ISSN:0975-9646

K. Vijay Kumar et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (5) , 2012,5231-5235

Content Based Image Retrieval System Consume Semantic Gap

K. Vijay Kumar[1] ,Dr.R.Rajeswara Rao[2] ,.V.Subba Ramaiah[3] ,.Jhansi Rani **kaka[4]**

[1]CSE, [4]ECE
Vivekananda Institute of Technology and Science SET,
Karimnagar, AP, India
[2,3]CSE, Mahatma Gandhi Institute of Technology ,
AP, India

Abstract-In order to improve the retrieval accuracy of content-based image retrieval systems, research focus has been shifted from designing sophisticated low-level feature extraction algorithms to reducing the 'semantic gap' between the visual features and the richness of human semantics. This paper attempts to provide a comprehensive survey of the recent technical achievements in high-level semantic-based image retrieval. Major recent publications are included in this survey covering different aspects of the research in this area, including low-level image feature extraction,similarity measurement, and deriving high-level semantic features. We identify five major categories of the state-of-the-art techniques in narrowing down the 'semantic gap': (1) using object ontology to define high-level concepts; (2) using machine learning methods to associate low-level features with query concepts; In addition, some other related issues such as image test bed and retrieval performance evaluation are also discussed. Finally, based on existing technology and the demand from real-world applications, a few promising future research directions are suggested.

Keywords: Content-based image retrieval, Semantic gap,Image segmentation, Gabor Filters

1.INTRODUCTION

With the development of the Internet, and the availability of image capturing devices such as digital cameras, image scanners, the size of digital image collection is increasing rapidly. Efficient image searching, browsing and retrieval tools are required by users from various domains, including remote sensing, fashion, publishing, medicine, architecture, etc. For this purpose, many general purpose image retrieval systems have been developed. There are two frameworks: text-based and content-based. The text-based approach can be tracked back to 1970s. In such systems, the images are manually annotated by text descriptors, which are then used by a database management system(DBMS) to perform image retrieval.

There are two disadvantages with this approach. The first is that a considerable level of human labour is required for manual annotation.The second is the annotation inaccuracy due to the subjectivity of human perception [1,2].To overcome the above disadvantages in text-based retrieval system, content-based image retrieval (CBIR) was introduced in the early 1980s. In CBIR,images are indexed by their visual content, such as color,texture, shapes.

The pictorial database consists of picture objects and picture relations. To construct picture indexes, abstraction operations are formulated to perform picture object clustering and classification. In the past decade, a few commercial products and experimental prototype systems have been developed,such as QBIC [4], Photobook [5], Virage [6], VisualSEEK[7], Netra [8], SIMPLIcity [9]. Comprehensive surveys in CBIR can be found in Refs. [10,11].

1.1. The semantic gap

The fundamental difference between content-based and text-based retrieval systems is that the human interaction is an indispensable part of the latter system. Humans tend to use high-level features (concepts), such as keywords, text descriptors, to interpret images and measure their similarity.While the features automatically extracted using computer vision techniques are mostly low-level features (color, texture, shape, spatial layout, etc.). In general, there is no direct link between the high-level concepts and the low-level features [2].

Though many sophisticated algorithms have been designed to describe color, shape, and texture features, these algorithms cannot adequately model image semantics and have many limitations when dealing with broad content image databases [12]. Extensive experiments on CBIR systems show that low-level contents often fail to describe the high level semantic concepts in user's mind [13]. Therefore, the performance of CBIR is still far from user's expectations.In Ref. [1], Eakins mentioned three levels of queries in CBIR.

Level 1: Retrieval by primitive features such as color, texture, shape or the spatial location of image elements. Typical query is query by example, 'find pictures like this'.

Level 2: Retrieval of objects of given type identified by derived features, with some degree of logical inference. For example, 'find a picture of a flower'.

Level 3: Retrieval by abstract attributes, involving a significant amount of high-level reasoning about the purpose of the objects or scenes depicted. This includes retrieval of named events, of pictures with emotional or religious significance, etc. Query example, 'find pictures of a joyful crowd'.

