ISSN 2229-5518

Indexing Techniques for Pictorial Databases

Muhammad Omer Mushtaq, Dr. Yasir Saleem, Maira Alvi, Muhammad Fuzail

Abstract— The emerging field of CBIR (content-based image retrieval) has enjoyed great interest and enthusiasm by researchers in the last decade. Hence making way for a large number of new techniques' systems and their associated fields. Following the same suit digital imagery has also expanded in all its dimensions. This paper focuses on explaining the significant influences on the progress of this field in the current decade. Multiple references have been deployed for this process. Some of the main challenges involved in building a useful system from existing image retrieval techniques is also discussed. The research paper is concluded by studying the trends in the research publications and the requirements for the future.

Index Terms— Image Retrieval, Image Annotations, Content-based, Relevance Feedback System, Indexing, Pictorial Databases, Information Retrieval.

1 Introduction

The low cost of image data holding systems and the growth of storage devices allows to store massive amount of digital images in pictorial databases. When time comes to retrieve the images from these database, lots of computations are performed to get the desired output image.

The problem we face is the interpretation of human visual system by means of query, which is very difficult. So many times we get erroneous results. With the advancements in various technologies users expect that when we are searching images from pictorial databases it must provide us with relevant results. But it is very difficult as human visual interpretation vary from individual to individual as well as it is also difficult to train machine about human visual system. It requires machine to establish connections between the visual effect of the image processing and the mathematical algorithms used in processing. Some databases follow linear searching for the query image and it results in too much time consumption and becomes very difficult for large image databases.

In order to overcome this problem images must be indexed, it results in database query optimization which speeds up query execution. The indexing of images help search engine to provide relevant results to user's query.

Traditional databases use some keywords, as labels which quickly access large text of data. In order to represent visual data with text label, is an overhead which involve lot of manual processing. In addition, results of retrieval might not be satisfactory as query is based on features but, not abstracted by its associated keywords [1].

In current image databases, the dominant retrieval techniques involve text annotations supplied by humans, which describes meaning of image. The basis of searching in these databases are these text annotations. Some freeware algorithms for settled text searching are available but problem still holds with them too. For example textual annotation for the same image may vary from individual to individual. A simple system explained in Figure 1.

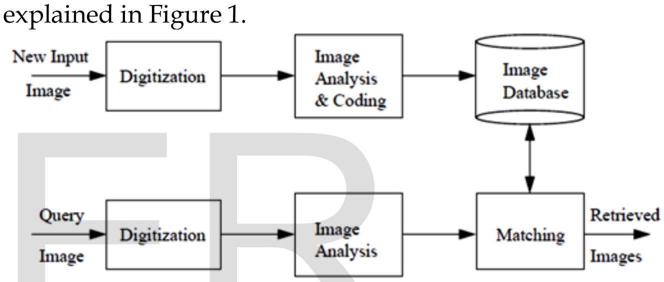

Fig.1 Schematic of an Image Storage and Retrieval System [1]

This makes awfully difficult for an algorithm to reliably answer to the user query. Furthermore entering textual annotations manually for large image databases becomes disproportionately expensive [2]. Image retrieval relevance basically relies on a number of factors. The factors are image and information extraction, image annotation and indexing, query manipulation and image truncation. Among the factors mentioned, the most influential in successful image retrieval is image annotation. Image annotation provides details that best describe an image. Currently, automatic image annotation through computer vision techniques is used to replace manual image annotation [3]. The text is human's creation, and images are replicas of what he has seen since birth, its physical description is relatively subtle. Human vision system grew natively over many centuries. It is hard to interpret what we see and hardest if it is to teach a machine. In the past decade, aspiring attempts have been made to make machines learn to understand, indexing and image annotation shows wide range of concepts with much progress. Content based image retrieval organizes digital picture archives by their visual contents [4].

The goal of this research is to discuss the challenges involved in image retrieval from pictorial databases and their possible solutions. This is an emerging field in which every day new

---

- ***Muhammad Omer Mushtaq is Lecturer at Pakistan Institute of Engineering and Technology, Multan, Pakistan. At the same time pursuing Master's degree in*** *Computer Engineering at UET Lahore, Pakistan. PH-00923006386194. E-mail: engr.omermushtaq@kgmail.com*
- ***Dr. Yasir Saleem is Associate Professor*** *at UET Lahore, Pakistan. E-mail: yasir@uet.edu.pk*
- ***Maira Alvi is pursuing Master's degree in Computer Engineering at UET Lahore, Pakistan. E-mail: malvimaira@gmail.com***
- ***Muhammad Fuzail*** *is Lecturer at Computer Science Department, National College of Business Administration and Economics, Multan, Pakistan. At the same time pursuing Master's Degree in Computer Engineering at UET Lahore, Pakistan, PH-00923457334224. Email:****m.fuzail@ymail.com***

ISSN 2229-5518

innovations and advancement are adding up. We discussed the image retrieval techniques used and how this field is emerging from the last decade. Various manual test are also performed by us to determine how much work is done in this field by counting number of research papers published in image retrieval domain by exploring IEEE Xplore, Springer's digital library and Google Scholar.
The rest of the paper is organized as follows: Section II describes the work relating to image retrieval techniques. Section III discusses the trade-off among different content based image retrieval techniques. A discussion on the publication trends within the field with respect to venues/journals is presented in Sec. 4. We conclude in Sec. 5 along with some of its future work.

