Image Retrieval Based On Feature Extraction

Gavara Jhansi Durga Gayathri
PG scholar, Department of MCA, CDNR collage, Bhimavaram, Andhra Pradesh.
B.S.Murthy
(Assistant Professor), Master of Computer Applications, DNR collage, Bhimavaram, Andhra Pradesh.

ABSTRACT The Image retrieval is progressively turning into a fascinating field of research as the pictures that clients store and procedure continue rising both in number and size particularly in advanced databases. The pictures are put away on convenient gadgets which clients have used to catch these pictures. The point of this project is to tackle the issues experienced by clients in picture recovery of advanced pictures put away in their gadgets, guaranteeing that pictures mentioned are recovered precisely from storage. The pictures are pre-processed to evacuate noise and pull together pictures to upgrade image content. The picture recovery depends on the substance (Color Based Image Retrieval) where pictures are coordinated in a database dependent regarding the matter color of the picture. Pictures are put in classes and pictures are recovered dependent on the client's input. Chi-square distance technique is utilized to decide the closest items, in this manner bringing about minimal number of pictures recovered by the framework. Color and surface highlights are utilized to produce the element lattices on which the picture correlation is made. For KNN calculation, various estimations of K will be tried to decide best an incentive for various classes of pictures. The outcomes got to show that the mix of shading, surface, and KNN in picture recovery brings about shorter calculation time when contrasted with the accurate output.

1. INTRODUCTION

1.1 MOTIVATION

Due to recent development in technology, there is an increase in the usage of digital cameras, smartphone, and Internet. shared and stored multimedia data are growing, and to search or to retrieve a relevant image from an archive is a challenging research problem fundamental need of any image retrieval model is to search and arrange the images that are in a visual semantic relationship with the query given by the user. Most of the search engines on the Internet retrieve the images on the basis of text-based approaches that require captions as input.

The user submits a query by entering some text or keywords that are matched with the keywords that are placed in the archive. The output is generated on the basis of matching in keywords, and this process can retrieve the images that are not relevant. difference in human visual perception and manual labeling/annotation is the main reason for generating the output that is irrelevant. It is near to impossible to apply the concept of manual labeling to existing large size image archives that contain millions of images. The second approach for image retrieval and analysis is to apply an automatic image annotation system that can label image on the basis of image contents.

The approaches based on automatic image annotation are dependent on how accurate a system is in detecting color, edges, texture, spatial layout, and shape-related information. Significant research is being performed in this area to enhance the performance of automatic image annotation, but the difference in visual perception can mislead the retrieval process. Content-based image retrieval (CBIR) is a framework that can overcome the abovementioned problems as it is based on the visual analysis of contents that are part of the query image.

1.2 PROBLEM DEFINITION

As the advances in computerized photography, storage limit and systems speed, putting away a lot of top-notch pictures has been made conceivable. Computerized pictures are utilized in a wide scope of utilizations, for example, therapeutic, virtual galleries, military and security purposes, and individual photograph collections. Be that as it may, clients experience issues in arranging and looking through huge quantities of pictures in databases, as the present business database frameworks are intended for content information and not appropriate for computerized pictures. Along these lines, an effective route for picture recovery is needed.

A picture recovery framework is a PC framework for perusing, looking and recovering pictures from a huge database of computerized pictures. Picture recovery is classified into two kinds of recovery are Text-Based Image Retrieval and Content-Based Image Retrieval. Content-Based Image Retrieval is having bad marks of proficiency, loss of data, progressively costly undertaking and tedious.

Color Feature is a most normal element of the picture. The shading pictures are having the standard Color is RGB shading. Color histograms are ordinarily utilized content-based picture recovery. Highlight implies the qualities of the article. Feature extraction has alluded that dimensionality decrease of that item. It assumes a significant job in picture handling.

Features are characterized into three sorts in picture preparing, that is low, center and high. Low level highlights are shading, surface and center level element is shape and elevated level component is semantic hole of items. These three Features are extricated by various methods. Color Feature is extricated by Color Histogram and Color correlogram.

Content-based image retrieval (CBIR) is a framework that can overcome the abovementioned problems as it is based on the visual analysis of contents that are part of the query image. To provide a query image as an input is the main requirement of CBIR and it matches the visual contents of query image with the images that are placed in the archive, and closeness in the visual similarity in terms of image feature vector provides a base to find images with similar contents.

