ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 5, March 2017

CONTENT BASED IMAGE RETRIEVAL USING MULTI DIMENSIONAL TEXTURE, EDGE ORIENTATION HISTOGRAM AND EDBTC COMPRESSION

Rekha.J 1* , B. Muthu Kumar 2 , Kavitha 3 , S.Prabhakaran 4

1, 3Department of Computer Science, Sakthi College of Arts and Science for Women, TN, India

2 Professor, Department of CSE, Syed Ammal Engineering College, Ramanathapuram, TN, India

4 Department of Information Technology, Kathir College of Engineering

*Corresponding Author: mailto: rekhamsc31@gmail.com

ABSTRACT

The graph based ranking model is proposed for information retrieval area. Our proposed technique concentrates on ranking the Data Manifold Model or Manifold Ranking (MR). The Content Based Image Retrieval (CBIR) gives an extra-ordinary result underlying geometrical structure for the given image database. MR is an expensive one. This approach limits its usefulness for the large database containing new queries. The name for Graph based Ranking model is Efficient Manifold Ranking (EMR). The scalable [8] graph construction and efficient ranking computation are the two views of MR. We specially develop an anchor graph on the database. To increase the speed of ranking adjacency matrix is introduced. EMR is a promising method for real-time retrieval applications.

Keywords: Content based image retrieval (CBIR), classification, segmentation.

1. INTRODUCTION

Graph based ranking model is mainly focused the problem of out of sample retrieval on large scale databases. Most of the image retrieval systems are based on the keyword search such as Google and Yahoo image search. The keyword is matched based on the image titles, manual annotation, web document etc. The problem occurs are shortage of text information and incomplete statement and image. CBIR [1] [3] is the best method to overcome the hurdles. The proposed system sets a trend set for low level features, global features can automatically extracted from images. The researches have been performed for designing more informative low-level features to represent images or better metrics. The performance is restricted due to the sensitive data and conditions. Traditional method concentrates on data features but they ignore the structure information. During the unknown label information at the time they focused the method structure information. The same semantic label is likely to share the

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 5, March 2017

neighbouring data points or belong to the same cluster or manifold. Our aim is to produce a good CBIR system for low level features as well as the intrinsic structure of the image database. The intrinsic geometrical structure collectively revealed by a large number of data. MR [9] [10] works on many applications and produce excellent performance and feasibility on a variety of data types such as the text [11], image [12] [13] and video [14]. He et al. [12] is the first person approached MR to CBIR and improved the performance of image retrieval.

MR has some of the drawbacks such as handle large scale databases cost wise high in both stages graph construction and ranking competition stages. In the existing system they don't know how to handle an out-of-sample query. The original MR [4] is inadequate for a real world CBIR system. We propose a novel framework named Efficient Manifold Ranking (EMR). The technique is based on two stages. The first one is Offline stage used for learning and the latter one is Online stage used for handling a new sample. EMR can handle a database with 1 million images. The online retrieval is used to retrieve within a short time.

The new contributions [13] of our proposed system are as follows:

1) To produce an excellent outsource for an out-of-sample retrieval proposed an efficient approximate method to compute ranking scores for a new query.
2) To test the efficiency of the proposed technique, 1 millions samples in three databases are included.
3) To overcome and design the local weight estimation problem of anchor graph, two various methods are compared.

In Section 2, we describes the related work and Section 3 overviews the MR algorithm. In Section 4, we proposed an EMR approach. Section 5 summarizes the experimental results on many real-time image databases. Section 6 concludes the paper.

2. RELATED WORKS

The problem of ranking has been overcome by information retrieval and machine learning areas. Conventional ranking models are content based models, like the Vector Space Model, BM25, language modelling [15], link structure based models [16] [17] [18] . The rank model [19] [20] is an important category to learn optimize a ranking function that incorporates relevance features and avoids tuning a large number of parameters empirically. The authors present a unified framework for jointly optimizing effectiveness and efficiency. [6] presented an automatic segmentation method which effectively combines Active Contour Model, Live Wire method and Graph Cut approach (CLG). The aim of Live wire method is to provide control to the user on segmentation process during execution. Active Contour Model provides a statistical model of object shape and appearance to a new image which are built during a training phase. In the graph cut technique, each pixel is represented as a node and the distance between those nodes is represented as edges. In graph theory, a cut is a partition of the nodes that divides the graph into two disjoint subsets. For initialization, a pseudo strategy is employed and the organs are

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 5, March 2017

segmented slice by slice through the OACAM (Oriented Active Contour Appearance Model). Initialization provides rough object localization and shape constraints which produce refined delineation. This method is tested with different set of images including CT and MR images especially 3D images and produced perfect segmentation results.