1.1 Related Work

Research in image retrieval expanded on two different angles, text-based and visual-based. Early research and development of image retrieval systems emphasized on text based image retrieval, in which the image is illustrated as text and then it was retrieved by using text based databases. This method did not work properly due to human individuality. Two persons do not observe a same picture alike rather same picture appears different to them. Owing to mismatches between query and indexing, the image retrieval cannot be done properly. In order to get rid of this problem, visual-based image retrieval systems were introduced [3, 5]. Figure 2 explains the architecture of initial prototypes.

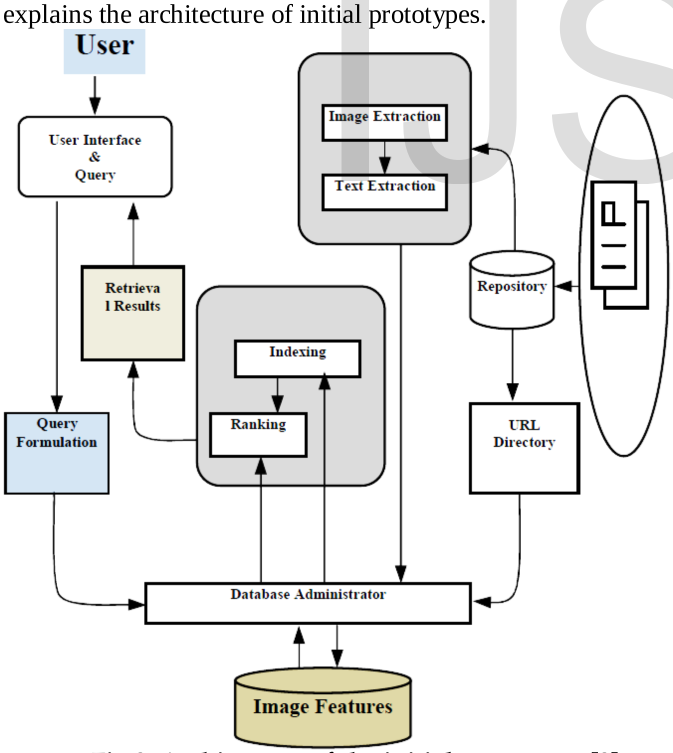

Fig 2. Architecture of the initial prototypes [3]

[1] also provides review of wavelet histogram technique where image is decomposed to M stages using wavelets. Figure 3 provides a description of histogram generation.

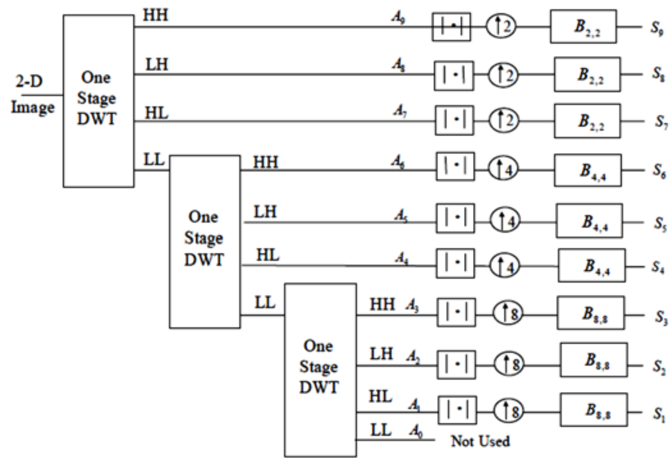

Figure 3. Wavelet Histogram Generation [1]

Several multidimensional indexing techniques for capturing low level features like features based or distance based techniques have been proposed, each of which can further divided as data partitioned [6] or space partitioned based algorithm[7 , 8].
In most of the systems feature extraction is performed as a pre-processing step, which extract global image features like color histogram or local descriptors like texture and shape [4]. A region based dominant color descriptor indexed in 3-D space along with their percentage coverage within the regions is proposed in [9]. It is shown to be computationally more efficient in similarity based retrieval in comparison with traditional color histograms. The research paper argues that this representation is more efficient than high dimensional histograms in terms of search and retrieval. There are certain drawbacks as well such as dimension reduction and color moment descriptors associated with earlier proposals have also been covered in it. In [10], a multi-resolution histogram is introduced that is capable of capturing spatial image information. It is shown to be efficient in retrieving textured images, while typical advantages of histograms remain there. In [11], Gaussian mixture vector quantization (GMVQ) extracts color histograms and seems to have better retrieval compared with uniform quantization and vector quantization with squared error. A set of color and texture descriptors which is well suited to natural images and videos, is described in [12]. These include histogram-based descriptors, dominant color descriptors, spatial color descriptors and texture descriptors used for retrieval and browsing purposes. By using generalized Gaussian distributions features based on texture can be modelled on marginal distribution on wavelet coefficients [13].
The key attribute of segmented image regions is shape. In image retrieval the efficient and robust representation of shape plays an important role. A shape similarity measure using discrete curve evolution to simplify contours, is discussed in [14].Doing this contour simplification helps to remove noisy and irrelevant shape features from consideration. A new shape descriptor for shape matching, referred to as shape context, has been proposed [15] which is fairly compact yet robust to a number of geometric transformations. A dynamic programming (DP) approach to shape matching has been proposed in [16]. One drawback of this approach is that moments and computation of Fourier descriptors is slow. While in Fourier descriptors, exploitation of both the amplitude and phase and using Dynamic Time Warping (DTW) distance instead of Euclidean distance has been shown to be a perfect shape matching technique in [17]. By discarding the phase information, the rotational and starting point invariance thus obtained is maintained here by adding compensation terms to the original phase, thus allowing its exploitation for better discrimination [4].

