novel feature selection criteria based on intra and inter cluster relations are proposed. A majority voting based method is adapted for efficient selection of features and feature combinations.

A novel Transform-Based Layered Query (TLQ) scheme designed for efficient handling of visual media retrieval is presented in Chapter 5. The TLQ scheme mainly aims at decreasing processing time and run-time memory consumption without degrading retrieval results semantically. The scheme is based on abstract layers in indexing and retrieval phases, where each indexing layer of TLQ corresponds to a retrieval layer. The layers are independent from the underlying indexing and retrieval methods, and mainly constructed using multimedia and feature data transformations for reducing data dimensions. A two-layer TLQ scheme is implemented and integrated into the MUVIS framework. A new feature dimension reduction method referred to as Mapping by Adaptive Threshold (MAT) is also proposed in this chapter. Theoretical and practical advantages of TLQ over existing methods are verified experimentally on image and video databases using MAT method for feature data and DCT-based downscaling for image data.

In Chapter 6, a novel study on system profiles and adaptation of parameters for the end-users of a CBIR system is presented. The main aim of the study is to improve the overall CBIR system performance in different hardware platforms having different technical capabilities and conditions. The study is composed of two main parts: system profiling and adaptation of indexing and retrieval parameters for each profile. The performance improvements based on the proposed parameters for each system profile are assessed over a large set of experiments. Improvements in semantic performance and gains in computational complexity and storage requirement are also expressed in the chapter.

Finally, the conclusions of the thesis are drawn in Chapter 7.

1.2. AUTHOR’S CONTRIBUTION

The main contributions brought by the author towards improving the overall performance of CBIR systems are summarized in the following list:

- Publication 1: A novel feature selection system, which aims at enhancing semantic image retrieval results, decreasing retrieval process complexity, and improving the overall system usability for end-users of multimedia search engines [42].
- Publication 2: A novel study on system profiles and adaptation of parameters for the end-users of a content-based image retrieval application.
- Publication 3 and 4: The design and implementation of a novel Transform-Based Layered Query scheme for efficient handling of visual media retrieval [39], [43]. The author carried out the main analysis and implementation of the work. MSc. Olcay Guldogan suggested some ideas on adapting the proposed scheme for video indexing and retrieval, which are excluded from this thesis.
- Publication 5: The design and implementation of a new feature dimension reduction method referred to as Mapping by Adaptive Threshold [36].
- Publication 6 and 7: A novel study for the evaluation of the DCT and DWT based downscaling effects in compressed domain content-based image retrieval [38], [41]. The author originated the main idea and performed all analysis and experiments. The evaluations of the experimental results are done together with MSc. Olcay Guldogan.
- Publication 8: A novel unsupervised elimination of irrelevant media items method for image retrieval systems, where the retrieval process consists of multiple steps [37].
- Publication 9: The author contributed to a joint work, which is a novel study for compression effects on color and texture image indexing and retrieval with MSc. Olcay Guldogan [40].
- Publication 10: Contribution to the implementation of a dynamic feature extraction framework within MUVIS system [45].
- Publication 11 and 12: Contribution to the implementation of DbsEditor application in MUVIS system [63], [64].

Chapter 2

MUVIS Framework

First MUVIS system was developed for indexing large image databases and retrieval based on visual features in the Institute of Signal Processing in Tampere University of Technology, and Pori School of Information Technology and Economics in 1998 [14], [20]. MUVIS aims to provide a generic solution for indexing, browsing and retrieval supporting various digital multimedia types on Windows operating system. The current version of the MUVIS system has been reformed to support a dynamic integration of feature extraction, spatial segmentation (SEG) and shot boundary detection (SBD) modules [94]. SBD module is responsible for extracting the list of shot boundaries and the list of key frames from a video clip. The purpose of the SEG module is to provide segmentation masks containing two or more segmented regions from an image or a key frame of a video clip.

Figure 1 represents the applications of MUVIS framework each of which has different responsibilities and functionalities. MUVIS framework provides three applications:

- AVDatabase (Audio/Video Database Creator),
- DbsEditor (Database Editor), and
- MBrowser (Media Browser).

Each application is responsible for specific tasks. More detailed description of the applications can be found in [61], [94].

AVDatabase is a real time audio/video database creator, capturing Audio/Video data from a peripheral device connected to a computer. In order to develop CBIR techniques that are independent from image file types, several file formats and codecs are supported for encoding and recording of audio/video sequences.

DbsEditor manages the indexing and editing tasks for the MUVIS databases. Feature extraction and editing is the primary task of DbsEditor application. It creates the features offline for the multimedia databases. Moreover, DbsEditor can also add and remove features to/from any type of database since MUVIS system supports querying based upon multiple features. A Graphical User Interface (GUI) of DbsEditor application is given in Figure 2.

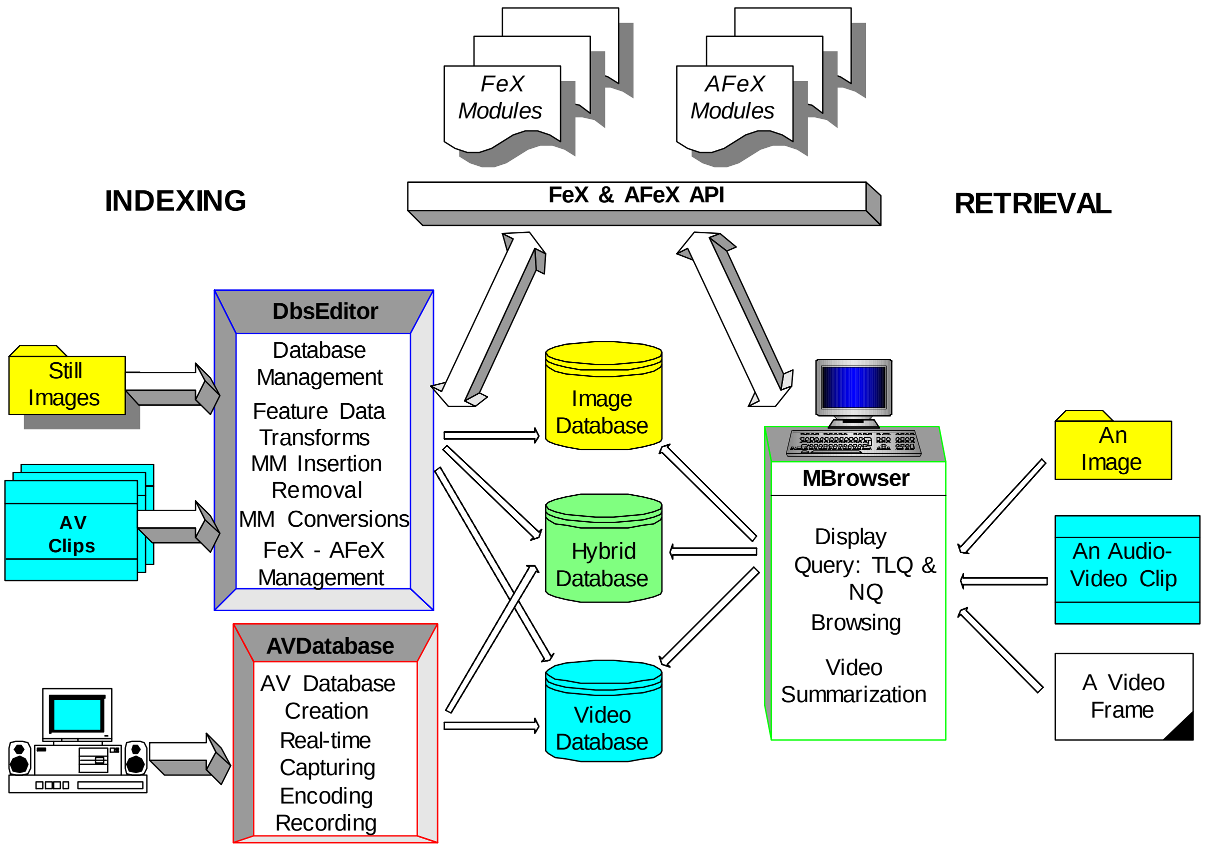


Figure 1: General Structure of MUVIS Framework.

Main functionalities of DbsEditor can be listed as follows:

- Appending new audio/video clips and still images to any MUVIS database and removing such multimedia items from the database,
- Dynamic integration and management of feature extraction (FeX) and audio feature extraction (AFeX) modules,
- Extracting new features or removing existing features of a database by using available FeX and AFeX modules,
- Converting of alien audio/video files into any MUVIS database,
- Preview of any audio/video clip or image in a database,
- Display statistical information of a database and/or items in a database,
- Hierarchical Cellular Tree (HCT) based visual and audio indexing [65].

MBrowser is intended for browsing and retrieving multimedia items. It supports any of the database, image and video types mentioned above. It has capabilities of a powerful multimedia player (or viewer) and a database browser. It allows users to browse multimedia easily, efficiently and in any of the display type that are defined for the video clips. Five types of video presentation methods are supported by MBrowser: single frame, shot frames (key frames), scene frames, a video segment and the entire video clip.

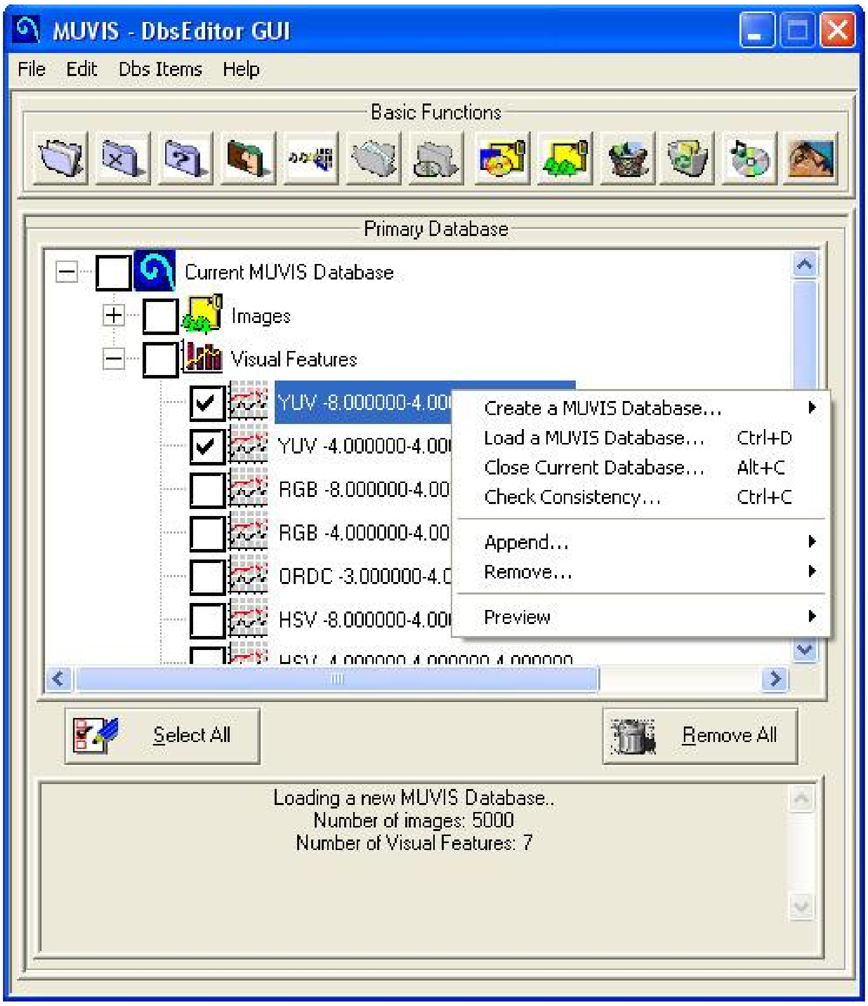


Figure 2: A Snapshot from DBS Editor Application.

MBrowser provides a search method to find multimedia items in any database that is similar to the query media. The query can be external or internal to the database, which is currently active in the application. Retrieval is based on comparing the query feature vector(s) with the feature vectors of multimedia items in the database. The comparison is based on the similarity measurement function implemented in the corresponding feature extraction module, which is integrated into MUVIS via FeX API. Sample retrieval results within MBrowser application are given in Figure 3. Additionally, MBrowser provides the following additional functionalities:

- Video summarization via scene cut detection and key frame browsing,
- Key frame browsing during video playback,
- Random access support for audio/video clips during video playback,
- Displaying information related to the active database (i.e. database features, parameters, status, etc.),
- Visualizations of feature vectors of the multimedia items, and
- Various browsing options such as random, forward or backward.

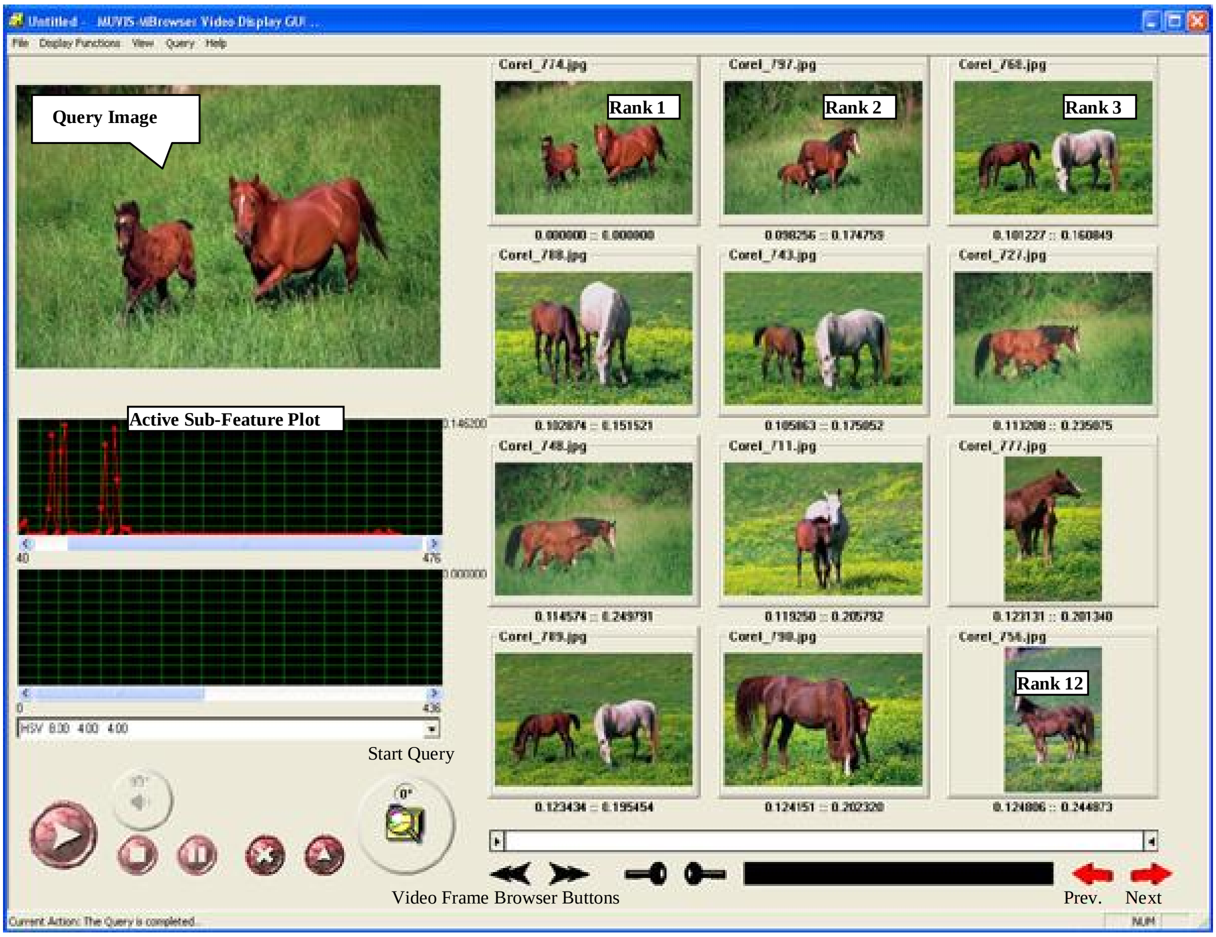


Figure 3: A Snapshot from MBrowser Application.

MUVIS Extended Framework [94] provides a generic and dynamic structure for development of external modules as DLLs to implement and test feature extraction, spatial segmentation and shot boundary detection methods. Explicitly implemented algorithms can be integrated into MUVIS Xt framework dynamically using defined APIs. There are four types of modules that can be dynamically integrated into MUVIS framework:

- Visual feature extraction (FeX),
- Aural feature extraction (AFeX),
- Spatial segmentation (SEG),
- Shot boundary detection in video sequences (SBD).

The following extended framework modules have been integrated into MUVIS framework:

FeX Modules: HSV, RGB and YUV Color Histograms [115], Dominant Color [88], Gabor Texture feature [84], Gray-level co-occurrence matrix [99], MPEG-7 Edge Histogram [88], 2D-Walking ant histogram [62], Canny Edge histogram [8].

AFeX Modules: Mel-Frequency Cepstral Coefficients (MFCC) Module [66].

Segmentation: K-Means with Connectivity Constraint (KMCC) module [90], J-Segmentation (JSEG) module [23], graph-based segmentation [32], and quad-tree split and merge, watershed.

2.1. INDEXING AND RETRIEVAL IN MUVIS

Improving overall CBIR performance is the main scope of the thesis and it is handled separately for indexing and retrieval processes. The proposed methods, developments and experiments are performed based on indexing and retrieval in MUVIS framework. Thus, in this section we briefly explain the indexing and retrieval concept in MUVIS framework.

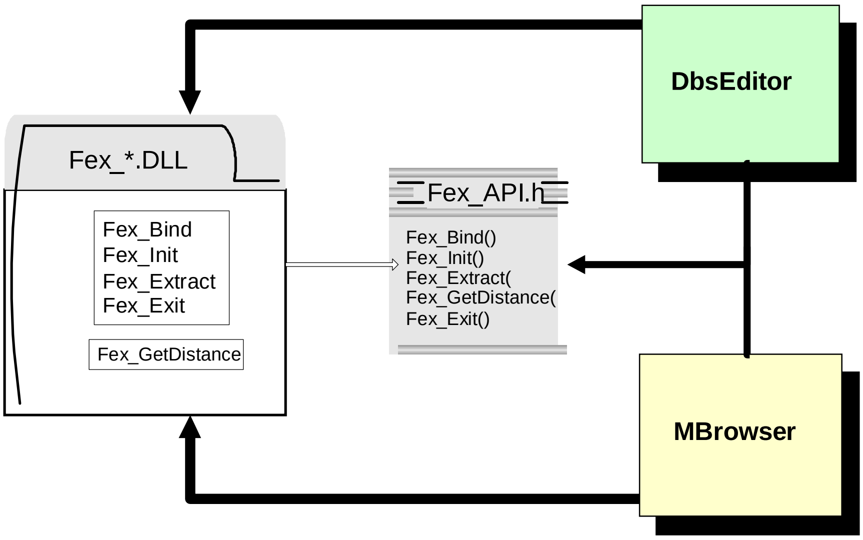


Figure 4: FeX Module Interaction with MUVIS Application.

DbsEditor application performs the indexing of a MUVIS database, where multimedia items are collected and sequentially numbered (indexed). Optionally, audio/video features are extracted using available FeX and AFeX modules in the system. MUVIS manages feature extraction in terms of explicit modules. Feature extraction algorithms can be implemented independently as modules, and then integrated into MUVIS framework via a specific interface called Feature Extraction Interface (FeX API), shown in Figure 4. Features of databases and media items in MUVIS are represented by normalized array of numbers. Furthermore, in order to support dynamic integration of exclusive feature extraction modules, features should be represented in a common and easily supportable format, which is vector representation.

DbsEditor version 1.6.2 has two optional processes for indexing the image databases to retrieve efficiently or to reduce indexing and/or retrieval process complexity. The first one is to extract features from downscaled images. Downscaling process is performed while decoding the images and scaled image data are used for feature extraction. Downscaling of images and its advantages are presented in Chapter 5 briefly.

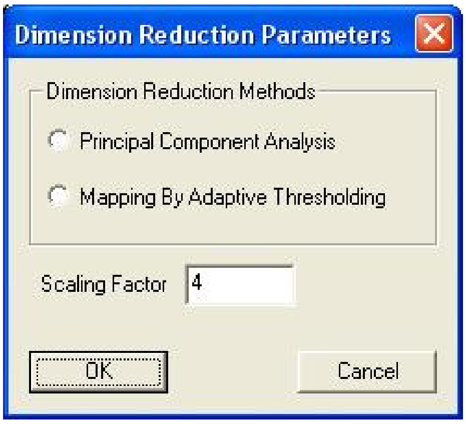


Figure 5: Dimension Reduction Parameters Window of DbsEditor Application.

Dimension reduction of the feature data is the second optional process in DbsEditor. It can be applied on existing feature files, which are created after the feature extraction process. Separate feature files are generated by the dimension reduction methods and saved as a MUVIS feature file to be used for retrieval purposes. DbsEditor version 1.6.2 supports two dimension reduction of feature data methods: Mapping by Adaptive Threshold (MAT) and Principle Component Analysis (PCA) as shown in Figure 5. These methods are explained in details in Chapter 5.

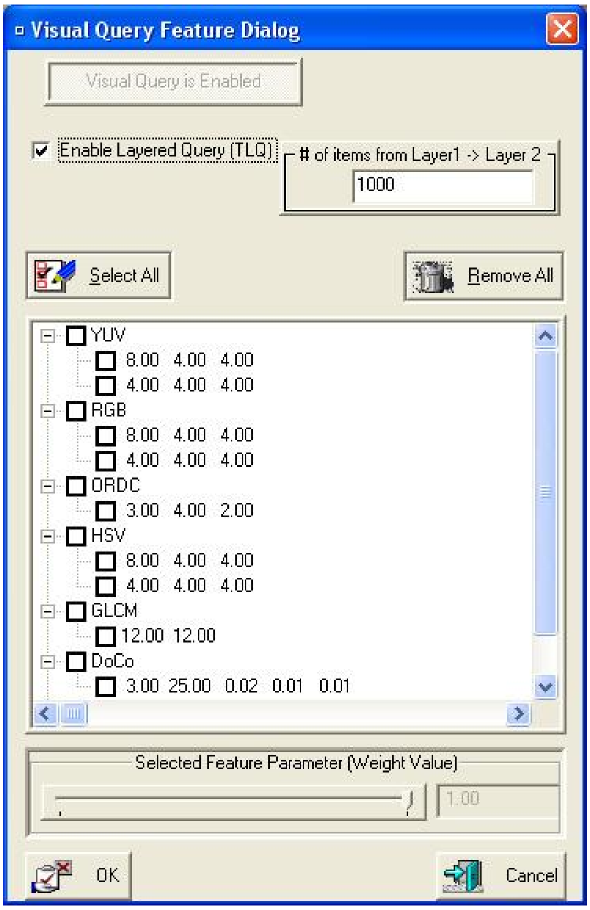


Figure 6: Visual Query and Feature Parameters Window of MBrowser Application.

The MBrowser application provides browsing and Query-by-Example (QBE) retrieval within a MUVIS database. Browsing is performed randomly or sequentially on database items. MBrowser v.1.6.2 supports two QBE methods: Normal Query and Transform-based Layered Query (TLQ) [39]. Query and feature parameters window of MBrowser application is given in Figure 6. Normal query is the basic QBE method that searches for multimedia items based on an example. Similarity distances are calculated and combined to obtain one similarity distance per image, using available visual features of the database. The query results are obtained by ranking the items according to their similarity distance to the queried item over the entire database. MBrowser is capable of merging multiple features for querying. The merging scheme used in MBrowser requires normalized feature vectors. The value of each item in a feature vector is divided by its theoretical maximum value for this purpose. Normal query is CPU and memory intensive process especially for large databases with multiple features. These are the important challenges we address and propose solutions for in this thesis.

2.2. SEMANTIC RETRIEVAL PERFORMANCE EVALUATION METHODS

CBIR systems retrieve a number of images similar to a query from a database, thus they differ distinctly from classification and recognition systems that can retrieve exact matches. Evaluating the performance of retrieval can be done by defining a quantitative objective metric. However, it is difficult to quantify semantic retrieval performance, since the similarities between images are subjective and feature dependent, and there are no common standard benchmark databases. In order to form a general standard evaluation process, three main issues should be addressed clearly: standard image databases, ground truth classification of the database, and performance metrics. Ground truth classification depends on human subjectivity and it is a time consuming, costly process especially with large image databases or databases that require specific expert knowledge. Several quantitative performance metrics are defined in the literature [77], [89].

Most common and basic metrics are precision and recall. Precision is defined as the ratio of relevant documents in the set of all documents returned by a query. Recall is the number of relevant documents retrieved as fraction of all relevant documents. They can be formulated as follows:

$$precision = \frac{Number\ of\ relevant\ items\ retrieved}{Number\ of\ items\ retrieved}$$

$$recall = \frac{Number\ of\ relevant\ items\ retrieved}{Total\ number\ of\ relevant\ items}$$

Generally, there is an inverse relationship between Precision and Recall, where it is possible to increase one at the cost of reducing the other yielding a trade-off between them. For example, precision starts to decrease while recall tends to increase as the number of retrieved items increases. In case of an ideal retrieval, precision value is equal to 1 regardless of the recall value. A sample precision-recall curve is illustrated in Figure 7.

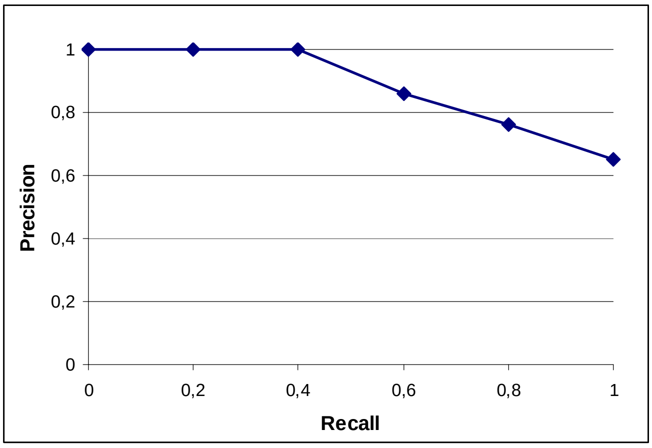


Figure 7: A Sample Precision-Recall Curve.