Levels 2 and 3 together are referred to as semantic image retrieval, and the gap between Levels 1 and 2 as the semantic gap [1]. More specifically, the discrepancy between the limited descriptive power of low-level image features and the richness of user semantics, is referred to as the 'semantic gap' [14,15].

Users in Level 1 retrieval are usually required to submit an example image or sketch as query. But what if the user does not have an example image at hand? Semantic image retrieval is more convenient for users as it supports query by keywords or by texture.

Therefore, to support query by high-level concepts, a CBIR systems should provide full support in bridging the 'semantic gap' between numerical image features and the richness of human semantics [13,15].

1.2. High-level semantic-based image retrieval

How do we relate low-level image features to high-level semantics? Our survey shows that the state-of-the-art techniques in reducing the 'semantic gap' categories: (1) using object ontology to define high-level concepts, (2) using machine learning tools to associate low-level features with query concepts, (3) introducing relevance feedback RF) into retrieval loop for continuous learning of users' intention, (4) generating semantic template (ST) to support high-level image retrieval, (5) making use of both the visual content of images and the textual information obtained from the Web for WWW (the Web) image retrieval.

Retrieval at Level 3 is difficult and less common. Possible Level 3 retrieval can be found in domain specific areas such as art museums or newspaper library. Current systems mostly perform retrieval at Level 2. There are three fundamental components in these systems: (1) low-level image feature extraction, (2) similarity measure, (3) 'semantic gap' reduction.

Relevance Feedback: In the process of searching for an image, a concept called Query by Example (QBE) is often employed in which the user will be able to identify which images are relevant and which ones are not. By taking into account of a user's feedback, it is possible to be more precise in the search of relevant images