For characterizing shape, reliable segmentation is critical within images. Without which the estimates of shape are completely meaningless. Although the general problem of segmentation in the context of human perception is far from being solved, some interesting new directions could also be found, one of the most important among them being segmentation based on the Normalized Cuts criteria [18]. This approach relies primarily on the spectral clustering theory. By using cues of contour and texture differences, it has been expanded to segmentation of textured image [19], and for the purpose of incorporating partial grouping priors into the process of segmentation by solving a constrained optimization problem [20]. The latter has potential for incorporating real-world application specific priors, e.g. in pathological images size and location cues of organs [4]. In case of medical imaging, 3D brain magnetic resonance (MR) images have been segmented using Expectation-Maximization (EM) and Hidden Markov Random Fields algorithm [21], and the spectral clustering approach is more successful in segmenting vertebral bodies from sagittal MR images [22]. Among other recent approaches proposed are segmentation based on the mean shift procedure [23], multi-resolution segmentation of low depth of field images [24], a Bayesian framework based segmentation involving the Markov chain Monte Carlo technique [25], and an EM algorithm based segmentation using a Gaussian mixture model [26], forming blobs suitable for image querying and retrieval. A sequential segmentation approach that starts with texture features and refine segmentation using color features is explored in [27].
In image understanding achieving good segmentation is a big step, some of the basic problems occurring in current techniques are speed considerations, segmentation's reliability, acceptable and a robust benchmark for assessment of the same. In the case of image retrieval, some of the ways of getting around this problem have been to reduce dependence on reliable segmentation [26], to obtain soft similarity measures every generated segment of image must involve in matching process [28], or by using block based multi resolution hidden Markov models spatial arrangement of color and texture can be characterized [29, 30], this technique has also been extended to segment 3D volume images [31].
Distance based indexing structures are formed based on distance or similarities between two data objects. Some famous distance based indexing structures are SS-Tree [32], M-Tree [33], vp-Tree [34], etc. among them only M-Tree guarantees a balanced structures as it is built in a bottom-up structure. In M-Tree, partitioning of objects is done based on their relative distances, measured by specific distance functions and these objects are stored in a fixed sized node [35]. Internal nodes store the routing object while leaf node stores the indexed objects. Histogram techniques are popular in indexing application due to their lower complexity [36].This technique usually performs well for natural images but it fails in case of the texture images. Fast wavelet histogram technique (FWHT) [1], improves the performance of image indexing system, the fast wavelet techniques provides a well performance compared to wavelet histogram technique (WHT) [37]. On the other hand its performance for the textures images is better as compared to the WHT with much reduced complexity.
Selection of content based image retrieval feature remains largely ad-hoc. In image retrieval semantic-sensitive feature selection has also been shown to be effective [28]. When a large number of image features are available, working with a feature subset is a way to improve generalization and efficiency. Shape is a key attribute of segmented image regions and its robust and efficient retrieval plays an important role. In [11] shape similarity measure using discrete curve evolution is discussed which helps to remove noisy and irrelevant shape features. A dynamic programming (DP) approach for shape matching is proposed in [14]. Drawback of this approach is that the computation of Fourier descriptors and moments is slow. With exploitation of both amplitude and phase, continuing with Fourier descriptors and using Dynamic Time Warping (DTW) has been shown to be accurate shape matching technique [16]. For characterizing shape within images, reliable segmentation is critical. Even though the general problem of segmentation in the context of human perception is far from being solved.