Ranking information is an important issue that should be addressed for evaluating the retrieval performance, however it is not considered within precision and recall metric. Therefore, Average Normalized Modified Retrieval Rank (ANMRR) is introduced by MPEG-7 as a retrieval performance metric. ANMRR is unbiased and limited metric defined for each query (q) and it considers precision, recall and ranking information as given in the following formulas [89]:

$$AVR(q) = \frac{\sum_{k=1}^{N(q)} R(k)}{N(q)} \text{ and } W = 2N(q) \tag{1}$$

$$NMRR(q) = \frac{2AVR(q) - N(q) - 1}{2W - N(q) + 1} \leq 1 \tag{2}$$

$$ANMRR = \frac{\sum_{q=1}^{Q} NMRR(q)}{Q} \leq 1 \tag{3}$$

where *N(q)* is the minimum number of relevant images in a set of *Q* retrieval experiments, *R(k)* is the rank of the k^{th} relevant retrieval within a window *W*, and *q* is a query.

The best retrieval performance can be achieved when *NMRR(q)=0*. On the other hand, in the worst case *NMRR(q)=1* ,which means that none of the relevant items can be retrieved among *W*. Thus, lower *NMRR* values represent successful retrieval results for the query *q*. Average NMRR (ANMRR) can be used as semantic retrieval performance criterion, if the number of query by example (QBE) experiments is high enough.

Chapter 3

Image Indexing and Retrieval Challenges

Content-based indexing can be defined as arranging the database items based on their content for further queries. The term 'content' in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. The content is usually represented with low-level and high-level features. In this thesis, we refer to low-level features, which represent color, texture and shape information of a multimedia item. Several low-level descriptors have been proposed recently in the domain of image indexing, retrieval and classification. In [70] novel shape descriptors are proposed for defect image retrieval, and color and texture descriptors are used in rock image classification [72]. High level features are also known as logical, semantic features. High level features involve various degrees of semantic existing in images, video and audio. They can be classified as objective or subjective features. The former concerns the object identification in images and action descriptions in video. Subjective features concern the abstract attributes. They describe the meaning and purpose of objects or scenes. The use of low-level features does not usually yield satisfactory retrieval results in many cases; especially, when high-level concepts in the user’s mind are not easily expressible in terms of low-level features.

Emphasizing one or more items from a structured collection via a process such as querying and browsing is referred to as *image retrieval*. Several retrieval techniques are introduced in CBIR systems such as Query by Example (QBE) and Query by Sketch (QBS). For example, query by example and query by sketch methods have different process steps for retrieval in image databases. Hence, retrieval improvements should be properly defined to the particular retrieval method. Consequently, in this thesis the design and implementation of the proposed methods are done regarding QBE systems using low-level features, such as MUVIS framework. Figure 8 illustrates such a content-based image indexing and retrieval system. In Figure 8, user selects a query image to search similar images in an existing database. Features of the query image are extracted in order to compare them with features of other images in the database. The comparison process is called similarity measurement, which quantifies numerically the distance between feature vectors. Finally, sorted results are displayed to the user.

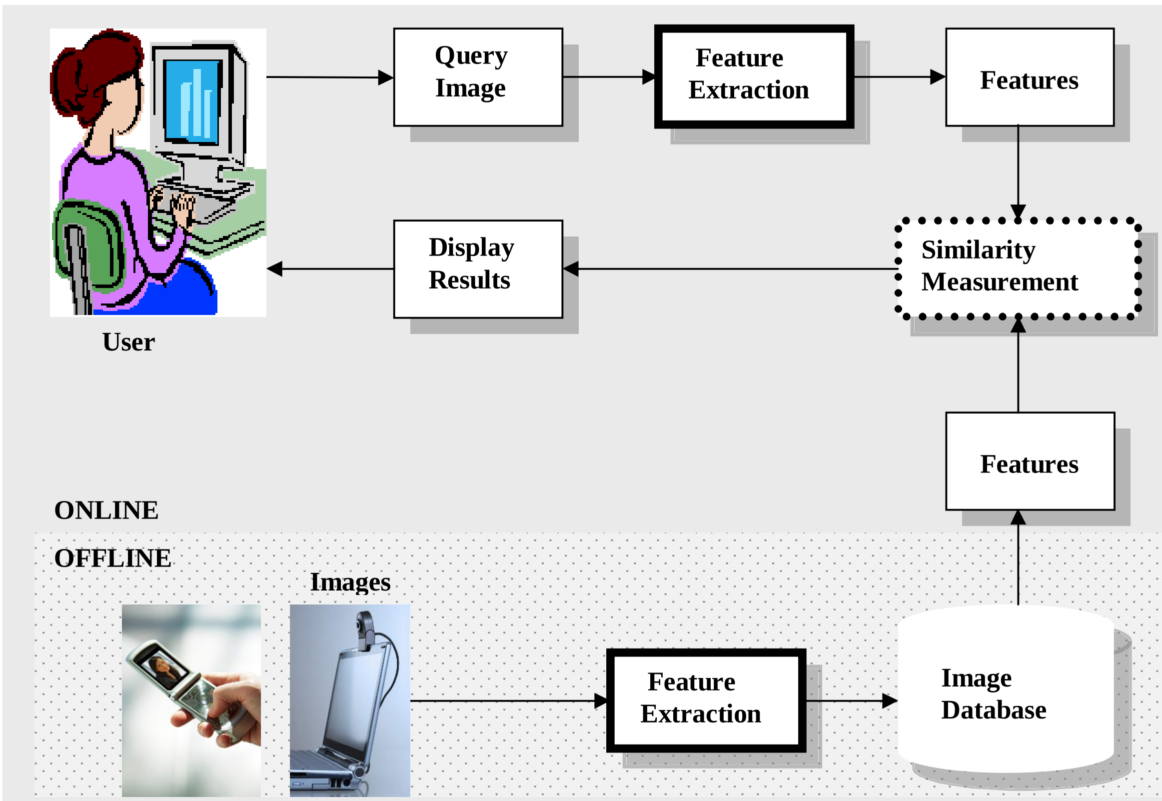


Figure 8: Content-Based Image Indexing and Retrieval Process.

In such systems, image indexing process steps are assumed as follows:

- Copying images to a certain database folder,
- Image Converting/Decoding,
- Feature extraction,
- Optional Post-processing (dimension reduction of feature data, image clustering etc.).

Retrieval process steps are:

- Feature extraction of query item,
- Similarity measurement,
- Ranking,
- Optional post-processing (i.e. relevance feedback).

Managing and analyzing indexing and retrieval processes categorized as above help to improve the overall CBIR performance. In other words, each step can be called a sub-process and it should be enhanced in order to improve the overall performance. Additionally, analyzing indexing and retrieval process steps separately help to observe potential challenges in CBIR systems. These challenges are grouped into four categories in this thesis: memory and disk space requirements, computational complexity, semantic gap, and usability complexity. Categories are determined while developing the MUVIS framework and reviewing the CBIR literature. Each indexing and retrieval step has considerable impact on one or more categories. In this thesis, we propose several methods regarding the indexing and retrieval steps mentioned above in order to introduce solutions to the categorized challenges.

The following four sections define the challenges in CBIR systems, give the state of art and briefly introduce the proposed solutions to the challenges.

3.1. MEMORY AND DISK SPACE REQUIREMENTS

Continuously developing imaging technologies bring along with large file sizes due to high quality images. Handling large scale image databases constructed with high quality images requires significant amount of disk space in CBIR systems. Moreover, any data that are saved to files for indexing and retrieval purposes may increase the demand for memory and disk space such as feature data files. Especially, run-time memory requirements may even affect the processing time due to lack of random access memory (RAM). In the scope of the indexing and retrieval steps defined above, we analyze possible solutions and proposals for disk space requirements in this thesis.

The most common solution for the memory and disk requirements is to work with compressed data for managing large scale image databases. There are many CBIR frameworks working with compressed multimedia items in the literature. Mandal and Liu in [87] proposed two image indexing techniques in JPEG-2000 framework. They generate the indexes from the JPEG2000 compressed image without decoding the images. Similarly, Irianto et. al. in [53] extracted the features from JPEG compressed images. They use a DC component of the image as a feature in order to decrease the storage needs and image retrieving process time in CBIR. Lu and Teng in [106] also used compressed data for indexing and retrieval scheme based on vector quantization. Lu et. al. in [82] extracted color, spatial and frequency-based features directly from DCT domain of JPEG-compressed images. Considering the wide use of compressed images in CBIR systems, the effects of compression on image retrieval has been studied in [40]. The details are not included in this thesis; however, the results that are used in the study, are explained in Chapter 6.

Image downscaling methods are also widely used methods in CBIR systems in order to reduce the memory and disk space usage [17]. In this thesis, DCT and DWT based image downscaling methods and their effects on the image retrieval results are studied. Chapter 5 discusses the effects of the methods and details. Additionally, Chapter 5 introduces an image retrieval scheme, which uses image compression and downscaling methods and shows their practical benefits in terms of memory and disk space requirements.

3.2. COMPUTATIONAL COMPLEXITY

Computational complexity is originally used to denote time or memory involved in the computation. In this section, we will refer to computational complexity as the elapsed time for computing the process rather than memory and disk space cost mentioned in the previous section.

The efficiency of a CBIR system becomes quite critical when usability of an application is evaluated. CBIR applications should establish its services in an acceptable time even with limited resources. Additional to satisfactory retrieval results, the system should provide the results in a reasonable time to the end-user.

We analyze the computational cost of CBIR systems by determining the computable functions or deterministic modules (procedures), which construct the cornerstones of the indexing and retrieval processes such as feature extraction. Several methods have been proposed in the literature to decrease computational cost and accordingly improve efficiency by decreasing access time of the images during indexing and retrieval processes. Mughal et. al. in [93] introduced a 3D hash function for image indexing and retrieval to speed up inserting and retrieving images from a database. Egas et. al. in [29] addressed the indexing problem by adapting the k-d trees to the problem of image retrieval in order to reduce indexing and access time.

There have been various proposed methods and studies on compressed domain based indexing and retrieval, where decompression of images is not required. Consequently, compared with the spatial domain based retrieval methods for Joint Photographic Experts Group (JPEG) images, the computational complexity can be greatly reduced. Irianto et. al. [53] proposed a simple method of DC (zero-frequency) feature extraction that enables to speed up the process and decrease storage need in image retrieving. They use the DC component of DCT coefficients of JPEG compressed image to decrease indexing and retrieval time in real applications primarily for large-scale image database. Lu et. al. [82] extracted color, spatial and frequency-based features directly from DCT domain of JPEG compressed images.

High dimensionality of feature vectors yields a high computational cost in indexing and querying especially during distance calculation for similarity retrieval. Several methods have been proposed to overcome these challenges in CBIR domain. Zhang et. al. in [130] used dominant colors in the histogram, and Wan and Kuo in [122] suggested a multiresolution color clustering to reduce the computational complexity in distance calculation. Similarly, several dimension reduction methods for feature data have been proposed to solve the complexity problem of the similarity measurement. Singular value decomposition (SVD) and Hilbert curve fitting are example methods used to reduce the dimensionality of feature vectors. However, these methods have their own drawbacks. Hafner et. al. in [46] performed SVD on the quadratic matrix of correlations between the color histogram bins. The resulting eigenvectors are not related to the feature data, and may result in significant errors when lower-dimensional transformed feature vectors are used to approximate the original feature vectors. The performance of curve fitting depends on the data distributions. Points that are close to each other in the original feature space might be far apart on the modeled curve. Ng and Sedighian [95] used principal component analysis (PCA) for dimension reduction of image features to reduce the dimensionality of the feature space and customize the selected multi-dimensional indexing structure to improve search performance.

Feature selection is another proposed solution for the computational cost problem of CBIR systems [18], [27], [118], [124]. It refers to selecting the most important features and their combinations for describing and querying items in the database in order to reduce retrieval complexity in terms of elapsed time while maintaining high retrieval performance. We propose a novel feature selection system, which aims at enhancing semantic image retrieval results, decrease retrieval process complexity, and improve the overall system usability for end-users of image search engines. The study is explained in more details in Chapter 4.

There have been several proposed methods and studies as a solution for computational cost problem of image indexing and retrieval. In this thesis, most of the studies aim to reduce the computational complexity of indexing and retrieval process that practically means to decrease the process time. We propose four new approaches in Chapter 5 to reduce the computational complexity while preserving the retrieval performance. Feature extraction time can be reduced proportionally by reducing the image sizes in the database. Thus image downscaling is the most effective way to reduce the computational costs of feature extraction. However, this process may cause significant information loss when representing the image content. In Chapter 5, we study

DCT and DWT based image downscaling effects on semantic image retrieval and show practical benefits of downscaling methods for indexing process complexity reduction. Additionally, a new feature data dimension reduction method is introduced in Chapter 5. The main goal of the study is to decrease the retrieval time whilst not affecting the semantic retrieval performance negatively. In Chapter 5, a new simple unsupervised image elimination method for indexing and retrieval systems is presented. The method aims at reducing the retrieval complexity by decreasing the number of images involved in the retrieval process. Moreover, Chapter 5 introduces an image retrieval scheme that uses several image downscaling and feature data dimension reduction methods to form a sample CBIR system with low computational costs and satisfactory semantic retrieval results.

3.3. SEMANTIC RETRIEVAL

Semantic refers to the meaning of the image content, which is a high-level concept, compared to simple low-level visual features. The users expect from an ideal CBIR system to find a meaningful result, when requests such as "find picture of cats" or even "find Van Gogh's paintings" are given. This type of queries is difficult for computers to understand and process because they refer to the semantic meaning of images. Generally, current CBIR systems use low-level features such as color, texture and shape. In this respect, one of the most important challenges in CBIR systems is how to bridge the semantic gap between low-level features and high-level semantics [73]. Visual features such as color, in general do not necessarily match perceptional semantics of images. To improve the semantic retrieval results, human perception subjectivity may be incorporated into the retrieval process by providing an opportunity for user to evaluate the results. This technique is called Relevance Feedback (RF) and has become common research study in CBIR area [25], [54], [58], [117]. Relevance feedback is an iterative process, which improves the performance of content-based image retrieval by modifying the query based on the user's feedback on the retrieval results. Long and Lew in [80] proposed an approach of increasing retrieval performance by improving the perceptual consistency of computational features and similarity measurements. They proposed a method for measuring the perceptual distances and constructing perceptual space based on relevance feedback.

Several frameworks use additional metadata with low-level features in order to bridge the semantic gap [7], [15], [31], [111]. Zhang et. al. in [128] investigated the role of user term feedback in interactive text-based image retrieval. Term feedback refers to the feedback from a user on specific terms regarding their relevance to a target image. Lee and Soo in [71] introduced an annotation guide agent to aid annotators since conducting consistent and complete image annotation is not a trivial task, especially when the domain is not unique. Jin et. al. in [57] used automatic annotation of database images for retrieval purpose by clustering image regions into region clusters.

Successful low-level features may lead to significantly high semantic retrieval performance in high-level. Recently, several CBIR systems based on low-level features reach satisfactory results [3], [85], [96], [113]. The semantic gap will be reduced with the assistance of these successful low-level features. Zhang in [129] presented a method combining both color and texture features for images to improve the retrieval performance. Given a query, images in the database are firstly ranked using color features. Then the top ranked images are re-ranked according to their texture features for improving retrieval performance. Mojsilovic and Rogowitz in [92] took another approach in overcoming the semantic gap and posed the following question: "Is it possible to find correlations between the high-level semantics and low-level descriptors and use them to capture the semantic meaning of an image?". They first identified broad semantic categories in the perceptual data, which are then modeled in terms of combinations of low-level image features.

Several low-level feature-based studies have been done in MUVIS frameworks, and satisfying retrieval performances are obtained [5], [62], [100]. Additionally, it was found out that optimum combinations of the low-level features improve the semantic retrieval results [126]. Feature selection and feature weighting methods define the most important features and their combinations for describing the images to improve the retrieval performance. In this thesis, we achieve satisfactory improvements in semantic retrieval performance with the proposed feature selection and feature weighting methods. Chapter 4 presents the feature selection and weighting methods in details.

3.4. USABILITY

Usability can be defined as perceived efficiency and ease of use of an application for achieving a particular task. The ISO/IS 9241-11 guidance on usability defines it as: "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". Usability of a CBIR framework is usually associated with the characteristics of the user interface. On the other hand, scalability and adaptability of a system are also important factors for the system usability. Scalability usually refers to the capability of a system to handle changing amount of media efficiently under an increased or decreased load. Additionally, adaptive systems are able to adapt themselves according to changes in the environment. Therefore, usability can be improved if a CBIR system is not affected by the changes on image databases, platforms, and devices with different hardware architectures. The system should work efficiently for every type of user having different levels of knowledge and on any platform having different technical hardware specifications.

The time or effort required for a user to accomplish a task is another important factor when evaluating the usability of a system. In addition to this, the degree of ease in using the graphical user interface for basic tasks also affects the usability.

One of the best criteria to evaluate content-based image retrieval performance is user satisfaction, which is related to the usability of a system. A user should be pleased with the retrieved results and usability of the application. The impacts affecting the satisfaction of the user start from the first moment of the application usage. Graphical user interface and presentation of the indexing and retrieval of the images are quite important factors for the user's opinions about the efficiency of CBIR systems [110]. Thus, the type of user interface (UI) is one of the key factors affecting the efficiency of the CBIR framework. There are four types of UI in retrieval systems: keywords searching, browsing categories, query by example (QBE) and query by sketch (QBS) [119]. Regardless of the UI types, current image retrieval interfaces are generally complicated for average users [6]. This fact decreases the overall system performance. Vermilyer in [119] described an enhanced CBIR system by the use of intelligent user interface agents. Kaster et. al. in [60] provided an overview of a project designed to investigate the requirements of a user interface for a CBIR system and presents preliminary results of an evaluation that assessed the usability of a query by example method.

The role of human factors can be considered when measuring image retrieval performance. Particularly, usability evaluation of retrieval depends on how the user performs the query, expectations and human factors. Jaimes in [55] studied levels of description, types of users, search strategies, image uses and human factors that affect the construction and evaluation of CBIR systems, such as human memory, context and subjectivity. Torres and Parkes in [116] discussed the need for user modeling and adaptability in effective retrieval systems. They form a model based on Bayesian networks and Bayesian user modeling, which can be applied to CBIR applications.

In this thesis, a novel study on system profiles and adaptation of parameters for the users are presented in Chapter 6. The main aim of the study is to improve the overall CBIR system performance on different hardware platforms having different technical capabilities and conditions. We define CBIR system profiles in terms of hardware and system platform properties and propose personalized CBIR parameters for each defined system profile to improve the system hardware scalability and adaptability.

Chapter 4

Feature Selection in CBIR

CBIR systems often analyze image content via so-called low-level features, such as color, texture, and shape [94], [96], [103], [113]. Recent systems tend to combine low-level features with high-level features that contain perceptual information for human to achieve significantly higher semantic retrieval performance [79], [132]. Nevertheless, such combinations increase time and memory requirements together with complexity of the feature extraction process.

CBIR has not been widely implemented on limited platforms, such as mobile devices due to high memory, and processing power requirements. However, multimedia capabilities of all computing devices are improving rapidly. Recently multimedia support became one of the key features of these devices for end-users. Consequently, the requirement of multimedia services running on these platforms has emerged, where image indexing and retrieval is one of the most important challenges. Improving the performance of indexing and retrieval processes plays a significant role for providing satisfactory CBIR services for such limited systems. Feature selection is one of the key challenges for advancing CBIR systems [18], [27], [104], [118], [124]. It refers to selecting the most important features and their combinations for describing and querying items in a database in order to reduce retrieval complexity while maintaining high retrieval performance. Moreover, it helps end-users by automatically associating proper features and weights for a given database.

Feature selection has been a common research topic in pattern recognition. It has been utilized in various research fields such as genomic data analysis, classification of network data, categorization of medical data, speech recognition, etc. [24], [56], [68], [78], [102], [126]. However, interpretations of feature performance and feature selection methods for CBIR have to be carried out in slightly different ways from classification and categorization. Decision errors are utilized for this purpose in classification. There is no perfect unsupervised method to evaluate retrieval results of a CBIR system for evaluating the semantic feature performance.

In this study, we mainly propose two criteria for feature evaluation and a method for feature selection that have not been addressed by earlier studies particularly in CBIR context:

- A new criterion based on categorized member relation within the same cluster from labeled training data to better understand the description power of the feature for each cluster,
- A new criterion based on the discrimination power of the features calculated using Pearson's Product Moment Correlation (PPMC) [101] for defining correlations between different classes.

Majority voting is utilized as a decision mechanism to select appropriate features based on the results of mutual information, inter-cluster and intra-cluster relations criteria.

The organization of the chapter is as follows: Section 4.1 presents relevant feature selection methods in details. A majority voting method is described in Section 4.2. Experimental results are given in Section 4.3, and finally Section 4.4 provides concluding remarks and discussions.

4.1. FEATURE SELECTION

Feature selection can be defined as selecting the combination of features among a given large set, which best describes a particular data collection. Feature selection has been a popular research topic since 1970's in pattern recognition and applied to several research fields [24], [56], [68], [78], [102].

Feature selection algorithms may be grouped into three categories in data mining: filters, wrappers, and hybrid methods. Filters use general characteristics of data independently from the classifier for the evaluation process. The evaluation process is classifier-dependent in wrapper methods. Finally, hybrid models use both filtering and wrapping methods for improving the performance of the selection process.

The key operation in the feature selection processes is to evaluate the discrimination power of the individual feature. Mutual information is the most common method, used for evaluating the discrimination power of a feature [24], [30], [48], [102]. Vasconcelos [118] used maximal divergence for feature selection in image retrieval, whereas Ding and Peng [24] used mutual information for feature selection from Microarray gene expression data.

The relations between clusters in the feature space are called inter-cluster relations. It has also been used for medical image feature data evaluation in [27]. Feature and data relations carry information about the characteristics of the features on the data. Mutual information is used for measuring the feature and data relations. Inter-cluster and intra-cluster affinity characterizes the relationship between features and classes, thus they are useful for evaluating the discrimination and description power of the feature, respectively. Consequently, different attributes of the feature-data relations are utilized in this study for wider specification.

The criteria, how and why they are used in our feature selection approaches are described in the following sub-sections.

4.1.1. Mutual Information (MI)

Mutual Information (MI) measures how much knowledge two variables represent about each other. The mutual information is the difference between the sum of the marginal entropies and their joint entropy and thus two independent items have always zero mutual information.

In [102], maximum dependency criterion based on mutual information is used for feature selection, and experimented with various data classification accuracies. Conditional mutual information is used for speech recognition in [30]. In this study, we use Shannon's entropy to calculate mutual information.

Definition:

Let X and Y be two random variables, $p(x)$ and $p(y)$ be their probability density functions and $p(x, y)$ be their joint probability density function. Then their mutual information is defined as follows:

$$I(X;Y) = \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \tag{4}$$

The relationship between entropy and mutual information can be described as follows:

Let $H(X)$ denote Shannon's entropy of X, then

$$H(X) = -\int_x p(x) \log(p(x)) dx \tag{5}$$

Then entropy is related to mutual information as follows:

$$I(X;Y) = H(X) - H(X \mid Y), \tag{6}$$

$$I(X;Y) = H(Y) - H(Y \mid X), \text{ or} \tag{7}$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{8}$$

where $H(X \mid Y)$ denotes conditional entropy and $H(X,Y)$ is joint entropy.

As a feature selection criterion, the best feature will maximize the mutual information $I(X;Y)$, where X is the image feature vector and Y is the class indicator, which represents an image cluster.

4.1.2. Pearson's Product Moment Correlation Between Data Clusters

Inter-cluster information is widely utilized in cluster analysis for classifying data using multiple features. The discrimination characteristics of a feature for a given data set can be represented by determining the attributes of affinity between clusters (inter-cluster). The sums of correlations or distances are compared for evaluating the features within clusters of the feature data space constructed for each feature individually. Class separability is also a common measure to evaluate features by their inter-cluster relations. It represents how the distances among the means of classes are maximized. Usually, intra-cluster and inter-cluster information are used together in feature selection approaches to achieve more reliable results.

In [27], information concerning cluster compactness and cluster separability is used for unsupervised feature selection for content-based medical image retrieval. Class separability is also used in [112] for feature selection in a handwritten character recognition system. Class correlation is one of the criteria proposed in this study for evaluating the discrimination characteristics of a feature for a given set of data. It measures how cluster means are scattered with respect to each other. Large distances between clusters means better cluster discrimination. We use the following correlation measure for evaluating the discrimination power of each feature separately:

$$S = \sum_{x=1}^{c} \sum_{y=x+1}^{c} \delta\ (x, y) \tag{9}$$

where c represents the number of classes and δ represents the correlation between clusters x and y.

We use Pearson's product moment correlation for defining the correlation between cluster means. The Pearson's product moment correlation coefficient is the most commonly used measure of correlation in machine learning. It is calculated by summing up the products of the score deviations from the mean. We will use the following expression for the cluster correlation;

$$\delta(x, y) = \frac{\sum_{i=1}^{N_x} (f_{xi} - \mu_x)(f_{yi} - \mu_y)}{N_x N_y \sigma_x \sigma_y} \tag{10}$$

where f_{xi} and f_{yi} represents the i^{th} item in the cluster x and y, μ_x and μ_y, σ_x and σ_y are the means and the standard deviations, N_x and N_y are the cardinality of clusters x and y, respectively. Clusters x and y are assumed to have equal number of elements in order to be compared by PPMC.

