ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

# Content Based Image Retrieval: A Survey on the Start of Art

**Duraisamy Yuvaraj *[1],K.R.Dinesh[2], Shanmugasundaram Hariharan[3]**

[*1]Associate Professor,Dept. of Information Technology M.I.E.T Engneering college, Trichy ,India,
[2] Student,Dept of Information Technology ,J.J.College of Engneering college,Trichy,India,
[3] Associate Professor ,Dept of Computer Science ,T.R.P Engneering college,Trichy,India.
yuva.r.d@gmail.com ,dineshkvr@gmail.com, mailtos.hariharan@gmail.com

## Abstract

Evolution of Internet has lead to the rapid increase in tremendous amount of information in the form of text or other multimedia. It is quite obvious, due to such availability, storing, browsing, searching and retrieval has become a great challenge. This paper addresses on the challenges and issues in the content based image retrieval (CBIR). Though there exist various approaches and methods for content based image retrieval proposed by researchers till date, the challenge on retrieving the desired image based on the query remains to be a challenge. We address the state of art on the several features, techniques that would result in a research prototype for designing a commercial CBIR system.

## Introduction

Due to evolution of Internet there is an expanded information search on a day to day life. Such evolution has brought a major shift among researchers to analyze the information available online in different formats namely text & images [14], relevant features [15], patent [3] and others [1, 2]. There exists several framework models [4, 10, 11] or novel algorithms on CBIR [7, 9].Traditionally text based retrieval systems have been more successful in the past and at the present [12]. However there is a huge gap between the other formats represented as above. Especially image database management combined with text has gained reasonable attention in many domains [6, 8] which has led to the analysis of such a detailed survey in this paper. Any CBIR systems attempts to retrieve an image based on the query fed by the user as single or joint fashion in a dynamic environment [5,13], user navigation patterns [16].

The above steps involves lower most processing which involves analysis of complex features like color , shape , texture and other multi modal features. Content based image retrieval (CBIR) aims at reduce textual description to provide appropriate images automatically based on the match between the query image in a computationally faster manner. In this paper, we provide a survey and propose a framework which would improve the retrieval accuracy and efficiency. A sample CBIR system is presented in fig 1. Image retrieval annotates images by text and then use text-based database management system to perform image retrieval.The basic steps involved in CBIR system is given below

1. Searching of image based on input
2. Retrieval of images from database collection
3. Display of retrieved images.

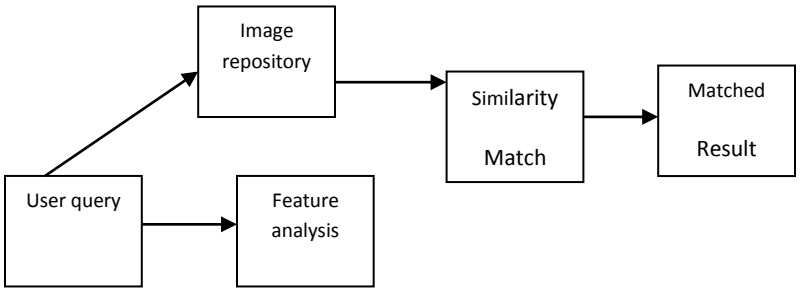


**Figure 1. Sample CBIR System**

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

There are enormous works being carried out on CBIR systems. The broad categories of the research being carried out on CBIR systems using the mechanisms listed below:

- Relevance feedback
- Color and texture based features
- Support Vector Machine (SVM)
- Ranking mechanisms
- Query expansion
- Similarity measure

Though there have been several investigations on CBIR, investigations on complex issues requires combined approach and on applying it several domains. We attempt to highlight on some of the areas where in attention is required. The rest of the paper is organized as follows. Section 1 presented some of the basics on content based image retrieval and techniques & application areas. Section 2 presents several segments that would enhance the performance of CBIR systems. Finally section 3 presents some considerations about the current trends and future directions on the subject investigated.

## Related Works

This section presents a survey on the work carried out on six broad categories of the work carried out on CBIR.