2 IMAGE RETRIEVAL TECHNIQUES

As there's no universally accepted method or algorithm for image retrieval by characterizing human vision, in the context of interpreting images. The continued effort is still there in this domain. We will be discussing various image retrieval techniques based on the content of images.
Content Based Image Retrieval faces two problems; (a) how to describe image mathematically, and (b) how to identify similarities between pair of images upon their abstracted description. Mathematical description of an image is referred to as signature of image. Their formulation is based upon the calculation of different measures of an image as their aspect ratios. There are some other motivating factors for designing similarity measures in certain way, which turns in construction of signatures.
The comparison with pre-2000 work in CBIR, the recent researches has been the diversity of image signatures [38]. There is advancement in the derivation of new features such as, shape the construction of signature is based on these features. These additional features add the richness in mathematical formulation of signatures on the basis of similarities defined by shape, aspect ratios etc. CBIR in recent years have emerged into the employment of statistical and machine learning technique in various aspects. The actively pursued direction of image retrieval is to engage human in this process that is, to include human in the loop. The involvement of user in the recent work has progressed toward more collaborative and iterative scheme by leveraging learning technique. The overhead for user to specify to specify what she is looking for at the beginning of search is much reduced.
Many system of CBIR have feature extraction as pre-processing step. After this, the visual features act as input to subsequent image analysis task for its similarity estimation, annotation and concept detection [38]. The interest of current decade is in region-based visual signature extraction, for which segmentation is the essential first step. To get a region-based signature of a given image the major step is to divide the image into different parts commonly referred as segments. In order to have our results to be accurate and reliable the segmentation process is critical for characterizing shapes in the given image. If the segmentation is not performed properly shape estimates are largely meaningless. The k means clustering enjoy the basic advantage of speed, but it is not the recent developed method. One of more new advances in segmentation employs the Normalized cuts criterion [39]. The problem of Normalized cuts is mapped to a weighted graph partitioning problem, where vertex set of a graph represent image pixels and edge weight represent some perceptual similarity between pixel pairs. The normalized cut segmentation method in [39] is extended to texture image segmentation by using cues of contour and texture differences [40], and to incorporate known partial grouping priors by solving a constrained optimization problem [41]. The latter has potential for incorporating real-world application specific priors, e.g., location and size cues of organs in pathological images.
Searching of medical images is an increasingly important research problem, due to high-throughput, and high-dimensional imaging introduced. In this domain, 3D brain magnetic resonance (MR) images have been segmented using Hidden Markov Random Fields and the Expectation-Maximization (EM) algorithm [42], and the spectral clustering

ISSN 2229-5518

approach has found some success in segmenting vertebral bodies from sagittal MR images [43]. Other approaches on segmentation are based on mean shift procedure [44], multi-resolution of low depth of field images [45],a Bayesian framework based segmentation involving the Markov chain Monte Carlo technique [46], and an EM algorithm based segmentation using a Gaussian mixture model [47], forming blobs suitable for image querying and retrieval. A sequential segmentation approach that starts with texture features and references segmentation using color features is explored in [48]. An unsupervised approach for segmentation of images containing homogeneous color, texture regions has been proposed in [49].

A feature captures certain visual properties of an image, either from entire image or locally for some portion of an image. Most common features are color, texture, shape and salient points in an image. While in global extraction, features are computed to capture overall characteristics of an image. For color layout approach, image is divided into a small number of sub-images and the average color components are red, green and blue intensities, are computed on every sub-image. The whole image is represented by a vector of color components where a particular dimension of a vector corresponds to a certain sub-image location. The global extraction is high speed in both, extraction and computing similarity. Global extraction is too rigid in representation of an image they can be over sensitive to location and hence fail to identify important visual characteristics. Local extraction is better as it work on small group of pixels and increase the robustness.

There are various methods for visual signature extraction that have some advantages and limitations. Global feature extraction gives big picture and local feature represents the details. Depending on the scale of the key pattern or contents and appropriate representation should be chosen. Hybrid representation might be sometimes more attractive but it has additional computational complexity. Segmentation recognize objects in a scene, precise segmentation is still a open problem. The alternative approaches to characterize structure may be more suitable. Among different approaches segmentation is often a trade-off between quality and complexity, which might lead to a difference in eventual search performance and speed. The choice for image signature should depend on the desirability of the system. The extracted features should be more coherent with human visual system. Figure 4 explains the image retrieval system.

In [38] the key factors which must be considered for designing a proposed image similarity measures are as follows