4.1.3. Intra-Cluster Relation (ICR) Based on First Principal Component

Cluster members and their relations among each other can be represented by intra-cluster information. Intra-cluster information is a widely used criterion in cluster analysis. The main objective of intra-cluster analysis is to understand the existing pattern in a given data space. Intra-cluster information is also used for feature selection [27], [91]. The most common intra-cluster information is “compactness”, which is a measure of the similarity and closeness of the elements in a cluster. We use intra-cluster information as a criterion for feature selection in this study. If the elements of a cluster are close to each other in the represented feature space, or if the cluster is tight and compact, then the feature is considered as descriptive for the cluster. On the other hand, it is often difficult to make assumptions about the cluster shape and distribution with the compactness criterion. Especially, irregular shape clusters are particularly problematic.

We propose a new measure for intra-cluster information, Intra-Cluster Relation (ICR), which represents the intra-cluster scatter information using the principal component information of the cluster. It is also related to the closeness of cluster elements similar to compactness.

ICR can be obtained by performing the following steps for a given set of feature vectors corresponding to a cluster of N elements:

Step 1. The aim of this step is to derive a feature vector (π) that represents the given cluster in terms of direction and characteristic by applying principal component analysis. It is utilized for identifying patterns and highlighting the relations of the cluster elements. The first principal component e_1 is the eigen vector corresponding to the largest eigen value of the covariance matrix of the cluster. The best representative feature vector π can be constructed using the following formula:

$$\pi = e_1^T X \tag{11}$$

where X is a matrix containing set of feature vectors ($x_1,..x_i,..x_N$) for a cluster, and x_i represents the feature vector corresponding to the i^{th} item in the cluster. x_{ij} is the j^{th} element of the feature vector corresponding to the i^{th} item of the cluster ($x_{11},..x_{ij},..x_{N\ell}$). “$e_1^T$” denotes the transpose of the eigen vector e_1.

Step 2. In this step, the distance (Δ) between the representing feature vector π and mean vector M are calculated in order to get information about the distribution of the cluster elements.

The elements of *M* are the mean values ($\mu_1,..\mu_j,..\mu_\ell$) of the feature vectors in the cluster calculated as follows:

$$M = \left\{ \mu_1, .. \mu_j, .. \mu_\ell \mid \mu_j = \frac{1}{N} \sum_{i=0}^{N-1} x_{ij} \right\} \tag{12}$$

where N is the number of items in the cluster.

The sum of distances Δ is calculated with the following formula:

$$\Delta = \frac{1}{\ell} \sum_{i=0}^{\ell-1} (\pi_i - \mu_i)^2 \tag{13}$$

where ℓ is the number of elements in the vectors π and M, and is also equal to feature vector dimension.

Step 3. In this step, ICR value is calculated by normalizing Δ in order to improve performance of the criterion on the clusters, where the cluster shape is not symmetric and the cluster distribution is not even. Finally ICR is obtained as follows:

$$ICR = \frac{\Delta}{\frac{2}{N(N-1)} \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} d(x_i, x_j)} \tag{14}$$

where d is the Euclidean distance between cluster members and N is the number of items within the cluster.

4.1.3.1 Comparison of ICR and Compactness

We have chosen three different compactness definitions S_{w1}, S_{w2}, and S_{w3} used in the field of feature selection and cluster analysis, in order to compare to them with the proposed method. The compactness criteria approximately measure how scattered the cluster members are from their cluster means. The following equations present the definitions:

$$S_{w1} = \sum_{i=0}^{N-1} \left\| x_i - \mu \right\|^2 \tag{15}$$

where x_i is the i^{th} member of the cluster and μ is the mean of the cluster [26].

$$S_{w2} = (\mu_d + \sigma_d + r) \tag{16}$$

where μ_d and σ_d are the mean and standard deviations of the Euclidean distances between the cluster members respectively. r is the covering radius which is the distance from the center to the furthest item in the cluster [65].

$$S_{w3} = trace(\prod \sum) \tag{17}$$

where Π is the probability of the cluster among the database and equal to one in this example, and Σ is the covariance matrix of the cluster [27].

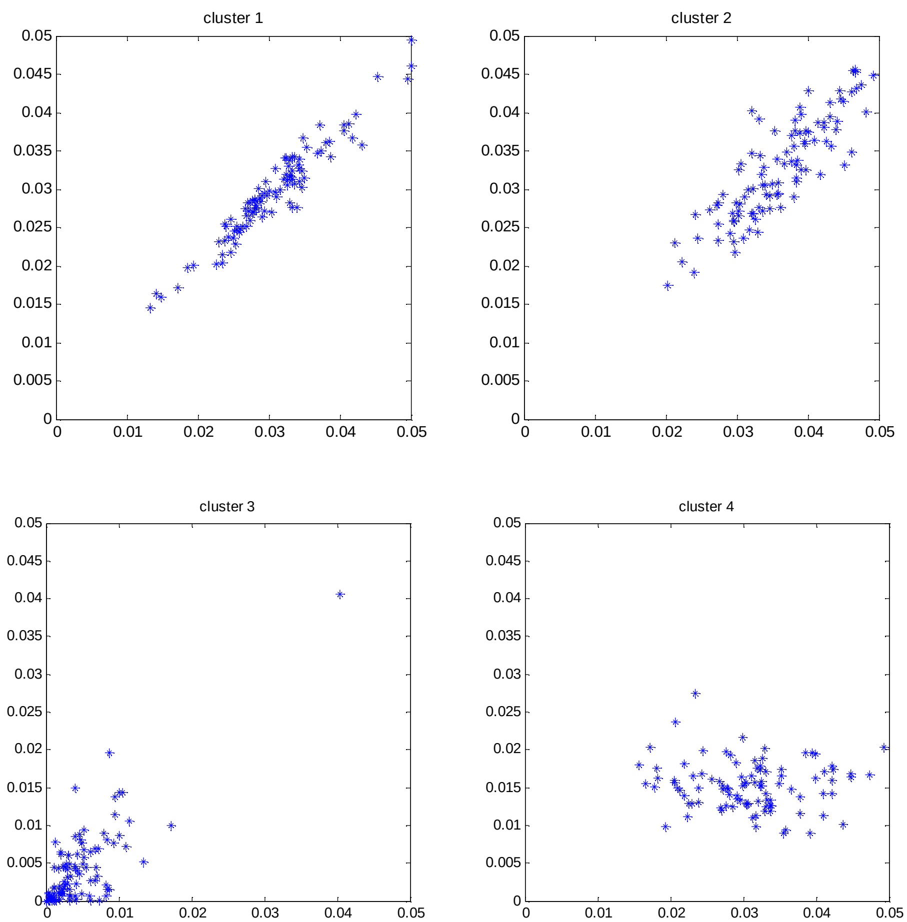


Figure 9: A Sample Data Set Represented with Four Different Sets of Features.

	Cluster 1	**Cluster2**	**Cluster 3**	**Cluster 4**
ICR	4.5306	7.8904	0.5669	11.2875
$\mathbf{S_{w1}}$	0.0112	0.0093	0.016	0.0077
$\mathbf{S_{w2}}$	0.0702	0.0620	0.11	0.0514
$\mathbf{S_{w3}}$	5.47 e-4	4.3 e-4	5.54 e-4	0.71 e-4

Table 1: ICR and Compactness Values of the Sample Data Clusters.

Figure 9 shows four different sets of two-dimensional features defined as Feature-1, 2, 3, and 4, and the clusters Cluster-1, 2, 3, and 4 are constructed respectively for each feature set for representing one sample data set containing subjectively similar images. Table 1 represents the ICR and three compactness values calculated for the sample data clusters given in Figure 9. The small values in the Table indicate the better cluster in terms of spatially closeness. x and y axes of Figure 9 represent the feature values, which are energy and entropy values of Gray Level Co-Occurrence Matrix (GLCM) texture features [99] and spatially closer cluster elements in the figure represent semantically related images selected from the Corel image database. In this respect, construction of Cluster-1 can be considered more successful than Cluster-2 and Cluster-4, since the elements in the Cluster-1 are spatially closer to each other than in Cluster-2 and Cluster-4. Comparing the ICR values ($ICR_{Cluster-1} < ICR_{Cluster-2}$) leads to the same conclusion; while, the other three compactness factors depict the opposite. Moreover, Cluster-3 elements can also be considered spatially closer than elements in Cluster-4. ICR values indicate the same ($ICR_{Cluster-3} < ICR_{Cluster-4}$); however, the other three compactness criteria show that Cluster-4 is the most compact cluster among the four clusters as given in Table 1. ICR gives better performance by the normalization of Δ value with the average distance between cluster elements in order to improve performance of the criterion on the clusters, where the cluster shape is not symmetric and distribution is not even. It can be observed from Figure 9 and Table 1 that ICR reveals better performance for describing the item distribution of the cluster.

4.2. MAJORITY VOTING FOR FEATURE SELECTION AND WEIGHTING

Voting is a widely used classifier combination technique used in various disciplines, particularly in multi-pattern recognition [75], [127]. Advantages of voting technique are generality, simplicity, and effectiveness. These advantages allow the method to be used in real-world applications as well as in social life as voting by majority. Additionally, voting may be used as a black box and it does not require additional internal information for the decision implementation.

In this study, we use the majority voting approach in the feature selection system. Majority voting selects the candidate having the largest amount of votes. In the feature selection system, different from the categorization problem, the output of the voting scheme is a feature list, which is sorted in descending order according to corresponding votes. The first feature in the output vote list represents the most important and powerful feature in discriminating the associated data.

Figure 10 represents the overview of the proposed feature selection system, where majority voting acts as a decision mechanism for the output feature list. Majority voting method for ranking the features works as follows:

The voting scheme gets the normalized numerical results from each individual criterion MI, ICR, and PPMC defined as υ_{MI}, υ_{ICR}, and υ_{PPMC}.

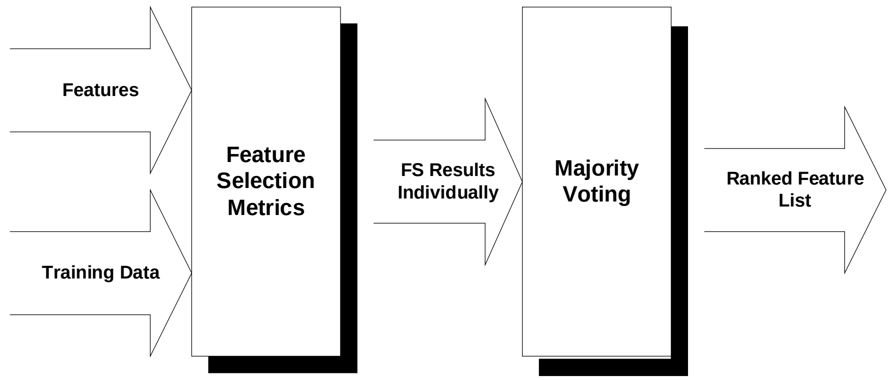


Figure 10: Overview of the Proposed Feature Selection System.

Votes are calculated for each feature using the following formula

$$\nu_{f_i} = \sum (\upsilon_{MI}(f_i) + \upsilon_{ICR}(f_i) + \upsilon_{PPMC}(f_i)) \tag{18}$$

The output Feature Selection Ranking List (FSRL) is constructed by sorting the features according to their ν_{f_i} values.

4.2.1. Feature Selection by Voting

The output of voting scheme is a list of features referred to as FSRL. It is used to define an appropriate number of features, which will be used automatically by the system, without the user's knowledge. Alternatively, the number of features can also be changed manually by the user according to the recommended features and FSRL list. In the ideal case, most definitive and discriminative features for querying a certain database are sorted and listed in the FSRL. The number of selected features is defined by a threshold determined from the FSRL list. The minimum gradient of the sorted list of votes in FSRL corresponding to the sharpest decrease can be used to define the threshold, where the features with higher votes are recommended for retrieval. The defined threshold gives the number of features that are recommended by the system to the user.

4.2.2. Feature Weighting by Voting

As mentioned before, most definitive and discriminative features for querying a certain database are sorted and listed in the FSRL in ideal case. Automatic weighting aims to improve semantic retrieval results by adjusting the weights of features according to their ranks in the FSRL list. Users of the feature selection system may require using all features in the FSRL instead of eliminating any of the existing features in the CBIR system. Appropriate feature combination should be given to the user in order to improve the semantic retrieval results of the CBIR system. The order of the features is given in FSRL list and an automatic weighting can be introduced to the user as follows:

Assume that the sum of the weights of the features is equal to 1. Thus,

$$\sum_{i=1}^{F} \alpha_i = 1 \tag{19}$$

where F is the number of features in the FSRL list. The weights of the features can be calculated as follows:

$$\alpha_i = \frac{(F - R_i) + 1}{\sum_{i=1}^{F} i} \tag{20}$$

where R_i represents the rank of the i^{th} feature in FSRL list, and F is the number of features.

4.3. EXPERIMENTAL RESULTS

4.3.1. Data Sets and Features

In the experimental studies, we utilized well-categorized Corel image data sets, which are widely used in the literature for training and testing purposes. Corel database with 10000 images are used for testing the method. These images are pre-assigned to 100 semantic classes each containing 100 images by a group of human observers. Some examples of the classes are autumn, balloon, bird, dog, eagle, sunset, and tiger. Another Corel image database including 1000 images categorized in 10 equal size classes is used for feature selection (training). In the first set of experiments, the following low-level color, shape, and texture features are used: YUV, RGB, and HSV color histograms with 128, 64, and 16 bins [115], Gabor wavelet texture feature [84], gray level co-occurrence matrix texture feature with parameters 12 and 6 [99], Canny edge histogram [8], and dominant color with 3 colors [88]. In the second set of experiments, the same training and test databases are used and only color features YUV, RGB, and HSV color histograms with 32, 16, and 8 bins are utilized.

Generally, feature selection systems need to use training data in order to decrease the computational cost due to large image databases used for CBIR purposes. Computational and storage complexity will be decreased if feature and class probabilities and class relations are obtained from the training data. On the other hand, construction method for the training data has considerable impact on the precision accuracy of the feature selection system. Usually, it is a challenging task to model general-purpose CBIR databases with training sets, since such databases contain random and irregular number of classes.

In this study, the training data are selected in a supervised way for evaluation and assessment of the methods. Corel database contains 100 classes, where 10 of them (Corel 1000 image database) are selected as training data. Feature subset selection is not employed since each feature is passed through the criteria individually.

4.3.2. Computation of Global Criteria

Global criteria values (υ_{MI} ,υ_{ICR} , and υ_{PPMC}) are the final output values from each criterion used as the input to the majority voting for final decision mechanism of the feature selection system. Each of the three criteria is applied separately for each individual feature in order to express clearly its effects on the data set for evaluation and global values are calculated as follows:

- Mutual information is calculated as shown in Equation 1 for each feature using the training database and the global mutual information value is the sum of these values.
- PPMC coefficients are calculated for every cluster combination for each feature, and the sum yields the global PPMC value.
- ICR criterion is applied to each cluster in the data set with each feature individually. The sum of the values for each feature gives the global ICR value.

4.3.3. Assessment of the Results

Evaluating the semantic performance of the CBIR system is a challenging task due to subjective nature of the semantic content of images. There are several methods commonly used in the literature as mentioned in Chapter 2. We have used average precision to assess the performance of the system. Precision is defined as the ratio of the number of relevant records over the total number of retrieved records. It is usually expressed as a percentage. Moreover, in order to express the accuracy and evaluate the usefulness of the proposed feature selection system, we have introduced a numerical assessment method described in the following steps:

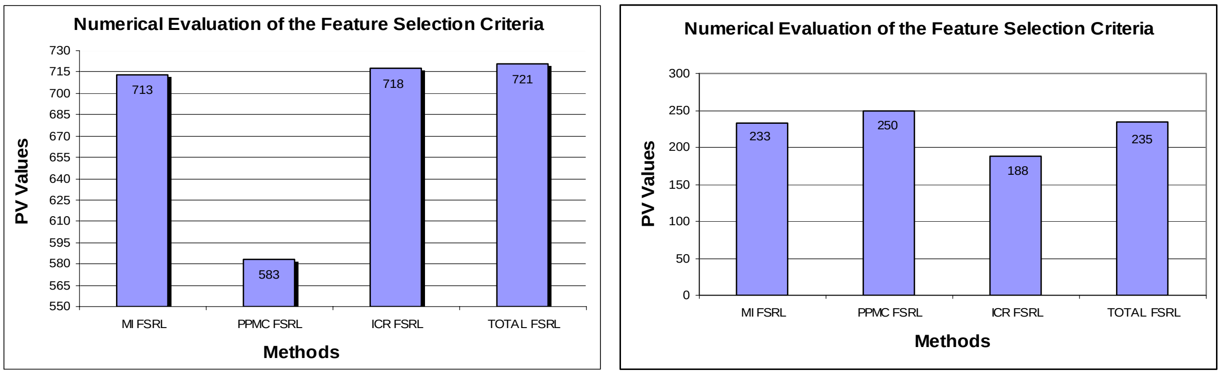
- Test retrieval experiments are performed using each feature separately on the test database, which has 100 classes each including 100 images for obtaining average precision values. Sample queries are performed with 500 images, 5 images from each class, and using each feature individually on the test database. In these experiments, 36 retrieved images are taken into account for calculating the average precision.

- Semantic retrieval performances for corresponding features are recorded according to the average precision values for verifying the feature selection method.
- Features are sorted according to the average precision values, equally by semantic retrieval performances. This step may also be expressed as sorting the features according to their representation level of the database. The sorted list is named as Semantic Performance Feature List (SPFL). SPFL is used for evaluating the results of a feature selection system.
- The output of a feature selection system is a list of features referred to as FSRL mentioned in Section 4.2. The proposed numerical assessment value, namely the Performance Value (PV) is calculated as follows:

$$PV = \sum_{i}^{N} \sum_{j}^{N} \omega_i . \omega_j \tag{21}$$

where ω_i $=$*N-{rank of item i in SPFL}+1*, represents the weight of item *i* in *SPFL;* and ω_j = *N-{rank of item j in FSRL}+1*, represents the weight of item *j* in *FSRL*.

Figure 11 (a) and (b) show the first and second set of experiments' performance values of each criterion calculated based on only their outputs referred as mutual information (MI FSRL), Pearson's product moment correlation (PPMC FSRL), and intra-cluster relation (ICR FSRL). Each criterion is separately calculated in order to compare them with the final feature selection ranking list results referred as TOTAL FSRL. It can be inferred from both figures that the performance value of the TOTAL FSRL is higher than other methods, which means that the final FSRL has high semantic retrieval performance and it is numerically closer to SPFL by the assessment method aforementioned. Note that, in the best case final FSRL should be equal to the SPFL, since it represents the semantic performance of the features sequentially. MI, PPMC and ICR criteria represent different characteristics of the features on the data. Each of these criteria may work better than the others in different cases, as it can be seen from the first and second experiments. The semantic effects of these results on image retrieval are presented in the following section.



(a) (b)

Figure 11 (a) and (b): Numerical Results of the Proposed Feature Selection Criteria for the First and Second Set of Experiments Respectively.

4.3.4. Comparisons of Compactness-ICR and Separability-PPMC with PV values

The proposed ICR and PPMC criteria are compared with well-known methods S_{w1} compactness (see section 4.1.3.) [26] and class separability, [27] by the numerical evaluation approach described in Section 4.3.3.

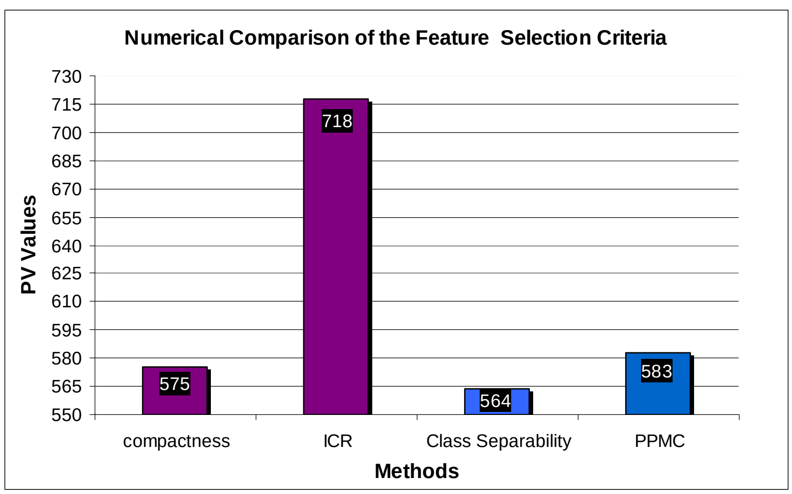


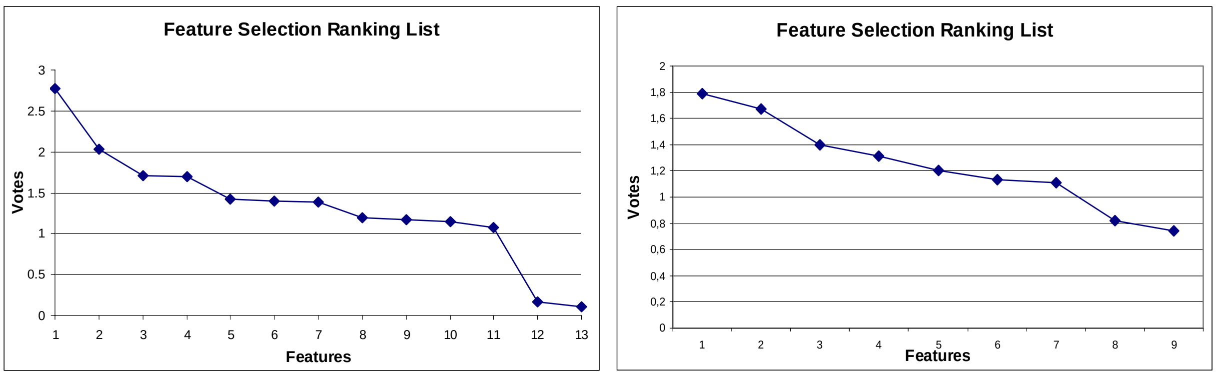
Figure 12: Numerical Comparisons of the Feature Selection Criteria.

Figure 12 presents the PV values of the retrieval results for each list of features constructed by the criteria. ICR outperforms compactness and PPMC outperforms class separability in terms of semantic retrieval performance based on the values in the figure.

4.3.5. Semantic Retrieval Results for Image Databases

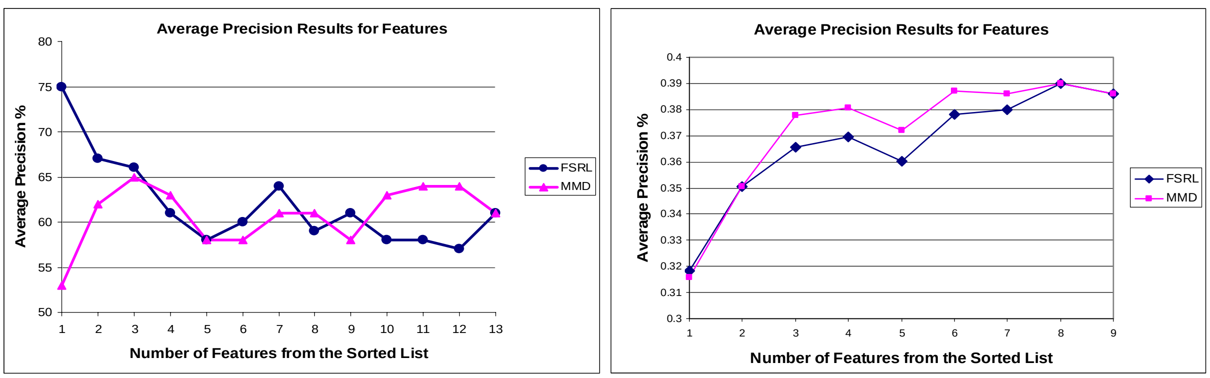
Semantic retrieval performance of the proposed system is evaluated with average precisions, which are obtained from experiments on Corel database including 10000 images (100 classes, each including 100 images). 500 queries using five images randomly selected from each class are considered while calculating the average precision values with visual features from FSRL list. Note that, FSRL list is a sequentially ordered feature list, which is sorted according to the representation and description properties of the features.

The proposed system is compared with the Maximum Marginal Diversity (MMD) method for feature selection of image retrieval systems presented in [118]. MMD selects the subset of features that leads to a set of maximally diverse marginal densities. We have constructed a feature list calculated based on MMD method similarly to the construction of FSRL. Majority voting with threshold approach mentioned in Section 4.2.1 is utilized to select the number of features for both methods in order to compare the criteria apart from the decision mechanism.



(a) (b)

Figure 13 (a) and (b): Votes of the Features in Final FSRL List for the First and Second Set of Experiments Respectively.

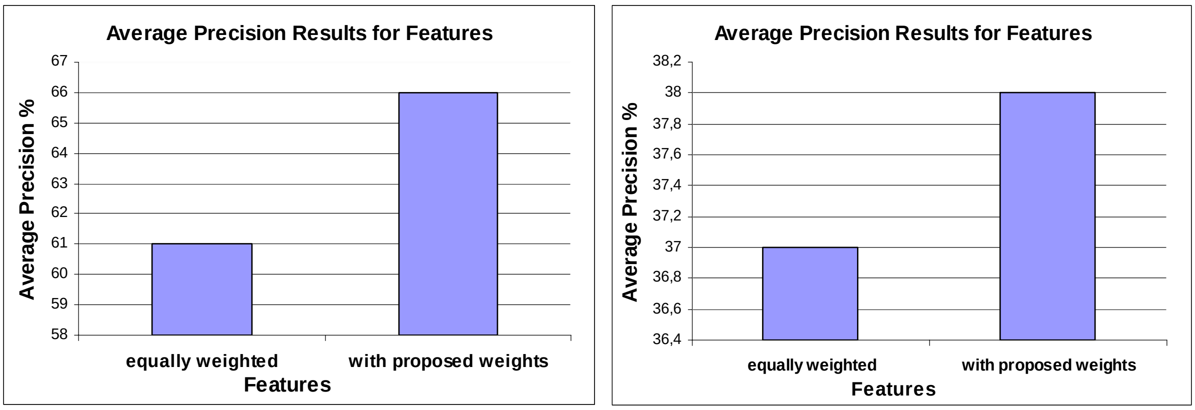


(a) (b)

Figure 14 (a) and (b): Average Precision Values of Features Employed in the Experiments for the First and Second Set of Experiments Respectively.

