This is the published version:


Available from Deakin Research Online:

http://hdl.handle.net/10536/DRO/DU:30009700

©2005 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Copyright : 2005, IEEE
Impact of image organizations on multimedia document retrieval

Nazmul Haque
Test Centre, ITS, Telstra,
Melbourne, Victoria 3000
Email: nazmul.haque@team.telstra.com.au

Morshed Chowdhury
Deakin University, Melbourne,
Victoria, Australia
Email: muc@deakin.edu.au

Syed Mahbubur Rahman
Minnesota State University, Minnesota, USA
Email: mahbubur.syed@mnsu.edu

Abstract—In this paper we compare ranking effectiveness of heterogeneous multimedia document retrieval when different image organizations are used for formulating queries. The quality of image queries depends on the organization of images used to make queries which in turn significantly impacts retrieval precision. CBIR (Content Based Information Retrieval) needs an effective and efficient organization of images including user interface which must be part of the configuration parameters of image retrieval research.

I. INTRODUCTION

Identification of configuration items for content based information retrieval (CBIR) research is crucial for comparing results of retrieval precisions. A huge number of CBIR publications are available where image organizations and user interfaces have been ignored as an item of retrieval research [5], [6], [8], [7].

In this paper, we identify image organization as a critical configuration parameter of the ranking experiment which cannot be ignored. We also assess the impact of image organizations on retrieval precision — especially for heterogeneous collections characterized by diverse source, themes and uncontrolled acquisition.

II. EXPERIMENTAL METHOD

We examined two collections of multimedia documents obtained from a random sampling of an Encyclopedia. These documents had digital still heterogeneous images, characterized by uncontrolled acquisitions and diverse themes and contents. We then conducted two experiments, keeping other configuration parameters unchanged: (i) users are shown images randomly and (ii) users are shown images organized in a tree view. We then compared the ranking precisions of both experiments at zero recall, to assess the impact of image organization.

Multimedia document retrieval algorithms can be divided into three interrelated algorithms: image featuring, feature matching and the combination of multiple evidence. Features are extracted, both from queries and multimedia documents, before matching the features of the query with the features of the documents. We only cover image feature extraction from a heterogeneous collection and a feature matching algorithm. In this paper, we shall consider only two image features: colour and texture.

III. FEATURE EXTRACTION

A. Colour

Numerous colour features can be utilized in ranking retrieval. Typical colour histograms are used for ranking which have frequency of occurrence on the y-axis and brightness of colour channel values on the x-axis.

Several different types of colour schemes are available to represent colour. In these colour schemes, three different measurements represent a colour when mixed together. These three measurements are known as primaries. RGB, HSI, YC₀C₁₀.CIE-L*a*b* etc. are examples of colour histograms. Among these colour spaces CIE-L*a*b* and CIE-L*u*v* are known as uniform colour spaces in which a change in any of the primaries would produce an equal amount of change in perceived colour. Both are similar but L*a*b* is slightly easier to convert. L*u*v* and L*a*b* are converted from RGB in two steps — first, they are converted to CIE-XYZ and then XYZ to the corresponding colour space [12]. CIE-XYZ can be represented by

\[
X(x, y) = 0.412 * G_r(x, y) + 0.357 * G_g(x, y) + 0.180 * G_b(x, y) \\
Y(x, y) = 0.212 * G_r(x, y) + 0.715 * G_g(x, y) + 0.072 * G_b(x, y) \\
Z(x, y) = 0.019 * G_r(x, y) + 0.119 * G_g(x, y) + 0.950 * G_b(x, y)
\]

\[
L* = \begin{cases} 
116(Y/Y_n)^{1/3} - 16 & \text{when } Y/Y_n > 0.008850 \\
903.3Y_n & \text{otherwise}
\end{cases} \\
A* = 500(f(X/X_n) - f(Y/Y_n)) \\
B* = 200(f(Y/Y_n) - f(Z/Z_n))
\]

where

\[
f(t) = \begin{cases} 
t^{1/3} & \text{for } t > 0.00856 \\
7.787t + 16/116 & \text{otherwise}
\end{cases}
\]
In the above formula $G_r, G_g, G_b \in [0, 1]$, $X_n, Y_n$ and $Z_n$ are the values of $X, Y$ and $Z$ for pure white colour represented by $G_{red}(x, y) = 255$, $G_{green}(x, y) = 255$, $G_{blue}(x, y) = 255$ and we get $X_n = 0.950456$ (from Formula 1), $Y_n = 1$ (from Formula 2) and $Z_n = 1.088754$ (from Formula 3).