Our aim is to give attention of ranking model, graph-based ranking. This approach successfully applied in link-structure analysis of the web [16] [17], social networks research and multimedia data analysis. Generally, a graph can be denoted as G = (V, E, W), where V is a set of vertices in which each vertex represents a data point, $E \subseteq V \times V$ is a set of edges connecting related vertices, and W is a adjacency matrix recording the pairwise weights between vertices. The importance of a vertex is based on local or global information drawn from the graph and it is achieved by graph-based ranking model.

The model data by a weighted graph, and incorporated this graph structure into the ranking function as a regularizer are recommended by Agarwal. Guan et al. proposed a graph-based ranking algorithm for interrelated multi-type resources to generate personalized tag recommendation. An automatically tag ranking scheme by performing a random walk over a tag similarity graph are proposed by Lieut al. The authors made the music recommendation by ranking on a unified hypergraph [7], combining with rich social information and music content. The graph-based model has been proposed a method namely, hypergraph. The MR has been currently spread over on several works. The authors partitioned the data into several parts and computed the ranking function by a block-wise way.

3. MANIFOLD RANKING REVIEW

In this section, we briefly review the manifold ranking algorithm and make a detailed analysis about its drawbacks. We start form the description of notations.

3.1 Notations and Formulations

Given a set of data $\chi = \{x_1, x2... xn\} \subset R^m$ and build a graph on the data (e.g., kNN graph).

$W \in R^{n \times n}$ denotes the adjacency matrix with element w_{ij} saving the weight of the edge between point i and j. Normally the weight can be defined by the heat kernel $w_{ij} = \exp[-d^2(xi,xj)/2\sigma^2)]$ if there is an edge linking x_i and x_j, otherwise $w_{ij} = 0$. Function d (xi,xj) is a distance metric of x_i and x_j defined on χ, such as the Euclidean distance. Let r: $\chi \rightarrow R$ be a ranking function which assigns to each point x_i a ranking score r_i. Finally, we define an initial vector $y = [y_1 \ldots y_n]^T$, in which $y_i = 1$ if x_i is a query and $y_i = 0$ otherwise.

The cost function associated with r is defined to be

$$O(r) = \frac{1}{2} \left(\sum_{i,j=1}^{n} w_{ij} \parallel 1/\sqrt{D}_{ij} r_i - 1/\sqrt{D}_{ij} r_j \parallel^2 + \mu \sum_{i=1}^{n} || r_i - y_i \parallel^2 \right) \quad \ldots..(1)$$

Where $\mu > 0$ is the regularization parameter and D is a diagonal matrix with $D_{ii} = \sum_{j=1}^{n} w_{ij}$.

The first term in the cost function is a smoothness constraint, which makes the nearby points in the space having close ranking scores. The second term is a fitting constraint, which means the ranking result should fit to the initial label assignment. With more prior knowledge about the relevance or confidence of

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 5, March 2017

each query, we can assign different initial scores to the queries. Minimizing the cost function respect to r results into the following closed form solution

$r* = (I_n-\alpha S)^{-1}y$ ….. (2)

where $\alpha = \frac{1}{1+\mu}$, I_n is an identity matrix with n×n, and S is the symmetrical normalization of W, $S=D^{-1/2}WD^{-1/2}$. In large scale problems, we prefer to use the iteration scheme:

$r(t+1) = \alpha Sr(t) + (1-\alpha)y$ …..(3)

During each iteration, each point receives information from its neighbors (first term), and retains its initial information (second term). The iteration process is repeated until convergence. When manifold ranking is applied to retrieval (such as image retrieval), after specifying a query by the user, we can use the closed form or iteration scheme to compute the ranking score of each point. The ranking score can be viewed as a metric of the manifold distance which is more meaningful to measure the semantic relevance.