Figure 13 (a) and (b) illustrate the votes for each feature in the FSRL. In the first set of experiments shown in Figure 13 (a), the database has 13 features given in Section 4.3.1, and only the first feature is selected for image retrieval based on the votes, since the first feature (Feature 1) gives the maximum gradient. In the second set of experiments shown in Figure 13 (b), 9 features are used, 7 features are selected for image retrieval based on the votes by calculating the maximum gradient.

Figure 14 (a) and (b) illustrate the results of retrieval using the features in the FSRL and MMD lists. The numbers in the x-axis of Figure 13 refer to the rank of the underlying feature, e.g. 1 is for Feature 1, 2 for Feature 2, etc. On the other hand, the same numbers in Figure 14 refer to the number of features from the beginning of the list, e.g. 1 for the first feature, 2 for first 2 features, etc. In the ideal case, a feature selection system should select the number of features, which correspond to the highest semantic retrieval performance. In other words, the number of features as x-axis values of Figure 14 should have highest average precision value on the recommended number of features and tend to decrease or remain constant along the x-axis after that. In the first example, the proposed system selects only one feature. It is verified by the results depicted in Figure 14 (a) as the highest precision is at the first feature and starts to decrease afterward. The results with MMD in the same figure suggest that using three features gives the highest precision. However, it is still lower than the precision obtained with the proposed feature selection system. In the second example given in Figure 14 (b), the proposed feature selection system selects 7 features, which have inconsiderably (% 0.05) lower average precision than the maximum value that can be obtained utilizing existing features in the system. It is verified by the results depicted in Figure 14 (b) as the precision slightly increases with the 8th feature and decreases after that. In Figure 14 (b), MMD suggests five features and the final MMD average precision is lower than the result of the proposed system, although there are higher precisions shown partially in the MMD results.



(a) (b)

Figure 15 (a) and (b): Average Precision Values of All Features with Different Weights for the First and Second Set of Experiments Respectively.

An alternative way of utilizing FSRL is an automatic setting of feature weights for image retrieval. The weights for each type of feature are calculated automatically by the system using FSRL as described in Section 4.2.2. In the experiments of Figure 15 (a) and (b), first all existing features in the system are combined with equal weights, then the weights of the features are replaced with the proposed values given automatically by the system. As shown in Figure 15 (a), the average precision of 61% is increased to 66% when the proposed feature weighting is employed and in Figure 15 (b) it is slightly increased from 37% to 38%.

4.4. SUMMARY

The proposed feature selection and weighting method is aimed to enhance semantic content-based image retrieval performance, to decrease retrieval process complexity, and to improve system usability for end-users. Three different criteria and a majority voting approach for the final decision making step, where each criterion represents different feature characteristics is utilized in the proposed system.

The novelties of the proposed feature selection approach in a CBIR system are:

- A new criterion based on categorized member relation information within the same cluster,
- A new criterion for defining the correlations of inter-clusters, which is based on Pearson’s Product Moment Correlation in order to determine the discrimination power of the feature,
- The three different criteria with majority voting approach as a decision making mechanism.

Mutual information criterion refers to the amount of information a feature carries about the data. MI gives a clue about the representation power of the feature on a particular image database. Intra-cluster relation criterion helps to show the description power of the feature for each labeled cluster, where members of a cluster are supposed to be similar and close to each other in the feature space. ICR is slightly similar to the compactness of clusters; however, its analysis accuracy is higher for irregular shape clusters. Pearson’s product moment correlation criterion represents the correlation between clusters, where each of them is uncorrelated in the ideal case. Instead of cluster separability, we use PPMC, since it expresses the discrimination power of the feature better. Majority voting is adopted for the use of feature listing in CBIR context. It generates a sorted vote list referred to as feature selection ranking list with associated feature names. FSRL may be utilized to obtain one of the two different outputs to the end-users as follows:

- Recommended set of features,
- Recommended weights for each feature.

Outputs of the proposed feature selection system may bring out various use-cases according to the computational power capacity of CBIR system platform. In case of automatic usage of the recommended set of features, semantic retrieval performance is increased by 22% as shown in the experimental results. On the other hand, all existing features with recommended weights can be utilized for retrieval to improve the semantic performance (approximately 10% as shown in the experimental results) without altering the complexity. Especially the first use-case may be applied on limited platforms having low capacity for computing features and all the features may not be used for CBIR.

Naturally, the proposed system in CBIR can be used automatically or in a supervised manner by an end-user. In the unsupervised case, the system internally uses feature combinations and weights automatically without users' interaction. In the supervised case, the user selects the features and weights to be used for retrieval based on the system's recommendation.

Forming an appropriate training data is a major challenge in this study. The proposed method should be utilized on a well representative training data for successful results in large image databases.

Flexibility and efficiency of the proposed approach allows it to be applied in various platforms and data. The success of the proposed approach is confirmed with the experimental results on image databases.

Chapter 5

Transform-Based Layered Query Scheme

Improving the performance of indexing and retrieval plays an important role for providing advanced CBIR services. Retrieval performance improvement is more visible for end-users of a CBIR system, although indexing may affect retrieval directly. An end-user may be directly interested in the query performance of a retrieval system, however higher user satisfaction can be achieved with improvements in the following challenges as mentioned in Chapter 3:

- Computational complexity,
- Run-time memory and disk space requirements,
- Semantic retrieval performance, and
- Usability.

Semantic retrieval performance is the most critical one among these problems and shall not be degraded significantly by any means. First and second problems may be solved using advanced algorithms in indexing and retrieval system implementations. However, a more generic solution is employing smart and efficient design schemes in the architecture level of the system. Dedicated pre- or post-processing components for image and feature data may be integrated into CBIR applications such as dimension reduction methods. Lower dimensions in both image and feature data lead to lower complexity, memory and disk space consumption. Transform-Based Layered Query (TLQ) Scheme is a new audio/visual media querying scheme that employs transformation and dimension reduction methods [39]. It is a progressive scheme consisting of multi-layers, which are constructed based on transformations. One of these transformation methods is Mapping by Adaptive Threshold (MAT), which is a new method proposed in this thesis for dimension reduction of feature data [36]. DCT- and DWT-based downscaling may also be used in TLQ for image dimension reduction [38], [41].

The organization of this chapter is as follows: Section 5.1 presents the overall TLQ scheme in details. A sample TLQ scheme integrated into the MUVIS content-based multimedia indexing and retrieval framework is described in the second chapter of the thesis. Section 5.2 presents dimension reduction methods, particularly focusing on the proposed MAT method, and DWT and DCT-based downscaling. Experimental results on the TLQ scheme and the practical benefits of the proposed MAT method are given in Section 5.4. Finally Section 5.5 provides concluding remarks and discussions.

5.1. TRANSFORM-BASED LAYERED QUERY (TLQ) SCHEME

As the name implies, a content-based image indexing and retrieval system contains two separate consecutive phases: Indexing and Retrieval. In the indexing phase, a given set of images are organized into one or more image databases based on their content information for efficient browsing and retrieval. This may involve feature extraction, sequential indexing, and clustering processes. The indexing phase is usually an offline process, since it may not require any real-time end-user interaction. Emphasizing one or more images from indexed image databases via a process such as querying or browsing constitutes the retrieval phase. One of the most common approaches for retrieval is query by example, which is referred shortly by the term “retrieval” or “query” in this thesis. Similarity measurements, ordering and rendition of the results are the main processes involved in this phase. The retrieval phase requires end-user interaction, thus it may be defined as an online process.

Transform-based layered query is an indexing and retrieval scheme proposed for reducing the overall computational complexity and memory consumption. The main elements of a TLQ scheme are two dedicated types of transformations, namely T1 and T2 as illustrated in Figure 16. T1 and T2 represent transformation methods for dimension reduction of image and feature data, respectively. Although a TLQ scheme does not directly depend on any specific transform method, the underlying framework and transformations have to follow the assumptions given below for achieving the following overall system targets:

- The indexing and feature extraction processes depend on the frame size in terms of memory usage, and computational complexity.
- Similarity measurement and feature extraction process depend on the number of images in terms of computational complexity.
- The query process depends on the dimensions of the feature data in terms of computational complexity.

T1 represents any image transformation method that can be used to reduce the size of an image, hence decreasing the complexity of the feature extraction process. Image downscaling is a good example of such a transformation. The most critical characteristic of T1 transformation is to preserve significant information content of an image. T1 transform should also be a low complex method for gaining in the overall processing time. The choice of the transformation and the features to be extracted should be considered together to achieve satisfactory semantic performance. Moreover, feature extraction methods should be able to extract significant information from the transformed multimedia data. Although T1 transformation improves the performance of the indexing phase in terms of feature extraction time, it may degrade the semantic retrieval performance. Thus, it is kept as an optional part of the scheme as indicated in Figure 16. Feature extraction methods may be applied on original images instead of the transformed data, if an advance in indexing phase is not required or cannot achieve satisfactory retrieval results.

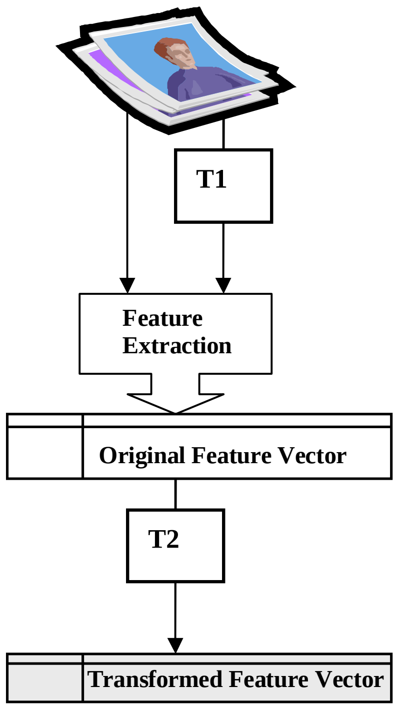


Figure 16: Overview of Indexing Phase of TLQ Scheme.

The second transformation in the TLQ scheme, namely T2, refers to any transformation method for dimension reduction of image feature data. It has direct impact on the retrieval process complexity and memory consumption. T2 should keep the significant feature data, which characterize the corresponding image distinctly among others in the database. Some transformation methods may have a filtering-like effect that may emphasize differences between feature data of an image, and may thus enhance the semantic retrieval performance. This transformation may also motivate the use of complex similarity measurements for the same purpose.

The indexing and retrieval phases of the proposed TLQ scheme utilize the so-called Indexing and Retrieval Layers, respectively. The indexing layers are constructed based on T2 transformation methods, and each indexing layer constitutes the base for the corresponding retrieval layer in the retrieval phase. These layers are further described in the following sub-sections for indexing and retrieval phases.

5.1.1. Indexing Phase of TLQ:

The most significant process in the indexing phase of the TLQ scheme is the construction of abstract indexing layers using T2 transformation methods. Figure 16 represents the overview of the indexing phase in TLQ scheme. Each abstract layer in TLQ scheme corresponds to feature vectors with different vector sizes according to the layer rank. A TLQ scheme contains at least two layers. The uppermost layer is indexed using the original feature data. The base layer uses smallest size feature data that likely to carry the least representative information. The original size feature data are extracted from the original or optionally T1 transformed image. Reduced size feature vectors are generated for lower layers by transforming the original size feature data using transformation T2. For each intermediate layer, a different T2 transformation method may be used in order to further reduce the size of the feature data. Reduced size feature data of each layer are used in the corresponding layer in the retrieval phase for retrieving intermediate results. Intermediate retrieval results of a given layer are expected to be more accurate than preceding (lower) layer's results.

Using T1 has certain benefits in the indexing phase of the database, such as feature extraction time and memory consumption reduction. However, it brings processing overhead at the same time. In most CBIR systems, such overheads are not relevant for the end-users since indexing is an offline process.

5.1.2. Retrieval Phase of TLQ:

The retrieval phase of a TLQ scheme is divided into L abstract retrieval layers where L is the number of indexing layers constructed in the indexing phase. Each retrieval layer corresponds to the indexing layer in the same rank. The whole retrieval process starts with the lowest (base) layer, which uses the lowest size feature data and the whole database. If the user proceeds with the higher layer, the retrieval process uses the corresponding layer feature data. The retrieval process in the next layer is performed only within the intermediate retrieval results of the current layer instead of the whole database. In the uppermost layer containing the original size feature vectors, the query process reveals the final retrieval results. The number of images in the intermediate retrieval results of each layer is based on a system- or user-defined parameter. The user may explicitly specify the number, define the distance threshold that will lead to a certain number in each layer, or use an automatic method for discarding certain number of images.

A novel method on unsupervised elimination of irrelevant media items for retrieval is presented in [37]. The elimination method uses first retrieval results to eliminate irrelevant items for further retrieval steps. The technique is only suitable for systems that have more than one query steps in order to improve the retrieval results progressively. Moreover, it helps to decrease UI complexity by decreasing the number of items that has to be displayed to the end-users. Experimental studies show that the elimination method gives satisfactory results. In the experiments, the error rate of missing relevant images in the eliminated part of image database is ~0.01%. Low complexity and successful accuracy allows the proposed method to be used in various platforms such as distributed and limited systems that will be defined later in this thesis.

In TLQ, an easier alternative for the user may be browsing the retrieval results of a layer and marking the relevant ones as the member of the corresponding layer's intermediate results.

Using T2 transformed feature data yield lower retrieval processing time, since querying heavily depends on feature vector comparisons and distance calculations. Additionally, elimination of irrelevant multimedia items in each layer leads to lower retrieval process time in the higher layer.

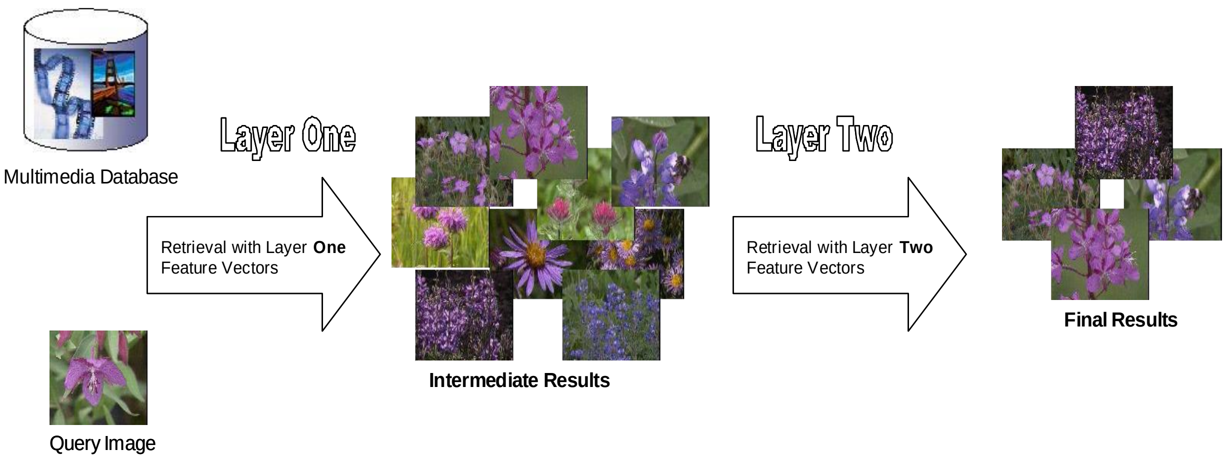


Figure 17: Retrieval Phase of a Two-Layer TLQ Scheme.

Whenever the user is satisfied with the intermediate results of a layer, the query may be halted as there is no need for proceeding with the higher layer. Moreover, if the underlying system allows relevance feedback, an experienced end-user may adjust query parameters (e.g. feature types and weights) between the layers for improving semantic retrieval performance. Figure 17 illustrates the retrieval phase of a two-layer TLQ scheme. In the sample layer one, a certain number of related items are retrieved from a large image database using layer one feature vectors. Layer two's retrieval is performed only within the intermediate results using layer two feature vectors in order to have refined final retrieval results.

5.2. TLQ SCHEME IMPLEMENTATION AND TRANSFORMATION METHODS

A two-layer TLQ scheme is implemented and integrated into the MUVIS framework for experimental studies. In each layer, MUVIS employs a sequential indexing and retrieval method that allows observing the benefits of TLQ more clearly in the experimental studies. DCT-based downscaling is used as T1 transform for image data due to its efficiency in JPEG image downscaling and retrieval performance [40]. The details of DCT- and DWT-based downscaling and their effects on image retrieval are presented in Section 5.2.2. On the other hand, two different transformations are used as T2 transforms for feature data dimension reduction. These are Principal Component Analysis (PCA) and a novel technique called Mapping by Adaptive Threshold (MAT) [36], [59]. T2 transform reveals some advantages for the indexing process in the implemented TLQ scheme. The dimension reduction methods are further described in Section 5.2.1.

5.2.1. Dimension Reduction Methods for Feature Data

In CBIR, dimension reduction methods are mostly utilized for reducing the dimension or size of the feature data. During the reduction process, distances between multimedia feature vectors should not be affected significantly for successful retrieval results. On the other hand, it may be beneficial to increase the distances for irrelevant items (with respect to the query item) while decreasing them for the most relevant ones to improve semantic results. In general, the main objective of utilizing a dimension reduction method in TLQ is decreasing the retrieval process run-time complexity and memory requirement. Thus, the dimension reduction method itself is expected to have a reasonable complexity, as not to offset the expected gain.

Principal Component Analysis (PCA) has been widely used as dimension reduction method for various purposes, as well as for CBIR features. PCA is further described in the following Section 5.2.1.1. Section 5.2.1.2 presents a new mapping by adaptive threshold based dimension reduction method that is employed in the two-layer TLQ scheme of MUVIS.

5.2.1.1 Principal Component Analysis (PCA)

Each multimedia item has a feature data that can be represented using an N-dimensional feature vector X. PCA is a well-known technique to map N-dimensional vectors into M-dimensional vectors, where $M < N$.

Principal components of an original feature vector X can be denoted as:

$$Y_{pca} = E^T X \tag{22}$$

where E is the N x M matrix containing M eigenvectors corresponding to the M largest eigenvalues of the data covariance matrix.

Choosing the M largest eigenvalues and their associated eigenvectors yields the minimum least mean square error of approximating the data [59]. PCA has been used for CBIR feature data dimension reduction, and may also be employed in the TLQ framework. However, computing the covariance matrix and its eigenvalue decomposition are computationally expensive processes. We proposed a new dimension reduction of feature data approach, which can be used in TLQ instead of PCA and is computationally less complex. This approach will be further described in the following section.

5.2.1.2 Mapping By Adaptive Threshold (MAT)

The proposed MAT method essentially consists of an adaptive threshold and non-overlapping window-based mapping functions. In the adaptive thresholding step, one of the feature vector values is selected as the threshold value according to a scaling factor defined by the indexing administrator. Then the thresholded feature vector is divided into windows having fixed size equal to the scaling factor. The mapping function assigns one representative value for each window similar to the vector quantization method (VQ) [35]. There are three main differences from VQ, however. First, the mapping function does not consider the distances between the elements in each window during this assignment. Second, the representative value is usually one of the values inside the window. Finally, the number of elements in each window is fixed and is equal to the scaling factor in the mapping function.

The MAT method reduces the dimensionality irrespective of any correlation that might exist among the elements of the vector. The following steps represent the proposed MAT method in details:

Step 1: Adaptive Threshold

Let X be the original N-dimensional feature vector, and M represents the target dimension, which equals to N/S where S is the user-defined scaling factor.

$$X = \{x_1, x_2, x_3 ... x_N\} \tag{23}$$

The values of vector X are sorted into X' in descending order.

$$X' = \{x_{(1)}, x_{(2)}, x_{(3)} ... x_{(N)}\} \tag{24}$$

where $x_{(i)}$ is the i^{th} largest sample of X.

The M^{th} value in X', $x_{(M)}$ is selected as the threshold, and X'' is constructed by thresholding the original feature vector X.

$$X'' = \left\{ x_i'' \mid x_i'' = \begin{Bmatrix} x_i, & if \quad x_i \geq x_{(M)} \\ 0 & otherwise \end{Bmatrix}, i = 1 \ldots N \right\} \tag{25}$$

Step 2: Mapping

The vector X'' is divided into M non-overlapping windows. A simple mapping is performed on these windows to construct the final decimated vector Y. A representative value is assigned for each window, where the representative value of the i^{th} window will be the i^{th} element of vector Y. Function G finds the representative value given a window of X''.

$$Y = \left\{ y_i \mid y_i = G\left(x''_{(i-1)S+1} \ldots x''_{iS}\right), i = 1..M \right\} \tag{26}$$

Some windows may contain only 0's, in this case the representative value of that window is also 0. Function G (.) may be the mean, maximum, or median operation. The following pseudo code expresses Step 2 more clearly, when G is the maximum operation.

```
Y = Zeros(M)    // Final feature vector
for i = 1:N   // Find window index
     w = round( i / S )
     // Maximum selection
    if ( X''[i] > Y[w] )
        Y[w] = X''[i]
    End if
End for
```

Figure 18 illustrates the original values of a HSV histogram feature vector, and Figure 19 illustrates the result of the proposed MAT method on the same feature vector when G is the maximum function and scaling factor is four. HSV color histogram feature vector represents the distribution of colors in an image, derived by counting the number of pixels of a given set of color range, illustrated as vector X in Figure 18. Similarly, the decimated feature vector Y shown in Figure 19 expresses the dominant color distribution in an image mapped from the original size feature vector.

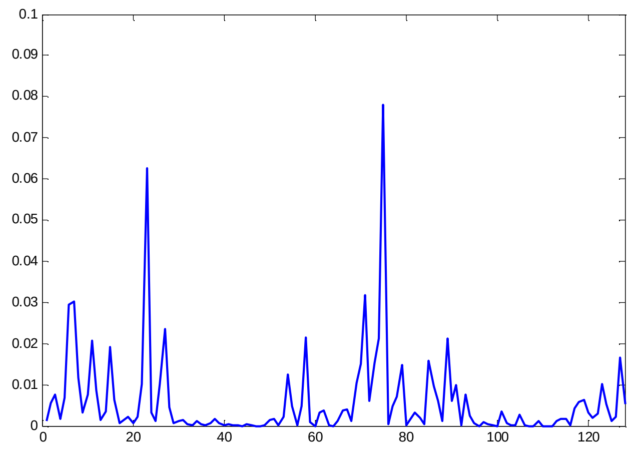


Figure 18: Original HSV Color Histogram Feature Vector X.

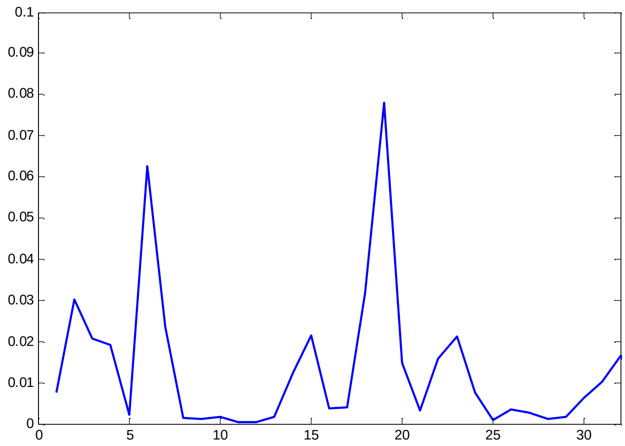


Figure 19: Decimated Feature Vector *Y* Obtained Using the MAT Method.

5.2.2. Image Downscaling Effects on CBIR Performance

Due to large space requirement of image databases, CBIR systems have the tendency to deal with compressed data. Using compressed domain feature extraction algorithms, it is even possible to extract information about the content directly from the compressed data. Image compression, filtering [83] and image transformation [97] techniques are the possible factors that may affect the image pixel values and hence image feature descriptors. Additionally, in certain cases compression methods may degrade image retrieval performance as presented in [40]. Alternatively, one can reduce image and video data sizes by downscaling the frame size [11]. The so-called thumbnails have been widely used for display purposes in retrieval terminals [17]. Besides, downscaling may also be useful for overcoming large memory and disk space as well as high processing capability requirements, particularly in limited applications and devices, e.g. mobile devices. Image downscaling may also have filtering effects on the color information due to smoothing the edges, as can be seen from the sample original and downscaled size images given in Figure 20. On the other hand, downscaling methods may cause crucial information loss, which in turn degrades the retrieval performance.

The main motivation for using image downscaling in TLQ scheme is decreasing the feature extraction process run-time complexity and memory requirement during indexing. DWT- and DCT-based downscaling are reasonably low-complex methods as a T1 transform [38], [41]. Furthermore, their limited effects on semantic retrieval performance were verified via image retrieval experiments using the MUVIS framework [38], [41]. In MUVIS framework, feature extraction process is performed on decoded and downscaled images. Image downscaling process is managed in image decoder (i.e. JPEG codec), which has an input parameter as scaling factor after decoding process.

Figure 20: Sample Original Size and Downscaled Images.

Image downscaling retrieval experiments are handled separately for color, texture, and edge based features. Experiments start with the original size image databases, and an average precision value (AP value) for each query image is obtained. Each calculated AP value is assumed to be the maximum that can be obtained for the corresponding image query. Thus, these AP values are used for normalizing further AP values that are obtained in the scaling factor experiment cases. We have utilized histograms in RGB, YUV, and HSV color spaces, Dominant Color, and Color Correlogram features as color features [49]. Gabor wavelet transform and gray level co-occurrence matrix features and block-based ordinal co-occurrence matrices are extracted for retrieval experiments [100]. These texture-based features are widely used for CBIR systems and readily available in the MUVIS framework. Canny and Multi-scale edge detection features are utilized in edge-based retrieval experiments [8], [62]. Both edge features are employed in MUVIS framework.

