Combining Semantics and Visual Content for Museum Information Retrieval

Manoj Kumar Sharma, Tanveer J Siddiqui

Abstract: ***Searching useful information from huge and unstructured museum multimedia data has been a difficult problem in information retrieval. In this paper, we propose the design of an image retrieval system that combines two important modalities- text and image – along with ontological concepts for retrieving digitized museum artifacts. The combination of text and domain ontology facilitates retrieval based on semantics. The evaluation of the proposed method has been done on a dataset comprising of 1200 images of artifacts displayed in various galleries of Allahabad museum and their textual descriptions. A domain ontology has been manually constructed and used as a source of knowledge to aid in the retrieval and browsing. The organization of the artifacts in the museum has been used to arrange domain concepts in a hierarchy.***

Index Terms: Artefacts; Framework; Feature Extraction; Ontology; Query Processing.

I. INTRODUCTION

The rapid growth in multimedia technologies has dramatically increased the amount of the multimedia data available in digital format. In multimedia retrieval, images play a big role in search and communication process. The two dominant approaches for image retrieval are: Content-based Image Retrieval (CBIR) and keyword-based image retrieval. CBIR is criticized due to lack of semantics and keyword-based image retrieval is constrained due to the need of manual annotation. In many retrieval applications, the database contains both the images and text. Museum information retrieval is one such application. A museum has different types of artifacts organized in different categories that are of interest of general public as well as historian, scientists, researchers. The displays in a museum are accompanied with useful textual descriptions which can help in retrieval process. Instead of using only one modality we can combine both in an automatic setting to overcome the limitation of CBIR and keyword-based image retrieval systems. In this paper, propose a method that combines the visual and textual features with ontological concepts to support retrieval of museum artifacts.

The rest of the paper is organized as follows. The next section discusses related works in the area of multimedia retrieval. In section III, we present the details of retrieval algorithm. Section IV presents the dataset used in the experimental solving. Results are discussed in section V. Finally, we conclude in section VI.

II. RELATED WORK

A lot of work has been done in field of multimedia retrieval by many researchers. The existing image retrieval system can be broadly categorized in to keyword based and content-based approaches. The keyword-based retrieval systems require a human being to describe images stored in the database using keywords. Though this approach can be useful in terms of semantic but it takes a lot of time and is subjective in nature [2]. Content based retrieval system use low level feature such as color, texture, shape, etc. in the retrieval process [1] [6]. The retrieval is based on visual similarity. A detail survey of CBIR system can be found in [9]. CBIR systems suffer from the semantic gap problem between low level image feature and high-level semantic concept. An approach explained to assign relevant words with image regions using probability theory, approach is generative in nature [15][16]. A framework proposed for combining words and pictures for information retrieval [2]. Early examples of CBIR systems include photobook [5], Virage framework [7], IBM's Query By image or content (QBIC) system [13]. In Google, search images are uploaded by user to find similar images and the retrieval process uses features such as points, colors, lines, and textures in images [8].

Efforts have been made to overcome the disadvantages of CBIR system including the development of automatic annotation technique, to combine textual and visual feature and use of domain ontology [4].

III. THE PROPOSED METHODOLOGY

The proposed system consists of five modules:

- User Interface Module (UIM)
- Query Processing Module (QPM)
- Indexing Module (IM)
- Retrieval and Ranking Module (RRM)
- Presentation Module (PM)

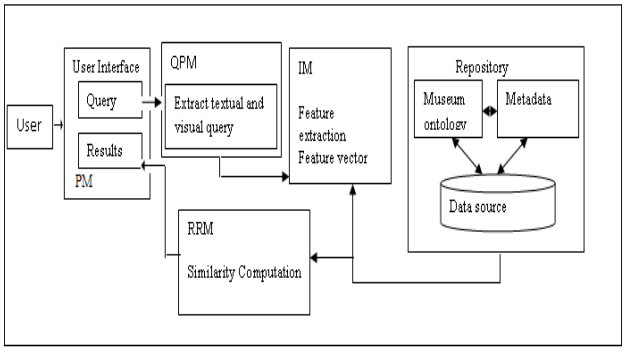
Fig. 1 shows these modules and their interaction. The UIM allows browsing or querying the database. The browsing mode offers a list of galleries to support navigation of the database. The query mode supports three modes of query: text only, image only and dual mode (combination of text and image). The QPM parses the query and prepares an appropriate query representation.