Hossein Nezamabadi pour and Ehsanollah Kabir[17] have viewed relevance feedback mechanism as a powerful technique for content-based image retrieval. They have used many parameter estimation approaches, most of them utilizing information of the relevant retrieved images or have not made great use of information of the irrelevant retrieved images. The work presents by them adopts a novel approach to update the inter-weights of integrated probability function by using the information of both relevant and irrelevant retrieved images. The results were effective and robust especially in the situation where there is no relevant retrieved images.

Chuen-Horng Lin et al [18] adopted three types of image features to describe the color and spatial distributions of an image. In these features, the K-means algorithm is used as a classifier for clustering images based on their colors. By measuring the spatial distance among the pixels in a same cluster, three types of color spatial distribution (CSD) features of the image is obtained and correspondingly appropriate retrieval methods are also provided. A fast filter is used to eliminate undesired images in advance and genetic algorithm is used to decide the most suitable parameters. Investigations carried out shows that the proposed image retrieval methods are simple and leads to provide impressive results.

Relevance feedback (RF) schemes based on support vector machines (SVMs) have been widely used in content-based image retrieval (CBIR). However, the performance of SVM-based RF approaches is often poor when the number of labeled feedback samples is small. This is mainly due to 1) the SVM classifier being unstable for small-size training sets because its optimal hyper plane is too sensitive to the training examples; and 2) the kernel method being ineffective because the feature dimension is much greater than the size of the training samples. Hence Jing Li at al[19] developed a new machine learning technique, multitraining SVM (MTSVM) combining the merits of the cotraining technique and a random sampling method in the feature space. Based on the proposed MTSVM algorithm, the above two problems can be mitigated. Experiments are carried out on a large image set of some 20 000 images, and the preliminary results demonstrate that the developed method consistently improves the performance over conventional SVM-based RFs in terms of precision and standard deviation, which are used to evaluate the effectiveness and robustness of a RF algorithm, respectively.

Deniz Kılınc and Adil Alpkocak [20] introduced expansion and reranking approach for annotation based image retrieval from Web pages, considering an image retrieval system using the surrounding texts nearby the image in a Web page as annotations. The authors have proposed a term selection approach, which first expands the document using WordNet, and selects descriptive terms among

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

them. Evaluation of the approach has lead to better results in a open database. Jurgen Assfalg at al [21] exploited images in different contexts, ranging from history of art, through medicine, to education. In general, existing querying paradigms are based either on the usage of textual strings, for high-level semantic queries or on 2D visual examples for the expression of perceptual queries. Semantic queries require manual annotation of the database images. Instead, perceptual queries only require that image analysis is performed on the database images in order to extract salient perceptual fea tures that are matched with those of the example. However, usage of 2D examples is generally inadequate as effective authoring of query images, attaining a realistic reproduction of complex scenes, needs manual editing and sketching ability. Investigation of new querying paradigms is therefore an important and marginally investigated factor for the success. A novel querying paradigm is presented by the authors based on usage of 3D interfaces exploiting navigation in virtual environments. The analysis has led to the development on user test retrieval efficiency and effectiveness, as well as on an evaluation of users' satisfaction.

Haiming Liu et al [22] focused on dissimilarity measurement in content-based image retrieval, where data objects and queries are represented as vectors in high-dimensional content feature spaces. The authors have summarized fourteen core dissimilarity measures and classified them into three categories. A systematic performance comparison is carried out to test the effectiveness of these dissimilarity measures with six different feature spaces and some of their combinations on the Corel image collection. From our experimental results, it is inferred that a number of observations and insights on dissimilarity measurement in content-based image retrieval would lay a foundation for developing more effective image search technologies.

### Analysis on Existing CBIR Systems

Based on the analysis made out, this section presents on some of the drawbacks on the existing CBIR systems. As discussed in section 1, though there exist several areas wherein people have explored research progress, there remains a semantic gap which has led down the results for CBIR. As far as image retrieval on the web in concerned query expansion, re-ranking has gained significant importance. Our proposal aims at improved CBIR systems on the web exploring the above two concepts adopting relevance feedback mechanisms and other significant features.