3.2 ANALYSIS

Manifold ranking has been used in various applications at the same time it has its own disadvantages to handle large scale databases. The first step is to construct graph. The kNN is suitable for MR, which has the ability to capture data's local structure. The construction cost for kNN graph is very high and it is represented as $O(n^2 \log k)$, MR ranking use the adjacency matrix W in their computation. The storage cost of W is denoted as O(kn). So the basic need is to construct a graph in both low cost and small storage memory.

The second, MR has very expensive computational cost because of the matrix inversion operation in equation (2). It is the main problem to apply in large databases. We can use the iteration algorithm temporarily to solve this problem.

4. EFFICIENT MANIFOLD RANKING

We present the disadvantages of original MR from to views. They are scalable graph construction [5] and efficient ranking computation.

4.1 COLOR BASED IMAGE

RGB Color Space

The RGB components are widely used in electronic media like CRT screens, LCDs, or phones that transmit to produce light. By using the type of color receptors in human eye can detect millions of colors. The receptor can relate with the RGB components. The color component value is stored in a byte and the values range between 0 and 255.

HSL & HSV Color Space

The RGB color space to form a cube is a challenging one. To represent a cube or color wheel is very poor. During working process with millions of colors at that time tint of color can't be aligned properly on RGB color space.

Step 1: Convert RGB color space image into HSV color space.

Step 2: Color quantization is carried out using color histogram by assigning 8 level each to hue, saturation and value to give a quantized HSV space with 8x8x8=512 histogram bins.

Step 3: The normalized histogram is obtained by dividing with the total number of pixels.

Step 4: Repeat step1 to step3 on an image in the database.

Step 5: Calculate the similarity matrix of query image and the image present in the database.

Step 6: Repeat the steps from 4 to 5 for all the images in the database.

Step 7: Retrieve the images

4.2 TEXTURE BASED IMAGE

The "uniform texture" is not well defined and the segmentation of texture is hard. The different segmentations are to provide a plausible at different scales of resolution. Each leaf in a tree is segmented as one scale so that the whole tree-top or the whole forest are considered to be individual regions at coarser scales. The retrieval error occurs when the mismatch between query and image descriptors in terms of resolution scales.

Retrieval based on comparing texture segments is usually sensitive to over- and under-segmentation. The changes in spatial texture appearance can cause single textures to be split into smaller segments (over-segmentation). The segmentation algorithm combines small regions of different textures (under-segmentation). The significant size is combined together and scattered image creates the problems in texture images. So that the images can be lost. An example of this phenomenon is an aerial view of a town in a richly vegetated area, in which both buildings and vegetation are made up of numerous but small texture patches.

Step 1: Read the Query input image name.

Step 2: Convert color image into grayscale image.

Step 3: Divide the Query image into 16 by 16 blocks.

Step 4: Calculate the query image feature as mean, max and min.

Step 5: Finally compute the covariance of the Query image.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 5, March 2017

Step 6: Load Covariance features of the image database.

Step 7: Compare the difference between Query image feature and the features of the image database.

Step 8: Sort the distance of the features of the image database.

Step 9: Retrieves the closest distance images from the image database.

Step 10: Display the retrieval images in the output panel.

4.3 SHAPE BASED IMAGE

The purpose of edge detection is to reduce the amount of data in an image. To preserve the structural properties is used to extract image processing. The aim of JFC was to develop an algorithm that is optimal with regards to the following criteria:

1. *Detection*: The probability of detecting real edge points should be maximized while the probability of falsely detecting non-edge points should be minimized. This corresponds to maximizing the signal-to-noise ratio.

2. Localization: The detected edges should be as close as possible to the real edges.

3. *Number of responses*: One real edge should not result in more than one detected edge (one can argue that this is implicitly included in the first requirement).

With JFC's mathematical formulation of these criteria, Canny's Edge Detector is optimal for a certain class of edges (known as step edges). The images used throughout this worksheet are generated using this implementation.

Step 1: Read the Query input image name.

Step 2: Convert color image into grayscale image.