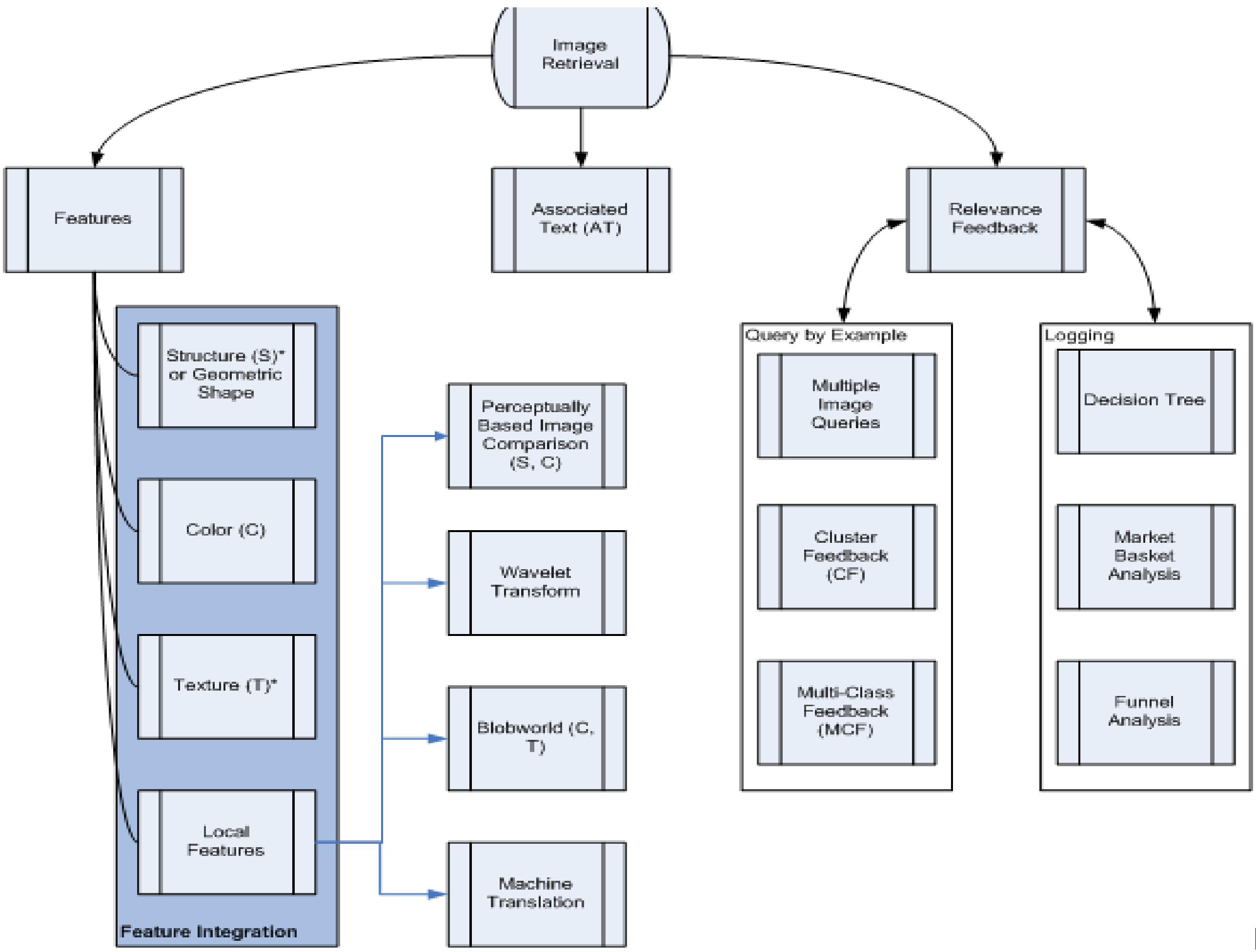

Fig 1:content based image retrieval

2.FEATURES

As noted in the above section, there are many different features associated to an image:

- Color
- Texture
- Shape

This section will provide a high level over view of these features.

2.1.Color

One of the primary components of image analysis for the purpose of content-based image retrieval is that of color analysis. As you may recall, color that is visible to the human eye represents a small range of the entire electromagnetic spectrum that represents everything from cosmic rays to x-rays to electric waves.

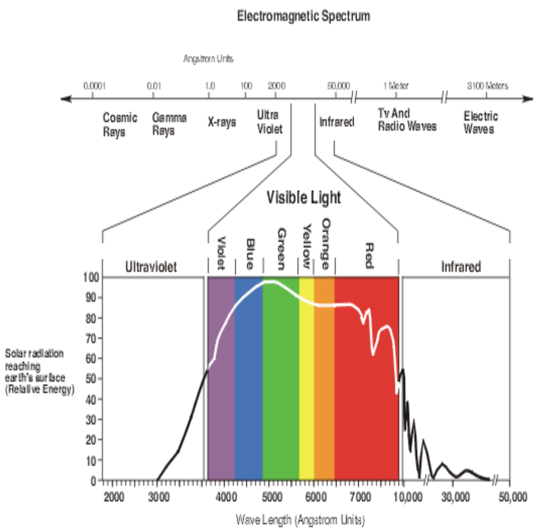

Figure.2: The Electromagentic Spectrum

As noted above, the color visible to the human eye range in wavelength from 4000 to 7000 angstroms respectively representing the colors violet and red and all of the colors in between. All other waves ranging from cosmic rays from the stars to the FM waves to our radios cannot be perceived by the human eye. It is this small range of the spectrum that is referred as human perceived color space.

2.1.1.Hue, Saturation, Value (HSV) Model

The HSV model represents color in its distinct components of hue, saturation, and value. To understand this model, we will first explore its components.

- **Hue**

The primary colors are identified as the primary set of colors that when combined together can create all of the other colors within the visible human spectrum. Similar to that of a computer monitor, the primary colors are that of red, green, and blue. Equal mixing of these colors produce what is known as the secondary colors of cyan, magenta, and yellow.

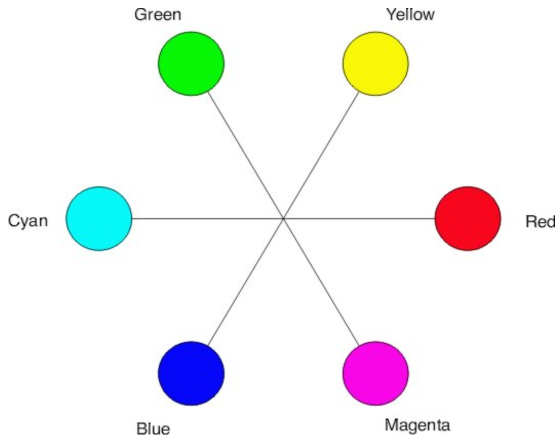

Figure.3 : Primary and Secondary Color Wheel

If we were to represent the primary and secondary colors within a color wheel, you will note that the secondary colors complement the primary colors. For example, the primary colors of red and blue mixed evenly will produce magenta, blue and green create cyan, and red and green create yellow. This process of inter-mixing colors will produce tertiary, quandary, eventually producing a solid ring of colors. This definition of color based on the combination of primary colors is also known as hue; note the color wheels above and below.