Many researchers have found uniform colour spaces to be the most suitable scheme for colour histogram representation [10], [9].

\[
h_{colour} = \begin{bmatrix} h_L(0), h_L(1), \ldots, h_L(C_{\text{Nb} \text{it}} - 1), \\
h_a(0), h_a(1), \ldots, h_a(C_{\text{Nb} \text{it}} - 1), \\
h_b(0), h_b(1), \ldots, h_b(C_{\text{Nb} \text{it}} - 1) \end{bmatrix}
\]  

(7)

**B. Edge Gradients**

Edge Gradients histogram can be represented from the run length histograms of differential images. A run-length is the length of the repetition of a given colour value along a direction. Horizontal and vertical run lengths are the commonly used directions [11]. The combination of a colour and run length can generate a huge number of variations which are difficult to compute. Hence, colour of run length is often dropped. A run length histogram contains frequencies of run lengths irrespective of their colours. Prewitt’s edge detector combines uniform smoothing in one direction with edge detection in the perpendicular direction to produce [1], [3] and hence is more suitable for image ranking. Prewitt’s edge is given by

\[
\begin{bmatrix} -1 & 0 & 1 \\
-1 & 1 & -1 \\
-1 & 0 & 1 \end{bmatrix}
\]  

(8)

The run length histogram collected from the differential image provides edge gradients that provide significantly higher effectiveness than texture represented by ordinary run length histograms.

**C. Parameters for Features**

Image feature extraction from heterogeneous image collection requires consideration of several parameters. The parameters used for the extraction of features have been presented in Table I.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values for Colour Histogram</th>
<th>Values for Run Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>File format</td>
<td>all GIF</td>
<td>All JPEG</td>
</tr>
<tr>
<td>Contrast</td>
<td>7 bits</td>
<td>7 bits</td>
</tr>
<tr>
<td>Segmentation</td>
<td>Two vertical Strips</td>
<td>None</td>
</tr>
<tr>
<td>Cropping</td>
<td>1% from all sides</td>
<td>1% from all sides</td>
</tr>
</tbody>
</table>

**IV. FEATURE MATCHING**

Feature vectors from query and target documents are matches to assess similarity, which is sorted to decide the presentation order of the documents in the list. Conceptually, similarity is an opposite concept of difference and hence differences are computed from the feature vectors of the target and query objects. Several different distance functions are used to compute similarity such as Euclidean, Canberra and Manhattan distance functions.

Euclidean distance is the most simple to conceive, and we also use it for our experiment. Euclidean distance is given by

\[
d_{\text{Euclidean}} = \sqrt{\sum_{i=0}^{M-1} (t_i - q_i)^2}
\]  

(9)

**V. EXPERIMENT**

There is no well known and wide accepted initiative available to compare image ranking algorithms across different research groups [17], [2], [13], [16], [15]. We modified the protocol used in TREC to suit image ranking. A protocol has three configuration items (i) the document collection, (ii) an image library with different browsing facilities to formulate image query term (iii) subjects of diverse background who would provide: (a) information need statements (b) queries (c) acceptance criteria to make a document relevant and (d) the actual relevance judgment.

Identification and sourcing of a suitable query image depends on both the richness of the image library and image organization. Our experiment shows that a significant improvement in quality of image query is possible when a well organised image database is used.

Relevant parameters of the test collection have been presented in Table II.

We used nine volunteers with different age, background and query skills, who provided information needs, formulated queries and provided relevance judgments, identifying relevant documents by browsing all documents in the collection.

**A. Procedure**

We pre-process images with parameters provided in section III-C and take colour histograms. We extract edge gradient histogram in the form of run length histogram taken from images preprocessed using Prewitt’s operator in addition to recommended processing in Section III-C. We them assess similarity with the Euclidean distance function.

**B. Measurement of Performance**

We used the same test collection to compare algorithms (existing and proposed). We used two commonly used matrices for assessing the effectiveness of ranking algorithms: precision and recall.

Several different measurements of effectiveness can be proposed [18] but recall and precision are commonly used.
Fig. 1. Precision-Recall curves for image retrieval using two different query methods.

(a) Image Query from random browsing

(b) Image Query from an organized collection

VI. CONCLUSION

Multimedia documents are retrieved using ranking algorithms containing feature extraction, computing similarities. This type of retrievals are based on inexact matching. Ranking is suitable for medical research and satellite image processing.

A significant improvement of ranking precision is possible when organized image database is used to facilitate querying. From this results we see that the user interface and the organization of image query system is also configuration items which must be considered in comparing multimedia algorithms. We also found that the use of organized images for query help user selecting desired image term much quickly which is approximately one-tenth of random image selection from random browsing of picture.

ACKNOWLEDGMENT

The authors would like to thank volunteers who helped in query formulation and provided relevance judgement.

REFERENCES