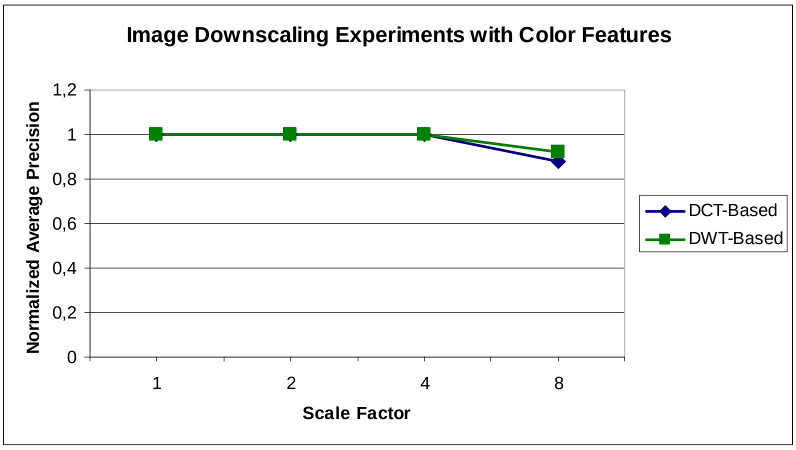


Figure 21: Results of Image Downscaling Experiments with Color Features.

In DCT-based image downscaling retrieval experiments, scaling the images does not affect the retrieval performance significantly for color visual features as shown in Figure 21.

The reason of this retrieval performance robustness can be explained with insignificant color information loss on the preserved image pixels after DCT-based downscaling. Considering the color features in general, downscaling by 8 achieves a satisfactory retrieval performance, where a satisfactory and efficient performance can be regarded as accomplishing the highest scaling factor without degrading retrieval performance significantly.

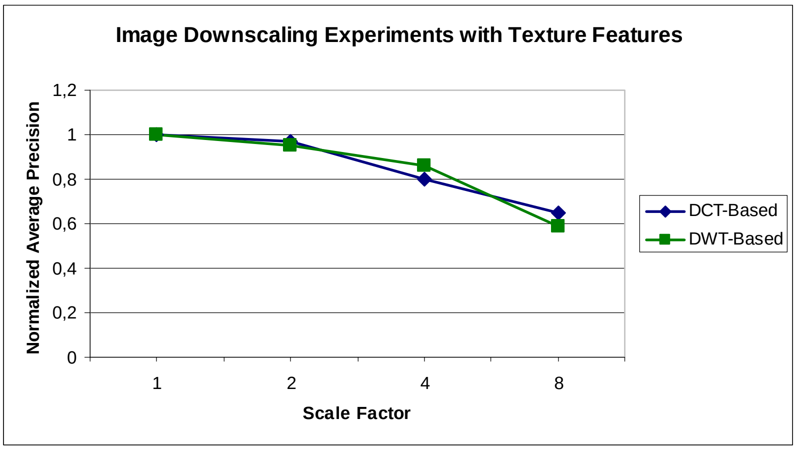


Figure 22: Results of Image Downscaling Experiments with Texture Features.

Figure 22 represents the retrieval results of the experiments with texture features. DCT-based image scaling tends to affect the texture-based retrieval performance drastically. It modifies the spatial information of image pixels, and it in turn affects the performance since texture features depend on spatial information. This effect can also be explained particularly considering the texture feature extraction methods. For example, each entry in GLCM corresponds to the number of occurrences of the pixels in a certain neighborhood. Hence, gray level co-occurrence matrix texture feature can be easily affected by downscaling.

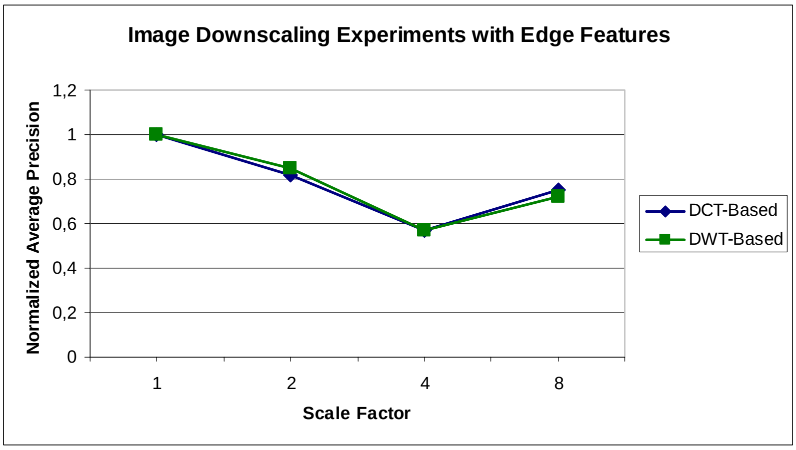


Figure 23: Results of Image Downscaling Experiments with Edge Features.

In color-based image retrieval experiments, DWT-based downscaling does not affect query performance significantly for any of the scale factors as represented in Figure 21. The reason of this retrieval performance robustness can be explained with insignificant color information loss on the preserved LL bands after DWT-based downscaling. Generally, scaling by 4 achieves a satisfactory retrieval performance for color visual features. Texture-based and edge-based image retrieval performances are given in Figure 22 and Figure 23, where they are negatively affected by DWT-based downscaling.

Most of the texture and edge information is preserved in the high frequency sub bands and they are eliminated by downscaling. This information loss causes a decrease in texture and edge-based retrieval performance. Image downscaling may enhance the edges between two objects having different colors in the image due to high frequency component loss. Therefore, this enhancement may cause slight (approximately 10%) increase in retrieval performance for high valued scaling factors.

5.2.3. Typical Retrieval Use Cases of Two-Layer TLQ Scheme

Generally, a retrieval scheme of a two-layer TLQ scheme has two typical use cases, in which performance benefits can be described clearly:

Use Case 1:

- User selects the query image, adjusts the query parameters and weights, and sets the number of items in the intermediate Layer One results.
- User starts the query for the first layer. Intermediate results are displayed for the user.
- User wishes to proceed with more advanced second layer query. If the underlying system allows, user can give relevance feedback and modify query parameters.
- User starts the query for the second layer. Final results are displayed.

Use Case 2:

- The first two steps are the same as in Use Case 1.
- User is satisfied with the intermediate results, and stops the query phase.

5.3. PROCESS TIME ANALYSIS OF TWO-LAYER TLQ SCHEME

The performance advantages of the two-layer TLQ scheme in terms of indexing and retrieval process time can be analyzed analytically and verified experimentally. For the former, we make the following assumptions:

- Feature extraction processing time is directly proportional to image dimensions,
- T1 and T2 transform methods have parameters k_{T1} and k_{T2} representing the scaling rates respectively, and
- The number of intermediate query results is $\frac{1}{k_{IQR}}$ of the total number of images.

Based on these assumptions, the total database indexing time can be estimated by a linear function:

$$F_p\left(C_p, n_p, m_p\right) \tag{27}$$

where *Cp* is the invariable factor that covers all factors affecting feature extraction time independent from database size and image size , n_p is the number of images in the database, m_p is the average number of pixels per image, and *p* refers to the underlying system ("*tlq*" for TLQ and "*o*" for Ordinary).

Since T1 and T2 transforms are involved in the indexing process, C_{tlq} is slightly higher than C_o. The overheads of these processes are considered as negligible compared to the feature extraction process itself.

$$C_{tlq} > C_o \;\; and \;\; C_{tlq} - C_o << F_p \tag{28}$$

If $k_{T1} = 4$, then

$$m_o = 16 * m_{tlq} \tag{29}$$

$$n_{tlq} = n_o \tag{30}$$

Finally,

$$F_o \cong 16 * F_{tlq} \tag{31}$$

In other words, the total indexing performance gain of the two-layer TLQ is approximately 90%.

Since T1 transform benefits only the offline indexing process, it may be discarded from TLQ scheme. In this case, F_{tlq} will be slightly higher than F_o.

Similar to F_p, the total query process is also estimated by a linear function:

$$Q_p\left(R_{p,} n_p, v_p\right) \tag{32}$$

where R_p is the invariable factor that covers all factors affecting query process time independent from database size and feature data size , n_p is the number of images, and v_p is the dimension of the feature vector.

$$Q_{tlq} = Q_{L1} + Q_{L2} \tag{33}$$

where *L1* and *L2* refer to the first and the second layer, respectively.

$$R_{L1} = R_{L2} = R_o \tag{34}$$

$$n_o = n_{L1} = k_i * n_{L2} \tag{35}$$

$$v_o = v_{L2} = k_{T2} * v_{L1} \tag{36}$$

If $k_i = 4$ and $k_{T2}=4$, then

$$Q_o = 4 * Q_{L1},\ \ Q_o = 4 * Q_{L2},\ \ and\ \ Q_o = 2 * Q_{tlq} \tag{37}$$

In other words, the performance gain is 50% for Use Case 1, and 75% for Use Case 2. These analytic results will be validated experimentally in the next section.

5.4. EXPERIMENTAL RESULTS

The implemented two-layer TLQ scheme is studied practically with retrieval experiments, in which the scaling rate is set equal to 4 (k_{T1} and k_{T2} = 4). The experiments are performed on a personal computer with Intel Pentium IV 2.8 GHz and 1 gigabyte Random Access Memory (RAM), running Microsoft Windows XP operating system. MUVIS Content-Based Indexing and Retrieval system is utilized as the experimental framework. Three Corel image databases are utilized for the experiments. Image databases labeled with numbers 1, 2, and 3 contain 10000, 30000, and 60000 images, respectively. Multiple low-level color, shape and texture features are extracted from the experimental databases to be used during the retrieval based on query by example. Features of database items are compared to corresponding features of the queried item, and the measured distances are merged using a weighted mean for final retrieval results.

Database / Query Type	Image Database #1	Image Database #2	Image Database #3
Ordinary Query Time	7.8	54	200
TLQ Layer 1 Query Time	2.4	12	42.5
Optional TLQ Layer 2 Time	2.2	14	44
Total TLQ Time = TLQ Layer 1 + Optional TLQ Layer 2	4.6	26	86.5

Table 2: Average Query Process Times in Seconds.

Table 2 presents the results of the TLQ experiments. Each number in the Table 2 refers to the average process time in seconds for 20 queries for the corresponding database and query type. The first row of the Table contains the times corresponding to ordinary query process not employing the TLQ method. The other rows give the times for the corresponding

TLQ layers and their sums. Table 2 also shows that, theoretical assumptions and practical experimental results point to approximately the same performance benefits. Moreover, it is likely to achieve further performance benefits in practice, since the user may already be satisfied with the intermediate layer one results. Figure 24 presents relatively successful layer one results for one of the image queries using only color features in MBrowser. Intermediate layer one retrieval results are improved with employing color, shape and texture features in layer two as presented in Figure 25.

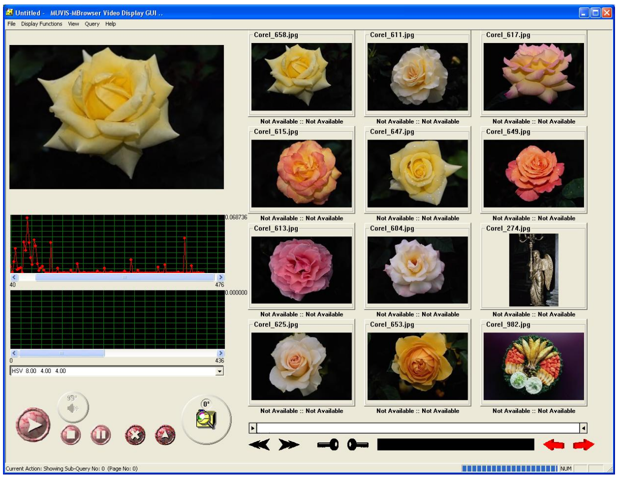


Figure 24: Sample Layer One Query Results.

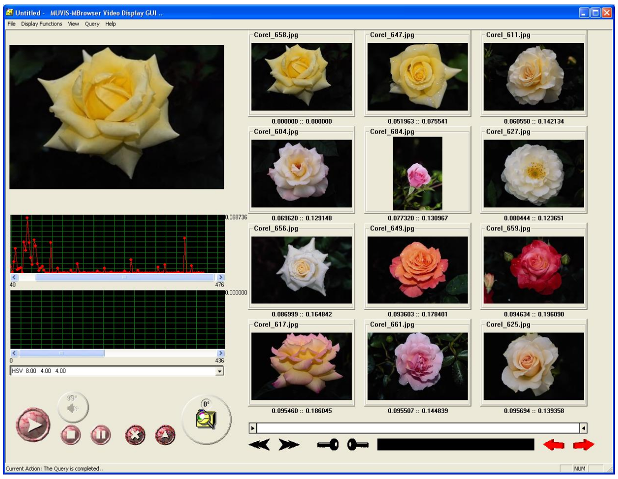


Figure 25: Sample Layer Two Query Results.

Additional color- and edge-based retrieval experiments are done in the same environment for the assessment of the proposed MAT method. Histogram features are utilized during the experiments since they often have high dimensions leading to high memory and processor usage in CBIR. HSV, YUV, and RGB color histogram features using various numbers of bins are extracted from Corel image database containing 1000 images for the color-based experiments. 400 images from the same database are also selected to create a constrained database for the edge-based experiments, where the selected images contain clear shapes that can be easily seen. Canny and Hough transform based edge histogram features are extracted from this database. Several randomly selected images are queried separately using PCA and the proposed MAT method with equal scaling factor for comparison. In the MAT experiments, the maximum function is utilized as the *G* function. The semantic retrieval performance of MAT and PCA methods are evaluated subjectively using ground-truth method.

Feature Data Size / Method	256 bins	512 bins	1024 bins
Principal Component Analysis (PCA)	1.3	7.8	46
MAT Method	0.2	0.62	2.1

Table 3: Execution Times for each Method in Seconds.

Table 3 presents total scaling process execution times of a database including 1000 images for each method on 256 (8, 8, and 4), 512 (8, 8, and 8), and 1024 (16, 8, and 8) bins HSV feature vectors using scaling factor 4. Considering the given process times, the MAT method is practically more feasible particularly in time-critical cases. PCA is more complex than the MAT method, since it considers the correlations between the media items in a database while the MAT method performs the dimension reduction process on a single feature vector.

Figure 26 and Figure 27 plot the Dimension Reduction Ratio-Average Precision curves for the average color- and edge-based image query results, respectively. The scaling factors used in these figures are 2, 4, 8 and 16. In order to display the loss in retrieval accuracy due to feature dimension reduction, the retrieval accuracy axis (in both plots) was normalized with respect to the average retrieval accuracy when the full feature vectors are used. These experimental results imply that the MAT method does not affect image retrieval results significantly. Furthermore, the results reveal that using scaling rate 4 (k_{T2} = 4) leads to satisfactory semantic and practical performance for MAT method as shown in Figure 26 and Figure 27. Therefore, MAT is a fast and flexible transform method that can be used for reducing processing time and memory requirements of CBIR systems, which employ feature data complying with the aforementioned assumptions.

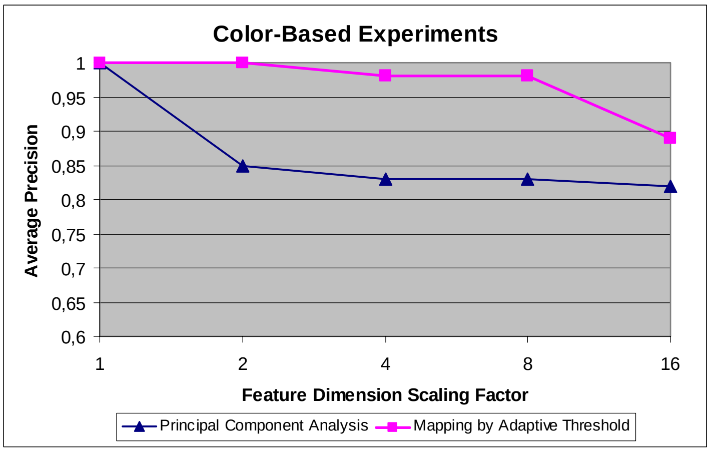


Figure 26: Dimension Reduction Ratio-Average Precision Curve for Color-Based Retrieval Results.

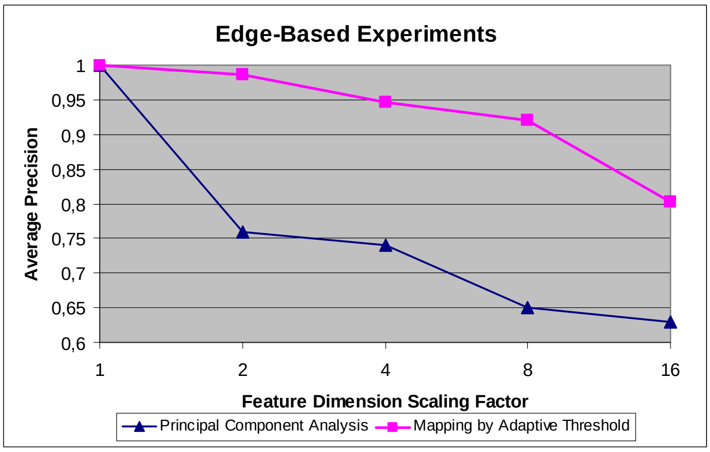


Figure 27: Dimension Reduction Ratio-Average Precision Curve for Edge-Based Retrieval Results.

5.5. SUMMARY

Transform-Based Layered Query Scheme is a novel retrieval framework addressing practical challenges in CBIR systems. It aims at particularly reducing the retrieval process time and memory consumption. Transformations applied on images and feature vectors play an important role in the proposed TLQ scheme for achieving these benefits. Suitable transforms need to be used in order to achieve a successful TLQ scheme. Specifically, they should be computationally efficient and less complex than the feature extraction processes. Transformations used for feature dimension reduction must retain the significant information in the reduced feature vectors.

The proposed TLQ scheme is flexible and can be used with any existing indexing and retrieval system, and transformation methods. Moreover, it allows various extensions such as relevance feedback between the layers. The scheme is also scalable due to multiple abstract layers approach. Such characteristics make the TLQ scheme feasible for integration into various platforms, such as mobile and distributed systems. The advantages of TLQ are expected to be more valuable when applied to limited platforms such as mobile devices.

A two-layer TLQ scheme is implemented and integrated into the MUVIS framework to assess its performance. Image data transform T1 is implemented via a DCT- and DWT-based downscaling. Mapping by Adaptive Threshold (MAT) and Principal Component Analysis methods are utilized as the feature data transform T2. The proposed MAT method is a simple and useful technique for reducing the dimension of feature data in CBIR context. The experimental results reveal the superiority of the MAT method over PCA in terms of processing time and retrieval performance. It is also shown that the MAT method does not affect query performance significantly. The expected theoretical performance of the integrated system is verified experimentally. The experiments reveal that scaling rate four (k_{T1} and k_{T2} = 4) is a practical value that leads to successful semantic retrieval results with low computational complexity, since it yields 50 to 80% query process time gain. Increasing k_{T1} and k_{T2} achieves higher gains at the expense of degrading the semantic performance. The studies also show that satisfactory semantic results can mostly be achieved by layer one intermediate query results.

Higher gains and better semantic results may be achieved by improving the integrated two-layer TLQ scheme with optimized transforms and indexing and retrieval methods. Another potential improvement is integrating a relevance feedback scheme between the retrieval layers of the TLQ scheme to achieve more meaningful retrieval results for end-users.

Chapter 6

System Profiles in CBIR

The use of content-based image indexing and retrieval systems have become widespread during the last decade on different hardware systems and platforms such as mobile phones and Internet [2], [44], [105]. Several CBIR applications have been developed for commercial and academic purposes. They often do not consider hardware architecture differences, and they are mostly not adaptable. Non-adaptable CBIR applications do not address effectively various user needs. User demands and expectations vary depending on the underlying hardware system (platform), which describes the set of hardware components of the device itself and can also be called hardware system in this thesis. Scalability and adaptability are desired attributes of a CBIR application. Scalability refers to the ability of handling growing amounts of data and adaptability refers to adapting itself effectively to changed platforms and situations [108], [131]. In this thesis, we address only hardware scalability. CBIR application would be hardware scalable in the sense that its performance remains suitably efficient and practical under changing database sizes and/or hardware capabilities. In addition, adaptability of a CBIR application to specific hardware architecture adds to the flexibility of the application. Adaptability here refers to the ability to select appropriate functionalities and suitable parameters in the CBIR application to fit the requirements of a given hardware system. CBIR evaluation workgroups effort stresses criteria such as the quality of adaptability of a CBIR application into a new domain [51]. They also express the importance of factors such as accuracy, speed and adaptability of CBIR applications. Datta et. al. in [22] discussed significant challenges involved in the adaptation of existing image retrieval techniques to build systems that can be useful in the real world. In order to improve the scalability and adaptability attributes of a CBIR application, different hardware systems hosting the application, their limitations, capabilities, and requirements have to be taken into account. Relatively little information about CBIR users and their hardware platforms are available in the existing literature. Therefore, the main motivation of this study is to widen the background knowledge on CBIR users and their hardware platforms.

System profiling is the baseline of this study as a step towards obtaining a complete set of CBIR parameters. System profiling is the process of acquiring knowledge about the hardware system of CBIR application users in order to provide enhanced services, adapt to specific requirements, and eventually improve the overall performance. This study does not consider user types in terms of their knowledge nor assumes any specific content of image databases. System profiling adapts CBIR applications to the specific hardware requirements of the system. In this study, we propose to use survey questionnaire method in order to define systems, specifications and requirements. With this method, demands of users from CBIR applications and technical hardware specifications are determined.

We propose a definition of system profiles and efficient CBIR parameters for each specified profile in order to have acceptable semantic performance with minimum computational complexity. Semantic performance in this study refers to the level of meaningful results of a retrieval process as perceived by human. One of the best criteria to evaluate content-based image retrieval performance is user satisfaction, which is proportional to the overall performance of a system. We evaluated the semantic retrieval performance of the proposed parameters by an objective evaluation technique mentioned in Section 2.2. The main objectives of the proposed study are:

- Tuning and defining CBIR application features and parameters for optimal overall performance (combination of semantic performance and complexity) in each hardware platform having different capabilities, capacities, and conditions.
- Improving the scalability of the CBIR application.

The rest of the chapter is organized as follows: The online survey and its outcomes are described in Section 6.1. It also presents analysis of the survey results. Specification of the system profiles and the proposed CBIR parameters are given in Section 6.2. Comparative and detailed experimental studies are presented in Section 6.3. Finally, Section 6.4 presents the concluding remarks, discussions and future work.

6.1. USER SURVEY QUESTIONNAIRE METHODS

Survey questionnaire methods are old and widely used methods in several studies [19], [28], [55], [114]. Especially, those systems involving user interactivity use survey methods for usability studies [19]. CBIR systems have also been the subject of several surveys. Jaimes [55] studied human factors, which influence automatic content-based retrieval systems, such as human memory, context and subjectivity. Eakins, Briggs and Burford [28] used online questionnaire method in order to improve the user interface of CBIR systems. Halvey and Keane in [47], studied log statistics of web-based video search engines to provide an analysis of user's interaction with video search engines. Kirk et. al. in [67] utilized interview and field observation methods to study the activities of digital image user activities such as searching and browsing. Frohlich et al. in [33] used interview and observation approach in order to understand the strengths and weaknesses of past and present technology of photo sharing. Rodden and Wood in [107] used interviews and questionnaires to find out how people organize and browse their digital photo collections and how these practices will compare to those they use at present, for their non-digital collections. In their conclusions, they claimed that CBIR would need to give more meaningful results to satisfy users, for example by providing face recognition. Catarci et. al. in [10] studied questionnaire-based approach to gather the user requirements for digital libraries.

User profiling has been utilized in various research domains in the literature. Kuniavsky in [69] gives answers for the questions “find out who your customers are, what they want and what they need”. Indeed, it is the starting point of designing and adapting a system according to the user’s requirements. Kuniavsky also explains the user profiling approach in the book and he expresses the importance of questionnaires for the profiling process in general. Weiss et. al. in [125] studied user-profile based personalization in order to select and recommend content with respect to users’ interest for automated online video or TV services.

Survey questionnaires are simply structured and carefully expressed to complete the purpose for which the survey is being conducted. Surveys are effective ways to collect information about users’ needs and choices and to identify the problem areas. They give users time to think about questions and their main advantage is low cost. When preparing the survey questions in this study, the following scientific strategy is followed [20]:

Preparing Survey Questionnaire Flow Chart

- Establish the goals of the project - What you want to learn
- Determine your sample - Whom you will interview
- Choose interviewing methodology - How you will interview
- Create your questionnaire - What you will ask
- Pre-test the questionnaire, if practical - Test the questions
- Conduct interviews and enter data - Ask the questions
- Analyze the data - Produce the reports

The main goals of the survey are to identify real world problems, to specify system requirements and to specify system limitations. The survey includes 38 questions organized in three categories in order to collect general information about the use of digital multimedia, the use of CBIR applications, and the use of CBIR features.

Survey questions are generated as a draft and interactively tested with 4 people. Direct observation approach using “think-aloud” protocol is employed for structuring the survey. The participants were observed during their first encounter with the questions, and they were encouraged to articulate their thoughts and opinions during the questionnaire. The improved survey is further assessed with 10 more people to obtain the final version for publishing. After the corrections and modifications, the online survey questionnaire is distributed by e-mail.

In the online survey, audiences answer the questions by selecting from the pre-defined list of choices, which are defined during the interactive test survey described above.

The survey includes 38 questions organized in three categories in order to collect general information about the use of digital multimedia, the use of CBIR applications, and the use of CBIR features. Finally, the responses are collected and further analyzed in the thesis.

