New Content based Image Retrieval using JEC and Lasso

Samiksha Jain M.Tech Scholar Sati Vldisha Satish Pawar Associate Professor Sati Vldisha

ABSTRACT

In current scenario with growing technologies as well as enhancement in the digital world, it has found itself surrounded by a huge quantity of data or information. To handle such huge amount of data/images will often creates difficulties while retrieving the data or images efficiently. One feasible solution to overcome such difficulties is data retrieval technique. Image retrieval is the process of retrieving images from a large database of digital images dataset. proposed method is done by three primitive methods namely through color, shape and texture other term process data based on color ,size and texture. In this paper extract the image from database based on size texture and color. The technique by which used on proposed Wavelet transform Joint equal contribution and lasso. For proposed method is firstly take query image and extract feature by using DWT and other algorithm and then match that data with exiting database and find similarity data from the database. In this paper is analysis different author work and proposed novel method to get better retrieval rate.

Keywords

Content based image retrieval, Joint equal contribution, low level features, High level features, Color histogram

1. INTRODUCTION

We World is growing with digital technology, all information and data in form binary data or image today. Image is basic digital data for information and reorganization. Day by day is increase data resource so data retrieval is very big challenge. Data retrieval in general or content-based image retrieval (CBIR) in particular are very knowing area and challenging of research. For data information management a large number of methods have been proposed and investigated but in which still no satisfying general solutions exist today, trying many method to find solution of similar type of data but not get satisfactory output. The amount of information available nowadays, in digital format, to ordinary people is breath taking. The fast growth of the amount of digital material available can be explained by the new models of digital media production, distribution and consumption. Information creation and distribution were restricted to experts and media companies. This changed radically with the appearance of the Internet and the first web browser; which revolutionized the distribution of information. The ease of creating new Web documents and linking them to existing ones caused exponential growth of the publicly available digital material. Therefore, even if we assume that in a specific context we can have a fixed set of keywords to annotate the images, using one word or another relies on the subjective judgment of the person performing the task. Due to the rich content of images, different persons may perceive them differently and annotate them with different keywords. To overcome these difficulties, query by content or content-based retrieval has recently been proposed as an alternative to text-based retrieval for retrieving media such as images, audio and video. It indexes the media documents based on features extracted from their content rather than by textual annotations. Content-based image retrieval (CBIR) has been an active area of research, promising to provide powerful tools for database management in the near future. CBIR captured the interest of both industrial as well as the academic research communities. The interest is large enough that the Motion Picture Experts Group (MPEG) dedicated a standard called MPEG-7: the "Multimedia Content Description Interface" to allow the interoperability between the devices and applications attempting to solve parts of this problem. In past many author is work get similar information or image from large data set but they are not get proper retrieval rate. So need to find appropriate algorithm which provide exact output as as query image and giving higher retrieval data rate. The importance of this task has increased with the growth of the digital images collections. An important amount of digital pictures is generated each year and thus there is a need for an efficient image management system that is capable to fast searching, browsing by topic (e.g. using Google Picasa or tagging images. Content-based Image Retrieval (CBIR) has been studied for several years. A possible approach is to obtain a textual description from the image and then use text retrieval for searching. A different approach is to combine two modalities for example text and visual features when indexing images. Image retrieval based on text is sometimes called Annotation Based Image Retrieval (ABIR) [Inoue (2004)]. The systems based on ABIR can have some draw-backs. Researchers working in CBIR have identified two limitations. The first limitation is that ABIR requires manual image annotation which is time consuming and costly. The second limitation is that human annotation is subjective and sometimes it is difficult to describe image contents by concepts. An AIA system can solve the first limitation.

2. PROPOSED METHODOLOGY

All The methodology for assessing the retrieval performance of the optimized feature extraction technique in Content Based Image Retrieval algorithm. The idea behind contentbased retrieval is to retrieve, from a database, media items (such as images, video and audio) that are relevant to a given query. Several steps are needed for this. First, the features from the media items are extracted and their values and indices are saved in the database. Then the index structure is used to ideally filter out all irrelevant items by checking attributes with the user's query. Relevancy is judged based on the content of media items.