## Conclusion & Future works

This paper analysis on several techniques and grey area where in there is possibility of improving the results. We propose a framework between query image and database images adopting semantics. One another research focus would be on query expansion, retrieval and ranking. The framework could be trained for different images in the database and tested on various real images. The proposal would be expected to give better as compared with the performance of conventional methods of content based image retrieval. We expect that, in the future, CBIR would be increasingly a challenging task when several aspects like content, structure, extraction and semantics are considered altogether or combined together.

## *References*

[1] *Michael S. Lew, Nicu Sebe, Chabane Djeraba and Ramesh Jain, "Content-based Multimedia Information Retrieval: State of the Art and Challenges", Journal ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), Vol. 2, Issue.1, pp. 1–19, Feb. 2006.*

[2] *Paresh Marwaha, Piyush Marwaha, and Shelly Sachdeva, "Content Based Image Retrieval in Multimedia Databases", International Journal of Recent Trends in Engineering, Vol. 1, No. 2, pp. 210-213, May 2009*.

[3] *Stefanos Vrochidis, Symeon Papadopoulos, Anastasia Moumtzidou, Panagiotis Sidiropoulos, Emanuelle Pianta and Ioannis Kompatsiaris, "Towards content-based patent image retrieval: A framework perspective", World Patent Information, Vol. 32, Issue. 2, pp.94–106, 2010.*

[4] *Guoyong Duana, Jing Yanga, Yilong Yanga, "Content-Based Image Retrieval Research", International Conference on Physics Science*

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

*and Technology (ICPST 2011), Physics Procedia, Vol. 22, pp.471 – 477, 2011.*

*[5] Nidhi Singh,Kanchan Singh and Ashok K. Sinha, "A Novel Approach for Content Based Image Retrieval", C3IT-2012, Procedia Technology, Vol.4, pp.245 – 250, 2012.*

*[6] Henning Müller, Nicolas Michoux, David Bandon and Antoine Geissbuhler, "A review of content-based image retrieval systems in medical applications- clinical benefits and future directions", International Journal of Medical Informatics (2004) 73, 1—23*

*[7] Kim-Hui Yap, and Kui Wu, "A Soft Relevance Framework in Content-Based Image Retrieval Systems", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 15, NO. 12, pp.1557- 1568, December 2005.*

*[8] Cordelia Schmid, "Constructing models for content-based image retrieval", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Vol. 2, pp.39- 45, 2001.*

*[9] Mourad Oussalah, "Content Based Image Retrieval: Review of State of Art and Future Directions", First Workshops on Image Processing Theory, Tools and Applications (IPTA 2008), pp.1-10, 2008.*

[10] *Rokia Missaoui, M. Sarifuddin and Jean Vaillancourt, "An Effective Approach Towards Content-Based Image Retrieval", Proceedings of Third International Conference CIVR 2004, Lecture Notes in Computer Science,Vol. 3115, pp 335-343,2004.*

*[11] Abolfazl Lakdashti, and M. Shahram Moin , "A New Content-Based Image Retrieval Approach Based on Pattern Orientation Histogram", Third International Conference MIRAGE 2007 ,Lecture Notes in Computer Science, Vol. 4418, pp 587-595, 2007.*

*[12] Ritendra Datta, Jia Li, and James Z. Wang, "Content-Based Image Retrieval - Approaches and Trends of the New Age", Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval (MIR'05), November 11-12, pp. 253 – 262, 2005.*

*[13] Wichian Premchaiswadi and Anucha Tungkatsathan, "On-line Content-Based Image Retrieval System using Joint Querying and Relevance Feedback Scheme", WSEAS TRANSACTIONS on COMPUTERS, Vol. 9, Issue. 5, pp.465-474, May 2010.*

