ISSN (Online) 2456-1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 1, Issue 1, May 2016

A review on bridging (Reducing) the SG (Semantic Gap) in CBIR and Annotation

[1]Tabassum Nahid Sultana, [2]Md Abdul Haleem Bahadur, [3]Dr. Ruksar Fatima,[4]Mohammed Shakeeb Ahmed
[1][2][3][4] Khaja Banda Nawaz College of Engineering, Gulbarga
[1]shakeebece@gmail.com, [2] er.abdulhaleem@gmail.com, [3]ruksarf@gmail.com

Abstract: **this paper has tried to identify the problems in content based image retrieval technique. The key obstacle of CBIR approaches is the SG i.e., the difference between the low level features to high level concept of human perception. To extend the image retrieval process beyond low level visual descriptors to high level image semantics. Therefore researches proposed different SBIR techniques to bridge the SG using various machine learning algorithms for the extraction of semantic images to increase the accuracy of the image retrieval process.**

Index Terms**— CBIR, SBIR, Semantic, SG**

I. INTRODUCTION

An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images [1][2]. Content base Image retrieval has been a very active research area since 1990, because of the emergence of large-scale image collections. CBIR is the computer vision technique to solve the problem of digital image search in large databases. CBIR is a tool which efficiently retrieves image visually similar to query image using low level features as shown in fig.1

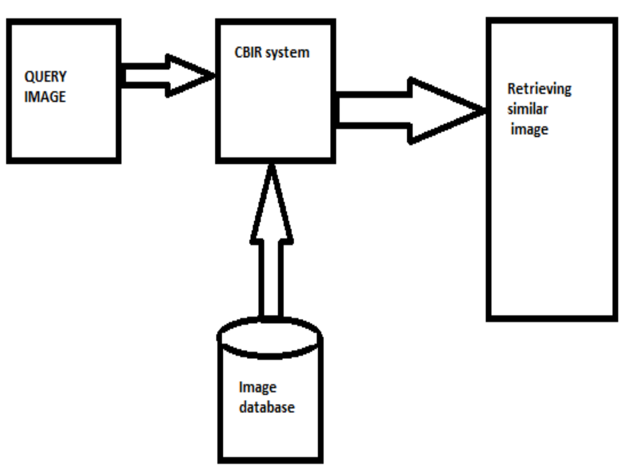

Fig.1. CBIR System

Feature extraction: CBIR extracts low level image features such as

Color: color is the feature of CBIR for retrieving color image. Due to its fast and easy computation helps in retrieving image from the database against the query image.
Texture: Texture is a property that represents the surface and structure of an image. The six visual texture properties were coarseness, contrast, directionality, line likeness, regularity and roughness.

Shape: shape can be defined as the description of an object regardless of its position, orientation and size. The shape feature will be region based and boundary based.
Low level features do not have a direct link to the high level concept and Low level visual features cannot fully capture high level concept of images. Besides, due to the performance of Image retrieval based on low level features are not satisfactory, there is a need for researchers to develop retrieval based on semantic meaning by trying to extract the cognitive concept of a human to map the low level image features to high level concept (semantic gap). In addition, representing image content with semantic terms allows users to access images through text query which is easier and initially preferred by the ultimate end user to express their mind compare with using images. For example, users' queries may be 'Find an image of sunset rather than 'find me an image contains orange and yellow colors'. General Framework of Semantic based Image Retrieval [3] is shown in Fig. 2.

International Journal of Science, Engineering and Management (IJSEM)
Vol 1, Issue 1, May 2016

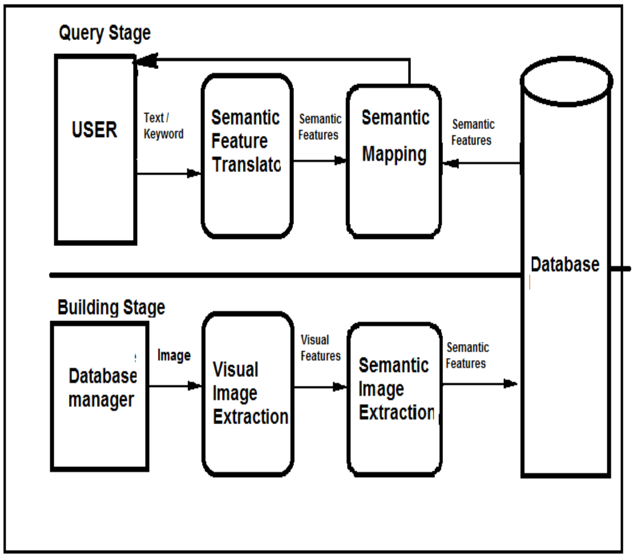

Fig.2. SBIR System

In sec.2 we describe the problem of CBIR is the semantic gap (SG).
In sec.3 the different image retrieval techniques are proposed to overcome SG. In sec 4 include concluding remarks.