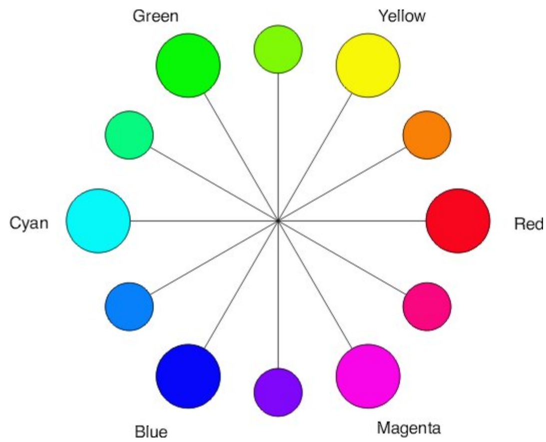

Figure.4: Primary, Secondary, and Tertiary Color Wheel

- **Color Representation**

The above descriptions describe the models used to quantify colors. For example, for the CIE-RGB model (the color model for computer monitors per the CIE color model), some numeric value is noted for each color component: R – Red, G – Green, B – Blue such as R:60, G:10, B:20. The same can be said for the HSV model in which numeric values are assigned to individual colors for the hue, saturation, and value.

Noting this, an image is composed of many pixels – many small segments that when put together piece together the image (i.e. think a puzzle except with many small square pieces instead of all of the weird pieces). It is then necessary to find a way to represent the numeric representation of color for the thousands of pixels that make up the image.

2.2.Texture

Another key component of image analysis is the analysis of the texture of an image – i.e. the perception of smoothness or coarseness of an object. Similar to the color histogram above, many of the current techniques for image texture analysis while quantified, lack the spatial information allowing one to compare the location of a coarse object within an image vs. a smooth object

- **Gabor Filters**

Similar to a Fourier transform, Gabor functions when applied to images convert image texture components into graphs similar to the ones below. There are many widely-used approaches to the usage of Gabor filters for text image characterization.

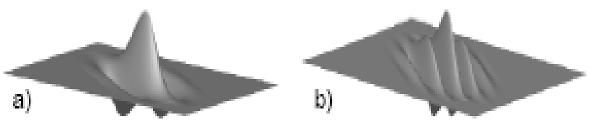

Figure.5: Gabor Filter Representation of Image Texture

The careful manipulation of these Gabor filters will allow one to quantify the coarseness or smoothness of an image. For example, within the above figure b) could indicate a more coarse texture than that of what was found in a). Note, the comparison of these images are performed against the mathematical representation of these graphs hence the CBIRS' ability to compare the textures of two different images.

2.3.Shape

Used in many CBIRS, shape features are usually described after the images have already been segmented or broken out [16]. While a good shape representation of an image should be handle changes in translation, rotation, and/or scaling; this is rather difficult to achieve. The primary difficulty is that images involve numerous geometric shapes that when numerically characterized will typically lose information.

A methodology that identifies information at too detail a level (down the individual colors and shapes of a Degas painting for example) will only be able to identify the color palette. For example, the above image has very few identifiable shapes that allow one to know what the entire image encompasses. But this shape found within the entire painting (as noted by the rectangle in the image below) will allow one to see the entire image of the ballerina dancing.

2.3.1. Image segmentation

Image segmentation is a difficult task. A variety Of techniques have been proposed in the past, Many existing segmentation techniques work well for images that contain only homogeneous color regions, such as direct clustering methods in color space [20]. These apply to retrieval systems working only with colors [21,22].

