SMART IMAGE BROWSER: MULTI-FORM IMAGE BROWSING APPLICATION ON MOBILE PHONES

Esin Guldogan, Ugur Kirbac and Moncef Gabbouj

Department of Signal Processing
Tampere University of Technology, Tampere, Finland
esin.guldogan@tut.fi, moncef.gabbouj@tut.fi

INTRODUCTION

Recent technology improvements along with the Internet growth have led to huge amount of digital multimedia during the recent decades. Various methods, algorithms and systems have been proposed addressing multimedia storage and management problems. Such studies revealed the indexing and retrieval concepts, which have further evolved to Content-Based Multimedia Indexing and Retrieval (CBMIR). Despite various successful systems, there is no perfect global solution for CBMIR in general.

CBMIR systems often analyze multimedia content via so-called low-level features for indexing and retrieval, such as color, texture and shape. Recent systems intend to combine low and high-level features for achieving significantly higher semantic performance [2], [3], [4].

The use of low-level features does not 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. This challenge is called "semantic gap" between low-level feature vector representation and semantic concept of image content.

Available image sharing, editing, classifying systems have various forms of associated information for the images. For example, Flickr claims to host more than 4 billion images (April 2010) and approximately all of them are tagged by users. Only during April 2010, 2.6 million geo-tagged images were uploaded to Flickr. Associated information of the images should be considered into indexing, browsing and searching mechanisms in order to improve the semantic accuracy and efficiency. Images could be grouped hierarchically and/or sequentially based on the following information:

- Date
- Location
- Content
- Tags

INNOVATION

Smart Image Browser is a novel application which utilizes four different modalities to help the user search/browse for images based four different categories. While this application has, to some extent, been available four PC users, our innovation brings this technology to the hands of mobile phone users with a greater searching capacity. The novelty of the application also includes significantly higher retrieval accuracy than the other existing content based search technologies. Smart Image Browser is the first flexible and powerful tool that can be utilized for personal image collections for searching and browsing.

MULTI-FORM STRUCTURES

Large image collections need to be organized and visualized efficiently. Various methods, algorithms and systems have been proposed addressing image categorization and indexing problems where further studies revealed to Content-based image indexing, browsing and retrieval applications (CBIR). CBIR systems often analyze image content via the so-called low-level features such as color, texture and shape. Usually, such low-level descriptors cannot completely express the semantic concepts of the image from the user's perspective. In order to achieve significantly higher semantic accuracy, recent systems combines low-level features with additional metadata information associated with the images. Consequently, entire information associated with the images should be combined appropriately, in order to improve the semantic accuracy and efficiency for the classification.

Image classification has been proposed to categorize images into semantically meaningful classes in order to fill the semantic gap between low-level visual features and human perception. There are various supervised and unsupervised approaches based on low-level image features for image clustering, such as Self-Organization Maps (SOM) and hierarchical cellular trees. However, these classification schemes are entirely based on visual information. Recent systems associate the words to images by unsupervised annotation [1]. In such systems, query process is based on natural-language on automatically annotated image database. Major weakness of this approach is to have limited defined vocabularies. Image categorization results highly impact the accuracy of annotation, which will proportionally affect the retrieval accuracy. Tags and annotations help the system to refer the high-level concepts in the image. However, they do not represent and express comprehensively the information lying in the

image individually. Similarly, geo-tags can be used to categorize images by their location, which can be useful when the task is based on location information. We propose that image browsing and searching applications should categorize images based on all available information separately. It should be noted that one image may belong to more than one category. In this study, we assume that current images systems (online sharing systems, personal photo collections etc.) have four forms of information available that can be listed as follows:

- Location
- Content
- Date
- Tags and annotations

Internal Structures of the Forms

Date: Images are ordered sequentially according to their dates in the database. The normalized distances between images are calculated based on time.
***Location*:** Images are hierarchically indexed according to their geo-tags.
***Content*:** Supervised image classification method is utilized for categorization of the image databases. The method is developed by MUVIS team in TUT. For more information please visit the webpage (http://muvis.cs.tut.fi/).
***Tags*:** Image database is grouped according to their common tags.
If the user is initiating the query with image (Query by example) as represented in Figure 1, user might be interested in finding:

a) Semantically similar images that contains same content
b) Images taken at the same day
c) Images taken at the same or nearby locations
d) Images that are tagged similarly by other users

Consequently, if the user is interested in finding images taken at the same place, there is no need to retrieve images taken at different places.