Figure 1 Block diagram of the content-based retrieval system

2.1 Feature Extraction

Feature extraction is one of the most important components in a content-based retrieval system. High-level semantics can be either annotated manually or constructed automatically from low-level features. In this chapter the general low-level visual features are described. In broad sense, features may be divided into low-level features (such as color, texture, shape, and spatial layout) and high-level semantics (such as concepts and keywords).

2.2 Color

Color is one of the most widely used visual attributes in image retrieval. Retrieval by color similarity requires using such models of color stimuli that distances in color space correspond to human perceptual distances between colors. Color histogram is the most commonly used presentation. The histogram reflects the statistical distribution, or the joint probability of the intensities of the three color channels. Color histogram is the most commonly used presentation. The color histogram is computed by discretizing the colors within the image and counting the number of pixels of each color. Color stimuli are commonly represented as points in threedimensional color spaces. Hue describes the actual wavelength of the color percept, saturation indicates the amount of white light present in the color and brightness (value) represents the intensity of the color.

2.3 Shape

The shape of an object is a binary image representing the extent of the object. Since the human perception and understanding of objects and visual forms relies heavily on their shape properties, shape features play a very important role in CBIR. In general the useful shape features can be divided into two categories, boundary-based and region-based. The former considers the shape being composed of a set of two-dimensional regions, while the latter presents the shape by its outline. Region-based feature vectors often result in shorter feature vectors and simpler matching algorithms. The idea in multi-resolution techniques is to decompose a planar curve contour into components at different scales so that the coarsest scale components carry the global approximation information while the finer scale components contain the local detail information.

2.4 Texture

Texture as an area containing variations of intensities, which form repeated patterns. Differences observed by visual inspection are difficult to define in quantitative manner, which leads to the necessity of defining texture using some features. Texture is a property of area, and therefore, its definition must involve gray values in a spatial neighborhood. The size of the neighborhood depends on the texture type, or the size of the primitives defining the texture. There are several properties, such as coarseness, contrast, and directionality, which play an important role in describing texture. Texture analysis is important, since texture is useful in various applications such as automated inspection, medical image processing, remote sensing, defect detection and, similarity evaluation.

Image retrieval is perform different section

- 1. Design dataset
- 2. Train Data set
- 3. Retrieval image

The feature extraction is an important step for improving the retrieval performance in any Content Based Image Retrieval system. It is performed on features by several ways to get information about the content of an image. The CBIR is also called Visual Information Retrieval. Content based image retrieval is categorizes features according to the levels which describe the contents of an image as:

- Low level features (human vision related)
- Middle level features (object related)
- High level features (semantic related)

Among them, Low level is widely used in image retrieval for its simplicity as compared to other feature. Low level features describe human vision perspective contents of an image such as color, texture and shape. so we used wavelet and jec for selection of feature.

Algorithm

procedure SEARCH(queryImg,index,dataset)

queryImg_feature=CBIR(queryImg)

for each i in index do

distance=compairdata(feature,queryImg_feature)

if distance<= threshold then

Add +same match result

end If

End For

return Results

end Procedure

3. RESULT ANALYSIS

It is widely acknowledged that image retrieval is an open and very difficult problem in computer vision. Solving this problem at the human level may, perhaps, require that the problem of scene understanding be solved first. However, identifying objects, events, and activities in a scene is still a topic of intense research with limited success.

In the absence of such information, most of the image retrieval methods have suggested modeling the joint distribution of feature and images to learn the association of image data and low-level image features such as color and texture. Most of these state-of-the-art techniques require elaborate modeling and training efforts. The goal of our work was to develop a new retrieval method but create a family of very simple and intuitive baseline methods for image retrieval.