- Computational efficiency
- Agreement with semantics

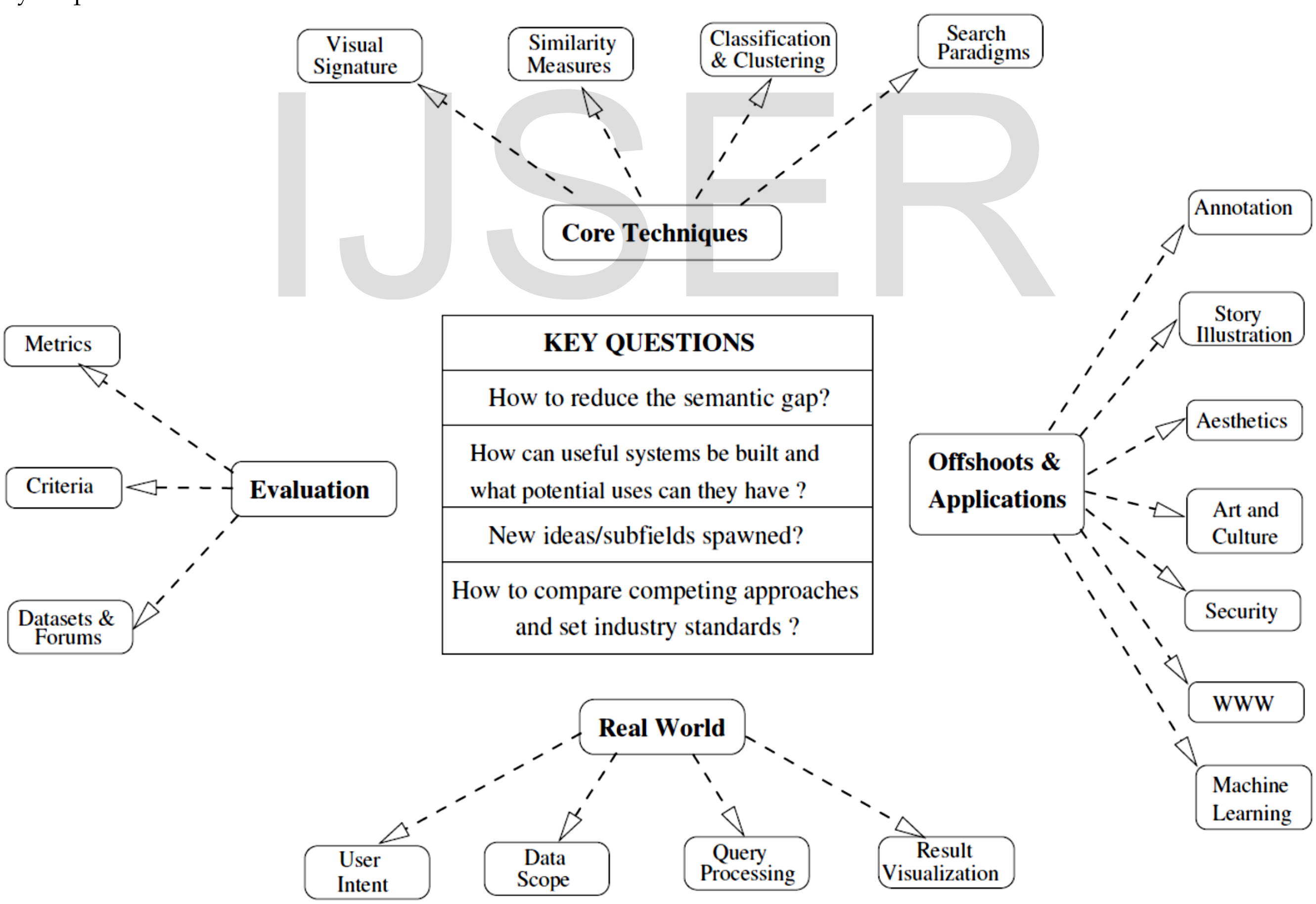

Fig 4. Over View of many facets of Image Retrieval [38]

- Robustness to the noise
- Local linearity
- Invariance to background

And the various techniques in [38] are grouped as follows

ISSN 2229-5518

- By the use of global based similarity, region based or a combination of both
- By treating features as ensembles, vectors or non-vectors
- Role of image segments in similarity computation
- Using supervised, unsupervised and semi-supervised learnings
- Similarity measures based on stochastic, deterministic or fuzzy logic

Let's start our discussion on the region-based techniques because of its wide spread use in recent years. Technically, region based signatures are emphasized due to its definition i.e distance between sets of vectors instead of distance between single vector. Instead of dividing an image into different segments, summaries of local feature like probability density function and codebook have been known to be used as signature. Codebooks are created by the quantization of vectors. Fitting a Gaussian mixture model [50] is an effective method to obtain density estimation. Disparity between distributions is measured by Kullback-Leibler.

The Earth Mover's Distance [51] (EMD) is also a matching technique in which image is represented in the form of a sets of vectors. In this image matching is treated as a moving component color histogram of images from one to the other, and thus the effort required is minimum. IRM (integrated region matching) distance [52] is another matching algorithm which uses the most similar highest priority (MSHP) principal to match regions. [53] explores the region based image retrieval, under the assumption of a hidden semantic concept underlying image generation. Region based frame work with much improved efficiency is proposed in [54],here from training image region, vector quantization is employed to build a region codebook. For the coarse foreground / background query segmentation on the user's query region-of interest (ROI), a hybrid approach is employed in [55]. In [56] by using the Kullback-Leibler (K-L), methods for texture retrieval have been proposed.

When the images are represented as single vectors, a problem is faced by many authors in measuring perceptual image distance by metrics in any linear feature space. On solution for this problem is to replace Euclidean distance by the geodesic distance. This idea is explored and applied to image similarity in [57, 58, 59, 60]. The advantage of a single vector representation is that the geometric and algebraic operations can be performed easily and much more efficiently. However complex image semantics and necessary details cannot be represented by having single vector approach.