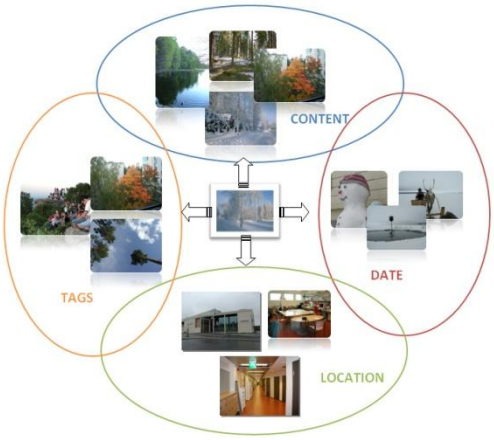

Figure 1: Example of multi-form representation of image browsing scheme

HOW MULTI-SEARCH WORKS

Single image tells more than hundred of words and often users just don't know what they search for. Every image has a story for its owner and every image reminds specific moments or emotions. While user surfs through in his/her own multimedia gallery, the search will be initiated by an image with which the user wants to see other associated images to the specific one based on a story that resembles. The user might be interested to find pictures taken in the same date or same place, or having same content or sharing common tags. Once the sets of images are formed as a path view, user can navigate through images by simply dragging them to the right or left. User may click other images in the groups for further querying. In that case, new images will be shown to the user associated with the query image.

MATERIALS

We have utilized public benchmark Flickr image database and Corel real-world image database for evaluating the accuracy of the proposed approaches. MIRFlickr [5] database contains 25,000 images with 223,500 tags, where the average number of tags per image is 8.94. The sample classes are: Sky, water, portrait, baby, animal etc. However, MIRFlickr database do not have available geo tags. For this reason, location and date information are added manually to the images for the experimental purposes.

DEVELOPMENT AND IMPLEMENTATION

The applicaton is developed by MUVIS team (http://muvis.cs.tut.fi/) in TUT. The MUVIS team is led by Academy Professor Prof. Moncef Gabbouj. The home Department of Signal Processing at TUT is an internationally well-known site of research in the fields of signal and multimedia processing (http://sp.cs.tut.fi/). The Signal Processing Algorithm Group has been selected as a Center of Excellence for the periods of 2000-2005 and 2006-2011 by the Academy of Finland.

COMMERCIAL POTENTIAL

Smart Image Browser has been developed with Qt-QML for Nokia devices and is currently running on Nokia N900. However, it could be easily integrated into Symbian S60 or MeeGo devices.

This work is supported by Devices and Interoperability Ecosystem - DIEM - project is part of the Finnish ICT SHOK program coordinated by TIVIT and funded by Finnish Funding Agency for Technology and Innovation.

REFERENCES

[1] 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(3), 2007, pp. 313-323.

[2] 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.

[3] A. W. M. Smeulders, M. Worring, S. Santini, *A.* Gupta, and. R. Jain, “Content-Based Image Retrieval At The End of The Early Years”, IEEE Transactions on Pattern Analysis and Machine. Intelligence, vol. 22*(*12*)*, 2000, pp.*1349–1380.*

[4] 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.

[5] M. J. Huiskes, M. S. Lew, ”The MIR Flickr Retrieval Evaluation”, ACM International Conference on Multimedia Information Retrieval (MIR'08),2008, Vancouver, Canada.