Step 3: Calculate the edge of the query input image.

Step 4: Compute the moment of the query input image.

Step 5: Integrate the moment features of given images row and column wise.

Step 6: Load moment features of the image database.

Step 7: Compare the difference between moment feature of the query image and moment features of the image database.

Step 8: Retrieve the closest difference images from the image database.

Step 9: Display retrieval images on the output panel.

4.4 EMR for CBIR

The EMR is implemented in CBIR to improve information and extend the data features. The extraction features first done with the low level features of images in the database. To utilize the data points in the graph. The representative points can be select as anchors and weight matrix denoted by Z-with a small neighborhood size s. The offline anchor selection doesn't affect the online process. We cannot frequently update the anchors for stable dataset. At the final stage of uploading images or query the user can specify and extract the low level features and update the weight matrix Z-directly in the ranking score.

4.5 Out of Sample retrieval

For the offline stage the data retrieval sample are constructed by using graph and the calculation are done by matrix inversion. The situation is handled differently for out of sample data. The main drawback of MR is hard to handle the new sample queries. The MR designed a fastest strategy for leaving the original graph unchanged and its used to add a new row and column to W (left picture). The new W is not used for the ranking process but it gives an efficient output. During the online stage new query is not acceptable because of its high computational cost.

The out of sample problem can be solved by listening the nearest neighbors query and that query can be used as a query points. If the database is static the query or not used in the graph. The query can be change initial semantic meaning in a large database. The cost wise is high for the linear search nearest neighbors

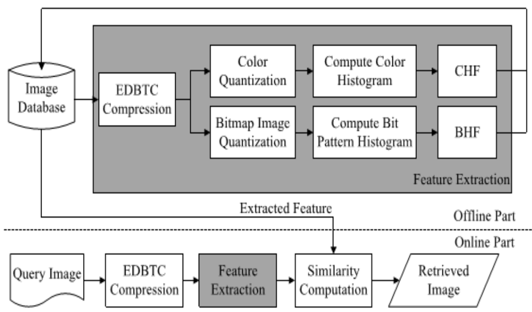
We use zt to denote the new column. Then, $D_t = z_t^t v$ and $h_t = z_t D_t^{-1/2}$, where h_t is the new column of H. As we have described, the main step of EMR is Eq.(11). Our goal is to further speedup the computation of Eq.(11) for a new query. Let

5. RESULT AND DISCUSSION

The average precision obtains the value for single query is by using Average Precision method and set of top k items to get the relevant item from the retrieved data. The existing systems does not concentrates on strong, content-specific geometric constraints among different visual words in an image. The large scale content based image cannot be retrieved from this method. The object of a graph-based ranking model has the priority to allocate the importance of a vertex, based on local or global information draw from the graph. The manifold ranking algorithm and make a detailed analysis about its drawbacks. MR can be used in large applications but it has some of the drawbacks such as to handle the large scale database.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 5, March 2017



In our proposed method overcome of run out-of-sample retrieval method for ranking scores. It produces an approximate result. The large scale databases can gives the result within a short time. The proposed method gives the outsource under the Corel [2] 100 and Corel 1000 databases. The data can be modelled by a weighted graph, and incorporated this graph structure into the ranking function as a regularizer. An automatically tag ranking scheme by performing a random walk over a tag similarity graph. A projected gradient based algorithm was proposed in to compute weight matrix and in our previous work, a kernel regression method was adopted.

REFERENCES

[1] Zhong Wu Tsinghua Univ., Ctr Adv Study “Scalable Face Image Retrieval with Identity-Based Quantization and Multi-Reference Re-ranking,” Asia Lab, 2012.

[2] C.P.Town AT&TLaboratoriesCambridge “Content Based Image Retrieval using Semantic Visual Categories”, Cambridge,England-2012.

[3] Behjat Siddiquie University of Maryland, College Park“ Image Ranking and Retrieval based on Multi-Attribute Queries” -2012.

[4] J. He, M. Li, H. Zhang, H. Tong, and C. Zhang, “Generalized manifold-ranking-based image retrieval,”IEEE Trans. Image Process., vol. 15, no. 10, pp. 3170–3177, Oct. 2006.