3 CURRENT RESEARCH TRENDS

We momentarily analyzed publication trends in image retrieval and annotations since year 2009. Our source was Google Scholars. The searching phrase was either, image or images or pictures or content-based or annotations or information retrieval or indexing or relevance feedback, for publications in [38] the journals- IEEE T. Pattern Analysis and Machine Intelligence (PAMI), IEEE T. Image Processing (TIP), IEEE T. Circuits and Systems for Video Technology (CSVT), IEEE T. Multimedia (TOM), J. Machine Learning Research (JMLR), International J. Computer Vision (IJCV), Pattern Recognition Letters (PRL), and ACM Computing Surveys (SURV) and conferences - IEEE Computer Vision and Pattern Recognition (CVPR), International Conference on Computer Vision (ICCV), European Conference on Computer Vision (ECCV), IEEE International Conference on Image Processing (ICIP), ACM Multimedia (MM), ACM SIG Information Retrieval (IR), and ACM Human Factors in Computing Systems (CHI). In relevant papers top 100 results were used for study. Google Scholar presents results roughly in decreasing order of citations. Limiting search to the top few papers translates to reporting statistics on work with noticeable impact. We gathered statistics on publishing venue/journal. These trends are reported in terms of (a) number of papers, and (b) total number of citations. Plots of these scores are presented in Fig. 4 and Fig. 5. Nevertheless, these plots convey general trends in the relative impact of scholarly work. CBIR is meaningful for its services to human users. At the same time it is difficult to identify user requirements as objective relevance based scores. Figure 5 and 6 explain conference/journal wise number of publication counts and conference/journal wise total number of citations respectively.

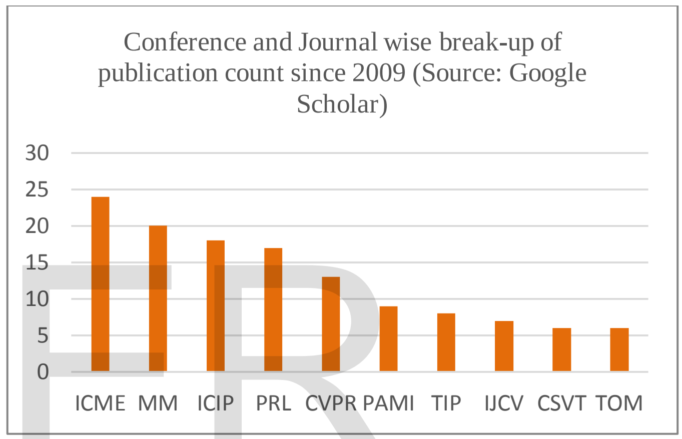

Fig 5. Conference/Journal wise No. of Publication Counts

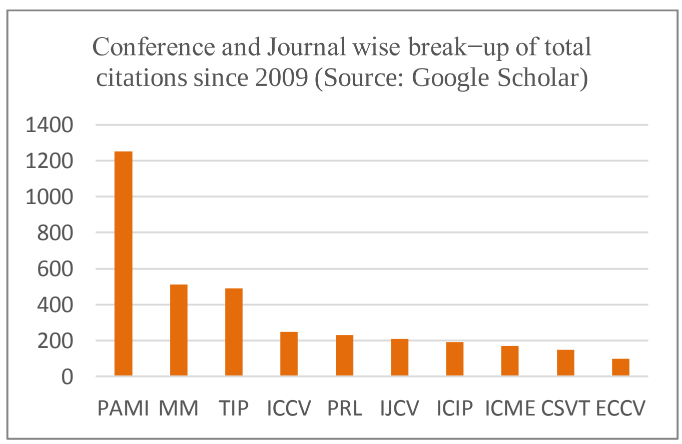

Fig 6. Conference/Journal wise Total No. of Citations

4 CONCLUSION AND FUTURE WORK

This research paper makes a brief survey on the new and exhilarating field of CBIR (content-based image retrieval) and the automated image annotation. It is foreseen that the field will go through a paradigm shift in the near future. Hence shifting the focus more on application-oriented, domain-specific work that will affect our everyday life. This research paper compiles research trends in this field through multiple citations. It can be perceived from the trends that while a lot of work has been done on feature extraction, relevance feedback and system, the same cannot be said for application-oriented aspects such as interface, visualization, scalability and evaluation. The future of this field depends on the fact that where the quest for robust and reliable image understanding technology must continue the application-oriented aspects should also be given equal importance. The effective growth of the field of CBIR and automated annotation requires the collective focus and overall progress in each aspect of image retrieval.

5 REFERENCES

[1] Mandal, M.K., T. Aboulnasr, and S. Panchanathan,"Fast wavelet histogram technique for image indexing" In Content-Based Access of Image and Video Libraries,1998. Proceedings. IEEE Workshop on. 1998.

[2] James Ze Wang, G.I., Oscar Firschein, Sha Xin We," Wavlet-Based Image Indexing Techniques with partial Sketch Reterieval Capablity" 1997, Standford University. P. 13-24

[3] Azizan, A. and M. Itaza Afiani, "The effects of surrounding text in image indexing" In Science and Social Research (CSSR), 2012 International Conference on.

[4] Ritendra Datta, J.L, James Z. Wang, "Content-Based Image Retrieval-Approaches and Trends of the New Age" 2005, The Pennsylvania State University, University Park,PA 16802, USA.