6.1.1. Survey Results and Analysis

In this study, we mainly focus on platforms utilized by end-users of CBIR applications. Thus, the target audiences for the survey are expert and non-expert CBIR users. 122 people contributed to the online survey, 27 female and 95 male including Computer, Software, Electronic, Telecommunication engineers, IT students, researchers and professors. Age distributions are 32% of 20-24, 61% of 25-35 and 7% of 36-50 years old.

Figure 28 shows the overall study for system profiling, where indexing and retrieval factors and parameters and system profiles will be defined in the thesis. A general CBIR application includes various indexing and retrieval factors and parameters, which should be tuned for each hardware platform to utilize the application efficiently. Analysis of the online survey results plays an important role for determining the system profiles and selecting suitable indexing and retrieval factors and parameters. Scalability of an application is not only the ability of functioning properly with large data but also utilizing the advantages of the modified environment efficiently. For example, a software application would be scalable if it could be ported to a new platform that has larger technical capacities and take full advantage of the larger system in terms of performance (user response time etc.). The relation between scalability and adaptability is quite close as described in [108]. In this respect, adaptation of the factors and parameters partially helps to improve hardware scalability of the application. The selection of appropriate factors and parameters is principally aimed to efficiently use the overall capacity of each system profile in order to maximize CBIR performance. The complete list of survey questions and results can be found in the appendix.

In this study, we employed heuristic methodology to interpret the results of the online survey. The study revealed distinct informative knowledge about the hardware specification of the users and their preferences about digital image management. The analysis method of the survey results can be classified into two categories: Direct answers from the question results and heuristic analysis of the relevant and associated survey questions. The system profiling study is based on an empirical approach using the latter method. The direct-answer method is employed for definitions and specifications of indexing and retrieval parameters for each profile. Each of the answers of the survey questions helps in the selection of factors and parameters and experimental case setup. For example, in the 16th question of the survey, 93% of the participants prefer to use JPEG image compression technique. Thus, we decide to use JPEG compressed images in the experiments.

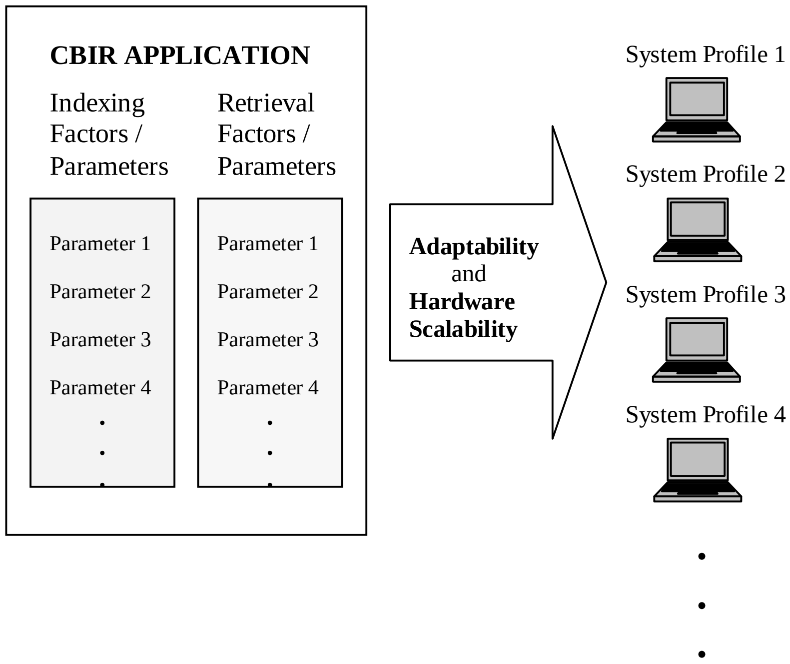
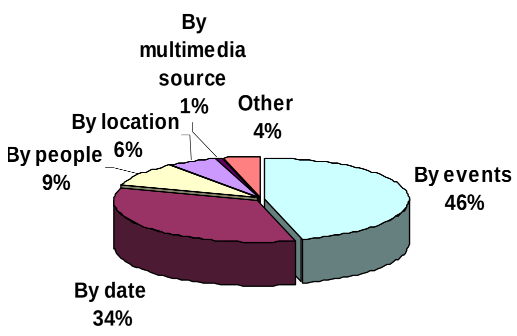


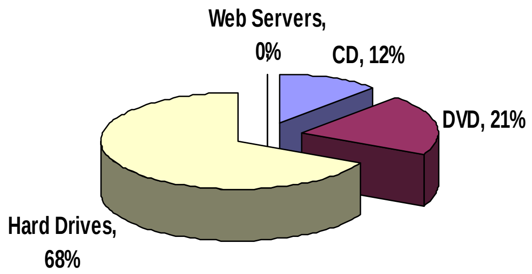
Figure 28: System Profiling for CBIR Applications.

Figure 29 I-IV are selected samples from online survey results. The first example represents the importance of CBIR applications, since 55% of the participants prefer to organize their multimedia files by events and people (46% by events and 9% by people), which may be partially considered as “content” of the image. The second chart helps us determine the storage space specification of the profiles, where we do not consider external storage spaces for image databases and utilize only hard drives of the concerned devices. The third chart reveals information about the approximate size of the image databases for the experimental studies. Finally, the fourth sample illustrates the need for use of image thumbnails in CBIR applications for browsing image databases.

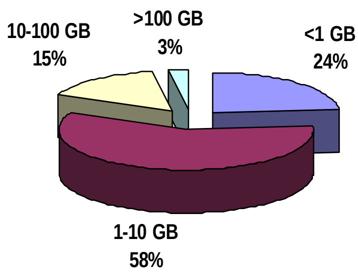
Answers of the participants for the questions of the survey are considered to define the requirements, capacities and conditions of the systems. The defined requirements, capacities and conditions help to determine the parameters of indexing and retrieval factors and system profiles. The defined system profiles are explained in details in the following section.



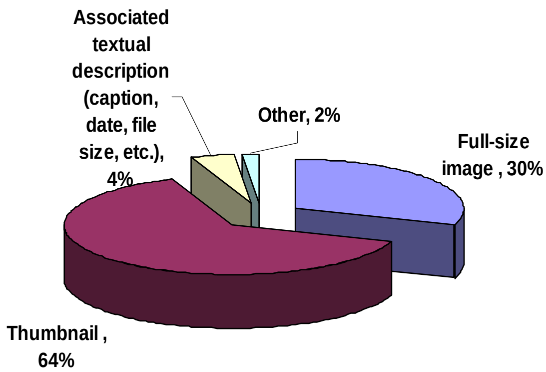
I) How would you prefer to organize your multimedia files?



II) Which of the following do you prefer to use for storing your multimedia files?



III) What is the approximate size of your digital multimedia collection?



IV) Which of the following do you prefer to see for each multimedia item when browsing?

Figure 29: Sample Survey Results Illustrations.

6.2. SYSTEM PROFILES AND SPECIFICATIONS

Forming the system profiles is a key issue for improving efficiency of CBIR on several platforms in terms of computational complexity and semantic retrieval performance. In this study, we define system profiles based on survey results. Especially, user needs and requirements are considered while forming the system profiles. Users' preferences may diverse depending on their CBIR system and hardware platform. For example, 56% of the users expect instantaneous response from a retrieval system when data and CBIR application are locally stored and running on the personal computer. On the other hand, this percentage decreases to 40%, when they are using web-based CBIR system.

In this thesis, we define four main groups of systems using CBIR: Limited, Distributed, Baseline and Powerful system profiles. Limited System Profile represents the limited platforms such as mobile phones. Systems having client-server architecture, e.g. web-based systems, are represented by Distributed System Profile. Typical CBIR users are grouped as a Baseline System Profile. Most of CBIR users belong to this profile. Powerful System Profile represents the powerful computer systems such as dedicated servers for professional use. Usually, professional users own such systems to use CBIR, for example, TV broadcasting archives, libraries, media companies etc.

After defining the system profiles, their technical attributes are specified in order to identify capacities of the system that will affect the CBIR performance. These technical attributes of the systems given in Table 4, have potential impact on CBIR usage and the overall performance. Network connection speed defines the data traffic between client and server in distributed CBIR systems, where data are expected to flow effectively in a short time interval. Multimedia codecs and storage space capacity help to specify limitations of multimedia databases that can be constructed on the existing hardware system. CPU power is one of the main factors to determine indexing and retrieval time. Finally, display size plays an important role in the interaction between CBIR systems and users. It will also increase the usability of the system, which increases the satisfaction of the user from the system. Table 4 is set according to the latest avaliable technologies (last quarter of 2007).

Considering the technical attributes given in Table 4 and relevant CBIR processes, CBIR parameters factors and parameters[1] can be divided into the following two categories to ease the adaptation of the parameters for each system profile:

- Indexing Factors/parameters:
  - Compression parameters
  - Scaling parameters
  - Feature type
- Retrieval Factors/parameters:
  - Dimension reduction of feature data parameters
  - Feature selection

These two categories are further split into sub-groups taking into account CBIR challenges and existing prior art. The indexing process includes feature extraction and database editing sub-processes. Consequently, the challenges presented by the indexing process may be divided into three main groups: run-time memory consumption, run-time computational complexity, and storage space requirement. The main indexing parameters are the JPEG quality factor used in the JPEG image compression [11], the scaling factor of the image downscaling, and the feature types. Image compression can be used to reduce the storage space and memory consumption. Additionally, image downscaling may be utilized to decrease the feature extraction process complexity and the storage space requirement. The

[1] Factors in this context may also refer to specific parameters when appropriate.

feature type factor also affects the feature extraction process complexity. For example, extracting shape-based features is often computationally more complex than extracting color-based features.

System Profiles / **Attributes**	**Limited Systems** 	**Distributed Systems** 	**Baseline Systems** 	**Powerful Systems** 
Connection Speed	114 Kb/s – 723 Kb/s	128 Kb/s – 1 Mb/s	128 Kb/s – 2 Mb/s	1 Mb/s – 100 Mb/s
Storage Space	160 MB-8 GB	10-20 GB client	~100GB	over 500 GB
Expected Retrieval Time	0-1 min	0-45 sec	0-30 sec	0-30 sec
CPU Power	Information Not Available	1-2 GHz	~2 GHz	~2x3.0 GHz +
Multimedia Codecs	Encoding: H263, MPEG4, AMR, AAC, JPEG Decoding: H263, MPEG4, AMR, AAC, MP3, JPEG2000, JPEG	Generally All	Generally All	Generally All

Table 4: Technical Features of System Profiles.

Retrieval factors include similarity measurement and sorting processes. The main factors controlling the response time of a retrieval process are run-time memory consumption and computational complexity. In this respect, two main factors are considered for the retrieval process: Dimension reduction of feature data and feature selection. Firstly, a dimension reduction method reduces the feature data size; and consequently they reduce proportionally the elapsed time of the retrieval process. Secondly, the feature selection process refers to the process of selecting the most important features and their combinations for describing images in the database in order to reduce the computational complexity while maintaining the retrieval performance.

Table 5 is constructed according to the aforementioned categories of CBIR factors and parameters. The degree of image compression can be adjusted by the quality factor defined in JPEG codec for allowing a selectable tradeoff between the storage size and the image quality.

Image downscaling represents the process of scaling the image down to a smaller size by a scaling factor, so-called image downscaling parameter in this chapter. The range of each CBIR parameter is specified based on the technical attributes given in Table 4 and the analysis of the survey results. The recommended CBIR factors and parameters for each system profile are assessed for validation. These experiments and results are given in the next section.

	System Profiles	**Limited Systems**	**Distributed Systems**	**Baseline Systems**	**Powerful System**
Indexing Factors	**Compression Parameters**	JPEG Compression with quality factor 50%	JPEG Compression with quality factor 75-50%	JPEG Compression with quality factor 75%	Uncompressed or JPEG Compression with quality factor 90%
	Image Downscaling Parameters	Image Scaling Factor (ISC) = 4 for Color features ISC=2 for texture and shape features	Image Scaling Factor = 4 for Color features ISC=2 for texture and shape features	Image Scaling Factor = 2 for Color features none for texture and shape features	Image Scaling Factor = 2 or none for Color features none for texture and shape features
	Feature Parameters	Use a feature selection method	Use a feature selection method	Optionally use a feature selection method	Optionally use a feature selection method
Retrieval Factors	**Dimension Reduction of Feature Data Parameters**	Scaling factor=4 or 8	Scaling factor=4	Scaling factor=2	None or Scaling factor=2
	Feature Selection and Combination Parameters	Use a feature selection method	Use a feature selection method	Optionally use a feature selection method	Optionally use a feature selection method

Table 5: Recommended CBIR Parameters for each System Profile.

6.3. EXPERIMENTAL STUDIES

Several experimental studies are accomplished in order to test the validity of the proposed system profiles and CBIR parameters. The main goal of the experimental studies is to indicate that the proposed minimum configuration of indexing and retrieval parameters for each system profiles may satisfy user needs in terms of complexity and retrieval accuracy. Parameters given in Table 5 are evaluated and compared for complexity and semantic performance analysis. MUVIS Content-Based Indexing and Retrieval system is utilized as the experimental framework. Online survey results are considered while selecting the experimental database properties. The specifications of indexing and retrieval parameters for each system profile are given in Table 5. Selection of the parameters follows a heuristic process based on previous experiments and online survey questions.

Image Compression: JPEG compression technique is selected for image compression since it is a commonly used method as validated by the results of the online survey. Uncompressed Corel database with 10,000 images that are pre-assigned to 100 semantic categories each containing 100 images by a group of human observers is utilized for experimental studies. Some examples of the categories are autumn, balloon, bird, dog, eagle, sunset, and tiger. 8 compressed and/or downscaled test databases are created from uncompressed Corel database by JPEG compression with quality factors 90, 75 and 50 for different tests on several platforms. JPEG quality setting constructs JPEG quantization tables appropriate for the indicated quality setting, which is expressed as 0 to 100. In the experiments, we utilized JPEG codec with quality setting function that implemented by Independent JPEG Group [52]. The effects of compression on CBIR system's semantic performance have been studied earlier in the literature. As shown in [40] and [86], the proposed quality factors for image compressions are the minimum compression parameters, which do not affect retrieval accuracy significantly.

Image Downscaling: Image downscaling parameters are selected according to the previous studies in [38] and [41]. DCT-based image downscaling effects on image retrieval performance were analyzed, and it was concluded that it does not affect color feature retrieval performance significantly. On the other hand, it affects retrieval accuracy with texture and shape features. The proposed downscaling parameters for color, texture and shape features have also been tested and compared with non-scaled image databases to express the effects clearly.

Feature Selection: Feature selection process affects both indexing and retrieval process complexity and semantic retrieval performance. In this study, we use mutual information, which is a widely used method for feature selection in various research fields such as genomic data analysis, classification of network data, categorization of medical data, and speech recognition [56], [68], [78], [102], [126]. Feature selection method based on mutual information and decision mechanisms in [102] is used for experimental studies in this chapter.

14 types of color, texture and shape features are utilized in all experimental studies. The following low-level color, shape, and texture features are used: YUV, RGB, and HSV color histograms with 128, 64, and 16 bins, Gabor Wavelet texture feature, Gray Level Co-Occurrence Matrix texture feature with parameters 12 and 6, Canny Edge Histogram, and Dominant Color with 3 colors.

Dimension Reduction of Feature Data: Dimension reduction of feature data process is a commonly applied method for reducing the retrieval process complexity. Mapping by Adaptive Threshold (MAT) based method is not affecting color-based retrieval performance. Therefore, MAT method and its scaling parameters are suggested according to the previous studies. Moreover, dimension reduction method for texture and shape features was not used, since the feature data sizes are comparatively smaller than those of color features.

Experimental Setup: Experimental databases are constructed according to the parameters proposed above. Several experimental studies have been performed in this study in order to validate decision results and present the advantages in terms of complexity and semantic performance.

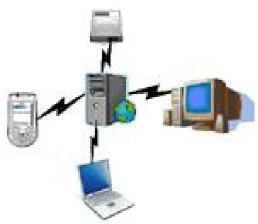
System Profiles / **Attributes**	**Mobile Phone** 	**Distributed System** 	**Baseline System** 	**Powerful System** 
Connection Speed	128/512 Kbit/s	128 Kb/s – 1 Mb/s	128 Kb/s – 2 Mb/s	1 Mb/s – 100 Mb/s
Storage Space	1 GB	60 GB client	120 GB	180 GB
CPU Power	Information Not Available	Intel Pentium 4 2.8 GHz	Intel Pentium 4 2.8 GHz	2x2.8 GHz
Multimedia Codecs	MPEG-4 , H.264/AVC , H.263/3GPP, MP3-, AAC-, eAAC- and eAAC	Generally All	Generally All	Generally All

Table 6: Technical Features of Sample System Profiles Utilized in the Experiments.

Complexity process is evaluated by indexing and retrieval process times for different platforms. Experimented platforms and their configurations are given in Table 6. In order to evaluate the semantic retrieval performance, Average Normalized Modified Retrieval Rank

(ANMRR) formulation is used. Equations (1), (2) and (3) in Section 2.2 are used to calculate the ANMRR values, where *N(q)* value is equal to 30 in the experiments.

The best retrieval performance can be achieved when *NMRR(q)=0*. On the other hand, the worst case *NMRR(q)=1* means none of the relevant items can be retrieved from *W*. Thus, lower *NMRR* values represent successful retrieval results for the query *q*. Average NMRR (ANMRR) can be used as semantic retrieval performance criterion, if the number of query by example (QBE) experiments is high enough.

6.3.1. Semantic Evaluation of Proposed CBIR Parameters for each System Profile

Experiments are performed on each sample system given in Table 6, using the recommended CBIR parameters in order to indicate the performance advances.

6.3.1.1 Powerful System Profile

Configurations given in Table 6 for Powerful Systems Profile (PSP) are used for the experiments. The Powerful Systems profile represents the special systems, which have dedicated servers and professional software applications. However, we have used a powerful personal computer to run these experiments to evaluate the semantic results of the proposed CBIR parameters. The semantic retrieval results are presented in ANMRR values aforementioned.

Table 7 shows the retrieval performance of the databases, which are created for each different indexing and retrieval parameters. It is clear from Table 7 that, image compression with quality factor 90% does not affect the retrieval accuracy, although it saves from the storage space. Since the retrieval accuracy is not significantly affected with this compression scheme, the compressed databases are employed in the system profiles according to their sizes. JPEG compressed image database with quality factor 90% is selected as Base 1 database for the Powerful System profile. Images are further downscaled by various scaling rates. Several combinations of experiments are performed to show the semantic retrieval accuracy and corresponding results are given in Table 7. The semantic retrieval accuracy of DCT-based downscaled image database is not affected and the feature extraction process on this database is 66% lower than on the original size image database, given in Table 7.

Afterwards, MAT dimension reduction technique is applied on the feature data in base database, and the results of the experiments with the reduced size feature data given in Table 7. MAT dimension reduction method enhances the semantic retrieval results due to its natural impact on histogram-based features as can be seen from the Table 7. The method thresholds the irrelevant details and emphasizes the higher peaks on histograms such as dominant colors in color histograms.

		ANMRR
Compression Parameters	**Original**	0.20
	JPEG Compressed with Quality Factor 90% (Base 1)	0.20
	JPEG Compressed with Quality Factor 75%	0.23
Image Downscaling Parameters (Base 1+)	**Color-based scaled by 2 Texture and shape-based none (Base 2)**	0.20
	Color-, texture-, and shape-based scaled by 2	0.21
Dimension Reduction of Feature Data Parameters (Base 1 + Base 2 +)	**Scaled by 2 (Base 3)**	0.15
	Scaled by 4	0.19

Table 7: ANMRR Results of Experiments on PSP.

Consequently, it is inferred from the experiments that the proposed CBIR parameter configurations for Powerful Systems profile achieve satisfactory semantic performance.

6.3.1.2 Baseline System Profile

The Baseline System Profile represents the hardware system of a typical CBIR user, who does not have particular strict requirements or high expectations from the system, while searching and browsing an image database. In this study, we employed a sample PC, which has the configuration given in Table 6.

Table 8 shows that the retrieval performances of the uncompressed database and the JPEG compressed databases, where ANMRR value of compressed database with quality factor 75% is degraded by 3% due to compression. Since the semantic retrieval results are slightly affected with this compression scheme, the compressed databases are employed in the system profiles according to their sizes on disk. JPEG compressed image database with quality factor 75% is selected as the base database and DCT-based image downscaling is applied prior to extracting features, which are presented in Table 8. Image downscaling may enhance the edges between two objects having different colors in the image due to high frequency component loss. Additionally, downscaling may also have filtering effects on the color information due to smoothing the colors. Therefore, these effects may cause slight (by

2-3%) increase in retrieval performance. As shown in Table 8, the ANMRR values increase by 3% due to the influence of image downscaling on color features.

The resulting feature data dimension is reduced using MAT method by 2, and 4 in order to assess the effects of the scaling parameters, and the corresponding results are presented in Table 8. The dimension reduction method improves the semantic retrieval results as explained in Section 6.3.1.1. It also reduces 66% of the run-time retrieval complexity as shown in Table 8. Consequently, the proposed CBIR parameters achieve a successful semantic performance, with feasible processing times for the Baseline System profile.

<table>
<tr><th></th><th></th><th>ANMRR</th></tr>
<tr><td rowspan="3">Compression Parameters</td><td>Original</td><td>0.20</td></tr>
<tr><td>JPEG Compressed with Quality Factor 90%</td><td>0.20</td></tr>
<tr><td>JPEG Compressed with Quality Factor 75%
(Base 1)</td><td>0.23</td></tr>
<tr><td rowspan="2">Image Downscaling Parameters
(Base 1+)</td><td>Color-based scaled by 2 & texture and shape-based none
(Base 2)</td><td>0.20</td></tr>
<tr><td>Color-, texture-, and shape-based scaled by 2</td><td>0.21</td></tr>
<tr><td rowspan="2">Dimension Reduction of Feature Data Parameters
(Base 1 + Base 2 +)</td><td>Scaled by 2</td><td>0.19</td></tr>
<tr><td>Scaled by 4 (Base 3)</td><td>0.19</td></tr>
</table>

Table 8: ANMRR Results of Experiments on BSP.

6.3.1.3 Distributed System Profile

Distributed System Profile (DSP) is a general system profile based on a client and server architecture. The online survey results reveal that web-based distributed systems are used widely among CBIR users. Note that, distributed systems may vary in the sense of different client and network capacities and capabilities. Ahmad in [1] studied compression and network effects on CBIR having client-server architecture. They tested a CBIR framework on different networks with various mobile devices. In these experiments, a personal computer was used as the server and a laptop computer as a client with Wireless Local Area network, which achieves the fastest query time in [1]. Corresponding configurations of the experimental devices are given in Table 6. Indexing process is employed at the server side while retrieval is handled at the client side of the distributed system.

Table 9 shows semantic retrieval results for queries performed on databases created with each of the proposed set of CBIR system parameters. The ANMRR results of JPEG compressed image databases with quality factor 75% and 50% show that the proposed compression parameters do not affect semantic performance considerably (by 3%). JPEG compressed image database with quality factor 75% is selected as Base 1 database for combining other CBIR parameters. DCT-based downscaled database with a scale factor of 2 is indexed using all features and retrieval results remain the same as Base 1 database. Additionally, color-based features are also extracted from downscaled image database by 4 in order to express the effects on ANMRR results. The retrieval results are slightly influenced by downscaling with scale factor 4.

		ANMRR
Compression Parameters	**Original**	0.20
	JPEG Compressed with Quality Factor 75% (Base 1)	0.23
	JPEG Compressed with Quality Factor 50%	0.23
Image Downscaling Parameters (Base 1+)	**Color, texture and shape-based scaled by 2 (Base 2)**	0.23
	Color-based scaled by 4 & texture and shape-based scaled by 2	0.24
Dimension Reduction of Feature Data Parameters (Base 1 + Base 2 +)	**Scaled by 2**	0.21
	Scaled by 4 (Base 3)	0.21
	Scaled by 8	0.25
Feature Parameters (Base 1 + Base 2 + Base 3 +)	**Using all Features (Base 4)**	0.21
	Using selected Features	0.34

Table 9: ANMRR Results of Experiments on DSP.

MAT-based dimension reduction method is performed on feature data extracted from compressed and DCT-based downscaled image database. Scaling feature data by 2 and 4 give the same semantic retrieval performance in terms of ANMRR values, which is slightly enhanced compared to the results of the original database.

The last two rows of Table 9 present the retrieval results of queries using all features and only a subset of selected features by the feature selection process on an image database created using the proposed parameters (Base databases).

In summary, JPEG image database compressed with quality factor 75% and DCT-based image database downscaled by a scaling factor of 2, and MAT-based dimension reduction method with scale factor of 4 with all features give successful semantic performance for Distributed System profile.

6.3.1.4 Limited System Profile

Limited System Profile (LSP) is for CBIR users whose platforms are hand-held devices, such as palms and mobile phones. In the experimental studies, a mobile phone with the properties given in Table 6 was used with Mobile MUVIS content-based multimedia indexing and retrieval system designed for mobile platforms running Symbian-based operating system [44]. The proposed CBIR system parameters are minimum configurations according to the capacities of current devices; however, the user may prefer to change the parameters with respect to the hardware platform.

		ANMRR
Compression Parameters	**Original**	0.20
	JPEG Compressed with Quality Factor 75%	0.23
	JPEG Compressed with Quality Factor 50% (Base 1)	0.23
Image Downscaling Parameters (Base 1 +)	**Color, texture and shape-based scaled by 4**	0.30
	Color-based scaled by 4 & texture and shape-based scaled by 2 (Base 2)	0.26
Dimension Reduction of Feature Data Parameters (Base 1 + Base 2 +)	**Scaled by 4 (Base 3)**	0.22
	Scaled by 8	0.23
Feature Parameters (Base 1 + Base 2 + Base 3 +)	**Using all features**	0.22
	Using selected features	0.38

Table 10: ANMRR Results of Experiments on LSP.