II. BRIDGING THE SEMANTIC GAP

The user may sometimes be interested in high-level concepts that he associates with images. The system must be able to extract such concepts from the content of the image or from the text surrounding it. In recent years, some attempts have been made in this direction. The authors narrow the semantic gap by training the system with hundreds of statistical models each representing a concept.

In the context of CBIR, it has been noted that user judgment on the similarity between images is subjective and depends on many factors.

Semantic mapping: In order to bridge the feature to concept gap effective extraction and selection of low level features are needed for correlating with high level semantics description of image according to end user. New low level features should be learned which are accurate in representing the semantic meaning of concepts [4]. To overcome the SG, between high level image semantics and low level features of large data base of CBIR is needed[5].Image semantics can be obtain through the process shown in fig.2. Image stored in the database will go in image extraction process, there low level features of images will be achieved next, low level features will undergoes segmentation process based on similar characteristics of the visual descriptors to form region or object representation in images.

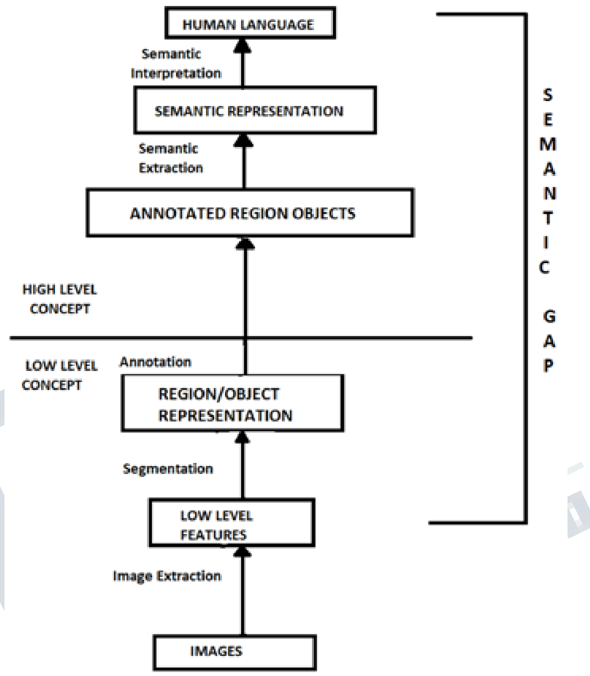

Fig.3

The region/object representation of images will be annotated i.e., labeled by text, keywords. The annotation will be done automatically by using automation concept annotation technique. The image then will be represented using high level semantics and image retrieving can be queried based on the human/user's level perception.

III. IMAGE RETRIEVAL TECHNIQUES

In this sec the problem CBIR is SG i.e., the semantic high level concept of human perception is discussed. Due to SG the accuracy performance of CBIR will be limited to increase the accuracy of the CBIR the different techniques of Semantic image retrieval are proposed.

In Content-based image retrieval, a database consisting of feature vectors of all images is created. Images in the database that are Nearest Neighbors to the query image according to similarity metric in the feature space are retrieved. To measure the similarity, a geometric based

ISSN (Online) 2456-1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 1, Issue 1, May 2016

approach is described; here a CBIR system is designed by selecting a suitable feature space so that images that are close in feature space are also perceptually close to the user. Content base retrieval image has to be conducted only in the visual feature space, but the performance is evaluated in the textual feature space. This motivates us to learn a new distance measure in the visual space to approximate the distance measure in the textual space. The learnt distance measure can be used in a content-based image retrieval system to retrieve more semantically relevant images. Distance metric learning is to learn a distance metric for the input space of data from a given collection of labeled points that preserves the distance relation among the training data. For a comprehensive survey please refer to [6]. In spite of many successful attempts, existing algorithms cannot be easily generalized to solve the problem of SG. Most of existing works learn distance metrics from well labeled training set, while we directly regard the surrounding text of Web images as their textual labels, which affirmatively have more noise. In fact, it is usually impossible to manually label or refine labels for millions of images. CBIR is essentially a ranking problem, rather than a classification problem as treated in traditional distance metric learning algorithms [6][7][8].Therefore Changhu wang et al, design a ranking-based distance metric learning algorithm is used in content-based image retrieval to retrieve more semantically relevant images.