Revised Manuscript Received on December 22, 2018.

Manoj Kumar Sharma, Department of Electronics & Communication, University of Allahabad,

Tanveer J Siddiqui, Department of Electronics & Communication, University of Allahabad,

The query representation depends on the mode of the query and uses features extracted from text and/or image part of the query and features extracted from ontology. The IM module prepares image and text feature vectors of the database. RRM takes the query representation computes similarity between feature vectors of query and database images and prepares a ranked list of database images based on the similarity value. The PM module takes the ranked list of images from RRM and communicates the results back to the user.



A. *Ontology construction*

In this work, we have used domain ontology to annotate the images. The concepts in this ontology are derived from the categorization used in the museum to display various artifacts. These categories correspond to galleries. The use of ontology helps in overcoming the well-known semantic gap problem of CBIR systems. The ontology is constructed using protégé tool and MySQl has been used for data storage. Ontology identifies objects of class and their behavior. The super-class of domain ontology is artefact class from which all other classes are derived. There are no standard methodology to model a domain [14]. A method adapted from [14] for ontology development is shown in fig. 2. The steps used in this work are listed below:

- Firstly, we define the domain and scope of the ontology.
- Next, we identify the important terms in the ontology. The ontology used in this work is build from scratch so scope of reuse doesn't arises. We use the existing categorization and accompanying textual descriptions to identify these terms.
- The salient concepts are identified and classes corresponding to them are created. These classes are generalized and specialized and arranged in a concept hierarchy.
- After defining classes, we describe the internal structure of concepts to identify class property. The property of each class is associated as slots to them.
- Next, we define features of the slots. This includes the type of values it takes, allowed values and its cardinality. A slot may have instance as its value type in which case list of allowed classes from which instances will come are also defined.
- Finally, we create individual instances of classes in the hierarchy.

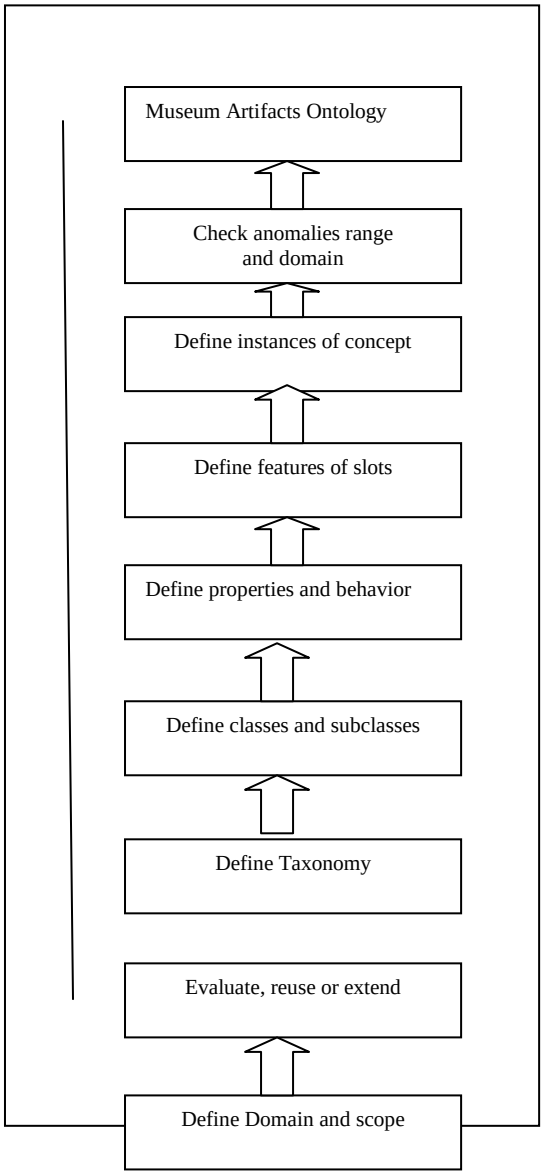


Fig. 2: Museum ontology construction

B. *Feature extraction*

The textual features used in this work are keywords extracted from textual description after stemming and stop word removal. A picture is interpreted by low level feature at computing level. These low-level features are color, texture, shape, spatial relationship, etc. We have used color and shape feature. The color feature is extracted using color histogram technique, which has been widely used in many applications [12].