However, natural scenes are rich in both color and texture, and a wide range of natural images can be considered as a mosaic of regions with different colors and textures. Texture is an important feature in defining high-level concepts. As stated in Ref. [23], texture is the main difficulty in a segmentation method. Many texture segmentation algorithms require the estimation of texture model parameterswhich is a very difficult task [23].'JSEG'segmentation

overcomes these problems. Instead of trying to estimate a specific model for texture region, it tests for the homogeneity of a given color-texture pattern. 'JSEG' consists of two steps. In the first step, image colors are quantized to several classes. Replacing the image pixels by their corresponding color class labels, we can obtain a class-map of the image. Spatial segmentation is then performed on this class-map which can be viewed as a special type of texture composition. The algorithm produces homogeneous color-texture regions and is used in many systems [16,24]

3. Reducing the 'semantic gap'

The state-of-the-art techniques in reducing the semantic gap can be classified in different ways from different point of view. For example, by considering the application domain, they can be classified as those targeting at artwork retrieval [21], scenery image retrieval [27],WWW images retrieval [12], etc. In this paper, we focus on the techniques used to derive high-level semantics and identify five categories as follows. (1) Using object ontology to define high-level concepts [21,13]. (2) Using supervised or unsupervised learning methods to associate low-level features with query concepts [2,24,27,28,]. (3) Introducing RF into retrieval loop for continuous learning of users'intention [16,11]. (4) Generating ST to support high-level image retrieval [29]. (5) Making use of both the textual

4. Object-ontology

In some cases, semantics can be easily derived from our daily language. For example, sky can be described as 'upper, uniform, and blue region'. In systems using such simple semantics, firstly, different intervals are defined for the low-level image features, with each interval corresponding to an intermediate-level descriptor of images, for example, 'light green, medium green, dark green'. These descriptors forma simple vocabulary, the so-called 'object-ontology' which provides a qualitative definition of high-level query concepts. Database images can be classified into different categories by mapping such descriptors to high-level semantics (keywords) based on our knowledge [28,17], for example, 'sky' can be defined as region of 'light blue'(color),'uniform' (texture), and 'upper' (spatial location).

5.Conclusion:

Research in content-based image retrieval (CBIR) in the past has been focused on image processing, low-level feature extraction, etc. Extensive experiments on CBIR systems demonstrate that low-level image features cannot always describe high-level semantic concepts in the users' mind. It is believed that CBIR systems should provide maximum support in bridging the ‘semantic gap’between low-level visual features and the richness of human semantics.

This paper provides a comprehensive survey of recent work towards narrowing down the ‘semantic gap’. We have identified five major categories of state-of-the-art techniques: (1) using object ontology to define high-level Focusing on the differences between CBIR with high-level semantics and traditional systems with low-level features, this paper also provides useful insights into how to obtain salient low-level features to facilitate ‘semantic gap’ reduction. In addition, current techniques in image similarity measure are described

REFERENCES:

[1] J. Eakins, M. Graham, Content-based image retrieval, Technical Report, University of Northumbria at Newcastle, 1999.

[2] I.K. Sethi, I.L. Coman, Mining association rules between low-level image features and high-level concepts, Proceedings of the SPIE Data Mining and Knowledge Discovery, vol. III, 2001, pp. 279–290.

[3] S.K. Chang, S.H. Liu, Picture indexing and abstraction techniques for pictorial databases, IEEE Trans. Pattern Anal. Mach. Intell. 6 (4) (1984) 475–483.

[4] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D.Petkovic, W. Equitz, Efficient and effective querying by image content, J. Intell. Inf. Syst. 3 (3–4) (1994) 231–262.

[5] A. Pentland, R.W. Picard, S. Scaroff, Photobook: content-based manipulation for image databases, Int. J. Comput. Vision 18 (3)

[6] A. Gupta, R. Jain, Visual information retrieval, Commun. ACM 40 (5) (1997) 70–79.

[7] J.R. Smith, S.F. Chang, VisualSeek: a fully automatic content-based query system, Proceedings of the Fourth ACM International Conference on Multimedia, 1996, pp. 87–98.