*[14] J.L. Martínez-Fernández, Julio Villena Román, Ana M. García-Serrano, and José Carlos González-Cristóbal, "Combining Textual and Visual Features for Image Retrieval", C. Peters et al. (Eds.): CLEF 2005, LNCS 4022, pp. 680 – 691, 2006.*

*[15] M.E. ElAlami, "A novel image retrieval model based on the most relevant features", Knowledge-Based Systems 24 (2011) 23–32.*

*[16] Ja-Hwung Su, Wei-Jyun Huang, Philip S. Yu and Vincent S. Tseng, "Efficient Relevance Feedback for Content-Based Image Retrieval by Mining User Navigation Patterns", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, Vol. 23, No. 3, pp.360-372, March 2011.*

*[17] Hossein Nezamabadi-pour and Ehsanollah Kabir, "Concept learning by fuzzy k-NN classification and relevance feedback for efficient image retrieval", Expert Systems with Applications 36 (2009) 5948–5954.*

*[18] Chuen-Horng Lin , Der-Chen Huang , Yung-Kuan Chan , Kai-Hung Chen and Yen-Jen Chang , "Fast color-spatial feature based image retrieval methods", Expert Systems with Applications 38 (2011) 11412–11420.*

*[19] Jing Li, Nigel Allinson, Dacheng Tao and Xuelong Li, "Multitraining Support Vector Machine for Image Retrieval", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 15, NO. 11, pp.3597-3601, NOVEMBER 2006.*

*[20] Deniz Kılınc, Adil Alpkocak, "An expansion and reranking approach for annotation-based image retrieval from Web", Expert Systems with Applications 38 (2011) 13121–13127.*

*[21] Jurgen Assfalg, Alberto Del Bimbo, and Pietro Pala, "Three-Dimensional Interfaces for Querying by Example in Content-Based Image Retrieval", IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL. 8, NO. 4, pp.305-318, OCTOBER-DECEMBER 2002.*

*[22] Haiming Liu, Dawei Song, Stefan Rüger, Rui Hu and Victoria Uren, " Comparing dissimilarity measures for content-based image retrieval", Proceedings of the 4th Asia information retrieval conference on*

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

*Information retrieval technology (AIRS2008), pp. 44-50,Jan 2008.*

## Author Biblography

**Duraisamy Yuvaraj** was born in 1977, Trichy, Tamilnadu, India. He completed his B.E (comp. science) at J.J.C College of Engineering and Technology, Bharathidasan University, Trichy. He awarded M.Tech (Computer Science & Information Technology) degree from Manonmanium Sundaranar University, Tirunelveli, Tamilnadu. He has nearly 13 years of experience in Teaching both UG and PG program and 1 years of Industrial experience.He is presently working as an Associate. Professor in Department of Information Technology in M.I.E.T Engineering College, Trichy. So far nearly 12 P.G students are guided by his guidance and also guided many projects
Email: yuva.r.d@gmail.com

**K.R.DINESH** received his B.Tech degree in Information Technology from J.J.College of Engineering and Technology, Trichy, Tamil Nadu in year 2014 from Anna University. His research areas include image processing, data mining and text analysis. He has published paper on ELSEVIER(International conference on CNC & CCPE) 2014 proceedings on the title Approaches and Trends in Content based image Retrieval.

Email: dineshkvr@gmail.com

**Dr.S.Hariharan** received his B.E degree specialized in Computer Science and Engineering from Madurai Kammaraj University, Madurai, India in 2002, M.E degree specialized in the field of Computer Science and Engineering from Anna University, Chennai, India in 2004. He holds his Ph.D degree in the area of Information Retrieval from Anna University, Chennai, India. He is a member of IAENG, IACSIT, ISTE, CSTA and has 9 years of experience in teaching. Currently he is working as Associate Professor in Department of Computer Science and Engineering, TRP Engineering College, India. His research interests include Information Retrieval, Data mining, Opinion Mining, Web mining. He has to his credit several papers in referred journals and conferences. He also serves as editorial board member and as program committee member for several international journals and conferences.

**E-mail:**
**mailtos.hariharan@gmail.com**