[5] W. Liu, J. He, and S. Chang, “Large graph construction for scalable semi-supervised learning,” in Proc. 27th Int. Conf. Mach. Learn.,Haifa, Israel, 2010, pp. 679–686 .

[6] Christo Ananth, S.Santhana Priya, S.Manisha, T.Ezhil Jothi, M.S.Ramasubhaeswari, “CLG for Automatic Image Segmentation”, International Journal of Electrical and Electronics Research (IJEER), Vol. 2, Issue 3, Month: July - September 2014, pp: 51-57

[7] J. Yu, D. Tao, and M. Wang, “Adaptive hypergraph learning and its application in image classification,” IEEE Trans. Image Process.,vol. 21, no. 7, pp. 3262–3272, Jul. 2012.

[8] W. Liu, J. Wang, and S. Chang, “Robust and scalable graph-based semisupervised learning,” Proc. IEEE, vol. 100, no. 9,pp. 2624–2638, Sept. 2012.

[9] D.Zhou,O.Bousquet,T.Lal,J.Weston,andB.Schölkopf, “Learning with local and global consistency,” inProc. Adv. NIPS, 2004, pp. 595–602.

[10] D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Schölkopf, “Ranking on data manifolds,”Proc.Adv.NIPS, vol. 16. Vancouver, BC, Canada, pp. 169–176, 2004.

[11] X. Wan, J. Yang, and J. Xiao, “Manifold-ranking based topic-focused multi-document summarization,” in Proc. 20th IJCAI, vol. 7. San Francisco, CA, USA, 2007, pp. 2903–2908.

[12] J. He, M. Li, H. Zhang, H. Tong, and C. Zhang, “Manifold-ranking based image retrieval,” inProc. 12th Annu. ACM Int. Conf. Multimedia, New York, NY, USA, 2004, pp. 9–16.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 5, March 2017

[13] B. Xu et al., "Efficient manifold ranking for image retrieval," in Proc. 34th Int. ACM SIGIR, New York, NY, USA, 2011, pp. 525–534.

[14] X. Yuan, X. Hua, M. Wang, and X. Wu, "Manifold-ranking based video concept detection on large database and feature pool," in Proc. 14th Annu. ACM Int. Conf. Multimedia, New York, NY, USA, 2006, pp. 623–626.

[15] J. Ponte and W. Croft, "A language modeling approach to infor-mation retrieval," nProc. 21st Annu. Int. ACM SIGIR, Melbourne, VIC, Australia, 1998, pp. 275–281.

[16] S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine,"Comput. Netw. ISDN Syst., vol. 30, no. 1–7, pp. 107–117, 1998.

[17] J. Kleinberg, "Authoritative sources in a hyperlinked environ-ment,"J. ACM, vol. 46, no. 5, pp. 604–632, 1999.

[18] J. Jeon, V. Lavrenko, and R. Manmatha, "Automatic image anno-tation and retrieval using cross-media relevance models," in Proc. 26th Annu. Int. ACM SIGIR, Toronto, ON, Canada, 2003, pp. 119–126.

[19] M. Tsai, T. Liu, T. Qin, H. Chen, and W. Ma, "FRank: A ranking method with fidelity loss," inProc. 30th Annu. Int. ACM SIGIR, Amsterdam, The Netherlands, 2007, pp. 383–390.

[20] W. Gao, P. Cai, K. Wong, and A. Zhou, "Learning to rank only using training data from related domain.

J.Rekha received her MSC(CS) degree from PSGR Krishnammal College for Women, Coimbatore, India in 2013. Now she is currently working toward the M.Phil degree at Sakthi Arts & Science College for women, Oddanchtram, India. Her main research interests include Image Processing, Network Security.

B. Muthu Kumar, Professor, Department of CSE, Syed Ammal Engineering College, Ramanathapuram, TN, India

Kavitha, Professor, Department of Computer Science, Sakthi College of Arts and Science for Women, TN, India

S.Prabhakaran, Department of Information Technology, Kathir College of Engineering, Now he is currently working in Ascox Techno Soft at Madurai, India. Him main research interests include Image Processing, Network Security, Cloud Computing.