[5] Y. Rui, T. S. Huang, and S. Chang, "Image retrieval: Current techniques, promising directions and open issues" J. Visual Communication and Image Representation, VOL. 10, no. 1, pp.39-62, March 1999.

[6] S. Berchtold, D. A. Keim, and H. Kriegel. "The x-tree:an index structure for high dimensional data." In Proceedings of the 22nd International Conference on Very Large Databases, pages 28–39, Bombay, India, September 1996.

[7] D. B. Lomet and B. Salzberg. "The hb-tree: A multi-attribute indexing method with good guaranteed performance." ACM Transactions on Database Systems, 15(4):625–658, 1990.

[8] J. Robinson. "The k-d-b-tree: A search structure for large multidimensional dynamic indexes". In Proceedings of the 1981 ACM SIGMOD International Conference on Management of Data, pages 10–18, Ann Arbor, Michigan, United States, April 1981.

[9] Y. Deng, B. S. Manjunath, C. Kenney, M. S. Moore, and H. Shin, "An Efficient Color Representation for Image Retrieval," IEEE Trans. Image Processing, 10(1):140–147, 2001.

[10] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar, "Multiresolution Histograms and Their Use for Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, 26(7):831–847, 2004.

[11] S. Jeong, C. S. Won, and R.M. Gray, "Image retrieval using color histograms generated by Gauss mixture vector quantization," Computer Vision and Image Understanding, 9(1–3):44–66, 2004.

[12] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and Texture Descriptors," IEEE Trans. Circuits and Systems for Video Technology, 11(6):703–715, 2001.

[13] M. N. Do and M. Vetterli, "Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler Distance," IEEE Trans. Image Processing, 11(2):146–158, 2002.

[14] L.J. Latecki and R. Lakamper, "Shape Similarity Measure Based on Correspondence of Visual Parts," IEEE Trans. Pattern Analysis and Machine Intelligence, 22(10):1185–1190, 2000.

[15] S. Belongie, J. Malik, and J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(4):509–522, 2002.

[16] E.G.M Petrakis, A Diplaros, and E. Milios,"Matching and Retrieval of Distorted and Occluded Shapes Using Dynamic Programming," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(4):509–522, 2002.

[17] Bartolini, P. Ciaccia, and M. Patella, "WARP: Accurate Retrieval of Shapes Using Phase of Fourier Descriptors and Time Warping Distance," IEEE Trans. Pattern Analysis and Machine Intelligence, 27(1):142–147, 2005.

[18] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," IEEE Trans. Pattern Analysis and Machine Intelligence, 22(8):888–905, 2000.

[19] J. Malik, S. Belongie, T. K. Leung, and J. Shi, "Contour and Texture Analysis for Image Segmentation," International Journal of Computer Vision, 43(1):7-27, 2001.

[20] S. X. Yu and J. Shi, "Segmentation Given Partial Grouping Constraints," IEEE Trans. Pattern Analysis and Machine Intelligence, 26(2):173–183, 2004.

[21] Y. Zhang, M. Brady, and S. Smith, "Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm," IEEE Trans. Medical Imaging, 20(1):45–57, 2001.

[22] J. Carballido-Gamio, S. Belongie, and S. Majumdar "Normalized Cuts in 3-D for Spinal MRI Segmentation," IEEE Trans. Medical Imaging, 23(1):36–44, 2004.

[23] D. Comaniciu and P. Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(5), 603–619, 2002.

[24] J. Z. Wang, J. Li, R. M. Gray and G.Wiederhold, "Unsupervised Multiresolution Segmentation for Images with Low Depth of Field," IEEE Trans. Pattern Analysis and Machine Intelligence, 23(1):85–90, 2001.

[25] Z. Tu and S.-C. Zhu, "Image Segmentation by Data-Driven Markov Chain Monte Carlo," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(5):657–673, 2002.

[26] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image Segmentation Using Expectation-maximization and Its Application to Image Querying," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(8):1026-1038, 2002.

[27] J. Chen, T.N. Pappas, A. Mojsilovic, and B. Rogowitz, "Adaptive image segmentation based on color and texture," Proc. IEEE International Conference on Image Processing, 2002.

[28] J.Z. Wang, J. Li, and G. Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE Trans. Pattern Analysis and Machine Intelligence, 23(9), 947–963, 2001.

[29] J. Li, R. M. Gray, and R. A. Olshen, "Multiresolution Image Classification by Hierarchical Modeling with Two Dimensional Hidden Markov Models," IEEE Trans. Information Theory, 46(5):1826–1841, 2000.

[30] J. Li and J. Z. Wang, "Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach," IEEE Trans. Pattern

Analysis and Machine Intelligence,25(9):1075–1088, 2003.

[31] J. Li, D. Joshi, and J. Z. Wang, “Stochastic Modeling of Volume Images with a 3-D Hidden Markov Model,” Proc. IEEE International Conference on Image Processing, 2004

[32] D. A. White and R. Jain. Similarity indexing with SS-tree. In Proceedings of the 12th International Conference on Data Engineering, pages 516–523, New Orleans, LA,United States, February 1996.