Table 10 represents the ANMRR results of the experiments, which are performed on Limited System profile. The semantic retrieval results employing JPEG compressed database with quality factor 50% are reasonable compared to the ones with uncompressed database. Thus, it is selected as Base 1 database for performing DCT-based downscaling process with scaling factors 2 and 4. Extracting texture- and shape-based features from downscaled image database by 2 gives better semantic performance than extracting them from downscaled image database by 4. On the other hand, color-based retrieval results are not affected considerably by DCT-based downscaling (only 3% from the Base 1 database), thus they are extracted from a downscaled image database by 4 in order to reduce 93% of the feature extraction complexity as shown in Table 11.

MAT-based dimension reduction method is performed on compressed and downscaled image database features (Base 1 and Base 2) and the semantic retrieval results are given in Table 10. It is observed that retrieval accuracy is improved by 4% using feature data scaled by 4. Therefore, dimension reduction of feature data with scale factor 4 is selected as Base 3 database.

The feature selection method is applied on the feature data, which are extracted from the compressed and downscaled database and further processed with MAT dimension reduction (using scale factor 4). Table 10 shows the ANMRR results of all the features and a subset of the selected features obtained by feature selection process on the image database created using the proposed parameters. The proposed CBIR parameters yield a successful semantic performance for the Limited System profile using all features, as given in Table 10.

6.3.2. Evaluation of Proposed CBIR Parameters In Terms of Complexity

Complexity evaluation experiments are performed on the databases used in the semantic performance experiments for each system profile. Sample systems and technical features are given in Table 6 and the methods aforementioned are utilized to calculate the process times.

Table 11 shows the elapsed times for the indexing and retrieval processes using different parameters in each profile. The elapsed time for the feature extraction process is the time required to extract the features from the original size and the downscaled images. The MAT-based dimension reduction technique is applied on the original size feature data, and the corresponding elapsed times from the Table indicate the feasibility of the proposed method on any platform. The query time with a downscaled feature data by 2 and 4 are also given in the same Table in order to emphasize the benefits of the proposed CBIR parameters on the computational complexity of the retrieval process. According to the experimental results, image downscaling reduces the time required for feature extraction by approximately 67% with a scaling factor 2 and 90% with a scaling factor 4. Note that although the same hardware system is used for the Baseline System Profile and the client of the Distributed System Profile in the experiments, the retrieval process times are higher in the Distributed System Profile due to the client-server communication overhead.

<table>
<tr><th colspan="3" rowspan="2"></th><th colspan="4">Process Times</th></tr>
<tr><th>Powerful System Profile</th><th>Baseline System Profile</th><th>Distributed System Profile</th><th>Limited System Profile</th></tr>
<tr><td rowspan="7">Indexing Process</td><td rowspan="5">Feature Extraction process</td><td>Original Database</td><td>3 hours</td><td>6 hours</td><td>6 hours</td><td>~65 hours</td></tr>
<tr><td>Color-based scaled by 2 & texture and shape-based none</td><td>2.5 hours</td><td>5.5 hours</td><td>5.5 hours</td><td>~65 hours</td></tr>
<tr><td>Color, texture and shape-based scaled by 2</td><td>1 hour</td><td>1.5 hours</td><td>1.5 hours</td><td>24 hours</td></tr>
<tr><td>Color-based scaled by 4 & texture and shape-based scaled by 2</td><td>50 min</td><td>1.2 hours</td><td>1.2 hours</td><td>13 hours</td></tr>
<tr><td>Color, texture and shape-based scaled by 4</td><td>18 min</td><td>25 min</td><td>25 min</td><td>4 hours</td></tr>
<tr><td rowspan="2">MAT-based Dimension Reduction of Feature Data</td><td>Scale by 2</td><td>13 sec</td><td>14 sec</td><td>14 sec</td><td>120 sec</td></tr>
<tr><td>Scale by 4</td><td>12 sec</td><td>12 sec</td><td>12 sec</td><td>90 sec</td></tr>
<tr><td rowspan="6">Retrieval Process</td><td rowspan="3">A query with all features</td><td>Original Size</td><td>9 sec</td><td>12 sec</td><td>100 sec</td><td>140 sec</td></tr>
<tr><td>Scaled by 2</td><td>5 sec</td><td>7 sec</td><td>50 sec</td><td>65 sec</td></tr>
<tr><td>Scaled by 4</td><td>3 sec</td><td>4 sec</td><td>25 sec</td><td>32 sec</td></tr>
<tr><td rowspan="3">A query with selected features</td><td>Original Size</td><td>5.5 sec</td><td>7 sec</td><td>90 sec</td><td>120 sec</td></tr>
<tr><td>Scaled by 2</td><td>3 sec</td><td>4 sec</td><td>45 sec</td><td>60 sec</td></tr>
<tr><td>Scaled by 4</td><td>2.5 sec</td><td>3 sec</td><td>22 sec</td><td>29 sec</td></tr>
</table>

Table 11: Indexing and Retrieval Process Times on each System Profile.

6.3.3. Summary of Recommended CBIR Parameters

The proposed CBIR parameters in this study are the minimum configurations according to the capability of each hardware platform; however, the user may change the parameters with respect to the hardware architecture used for specific CBIR applications.

Table 12 shows the recommended CBIR parameters for each system profile. The aforementioned parameters give satisfactory CBIR indexing and retrieval results for each platform users as shown in the experiments.

Image compression with quality factor 75% gives successful retrieval results and hence can be utilized in Distributed System Profile, Baseline System Profile, and Powerful System Profile. However, it may be better to use compression with quality factor 90% in Powerful System Profile in order to improve the retrieval performance slightly, since there are no memory and storage space consumption problems in such systems. Additionally, due to memory and storage space capacities of limited systems, image databases should be compressed with a minimum quality factor of 50% for obtaining satisfactory retrieval results.

		Limited Systems 	**Distributed Systems** 	**Baseline Systems** 	**Powerful Systems** 
Indexing Parameters	**Compression Parameters**	JPEG Compression with quality factor 50%	JPEG Compression with quality factor 75%	JPEG Compression with quality factor 75%	JPEG Compression with quality factor 90%
	Image Downscaling Parameters	Image Scaling Factor (ISF) = 4 for Color features ISF=2 for texture and shape features	ISF=2 for color, texture and shape features	ISF = 2 for Color features none for texture and shape features	ISF = 2 for Color features none for texture and shape features
	Feature Parameters	Use a feature selection method	Use a feature selection method	none	none
Retrieval Parameters	**Dimension Reduction of Feature Data Parameters**	Scaling factor=4	Scaling factor=4	Scaling factor=4	Scaling factor=2
	Feature Selection and Combination Parameters	Use a feature selection method	Use a feature selection method	none	none

Table 12: Recommendation for CBIR Parameters.

Color features extracted from downscaled images do not affect the CBIR retrieval results. Thus, they can be utilized on every platform. On the other hand, retrieval results using texture and shape features are slightly affected by image downscaling. Hence, they can be employed on platforms that have limited processing power capacities such as distributed systems and limited systems.

Dimension reduction of feature data tends to reduce the retrieval complexity similar to feature selection methods. It can be used on every platform, and it does not affect the semantic performance significantly. Feature selection methods can also be employed on every system profile. However, they may affect the retrieval performance depending on the feature selection method existing in the system. Baseline System Profile and Powerful System Profile have high processing power capabilities and do not have to utilize any feature selection method for preserving the retrieval performance.

Distributed and limited platforms have low resources in terms of processing power, and storage space. Thus, scaling factor for feature data dimension reduction may be higher in those profiles to reduce the overall complexity.

6.4. SUMMARY

In this chapter, a novel study on CBIR system profiling and adaptation of indexing and retrieval parameters is presented. The main goals of the study are defining the major system resources and conditions, and adapting CBIR systems for different user platforms.

Specifying a system profile is an important factor in CBIR studies, which may help to make the system scalable and adaptable for existing hardware platforms. System profiles allow systems to answer demands of different users. CBIR applications often consist of complex processes; therefore, the underlying main factors and parameters need to be adapted according to the limitations of the platforms. In this study, appropriate CBIR parameters are proposed for each of the defined system profiles.

An online user survey is used for determining the systems of the end-users by heuristic definitions inferred from the survey results. The proposed system profiles are the Powerful System Profile, the Baseline System Profile, the Distributed System Profile, and the Limited System Profile.

CBIR parameters are handled by grouping them into two parts: the indexing and the retrieval. Experimental studies have been conducted for each of the proposed CBIR system profile. It was shown that the proposed factors and parameters for each system profile yield satisfactory semantic retrieval performance. On the other hand, they lead to substantial gains in computational complexity and storage space requirements. The gains with regard to the retrieval process complexity are 45% for the Powerful System Profile, 42% for the Baseline

System Profile, 78% for the Distributed System Profile, and 78% for the Limited System Profile. It is however important and required to adapt CBIR applications for the Limited System Profile. Users of the latter profile may have approximately the same level of semantic performance with those of the Baseline System Profile by appropriately modifying the parameters according to the underlying hardware architecture.

The proposed CBIR parameters are appropriate configurations to improve the efficiency of CBIR applications on different hardware platforms. However, trade-offs between parameters should be considered in case modifications or adjustments are required. For example, a Limited System Profile user may utilize an uncompressed image database instead of the compressed one for better image quality; however, in this case the storage space capacity of the device should be considered for the database size.

Finally, this study may be extended and supplemented by additional experiments especially for future CBIR applications and user platforms which are expected to change the proposed profiles and the proposed parameters due to advances in technology. Future work may also include investigating user satisfaction for the proposed system profiles and CBIR parameters using online surveys and further analysis.

Chapter 7

Conclusions

The overall size of digital image collection is rapidly increasing with the development of Internet and digital image sources such as digital cameras and image scanners. This improvement in digital image technology reveals the need for efficient image searching, browsing and retrieval. Content-based image indexing and retrieval studies have been started with the motivation of these needs and have become widespread applications on different platforms such as Internet. Generally, current CBIR systems do not consider variety of user system architectures and environments. Extensive use of CBIR and variety of users in terms of knowledge and their devices having different hardware architectures create a demand for scalable and adaptable CBIR systems. Therefore, in order to obtain a generic CBIR system, which provides efficient services independent from its environment and user, adaptation of the overall CBIR system performance is required.

CBIR systems usually succeed in meeting the needs of users in terms of visual content retrieval. However, extracting the high-level semantic from the content of the image is still a challenging task for researchers. Narrowing the semantic gap between high-level semantic and low-level features is the most important problem that should be considered for content-based retrieval performance. There are several factors besides semantic performance, to be improved in this regard. These factors can be categorized into four main groups of challenges: computational complexity, memory and disk space requirements, semantic retrieval performance, and usability. These factors and the needs for improvement in scalability and adaptability characteristics of CBIR systems construct the main motivation of this thesis. The contributions of the thesis can be summarized into five parts: Transform-based layered query scheme, DCT- and DWT-based image downscaling effects on CBIR performance, mapping by adaptive threshold method for dimension reduction of feature data, feature selection system, and user system profiles in CBIR.

Transform-Based Layered Query Scheme is a novel retrieval approach aiming to reduce query process time and memory consumption. It also involves an unsupervised method for decimating the underlying database by automatically eliminating irrelevant images. Unsupervised elimination method utilizes gradient of distance series that are the results of a query. A two-layer TLQ scheme is implemented and integrated into MUVIS framework for testing theoretical and practical outcomes of the proposed approach. Image transforms are implemented with DCT- and DWT-based downscaling in order to decrease indexing time and run-time memory cost. Similarly, feature data transforms are implemented with Principal Component Analysis and MAT method for reducing retrieval process complexity. Expected theoretical performance gain of the integrated system is validated with experimental results.

The proposed TLQ scheme is flexible due to its independence from the underlying methods for indexing and retrieval. Moreover, it allows various extensions such as relevance feedback between the layers. TLQ scheme improves the scalability of the system, since it has multiple abstract layers approach. Such characteristics make the TLQ scheme feasible for integration into various platforms, such as mobile and distributed platforms. The advantages and advances of TLQ would be more valuable when applied on limited platforms.

In this thesis, the effects of DCT- and DWT-based image downscaling are also studied. Color, texture, and edge based image retrieval experiments are performed separately in order to study the effects of each low-level feature type. Practical benefits of downscaling and compression in CBIR systems are clearly visible, especially for limited and distributed systems. Experimental studies show that image downscaling is an efficient solution to overcome various strict system requirements such as storage space and computational complexity, and does not have considerable negative impact on the semantic image retrieval performance.

Dimension reduction methods contribute in reducing memory consumption and process execution time, particularly in retrieval. The use of a simple technique referred to as Mapping by Adaptive Threshold for reducing the dimension of feature data is investigated and integrated into the indexing and retrieval scheme for experimental studies. Similarly, PCA is also used in the experiments for comparison. The experimental results reveal the superiority of the proposed method over PCA in terms of process time with better semantic retrieval performance. However, the method may not be suitable for other indexing and retrieval purposes, such as clustering. Low computational complexity and high semantic performance of the proposed MAT method makes it applicable in time-critical limited systems.

In this thesis, we also explore the use of feature selection within CBIR context. Two novel feature selection criteria based on intra-cluster and inter-cluster relations are proposed, and an efficient majority voting based method is implemented for the selection and combination of features. The proposed method includes three main criteria for feature-data relation. These criteria produce results for each feature that are fed to majority voting as input. Each criterion is based on different associations of feature-data affinity in order to define the best discriminative and representative feature of the database. Two proposed criteria are compared with other state-of-the-art criteria through dedicated experiments, which show that the proposed methods improve retrieval performance. The proposed feature selection system is implemented as a black-box approach that gives flexibility for using it in different platforms. It may be applied on various types of databases and sets of features.

A novel study on CBIR system profiling and adaptation of indexing and retrieval parameters is presented in this thesis. The main goals of the study are defining the major system resources and conditions and adapting CBIR systems for different user platforms. Definition of a system profile is a key factor in CBIR studies to make the system hardware scalable and adaptive for existing hardware platforms. System Profiles allow systems to answer demands of different users. CBIR systems have various parameters to be tuned according to the limitations of the platforms. In this study, appropriate CBIR parameters are proposed for each system profile. An online user survey is used for determining the systems of end-users by heuristic definitions inferred from the survey results. The introduced system profiles are: Powerful System Profile, Baseline System Profile, Distributed System Profile, and Limited System Profile. Potential improvements and adaptations are studied and justified by previous related studies in the literature for indexing and retrieval parameters. The proposed CBIR parameters for each system profile are tested for various possible cases and the results are illustrated in the thesis.

In this thesis, several techniques and methods are introduced to provide generic and feasible solutions for CBIR challenges. The proposed improvement approaches for overall CBIR performance can be considered as an important and relevant contribution to the CBIR literature. While the challenges determined in this thesis are still not completely solved, the presented studies represent a significant step towards the solution. All the techniques within the context of improving overall CBIR performance promote several possibilities and options for further research studies.

Bibliography

[1] I. Ahmad and M. Gabbouj, “Compression And Network Effect on Content-Based Image Retrieval on Java Enabled Mobile Devices,” *In Proceedings of the 2005 Finnish Signal Processing Symposium (FINSIG'05)*, Kuopio, Finland, August 25, 2005.

[2] I. Ahmad, S. Kiranyaz and M. Gabbouj, “An Efficient Image Retrieval Scheme on Java Enabled Mobile Devices,” *In Proceedings of IEEE MMSP05*, Shanghai, China, October 2005, pp. 1-4.

[3] J.R. Bach, C. Fuller, A. Gupta, “The Virage Image Search Engine: An Open Framework for Image Management,” *In Proceedings of the SPIE*, vol. 2670, February 1996, pp. 76–87.

[4] E. Bingham and H. Mannila, “Random Projection in Dimensionality Reduction: Applications to Image and Text Data,” *In Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD-2001, August 2001, USA, pp. 245-250.

[5] M. Birinci, S. Kiranyaz and M. Gabbouj, “Image Color Content Description Utilizing Perceptual Color Correlogram,” *In Proceedings of the IEEE Workshop on Content Based Multimedia Indexing (CBMI'08)*, London, UK, June 2008, pp. 200-207.

[6] C. Breiteneder and H. Eidenberger, “Content-Based Image Retrieval and Digital Libraries,” *In Proceedings of International Conference on Digital Libraries*, Tokyo, Japan, 2000, pp.67-74.

[7] D. Cai, X. He, Z. Li, W.Y. Ma, J.R. Wen, “Hierachical Clustering Of WWW Image Search Results Using Visual, Textual And Link Information,” *In Proceedings of the ACM International Conference on Multimedia*, New York, USA, 2004, pp. 952-959.

[8] J. Canny, “A Computational Approach to Edge Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, November 1986, pp. 679-698.

[9] C. Carson, S. Belongie, H. Greenspan, J. Malik, “Blobworld: Image Segmentation Using Expectation-Maximization and its Application to Image Querying,” *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, 2002, Vol. 24, pp. 1026-1038.

[10] T. Catarci, S. Kimani, E. Panizzi, M. Antona, "Digital Library Requirements: a Questionnaire-based Study," *In Proceedings of the 1st Italian Research Conference on Digital Library Management Systems*, IRCDL 2005, DELOS Network of Excellence, January 2005, pp. 65-74.

[11] CCITT Recommendation T.81, "Information Technology - Digital Compression and Coding of Continuous-tone Still Images," September 1992.

[12] G. Cha and C. Chung, "Multi-Mode Indices for Effective Image Retrieval in Multimedia Systems," *In Proceedings of IEEE Multimedia Computing Systems*, 1998, pp.152–159.

[13] S.F. Chang, W. Chen, J. Meng, H. Sundaram and D. Zhong, "VideoQ: An Automated Content Based Video Search System Using Visual Cues," *In Proceedings of ACM Multimedia*, Seattle, 1997, pp. 313-324.

[14] F.A. Cheikh, "MUVIS: A System for Content-Based Image Retrieval," *PhD. Thesis*, Tampere University of Technology, Tampere, Finland, April 2004.

[15] F.A. Cheikh, B. Cramariuc, M. Gabbouj, "Relevance Feedback For Shape Query Refinement," *In Proceedings of IEEE International Conference on Image Processing*, ICIP 2003, Barcelona, Spain, Vol. 1, September 2003, pp. 745-748.

[16] J.Y. Chen, C. Taskiran, A. Albiol, E. J. Delp and C. A. Bouman, "ViBE: A Video Indexing and Browsing Environment," *In Proceedings of the SPIE Conference on Multimedia Storage and Archiving Systems IV*, September 20-22, 1999, Boston, Vol. 3846, pp. 148-164.

[17] A.H. Cohen, "Retrieval and Browsing Image Database Using Image Thumbnails," *Journal of Visual Computing and Image Representation*, Vol. 8, No. 2, 1997, pp 226-234.

[18] R.T. Collins, L. Yanxi, M. Leordeanu, "Online Selection Of Discriminative Tracking Features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, Issue 10, 2005, pp. 1631- 1643.

[19] D.T. Covey, "Usage and Usability Assessment: Library Practices and Concerns," *Digital Library Federation, Council on Library and Information Resources reports*, January 2002.

[20] B. Cramariuc, "Image Analysis and Understanding Techniques with Application to Content Based Image Manipulation," *PhD. Thesis*, Tampere University of Technology, Tampere, Finland, December 2005.

[21] R. Czaja, J. Blair, "Designing Surveys: A Guide to Decisions and Procedures," 2nd edition. London: Sage, 300 p., 2005.

[22] R. Datta, J. Dhiraj, L. Jia, Z.W. James, "Image Retrieval: Ideas, Influences, and Trends of the New Age," *In Proceedings of ACM Computing Surveys*, Vol. 40, No. 2, 2008.

[23] Y. Deng, and B.S. Manjunath, "Unsupervised Segmentation of Color-Texture Regions in Images and Video," *IEEE Transactions on PAMI*, Vol. 23, August 2001, pp. 800-810.

[24] C. Ding, H. Peng, "Minimum Redundancy Feature Selection from Microarray Gene Expression Data", *In Proceedings of the IEEE Computer Society Conference on Bioinformatics*, August 11-14, 2003, pp.523-528.

[25] D. Djordjevic, E. Izquierdo, "An Object- and User-Driven System for Semantic-Based Image Annotation and Retrieval," *IEEE Trans. Circuits Syst. Video Techn.*, Vol. 17, No. 3, 2007, pp. 313-323.

[26] R.O. Duda, P.E. Hart, and D.G. Stork, "Pattern Classification", *New York John Wiley & Sons, Inc.*, Second Edition, 2001.

[27] J.G. Dy, C.E. Brodley, A.C. Kak, L.S. Broderick, A.M. Aisen, "Unsupervised Feature Selection Applied to Content-Based Retrieval of Lung Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, Issue 3, 2003, pp. 373-378.

[28] J.P. Eakins, P. Briggs, B. Burford, "Image Retrieval Interfaces: A User Perspective," *In Proceedings of Third International Conference on Image and Video Retrieval*, CIVR 2004, Proceedings of Lecture Notes in Computer Science 3115, Dublin, Ireland, July 21-23, 2004, pp. 628-637.

[29] R. Egas and D.P. Huijsmans and M.S. Lew and N. Sebe, "Adapting k-d Trees to Visual Retrieval," *In Proceedings of Visual 99*, Springer-Verlag, Berlin, 1999, pp. 533-540.

[30] D.P.W. Ellis and J.A. Bilmes, "Using Mutual Information to Design Feature Combinations," *In Proceedings of International Conference on Spoken Language Processing*, ICSLP-2000, Beijing, Vol. 3, October 2000, pp. 79-82.

[31] P.G.B. Enser, "Progress in Documentation: Pictorial Information Retrieval," *Journal of Documentation*, Vol. 51, 1995, pp. 126-170.

[32] P. Felzenszwalb and D. Huttenlocher, "Efficient Graph-Based Image Segmentation," *International Journal of Computer Vision*, Vol. 59, September 2004, pp. 167-181.

[33] D. Frohlich, A. Kuchinsky, C. Pering, A. Don, S. Ariss, "Requirements for Photoware," *In Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, New Orleans, Louisiana, USA, 2002, pp.166-175.

[34] K. Fukunaga, "Introduction to Statistical Pattern Recognition," *Academic Press*, 1990.

[35] A. Gersho and R. M. Gray, "Vector Quantization and Signal Compression," Boston, *Kluwer*, 1992.

[36] E. Guldogan and M. Gabbouj, "Mapping By Adaptive Threshold Method for Dimension Reduction of Content-Based Indexing and Retrieval Features," *In Proceedings of EUSIPCO 2005, European Signal Processing Conference*, Antalya, Turkey, September 2005.

[37] E. Guldogan and M. Gabbouj, "Unsupervised Elimination of Media Items in Content-Based Image Retrieval," *In Proceedings of the IEEE International Conference on Signal Processing and Communications*, ICSPC 2007, Dubai United Arab Emirates (UAE), 24–27 November 2007, pp. 5-8.

[38] E. Guldogan, O. Guldogan and M. Gabbouj, "DCT-Based Downscaling Effects On Color And Texture-Based Image Retrieval," *In Proceedings of the EWIMT 2004, IEE European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology*, London, U.K., November 2004, pp. 79-86.

[39] E. Guldogan, O. Guldogan and M. Gabbouj, "Visual Media Retrieval Using Transform-Based Layered Query Scheme," *In Proceedings of the ICIP 2005 International Conference on Image Processing*, Genoa, Italy, September 2005, pp. 521-524.

[40] E. Guldogan, O. Guldogan, S. Kiranyaz, K. Caglar and M. Gabbouj, "Compression Effects on Color and Texture Based Multimedia Indexing and Retrieval," *In Proceedings of IEEE International Conference on Image Processing, ICIP 2003*, Barcelona, Spain, 14-17 September 2003. 2. pp. 9-12.

[41] E. Guldogan, O.Guldogan and M. Gabbouj, "Efficient Image Retrieval with JPEG 2000 and DWT-Based Downscaling," *In Proceedings of the International Workshop*

on Spectral Methods and Multirate Signal Processing, SMMSP 2006, Florence, Italy, Vol. 34, September 2006, pp. 199-203.

[42] E. Guldogan and M. Gabbouj, "Feature Selection for Content-based Image Retrieval," *Springer Journal on Signal, Image and Video Processing*, Vol. 2, No. 3, September 2008, pp. 241-250.

[43] E. Guldogan and M. Gabbouj, "Transform-Based Layered Query System for Image Indexing and Retrieval," *In Proceedings of International Workshop on Nonlinear Signal and Image Processing*, NSIP 2005, Sapporo, Japan, May 2005, pp.381-384.

[44] O. Guldogan, M. Gabbouj, "Content-Based Image Indexing and Retrieval Framework on Symbian Based Mobile Platform," *In Proceedings of 13. European Signal Processing Conference*, EUSIPCO, Antalya, Turkey, 4-8 September 2005.

[45] O. Guldogan, E. Guldogan, S. Kiranyaz, K. Caglar and M. Gabbouj, "Dynamic Integration of Explicit Feature Extraction Algorithms into MUVIS Framework," *In Proceedings of the FINSIG 2003, Finnish Signal Processing Symposium*, Tampere, Finland, May 2003, pp. 120-123.

[46] J. Hafner, H. S. Sawhney, W. Equitz, M. Flickner, and W. Niblack, "Efficient Color Histogram Indexing for Quadratic Form Distance Functions," *IEEE Transactions on Pattern Anal. Machine Intelligence*, vol. 17, July 1995, pp. 729–736.

[47] M.J. Halvey, and M.T. Keane, "Analysis of Online Video Search and Sharing," *In Proceedings of the 18th ACM Conference on Hypertext and Hypermedia*, Manchester, UK, 2007, pp. 217-226.

[48] S. Hariri, M. Yousif, G. QuA, "New Dependency and Correlation Analysis for Features," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17, Issue 9, September 2005, pp. 1199-1207.

[49] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, R. Zabih, "Image Indexing Using Color Correlograms", *In Proceedings of Computer Vision and Pattern Recognition*, 1997, pp.762- 768.