The shape of an object is an important feature. Shape feature description and representation work are broadly categorized into edge-based and region-based technique. In this paper work region- based shape features are extracted. Shape provides numerical information of an image, which do not change even when the position, size and direction of objects are changed [3] [10].

IV. EXPERIMENTS AND RESULTS

A. Dataset

We have used a dataset of Museum artifacts for evaluating the proposed system. The dataset consists of twelve hundreds images collected from different galleries of the Allahabad museum. The accompanying textual descriptions of these images are also collected. Table 1 lists the number of images taken from each of the 12 galleries. The images are pre-processed to using Adobe Photoshop CS. The pre-processing step involves resizing, contrast adjusted and noise reduction. The pre-processed images and their descriptions are stored in database. We have used .jpg format for images.

Table 1: Allahabad Museum Gallery

Name of Gallery	No. of images
Sumitra nandan	35
Freedom fighters	70
Stone	62
Natural things	80
Jawahar art	190
Early Sculpture	51
Mahatma Gandhi	125
Ancient Jewellery	37
Terracotta	78
Medieval	73
Archaeological	56
Others	342

B. Experiments and Results

In order to evaluate our methodology, we performed retrieval experiments using three different query options: text only query, image only query, and query using both image and text. The keyword and visual features are mapped to domain concepts during retrieval. The evaluation is done in terms of precision and recall. Precision is the ratio of the images retrieved that are relevant to the user's information need. Recall is the ratio of the relevant images retrieved to total number of relevant images in the database. More formally,

$$Precision = \frac{A}{B} \quad (1)$$

$$Recall = \frac{A}{C} \quad (2)$$

Where, A is the number of relevant images retrieved, B is the total number of images retrieved and C is the total number of images relevant to a given query.

The retrieval results corresponding to (i) text only query "charkha", (ii) using the image of charkha and (iii) using both are shown in Fig. 3. Out of 5 relevant images, we are able to retrieve 3 images using only text and only image as a query. When both text and image is used 4 relevant images are retrieved. Table II shows the precision and recall averaged over queries using the three query modes.

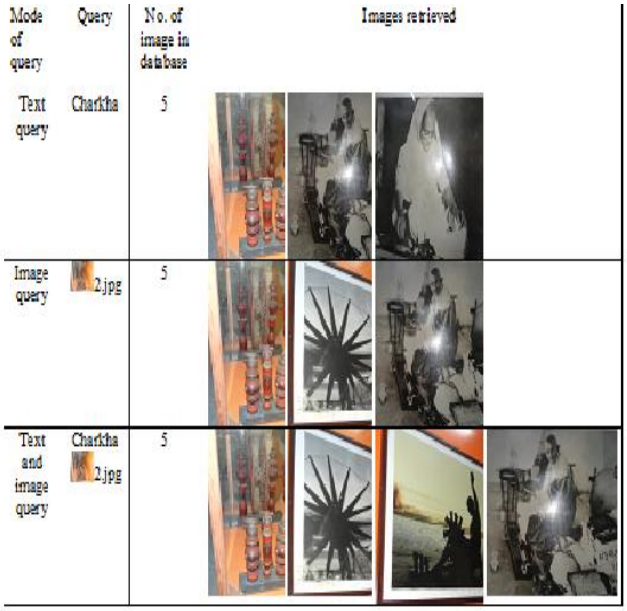


Fig. 3: Retrieval results of images from museum dataset

Table 2: Precision and Recall

Mode of query	No. of relevant image in database	Images retrieved	Relevant	Precision	Recall
Text query	5	3	2	0.66	0.4
Image query	5	3	2	0.66	0.4
Text and image query	5	4	4	1	0.8

The experimental results demonstrate the efficacy of the proposed method. By using both the text and image we are able to retrieve more relevant images compare to search based only on text or sample image. Numbers of experiments are performed on collected dataset, precision recalls are calculated for individual query and mean average precision of 15 query is .422249.