[33] P. Ciaccia, M. Patella, and P. Zezula. M-tree: An efficient access method for similarity search in metric spaces. In Proceedings of the 23rd VLDB International Conference, pages 426–435, Athens, Greece, August 1997.

[34] P. N. Yianilos. Data structures and algorithms for nearest neighbor search in general metric spaces. In Proceedings of the 3rd Annual ACM-SIAM Symposium on Discrete Algorithms, pages 311–321, Philadelphia, PA, United States, January 1993.

[35] P. Ciaccia, M. Patella, and P. Zezula. M-tree: An efficient access method for similarity search in metric spaces. In Proceedings of the 23rd VLDB International Conference, pages 426–435, Athens, Greece, August 1997.

[36] M. J. Swain and D. H. Ballard, “Color indexing,” International Journal of Computer Vision, Vol. 7, No. 1,1991.

[37] J. R. Smith and S. F. Chang, "Automated binary texture feature sets for image retrieval," Proc. of ICASSP,Atlanta, May 1996.

[38] Ritendra Datta, D.J.,Jia Li, and James Z. Wang, Image Retrieval: Ideas, Influences,and Trends of the New Age. ACM Transactions on Computing Surveys,2008.4(2).

[39] Shi, J. and Malik, J. 2000. Normalized cuts and image segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence 22, 8, 888-905.

[40] Malik, J., Belongie, S., Leung, T. K., and Shi, J. 2001. Contour and texture analysis for image segmentation. Intl. J. Computer Vision 43, 1, 7-27.

[41] Yu, S. X. and Shi, J. 2004. Segmentation given partial grouping constraints. IEEE Trans. Pattern Analysis and Machine Intelligence 26, 2, 173-183.

[42] Zhang, Y., Brady, M., and Smith, S. 2001. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. IEEE Trans. Medical Imaging 20, 1, 45-57.

[43] Carballido-Gamio, J., Belongie, S., and Majumdar, S. 2004. Normalized cuts in 3-d for spinal mri segmentation. IEEE Trans. Medical Imaging 23, 1, 36-44.

[44] Comaniciu, D. and Meer, P. 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans. Pattern Analysis and Machine Intelligence 24, 5, 603-619.

[45] Wang, J., Li, J., and Wiederhold, G. 2001. Simplicity: Semantics-sensitive integrated matching for picture libraries. IEEE Trans. Pattern Analysis and Machine Intelligence 23, 9, 947-963.

[46] Tu, Z. and Zhu, S.-C. 2002. Image segmentation by data-driven markov chain monte carlo. IEEE Trans. Pattern Analysis and Machine Intelligence 24, 5, 657-673.

[47] Carson, C., Belongie, S., Greenspan, H., and Malik, J. 2002. Blobworld: Image segmentation using expectation maximization and its application to image querying. IEEE Trans. Pattern Analysis and Machine Intelligence 24, 8, 1026-1038.

[48] Chen, J., Pappas, T., Mojsilovic, A., and Rogowitz, B. 2002. Adaptive image segmentation based on color and texture. In Proc. IEEE ICIP.

[49] Deng, Y. and Manjunath, B. 2001. Unsupervised segmentation of color-texture regions in images and video. IEEE Trans. Pattern Analysis and Machine Intelligence 23, 8, 800-810.

[50] Hastie, T., Tibshirani, R., and Friedman, J. 2001. The Elements of Statistical Learning. Springer-Verlag.

[51] Rubner, Y., Tomasi, C., and Guibas, L. J. 2000. The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vision 40, 99-121.

[52] Li, J., Najmi, A., and Gray, R. M. 2000. Image classi_cation by a two dimensional hidden markov model. IEEE Trans. Signal Processing 48, 2, 527-533.

[53] Zhang, R. and Zhang, Z. 2004. Hidden semantic concept discovery in region based image retrieval. In Proc. IEEE CVPR.

[54] Jing, F., Li, M., Zhang, H.-J., and Zhang, B. 2004a. An e_cient and e_ective region-based image retrieval framework. IEEE Trans. Image Processing 13, 5, 699-709.

[55] Dagli, C. and Huang, T. S. 2004. A framework for grid-based image retrieval. In Proc. IEEE ICPR.

[56] Do, M. N. and Vetterli, M. 2002. Wavelet-based texture retrieval using generalized Gaussian density and kullback-leibler distance. IEEE Trans. Image Processing 11, 2, 146-158.

[57] Vasconcelos, N. and Lippman, A. 2005. A multiresolution manifold distance for invariant image similarity. IEEE Trans. Multimedia 7, 1, 127-142.

[58] He, X. 2004. Incremental semi-supervised subspace learning for image retrieval. In Proc. ACM Multimedia.

[59] He, J., Li, M., Zhang, H.-J., Tong, H., and Zhang, C. 2004a. Manifold-ranking based image retrieval. In Proc. ACM Multimedia.

[60] Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A. 2003. Face recognition: A literature survey. ACM Computing Surveys 35, 4, 399-458.