Comparing existing retrieval techniques with the proposed baseline methods helps us better understand the utility of the elaborate modeling and training steps employed by the existing techniques. Our proposed baseline methods combine distance measures over very global color and texture features. Furthermore, a combination of the Euclidian distances (JEC), or a combination trained on noisy labeled data (Lasso), outperforms the best state-of-the-art methods on three different datasets. These somewhat surprising results make a case for revisiting the state-of-the-art methods and carefully analyzing their different modeling and training steps to understand why they fail to achieve performance at the level of these simplistic baseline methods.

Given the general performance level of current retrieval methods as well as our proposed baselines, it is clear there is much room for improvement in the state-of-the-art. Our hope is that the existence of such baseline methods as proposed in this work will spur the development of more powerful retrieval techniques in the future by providing an effective evaluation platform.

In figure 3 is show proposed architecture which having 7 different section to retrieve image from data set. First section is data base creation and loading section, it is used to design database here we collect the image and train as show on figure 4. We also load our desire database for retrieval as show on figure 5. Second section is design to select image for processing, it is user section to select image easy.

Third section is showing processing image as show image 6. This three is initial section, after it 4th section is preprocessing section of image. Here is preprocess of data we change color and also apply histogram to eliminate of noise as show on figure 7 and figure 8. The extract feature like size, color and texture as show on figure 9. After is extracting all feature we collect entire value and calculated distance by using JEC and lasso of all data with dataset image and arrange it and showing better match of retrieval image as figure 10.



Figure.3. Final Project Model



Figure.4.: Select Data Set



Figure.5. Load Data set

International Journal of Computer Applications (0975 – 8887) Volume 175– No. 21, September 2020



Figure.6. Select Image



Figure.7. Apply Median Filter



Figure.8. Apply Histogram equalization



Figure.9. Apply Select Feature Extraction



Figure 10. Result Retrieved Image As part of Image

Comparision table on the bases of Previous work and our proposed work:

Method	Precision(P%)	Recall(R%)
MBRM	24	23
CRM	16	19
SML	23	29
Proposed	80	32

Table 1: Result Comparison Table



Figure 11: Comparison Previous method and proposed

Here P% and R% denote the mean precision and mean recall values. The table is based on the mean precision - recall value of some retrieval images using Lasso, Gabor and JEC. Lasso, Gabor and JEC perform significantly better than most of the published methods. Here we have compared our result with three other method to clarify the efficiency of occupied result.

This observation table is on the bases of precision recall values of retrieved images. Here precision is denoting as true positive value and recall is denoting false positive value.

While finding the precision and recall values :

For first instance the true positive values of retrieval is 80 and false positive value is 32, so like wise for every instance of search query we get these certain values and according to these values we have plotted a graph.

Some more result



(a)Rose



(b)Duck



(c) Dinosaur



(d)Sun Rise



(e)Hill

Figure 12: Some more example

Evaluation of Image Retrieval Results:

Table 2: Observation Table of graph

	Precision	Recall
Bus	80	32
Hill Station	76	38
Butterfly	83	44
Flower	87	37
Horse	68	44
Sunrise	67	47
Duck	73	43
Dinosaur	78	34
Tiger	81	37



Figure 13: Precision and Recall Graph

4. CONCLUSION

The task of automatic image retrieval is of great interest because it can play a crucial role in building an effective engine for image annotation. Image retrieval is main part of data acknowledgement and searching of a particular images from a large database. In this work HSV histogram apply to separate the color. Gabor wavelet extract feature and with the help of JEC and L1, L2 we get best retrieved image of the data set. In the proposed system we retrieve the relevant images from the large collection of images on the bases of texture, shape and color. By using the combination of Gabor, Lasso and JEC we find the increased efficiency in retrieval of images.