V. DISCUSSION AND CONCLUSION

In this paper, we proposed and evaluate an image retrieval system for digital museum. The proposed system combines visual and textual features to overcome the semantic gap problem. The textual features are mapped to domain concepts using Museum ontology and used in the retrieval process.

The evaluation is done on a database created from images and text captured from Allahabad museum. The experimental results demonstrate the effectiveness of using semantics in the retrieval process.

ACKNOWLEDGMENT

We are very grateful to the Director of Allahabad Museum for permitting us to capture these images to create a dataset. The authors gratefully acknowledge the support of the SERB, Govt. of India under Grant No SR/FTP/ETAY36/2011.

REFERENCES

1. A. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain," Content based Image retrieval at the end of the early years", IEEE, Vol.22, pp.1349, 2000.
2. T. J Siddiqui ,U.S. Tiwary, , "Words and Pictures", An HCI Perspective, Proceeding of the First International Conference on Intelligent Human Computer Interaction, pp. 59-70, 2009.
3. J. F. Canny , "A Computational Approach to Edge Detection", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol8, Nov 1986.
4. S. Sarwar, Z.UI. Qayyum, S. Majeed, "Ontology based image retrieval framework using qualitative semantics image descriptions", In 17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems(KES), Procedia Computer Science 22, 285-294, 2013.
5. A. Pentland, Picard WR, S. Sclaroff, "Photobook: Content based manipulation of image database", Int J Computing Vis 18, 233-54, 1996.
6. B .Holt, L. Hartwick, "Retrieving art images by image content: the UC Davis QBIC project," Proc first international conference on Electronic library and visual information system Research, de mohtfor University, Mitton keyners, pp.93-100, 1994.
7. R.J. Bach , C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, "Virage image search engine: an open framework for image management", In: Sethi IK,Jain RC Eds. Proceedings of SPIE 2670 (1), 76-87, 1996.
8. G. Chechik, V. Sharma, U. Shalit, Bengio US,"Large scale online learning of image similarity through ranking", J Mach Learn Res 11, 1109-35, 2010.
9. Y. Liu, D. Zhang, G. Lu, and W.Y. Ma, "A survey of content based image retrieval with high level semantics", Pattern recognition, vol. 40, pp. 262-282, 2007.
10. F. A. Mahdi and A. Ibadi, "MIRS: Museum Image Retrieval System using Most Appropriate Low-Level Feature Descriptors," International Journal of Computer Science, vol. 11, Issue 5, No.2, pp. 1694-0784, September 2014.
11. G. Chechik, V. Sharma, U. Shalit, Bengio US,"Large scale online learning of image similarity through ranking", J Mach Learn Res 11, 1109-35, 2010.
12. W.Y Ma, H. J. Zhang,"Content Based Image Indexing and Retrieval," In: Handbook of Multimedia Computing,(Ed.) Furht, B.crc Press, Boca Raton,1998.
13. M. Flickner, H Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafne, D. Lee, D. Petkovic, D. Steele and P. Yanker, "Query by image and video content the QBIC system", IEEE Computer, 23-32 1995.
14. Natalya F. Noy,Deborah L., McGuinness,"Ontology Development 101: A guide to creating your first ontology",Stanford University,Stanford,,CA,94305.
15. K. Barnard,D. Forsyth,"Learning the semantics of words and pictures",International conference on computer vision 2,University of California,Berkeley,2001.
16. K. Barnard, P. Duygulu, N. de Freitas, D.Forsyth, D, Blei , M. I. Jordan,"Matching words and pictures",Journal of Machine Learning Research,3:1107-1135,2003.