A novel content-based image retrieval framework is introduced. First, to reflect the scalability of the proposed framework, millions of images together with rich textual information have been crawled from the Web for experiments. This image collection is indexed based on K-means-based indexing method [9] using visual features. Second, to alleviate the noises in the unbalanced labels of images and fully utilize the textual information, a *tf-idf* based term-level text model i.e. introduced to define pairwise semantic similarity between any two images and build semantic maps for each local model. Finally, to bridge the semantic gap using the constructed semantic maps, a ranking-based distance metric learning algorithm is proposed. In particular, different from traditional classification-based distance metric learning algorithms, to learn each local Mahalanobis distance metric, the algorithm directly minimizes the leave-one-out retrieval cost on the training set, based on a ranking-based probabilistic model. Existing distance metric learning algorithms cannot be directly generalized to solve the problem we are facing, due to the scalability problem and the noises in the rich textual labels. However, to evaluate the proposed ranking-based distance metric learning algorithm, we have developed a localized Neighbourhood Components Analysis (LNCA) algorithm, which is the same as the proposed algorithm with term-level model, except that the local transformation matrix A is learnt by the Neighbourhood Components Analysis (NCA) algorithm [11].

Automatic concept annotation: Annotation allows image search through the use of text, automated annotation will be more practical for large data sets of CBIR. Annotation is subset of concept detection i.e., same concept can be described in different way based on instance of the concept. Concept detection through supervised learning method has achieved high accuracy in CBIR. The problem of annotation is treated as translation from a set of image segments to a set of words is irrelevant to translation.

Relevance feedback: An overview of research development of RF is shown. Relevance feedback (RF) is an automatic process. In CBIR, researchers soon felt the need to integrate RF in order to overcome some difficulties. First, it is not always easy for the user to express his needs using an example-based query. This may be because none of the available images match what user is looking for, or merely because user cannot translate the query or what the user has in mind into a combination of example images. Second, the retrieval system may fail in translating the user’s needs into image features and similarity measures. It can be applied to identify the ideal query that is in the user’s mind. It also makes it possible to automatically capture the user’s needs in terms of image features and apply this information in assigning a degree of importance to each feature. This is done by enhancing the importance of those features that help in retrieving relevant images and reducing the importance of those that do not [7]. The results of RF technique for the end user’s are only a small number of labeled images pertaining to each high level concept. Methods for performing RF using the visual features as well as keywords(semantics) in unified framework have been reported in[12] [13].One problem with RF is that after every round of user interaction usually the top results w.r.t the query have to be recomputed using a modified similarity measure. A way to speed up this nearest neighbor search has been proposed in[14].Another issue is the user’s patience in supporting multi-round feedbacks. A way to reduce the user’s interaction is to incorporate logged feedback history into the current query.

Object Ontology: The use of ontology is to understand the relationship between desperate types of information in order to more accurately analyze and retrieve images. The ontology reasoning is based on the semantic association between keywords. This is achieved by finding which concepts in the ontology relate to a keyword and retrieving information about each of these concepts. By this module the ontology is used for quickly locating the relevant semantic concept and a set of images that are semantically related to the user query are returned.

IV. USER QUERY

Query mechanisms are helpful in the role of bridging the semantic gap between users and retrieval systems [15]. The user query try to explain the need of user's information according to human perception. The quality of queries submitted to information retrieval (IR) systems directly affects the quality of search results generated [16]

A. *Query by Visual Example*

Querying by visual example [17,18,19] is a program that refers to human perceptual features w.r.t low level features of images.

B. *Query by Texts*

User usually prefers using keywords to indicate what they want. [20,21]. The textual words need to be translated automatically to semantic meaning and representation that are matched in the images semantic representation in database in order to have fully and precisely understand the user input.

V. CONCLUDING REMARKS

We present a strong survey on reducing the SG by describing the detail description of image retrieval techniques, by narrowing the SG between high level concept and low level features, the accuracy of the image retrieval can be increased. Image with high level semantics can be achieved in future.

REFERENCES

[1] R. Priyatharshini , S. Chitrakala,"Association based Image retrieval: A survey,"Springer-Verlag Berlin Heidelberge, pp. 17-26, 2013

[2] Vaishali D. Dhale , A. R. Mahajan, Uma Thakur, "A Survey of Feature Extraction Methods for Image Retrieval," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 10, October 2012 ISSN: 2277 128X.

[3] Alaa M. Riad, Hamdy.K. Elminir, SamehAbd-Elghany, "A Literature Review of Image Retrieval based on Semantic Concept," International Journal of Computer Applications (0975 – 8887) Volume 40– No.11, February 2012.