[8] W.Y. Ma, B. Manjunath, Netra: a toolbox for navigating large image databases, Proceedings of the IEEE International Conference on Image Processing, 1997, pp. 568–571.

[9] J.Z. Wang, J. Li, G. Wiederhold, SIMPLIcity: semantics-sensitive integrated matching for picture libraries, IEEE Trans. Pattern Anal.Mach. Intell. 23 (9) (2001) 947–963.

[10] F. Long, H.J. Zhang, D.D. Feng, Fundamentals of content-based image retrieval, in: D. Feng (Ed.), Multimedia Information Retrieval and Management, Springer, Berlin, 2003.

[11] Y. Rui, T.S. Huang, S.-F. Chang, Image retrieval: current techniques,promising directions, and open issues, J. Visual Commun. Image Representation 10 (4) (1999) 39–62.

[12] A. Mojsilovic, B. Rogowitz, Capturing image semantics with low-level descriptors, Proceedings of the ICIP, September 2001,pp. 18–21.

[13] X.S. Zhou, T.S. Huang, CBIR: from low-level features to high-level semantics, Proceedings of the SPIE, Image and Video Communication and Processing, San Jose, CA, vol. 3974, January 2000, pp. 426–431.

[14] Y. Chen, J.Z. Wang, R.Krovetz, An unsupervised learning approach to content-based image retrieval, IEEE Proceedings of the International Symposium on Signal Processing and its Applications, July 2003, pp. 197–200.

[15] A.W.M. Smeulders, M. Worring, A. Gupta, R. Jain, Content-based image retrieval at the end of the early years, IEEE Trans. Pattern Anal. Mach. Intell. 22 (12) (2000) 1349–1380.

[16] F. Jing, M. Li, L. Zhang, H.-J. Zhang, B. Zhang, Learning in region-based image retrieval, Proceedings of the International Conference on Image and Video Retrieval (CIVR2003), 2003, pp. 206–215.

[17] H. Feng, D.A. Castanon, W.C. Karl, A curve evolution approach for image segmentation using adaptive flows, Proceedings of the International Conference on Computer Vision (ICCV’01), 2001, pp. 494–499.

[18] W.Y. Ma, B.S. Majunath, Edge flow: a framework of boundary detection and image segmentation, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1997, pp. 744–749.

[19] J. Shi, J. Malik, Normalized cuts and image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) 22 (8) (2000) 888–905.

[20] D. Comaniciu, P. Meer, Robust analysis of feature spaces: color image segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 750–755.

[21] P.L. Stanchev, D. Green Jr., B. Dimitrov, High level color similarity retrieval, Int. J. Inf. Theories Appl. 10 (3) (2003) 363–369.

[22] K.A. Hua, K. Vu, J.-H. Oh, SamMatch: a flexible and efficient sampling-based image retrieval technique for large image databases, Proceedings of the Seventh ACM International Multimedia Conference (ACM Multimedia’99), November 1999, pp. 225–234.

[23] Y. Deng, B.S. Manjunath, Unsupervised segmentation of color-texture regions in images and video, IEEE Trans. Pattern Anal. Mach. Learn. (PAMI) 23 (8) (2001) 800–810.

[24] H. Feng, T.-S. Chua, A boostrapping approach to annotating large image collection, Workshop on Multimedia Information Retrieval in ACM Multimedia, November 2003, pp. 55–62.

[25] Y. Liu, D.S. Zhang, G. Lu, W.-Y. Ma, Region-based image retrieval with perceptual colors, Proceedings of the Pacific-Rim Multimedia Conference (PCM), December 2004, pp. 931–938.

[26] C. Carson, S. Belongie, H. Greenspan, J. Malik, Blobworld: image segmentation using expectation-maximization and its application to image querying, IEEE Trans. Pattern Anal. Mach. Intell. 8 (8) (2002)1026–1038.

[27] R. Shi, H. Feng, T.-S. Chua, C.-H. Lee, An adaptive image content representation and segmentation approach to automatic image

[28] C.P. Town, D. Sinclair, Content-based image retrieval using semantic visual categories, Society for Manufacturing Engineers, Technical Report MV01-211, 2001.

[29] J.R. Smith, C.-S. Li, Decoding image semantics using composite region templates,