[50] T. Huang, S. Mehrotra, K. Ramchandran, "Multimedia Analysis and Retrieval System (MARS) Project," *In Proceedings of 33rd Annual Clinic on Library Application of Data Processing, Digital Image Access and Retrieval*, 1996.

[51] ImagEval Webpage: http://www.imageval.org/.

[52] Independent JPEG Group: http://www.ijg.org/.

[53] Y.S. Irianto, J. Jiang, and S.S. Ipson, "An Efficient Image Retrieval through DC Feature Extraction," *In Proceedings of (531) Circuits Signals and Systems CSS*, San Francisco, California, USA, 2006.

[54] E. Izquierdo and D. Djordjevic "Using Relevance Feedback to Bridge The Semantic Gap," *Lecture Notes in Computer Science*, Vol. 3877, 2006, pp. 19-34.

[55] A. Jaimes, "Human Factors in Automatic Image Retrieval System Design and Evaluation," *invited paper, IS&T/SPIE Internet Imaging 2006*, San Jose, CA, January 2006.

[56] A. Jain, D. Zongker, "Feature Selection: Evaluation, Application, and Small Sample Performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 2, February 1997, pp. 153-158.

[57] W. Jin, R. Shi, T.-S. Chua, "A Semi-naı¨ve Bayesian Method Incorporating Clustering with Pair-Wise Constraints for Auto Image Annotation," *In Proceedings of the ACM Multimedia*, 2004, pp. 336-339.

[58] F. Jing, M. Li, H.-J. Zhang, B. Zhang, "Relevance Feedback in Region-based Image Retrieval," *IEEE Transaction on Circuits and Systems for Video technology*, Vol. 14, Issue 5, May 2004, pp. 672 - 681.

[59] I.T. Jolliffe, "Principal Component Analysis," *Springer-Verlag*, New-York, 1988.

[60] T. Kaster, M. Pfeiffer, C. Bauckhage, "Usability Evaluation for Image Retrieval Beyond Desktop Applications," *In Proceedings of IEEE International Conference on Multimedia and Expo*, 9-12 July 2006, pp. 385-388.

[61] S. Kiranyaz, "Advanced Techniques for Content-Based Management of Multimedia Databases," *PhD. Thesis*, Tampere University of Technology, Tampere, Finland, June 2005.

[62] S. Kiranyaz, M. Ferreira, M. Gabbouj, "A Generic Shape/Texture Descriptor over Multi-Scale Edge Field: 2D Walking Ant Histogram," *in IEEE Transactions on Image Processing*, vol. 17, no. 3, 2008, pp. 377-391.

[63] S. Kiranyaz, K. Caglar, O. Guldogan, and E. (Karaoglu) Guldogan, "MUVIS: A Multimedia Browsing, Indexing and Retrieval Framework," *In Proceedings of Third*

International Workshop on Content Based Multimedia Indexing, CBMI'03, France, September 2003, pp. 405-412.

[64] S. Kiranyaz, K. Caglar, E. Guldogan, O. Guldogan and M. Gabbouj, "MUVIS: A Content Based Multimedia Indexing and Retrieval Framework," *In Proceedings of the ISSPA 2003 International Symposium on Signal Processing and its Applications*, July 2003, pp.1-8.

[65] S. Kiranyaz, M. Gabbouj, "Hierarchical Cellular Tree: An Efficient Indexing Scheme for Content-based Retrieval on Multimedia Databases," *IEEE Transactions on Multimedia*, vol. 9, 2007, pp. 102-119.

[66] S. Kiranyaz, A.F. Qureshi, and M. Gabbouj, "A Generic Audio Classification And Segmentation Approach For Multimedia Indexing And Retrieval," *IEEE Transactions on Audio, Speech and Language Processing*, 2006, pp. 1062-1081.

[67] D. Kirk, A. Sellen, C. Rother, K. Wood, "Understanding Photowork," *In Proceedings of the Conference on Human Factors in Computing Systems*, Canada, 2006, pp.761-770.

[68] D. Koller and M. Sahami. "Toward Optimal Feature Selection," *In Proceedings of the 13th International Conference on Machine Learning*, 1996, pp. 284-292.

[69] M. Kuniavsky, "Observing the User Experience: A Practitioner's Guide to User Research," *Published by Morgan Kaufmann*, 2003, 560 p, pp. 129-155.

[70] I. Kunttu, "Shape and Gray Level Descriptors for Surface Defect Image Retrieval and Classification," *PhD. Thesis*, Tampere University of Technology, Tampere, Finland, November 2005.

[71] C. Lee and V. Soo, "Modeling Annotators To Improve Semantic Annotation and Retrieval of Images," *In Proceedings of IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, 19-22 Sept. 2005, pp. 254- 257.

[72] L. Lepistö, "Colour And Texture Based Classification of Rock Images Using Classifier Combinations," *PhD. Thesis*, Tampere University of Technology, Tampere, Finland, April 2006.

[73] M. Lew, N. Sebe, C. Djeraba, R. Jain, "Content-based Multimedia Information Retrieval: State of the Art and Challenges," *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2006, pp. 1-19.

[74] X.Q. Li, I. King, "Information Retrieval Using Local Linear PCA," *Proc. of 6th International Conference on Neural Information Processing*, ICONIP '99, 1999, Vol. III, pp. 867-872.

[75] X. Lin, S.M. Yacoub, J. Burns, S.J. Simske, "Performance Analysis Of Pattern Classifier Combination By Plurality Voting," *Pattern Recognition Letters*, Vol. 24, Issue 12 , 2003, pp. 1959-1969.

[76] V.T. Linh, L. Reiner "PCA-Based Representation Of Color Distributions For Color Based Image Retrieval," *In Proceedings of the IEEE International Conference on Image Processing*, ICIP 2001, vol.2, pp 697–700.

[77] W. Liu, Z. Su, S. Li, H.J. Zhang, "A Performance Evaluation Protocol for Content-Based Image Retrieval Algorithms/Systems," *In Proceedings of IEEE CVPR Workshop on Empirical Evaluation in Computer Vision*, Kauai, USA, December, 2001

[78] H. Liu, L. Yu, "Toward Integrating Feature Selection Algorithms for Classification and Clustering," *IEEE Transactions on Knowledge and Data Engineering,* Vol. 17, Issue 4, 2005, pp.491-502.

[79] Y. Liu, D. Zhang, G. Lu, W.-Y. Ma, "A Survey of Content-Based Image Retrieval with High-Level Semantics," *Pattern Recognition*, Vol. 40, No.1, January 2007, pp.262-282.

[80] H. Long and W.K. Leow, "Perceptual Consistency Improves Image Retrieval Performance," *In Proceedings of ACM SIGIR*, 2001, pp. 9-12.

[81] G. Lu and S. Teng. "A Novel Image Retrieval Technique Based on Vector Quantization," *In Proceedings of International Conference on Computational Intelligence for Modeling, Control and Automation*, 1999, pp. 36-41.

[82] Z. Lu, S. Li and H. Burkhardt, "A Content-Based Image Retrieval Scheme In Jpeg Compressed Domain," *International Journal of Innovative Computing, Information and Control ICIC*, Vol. 2, Number 4, August 2006, pp. 831-839.

[83] V.V. Lukin, R. Oktem, N. Ponomarenko, K. Egiazarian, "Image Filtering Based on Discrete Cosine Transform," *Telecommunications and Radio Engineering*, Vol. 66, No 18, 2007, pp. 1685-1701.

[84] W.Y. Ma, B.S. Manjunath, "Texture Features for Browsing and Retrieval of Image Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, Aug. 1996, pp. 837-842.

[85] W.Y. Ma, B.S. Manjunath, “Netra: A Toolbox For Navigating Large Image Databases,” *Multimedia Systems*, Vol. 7, No. 3, 1999, pp. 184-198.

[86] A.J. Maeder, “Lossy Compression Effects On Digital Image Matching,” *In Proceedings of 14th International Conference on Pattern Recognition*, Australia, 1998, Vol. 2, pp. 1626-1629.

[87] M.K. Mandal and C. Liu, “Efficient Image Indexing Techniques In The JPEG2000 Domain,” *Journal of Electronic Imaging*, January 2004, Vol. 13, Issue 1, pp. 182-190.

[88] B.S. Manjunath, J.-R. Ohm, V.V. Vasudevan, A. Yamada, “Color and Texture Descriptors,” *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 6, June 2001, pp. 703 -715.

[89] B.S. Manjunath, P. Salembier, and T. Sikora Eds., “Introduction to MPEG-7, Multimedia Content Description Interface,” *John Wiley & Sons Ltd*, England, 371 p., 2002.

[90] V. Mezaris, I. Kompatsiaris and M.G. Strintzis, “Still Image Segmentation Tools for Object-based Multimedia Applications,” *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 18, June 2004, pp. 701-725.

[91] P. Mitra, C.A. Murthy and S.K. Pal, “Unsupervised Feature Selection Using Feature Similarity”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 3, 2002, pp.301-312.

[92] A. Mojsilovic, B. Rogowitz, “Capturing Image Semantics with Low-Level Descriptors,” *In Proceedings of the IEEE International Conference on Image Processing ICIP*, September 2001, pp. 18–21.

[93] M.S. Mughal, M. Nawaz, F. Ahmad, S. Shahzad, A.K. Bhatti, S. Mohsin, “A 3D-Hash Function for Fast Image Indexing and Retrieval,” *In Proceedings of Computer Graphics, Imaging and Visualization (CGIV 2007)*, 2007, pp. 341-348,

[94] MUVIS: A System for Content-Based Multimedia Indexing and Retrieval in Multimedia Databases: http://muvis.cs.tut.fi/.

[95] R. Ng and A. Sedighian, “Evaluating Multi-Dimensional Indexing Structures For Images Transformed By Principle Component Analysis,” *In Proceedings of the SPIE*, 2670, 1994, pp. 50-61.

[96] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, “The QBIC Project: Querying Images by Content Using Color, Textures and Shape,” *In Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, vol. 1908, April 1993, pp.173-187.

[97] R. Oktem, L. Yaroslavsky, K. Egiazarian, and J. Astola, “Transform Domain Approaches For Image Denoising,” *Journal of Electronic Imaging*, 2002, pp. 149-156.

[98] A.V. Oppenheim, R.W. Schafer, and J.R. Buck, “Discrete-Time Signal Processing,” second edition, *Prentice-Hall*, New Jersey, 1999.

[99] M. Partio, B. Cramariuc, M. Gabbouj, A. Visa, “Rock Texture Retrieval Using Gray Level Co-occurrence Matrix,” *In Proceedings of 5th Nordic Signal Processing Symposium*, October 2002.

[100] M. Partio, “Ordinal Methods for Similarity Evaluation Applied to Texture and Shape Retrieval,” *PhD. Thesis*, Tampere University of Technology, Tampere, Finland, October 2007.

[101] K. Pearson, “Mathematical Contributions to the Theory of Evolution: III. Regression, Heredity and Panmixia,” *Philosophical Transactions of the Royal Society of London Ser. A*, vol. 187, 1896, pp. 253-318.

[102] H. Peng, F. Long, and C. Ding, “Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 8, 2005, pp.1226-238.

[103] A. Pentland, R.W. Picard, S. Sclaroff, “Photobook: Tools for Content Based Manipulation of Image Databases,” *In Proceedings of SPIE Storage and Retrieval for Image and Video Databases II*, 1994, pp. 34-47.

[104] N.N. Ponomarenko, V.V. Lukin, S.K. Abramov, A.A. Zelensky, “Image Similarity Measures For Search Systems,” *Radioelectronic and Computer Systems*, Kharkov, Ukraine, 2007, No. 2, pp. 49-58.

[105] T. Quack, U. Monich, L. Thiele, B.S. Manjunath, “Cortina: A System for Large-scale, Content-based Web Image Retrieval,” *In Proceedings of the 12th annual ACM International Conference on Multimedia*, 2004, pp.508-511.

[106] Y. Rao, P. Mundur, Y. Yesha, "Fuzzy SVM Ensembles for Relevance Feedback in Image Retrieval," *In Proceedings of International Conference on Image and Video Retrieval CIVR*, 2006, pp. 350-359.

[107] K. Rodden, K. R. Wood, "How do People manage their Digital Photographs?," *In Proceedings of the Conference on Human Factors and Computing Systems*, CHI 2003, ACM Press, Florida, April 2003, pp. 409-416.

[108] A.M. Ross, D.H. Rhodes, D.E. Hastings, "Defining Changeability: Reconciling Flexibility, Adaptability, Scalability, Modifiability, and Robustness For Maintaining System Lifecycle Value," *Journal of Systems Engineering*, Vol. 11, Issue 3, 2008, pp. 246 – 262.

[109] Y. Rui, T.S. Huang, S.F. Chang, "Image Retrieval: Current Techniques, Promising Directions and Open Issues," *Journal of visual communication and image representation*, 1999, No. 10, pp. 39-62.

[110] D. Shaw, "The Human-Computer Interface for Information Retrieval," *Annual Review of Information Science and Technology, ASIS*, Vol. 26, 1991, pp.155-195.

[111] A. Sheth and W. Klas, "Multimedia Data Management: Using Metadata to Integrate and Apply Digital Media," *McGraw-Hill*, 1998, San Francisco.

[112] D. Shi, W. Shu, H. Liu, "Feature Selection for Handwritten Chinese Character Recognition Based on Genetic Algorithms," *In Proceedings of IEEE Int. Conference on Systems, Man, and Cybernetics*, Vol.5, 1998, pp. 4201-4206.

[113] J.R. Smith and Chang, "Visualseek: A Fully Automated Content-Based Image Query System," *In Proceedings of ACM Multimedia*, Boston, November 1996, pp.87-98.

[114] B. Smyth, C.P. Lam, X. Chen, V. Maxville, "Heuristics for Image Retrieval Using Spatial Configurations," *In Proceedings of Digital Image Computing: Techniques & Applications*, DICTA 2003, Sydney, 2003, pp. 909-918.

[115] M. J. Swain, D.H. Ballard, "Color Indexing," *International Journal of Computer Vision*, Vol.7, No. 1, 1991, pp. 11-32.

[116] J.M. Torres, A. Parkes, "User Modeling and Adaptivity in Visual Information Retrieval Systems," *In Proceedings of the Workshop on Computational Semiotics for New Media*, 2000.

[117] N. Vasconcelos and A. Lippman, "Learning from User Feedback in Image Retrieval Systems," *In Proceedings of Neural Information Processing Systems* 12, Denver, Colorado, 1999, pp. 977-986.

[118] N. Vasconcelos and M. Vasconcelos, "Scalable Discriminant Feature Selection for Image Retrieval and Recognition," *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, San Diego, CA, USA, Vol. 2, 2004, pp. 770-775.

[119] R. Vermilyer, "Intelligent User Interface Agents in Content-Based Image Retrieval," *In Proceedings of the IEEE Conference SoutheastCon*, 2006, pp. 136-142.

[120] R. Vieux and J.-P. Domenger, and J. Benois-Pineau, and A. Braquelaire, "Image Classification with User Defined Ontology," *In Proceedings of European Signal Processing Conference EUSIPCO 2007*, Poznan, Poland, 2007.

[121] V.V. Vinod (ed.), "Description of Core Experiments for MPEG-7 Color/Texture Descriptors," ISO/IEC JTC1/SC29/WG11 (MPEG) document no. N2929, October, 1999.

[122] X. Wan, C.J. Kuo, "A Multiresolution Color Clustering Approach To Image Indexing And Retrieval," *In Proceedings of IEEE International Conference Acoustics, Speech, Signals Processing*, vol. 6, 1998 pp. 3705–3708.

[123] J.Z. Wang, J. Li, G. Wiederhold, "SIMPLIcity: Semantics-sensitive Integrated Matching for Picture Libraries," *In Proceedings of IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 9, 2001, pp. 947-963.

[124] J. Wei, E. Guihua, D. Qionghai, G. Jinwei, "Similarity-Based Online Feature Selection In Content-Based Image Retrieval," *IEEE Transactions on Image Processing*, Vol. 15, Issue: 3, March 2006, pp.702-712.

[125] D. Weiss, J. Scheuerer, M. Wenleder, A. Erk,, M. Gülbahar, and C. Linnhoff-Popien, "A User Profile-Based Personalization System for Digital Multimedia Content," *In Proceedings of The 3rd International Conference On Digital Interactive Media In Entertainment And Arts*, DIMEA '08, Athens, Greece, September, 2008, vol. 349, pp. 281-288.

[126] E.P. Xing, M.I. Jordan and R.M. Karp, "Feature Selection For High-Dimensional Genomic Microarray Data," *In Proceedings of the Eighteenth International Conference on Machine Learning*, 2001, pp. 601-608.

[127] L. Xu, A. Krzyzak, C.Y. Suen, “Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition,” *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 22, Issue: 3, 1992, pp. 418-435.

[128] C. Zhang, J. Chai, and R. Jin, “User Term Feedback in Interactive Text-based Image Retrieval,” *In Proceedings of 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2005)*, Salvador, Brazil, August 15-19, 2005, pp. 51-58.

[129] D. Zhang, “Improving Image Retrieval Performance by Using Both Color and Texture Features,” *In Proceedings of Third International Conference on Image and Graphics (ICIG'04)*, 2004, pp. 172-175.

[130] H. Zhang, Y. Gong, C.Y. Low, and S.W. Smoliar, “Image Retrieval Based on Color Features: An Evaluation Study,” *In Proceedings of SPIE Digital Image Storage Archiving Systems*, vol. 2606, 1995, pp. 212-220.

[131] Z. Zhang, Z. Guo, C. Faloutsos, E.P. Xing, J.Y. Pan, “On the Scalability and Adaptability for Multimodal Retrieval and Annotation,” *In Proceedings of 14th International Conference on Image Analysis and Processing Workshops*, 2007, ICIAPW 2007, pp. 39-44.

[132] X.S. Zhou and T.S. Huang, “CBIR: From Low-Level Features to High-Level Semantics,” *In Proceedings of SPIE Image and Video Communication and Processing*, 2000, pp. 24-28.

APPENDIX

ONLINE USER SURVEY ON USE OF MULTIMEDIA AND CBIR

General Information on Audiences:

122 Persons

Female: 27 persons

Male: 95 persons

Age distributions are 32% of 20-24, 61% of 25-35, and 7% of 36-50 years old.

Profession: Computer, Software, Electronic, Telecommunication Engineers, IT students, Researchers and Professors.

First Part: General Information on Use of Digital Multimedia

1- Which of the following multimedia devices do you use?

a.	Digital Photo Camera	**85 %**
b.	Digital Video Camera	**42 %**
c.	Personal Computer	**95 %**
d.	Mobile Phone	**98 %**
e.	Palm	**10 %**
f.	Other	**17 %**

2- How often do you take digital pictures?

a.	almost every day	**19 %**
b.	once in a week	**36 %**
c.	once in a month	**36 %**
d.	a few times a year	**16 %**
e.	less frequently	**7 %**

3- Which of the following devices do you use to take pictures?

a.	Digital camera	**87 %**
b.	Mobile Phone	**47 %**
c.	Analog Camera	**12 %**
d.	Web-cam	**14 %**
e.	other	**0 %**

4- What is the approximate size of your digital image database/collection?

a. <1GB **24 %**
b. 1-10 GB **58 %**
c. 10-100 GB **15 %**
d. >100 GB **3 %**

5- Which of the following do you prefer to use for storing your multimedia files?

a. CD **12 %**
b. DVD **21 %**
c. Hard-drive **68 %**
d. Web-servers **0%**
e. Other **0%**

6- Do you take pictures with your mobile phone?

a. Yes **61 %**
b. No **39 %**

7- Do you take video clips with your mobile phone?

a. Yes **41 %**
b. No **59 %**

8- Do you prefer to store your images/videos in your mobile phone/device?

a. Yes **31 %**
b. No **69 %**

9- How frequently do you access your digital image database/collection?

a. almost every day **12 %**
b. once in a week **45 %**
c. once in a month **32 %**
d. a few times a year **8 %**
e. less frequently **3 %**

10- How would you prefer to organize your multimedia files?

a. by events **46 %**
b. by date **34 %**
c. by people **9 %**
d. by location **6 %**
e. by multimedia source **1 %**
f. other **4 %**

11- Do you use any automated/advanced tools for organizing your personal multimedia files?

a.	Yes	**21 %**
b.	No	**79 %**

12- Which of the following do you prefer to see for each multimedia item when browsing?

a.	Full-size image	**30 %**
b.	Thumbnail (decimated version of the image)	**64 %**
c.	Associated textual description (caption, date, file size, etc.)	**4 %**
d.	Other	**2 %**

13- What is your preferred image size/resolution for browsing?

a.	Large: over 3 megapixels	**22 %**
b.	Medium: 1-3 megapixels	**46 %**
c.	Small: ~1 megapixels	**19 %**
d.	Very Small: less than 1 megapixels	**13 %**

14- Do you prefer to use compression for your image and video files?

a.	Yes	**65 %**
b.	No	**35 %**

15- Which type of compression do you prefer to use for compressing your multimedia files?

a.	Lossy	**40 %**
b.	Lossless	**60 %**

16- Which of the following codecs do you prefer to use for compressing your image files?

a.	Jpeg	**93 %**
b.	Jpeg-2000	**17 %**
c.	Bmp	**12 %**
d.	Gif	**20 %**
e.	Tif	**10 %**
f.	Other	**5 %**
g.	None	**2 %**

17- Which of the following codecs do you prefer to use for compressing your video files?

a.	Mpeg	**34 %**
b.	Mpeg-2	**15 %**
c.	Mpeg-4	**34 %**
d.	Divx	**53 %**
e.	H263	**5 %**
f.	H264	**11 %**
g.	Other	**7 %**
h.	None	**7 %**

18- Do you use web-browsers for searching image and video files?

a.	Yes	**67 %**
b.	No	**33 %**

19- Which network speed do you have while using multimedia services?

a.	Fast: over 2 Megabit per sec	**45 %**
b.	Medium: 1-2 Megabit per sec	**35 %**
c.	Slow: 128-1024 Kilobit per sec	**17 %**
d.	Very Slow: less than 128 Kilobit per sec	**2 %**

Second Part: Use of Content-Based Multimedia Indexing and Retrieval

20- If you were given the following choices, which one would you use to classify/organize your multimedia database/collection

a.	Color Content	**15 %**
b.	Object/Shape content	**32 %**
c.	Metadata (such as date, caption etc.)	**61 %**
d.	Short Text Description	**48 %**
e.	Texture Content	**5 %**
f.	Other	**1 %**

21- Do you ever search a specific/certain multimedia item in your digital media collection?

a.	Yes	**62 %**
b.	No	**38 %**

22- Do you search a multimedia item resembling/similar to a reference/example multimedia item in your digital media database/collection?

a.	Yes	**37 %**
b.	No	**63 %**

23- How would like to search your media files?

a.	By example	**28 %**
b.	By text	**58 %**
c.	By sketch	**8 %**
d.	Other	**5 %**

24- Do you make an image/video search from the web?

a.	Yes	**81 %**
b.	No	**19 %**

25- How do you search an image/video from the web?

a.	By text	**77 %**
b.	By content	**18 %**
c.	By color	**1 %**
d.	By texture	**2 %**
e.	By shape	**0 %**
f.	By example	**1 %**
g.	Other	**0 %**

26- How would you prefer to search an image/video from the web?

a.	By text	**39 %**
b.	By content	**39 %**
c.	By color	**4 %**
d.	By texture	**3 %**
e.	By shape	**3 %**
f.	By example	**12 %**
g.	Other	**1 %**

27- Which of the following environment/system would you use to search your multimedia files?

a.	Mobile systems	**3 %**
b.	Web-based systems	**36 %**
c.	Personal Computer	**61 %**
d.	Other Distributed Systems	**0 %**
e.	Other	**0 %**

28- Which of the following has first priority for you while searching an image/video from the database/collection?

a.	Content	**73 %**
b.	Date	**13 %**
c.	Size	**3 %**
d.	Other	**1 %**

29- How would an automatic search mechanism contribute in managing/handling your image database/collection?

a.	Strongly oppose	**3 %**
b.	Oppose	**11 %**
c.	Support	**51 %**
d.	Strongly support	**20 %**
e.	Neutral	**15 %**

30- What is the reasonable waiting time in your opinion to see the results of an image/video search on the Internet?

a.	Instantaneous	**40 %**
b.	approximately 30 seconds	**48 %**
c.	between 30 sec and 1 min	**6 %**
d.	1-3 mins	**5 %**
e.	more than 3 mins	**1 %**

31- What is the reasonable waiting time in your opinion to see the results of an image/video search on your personal computer?

a.	Instantaneous	**56 %**
b.	approximately 30 seconds	**31 %**
c.	between 30 sec and 1 min	**7 %**
d.	1-3 min	**5 %**
e.	more than 3 min	**2 %**

32- What would you like to see as a result of an image search?

a.	A certain image that exactly matches your criteria	**15 %**
b.	A set of relevant images ordered according to their relevancy	**73 %**
c.	Categorization of image collection	**9%**
d.	Other	**3%**

33- Do you prefer to save your image database/collection in your web-server?

a.	Yes	**41 %**
b.	No	**59 %**

34- Which of the following attribute is the most critical one for an image retrieval system?

a.	High Speed	**30 %**
b.	Accurate results	**70 %**

Third Part: Use of Features

35- What is your knowledge about image features?

a.	Excellent	**20 %**
b.	Good	**40 %**
c.	Fair	**27%**
d.	Poor	**13%**

36- If you are given a retrieval system to use, would you be interested in adjusting its feature settings for each retrieval in order to potentially achieve higher accuracy?

a.	Yes	**77 %**
b.	No	**23 %**

37- If you are given a retrieval system to use, would you prefer it to adjust the feature settings automatically with a reasonable error margin for each retrieval?

a.	Yes	**70 %**
b.	No	**30 %**

38- Which of the following do you prefer for a retrieval system?

a.	Interactive retrieval system where you can influence the results on the fly or between multiple steps	**72 %**
b.	System that does not require any interaction and running at once	**28 %**