[4] Chuan yu chang, hung jen wang, chi fang li. "Semantic analysis of real world images using support vector machine".Expert systems with application : An interntation journal 2009.ACM

[5] Wei Jiang, Kap luk chan, Mingjing Li. „Mapping low level features to high level semantic concept in region based image retrieval". IEEE CVPR05. 2005

[6] Yang, L. and Jin, R. (2006). Distance metric learning: A comprehensive survey.
[7] Goldberger, J., Roweis, S., Hinton, G., and Salakhutdinov, R. Neighbourhood components analysis. In NIPS, 2005.

[8] Torresani, L. and Lee, K. Large margin component analysis. In NIPS, 2007.

[9] Chad Carson and Virginia E. Ogle. "Storage and retrieval of feature data for a very large online image collection". IEEE Computer Society Bulletin of the Technical Committee on Data Engineering, 19(4):19–27, December 1996.

[10] Blei, D. M., Ng, A. Y., and Jordan, M. I. 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3 (Mar. 2003).

[11] Goldberger, J., Roweis, S., Hinton, G., and Salakhutdinov, R.Neighbourhood components analysis. In NIPS, 2005

[12] Y. Lu, C. Hu, X. Zhu, H.J. Zhang, and Q. Yang, "A Unified Framework for Semantics and Feature Based Relevance Feedback in Image Retrieval Systems," Proc. ACM Multimedia, 2000

[13] X. S. Zhou and T. S. Huang, "Unifying Keywords and Visual Contents in Image Retrieval," IEEE Multimedia, 9(2):23–33, 2002.

[14] P. Wu and B. S. Manjunath, "Adaptive Nearest Neighbor Search for Relevance Feedback in Large Image Databases," Proc. CM Multimedia, 2001

[15] V. Gudivada and V. Raghavan. "Content- based image retrieval systems". IEEE Computer, 28(9):18–22, 1995.

[16] Croft, W. B., & Thompson, R. H.. "I3R: A new approach to the design of document retrieval systems". Journal of the

American Society for Information Science, 38(6), 389–404. 1987.

[17] I. Cox, M. Minka, T. Minka, T. Papathomas, and P. Yianilos. “The bayesian image retrieval system”, PicHunter. IEEE Transactions on Image Processing, 9(1):20–37, 2000.

[18] P. Kelly, T. Cannon, and D. Hush. “Query by image example:The comparison algorithm for navigating digital image databases(CANDID) approach”. In Storage and Retrieval for Imageand Video Databases (SPIE), pages 238–248, 1995.

[19] K. Tieu and P. Viola. “Boosting image retrieval”. In Proc. International Conference on Computer Vision, 2000.

[20] M. L. Kher_, D. Ziou, and A. Benardi. “Image retrieval from the World Wide Web: Issues, techniques and systems”. ACM Computing Surveys, 36(1):35{67, 2004. 1993

[21] Hong Fu. “Attention driven image interpretation, annotation and retrieval”. PhD thesis. The Hong Kong Polytechnic University

[22] J. Li and J. Z.Wang, “Automatic linguistic indexing of pictures by a statistical modeling approach,” IEEE Trans. Pattern Anal. Machine Intell., vol. 25, no. 9, pp. 1075–1088, Sep. 2003.

[23] M. L. Kherfi, D. Ziou, and A. Bernardi, “Image retrieval from the World Wide Web: Issnums, techniques, and systems,” ACM Comput. Surv., vol. 36, no. 1, 2004.

[24] M. L. Kherfi, D. Ziou, and A. Bernardi, “Combining positive and negative examples in relevance feedback for content-based image retrieval,” J. Vis. Commun. Image Represent., vol. 14, no. 4, pp. 428–457, 2003.

[25] V. Mezaris, I. Kompatsiaris, and M. G. Strintzis, “An Ontology Approach to Object-Based Image Retrieval”, in Proc. IEEE ICIP, 2003.

[26] O. Murdoch, L. Coyle and S. Dobson, “Ontology-Based query Recommendation as A Support to Image Retrieval”. In Proceedings of the 19th Irish Conference in Artificial Intelligence and Cognitive Science. Cork, IE. 2008.

[27] H. Wang, S. Liu, L.T. Chia, “Does Ontology Help in Image Retrieval ?: A Comparison Between Keyword, Text Ontology and Multi-Modality Ontology Approaches”, Proceedings of the 14th annual ACM International Conference on Multimedia, 2006, pp. 23-27.