5. REFERENCES

- Ganar, Apurva N., C. S. Gode, and Sachin M. Jambhulkar. "Enhancement of image retrieval by using colour, texture and shape features" In Electronic Systems, Signal Processing and Computing Technologies (ICESC), 2014 International Conference on, pp. 251-255. IEEE, 2014.
- [2] Pujari, Jagadeesh, and P. Hiremath. "Content based image retrieval based on color texture and shape features using image and its complement." International Journal of Computer Science and Security 1, no. 4 (2007): 25-35.
- [3] Huang, Zhi-Chun, Patrick PK Chan, Wing WY Ng, and Daniel S. Yeung. "Content-based image retrieval using color moment and gabor texture feature" In Machine Learning and Cybernetics (ICMLC), 2010 International Conference on, vol. 2, pp. 719-724. IEEE, 2010.
- [4] Kekre, H. B., Ms Tanuja K. Sarode, and Sudeep D. Thepade. "Image retrieval using color-texture features from DCT on VQ codevectors obtained by Kekre's fast codebook generation." ICGST-International Journal on Graphics, Vision and Image Processing (GVIP) 9, no. 5 (2009): 1-8.
- [5] Bama, B. Sathya, S. Mohana Valli, S. Raju, and V. Abhai Kumar. "Content based leaf image retrieval (CBLIR) using shape, color and texture features." Indian Journal of Computer Science and Engineering 2, no. 2 (2011): 202-211.

- [6] Iqbal, Kashif, Michael O. Odetayo, and Anne James. "Content-based image retrieval approach for biometric security using colour, texture and shape features controlled by fuzzy heuristics." Journal of Computer and System Sciences 78, no. 4 (2012): 1258-1277.
- [7] Wang, Xiang-Yang, Yong-Jian Yu, and Hong-Ying Yang. "An effective image retrieval scheme using color, texture and shape features." Computer Standards & Interfaces 33, no. 1 (2011): 59-68.
- [8] Murala, Subrahmanyam, R. P. Maheshwari, and R. Balasubramanian. "Local tetra patterns: a new feature descriptor for content-based image retrieval." IEEE Transactions on Image Processing 21, no. 5 (2012): 2874-2886.
- [9] Kekre, Dr HB, Sudeep D. Thepade, Priyadarshini Mukherjee, Shobhit Wadhwa, Miti Kakaiya, and Satyajit Singh. "Image retrieval with shape features extracted using gradient operators and slope magnitude technique with BTC." International Journal of Computer Applications 6, no. 8 (2010).
- [10] Kekre, H. B., Sudeep D. Thepade, Tanuja K. Sarode, and Vashali Suryawanshi. "Image Retrieval using Texture Features extracted from GLCM, LBG and KPE" International Journal of Computer Theory and Engineering 2, no. 5 (2010): 695.
- [11] Metzler, D., Manmatha, R.: An inference network approach to image retrieval. In: Image and Video Retrieval, Springer (2005) 42–50
- [12] Hare, J.S., Lewisa, P.H., Enserb, P.G.B., Sandomb, C.J.: Mind the gap: Another look at the problem of the semantic gap in image retrieval. Multimedia Content, Analysis, Management and Retrieval (2006).
- [13] Frome, A., Singer, Y., Sha, F., Malik., J.: Learning globally-consistent local distance functions for shapebased image retrieval and classification. In: Proceedings of the IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil. (2007)
- [14] Tibshirani, R.: Regression shrinkage and selection via the Lasso. J. Royal Statistical Soc., B 58 (1996) 267–288
- [15] Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: Ideas, influences, and trends of the new age. ACM Computing Surveys (2008)
- [16] Mori, Y., Takahashi, H., Oka, R.: Image-to-word transformation based on dividing and vector quantizing images with words. In: First International Workshop on Multimedia Intelligent Storage and Retrieval Management (MISRM). (1999)
- [17] Jin, R., Chai, J.Y., Si, L.: Effective automatic image annotation via a coherent language model and active learning. In: ACM Multimedia Conference. (2004) 892– 899
- [18] D. Zhang, Md. M. Islam, G. Lu, 2012. "A review on automatic image annotation techniques", Pattern Recognition, vol. 45, no. 1,pp.346–362.