In CBIR, low-level visual features (e.g., color, shape, texture, and spatial layout) are computed from the query and matching of these features is performed to sort the output. According to the literature, Query-By-Image Content (QBIC) and Simplicity are the examples of image retrieval models that are based on the extraction of low-level visual semantic. After the successful implementation of the abovementioned models, CBIR and feature extraction approaches are applied in various applications such as medical image analysis, remote sensing, crime detection, video analysis, military surveillance, and textile industry. It provides an overview of the basic concepts and mechanism of image retrieval . The basic need for any image retrieval system is to search and sort similar images from the archive with minimum human interaction with the machine.

According to the literature, the selection of visual features for any system is dependent on the requirements of the end user. The discriminative feature representation is another main requirement for any image retrieval system. To make the feature more robust and unique in terms of representation fusion of low-level visual features, high computational cost is required to obtain more reliable results. However, the improper selection of features can decrease the performance of image retrieval model. The image feature vector can be used as an input for machine learning algorithms through training and test models and it can improve the performance of CBIR.

1.3 OBJECTIVE

A machine learning algorithm can be applied by using training-testing (either through supervised or through unsupervised) framework in both cases. The recent trends for image retrievals are focused on deep neural networks that are able to generate better results at a high computational cost.

In this paper, we aim to provide a compressive overview of the recent research trends that are challenging in the field of CBIR and feature representation. The basic objectives of this research study are as follows:

(1) How the performance of CBIR can be enhanced by using low-level visual features?
(2) How semantic gap between the low-level image representation and high-level image semantics can be reduced?
(3) How important is image spatial layout for image retrieval and representation?
(4) How machine learning-based approaches can improve the performance of CBIR?

In this review, we have conducted a detailed analysis to address the abovementioned objectives. The recent trends are discussed in detail by highlighting the main contributions, and upcoming future challenges are discussed by keeping the focus of CBIR and feature extraction. The structure of the paper is about color feature, texture features, shape features, spatial features, low-level feature fusion, local feature, commonly used dataset for CBIR and overview to basic machine learning techniques, Machine learning-based CBIR.

2. LITERATURE SURVEY
2.1. CBIR OVERVIEW

Initially it working based on text of image. Text Based Image Retrieval is having demerits of efficiency, loss of information, more expensive task and time consuming. Overcome these problems by using Content Based Image Retrieval (CBIR) system for image retrieval. CBIR working by using features of the content of the image is known as Content-Based Image Retrieval (CBIR). The image retrieval system acts as a classifier to divide the images in the image database into two classes, either relevant or irrelevant. In this sense, an annotated image can be represented by a feature vector x, e.g., a set of image features or eigen features, and its label y that is either relevant or irrelevant and without label image is called unannotated or unlabeled image. These images are not considered during retrieval process.
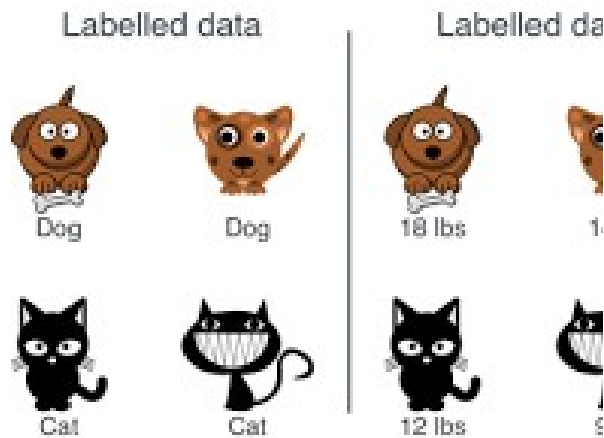
Unlabeled image is most useful to getting the better result of retrieval. Unlabeled images are identified when classifying the image database by using classifiers. The classifiers are working based on various classification properties of image database. The begin stage of image database classification is supervised.

The supervised classification is finding the labeled images very well. The unsupervised classification is used to finding the unlabeled images very well. Content based image retrieval is working with different types of image database. All databases are having two types of images like labelled and unlabeled.

A. **Example of Labelled and Unlabeled images Fig 1,2 mentioned below :**

clouds, sky, building, people   without label of image

Labeled image   Label for image   Unlabelled image

```
query: shire-query.png
    1. images\Shire-004.png : 1.0773438436122535
    2. images\Shire-003.png : 1.113796675484044
    3. images\Shire-001.png : 1.2779495715356912
    4. images\Shire-002.png : 1.3763701229130474
    5. images\Shire-005.png : 1.778740775651079
    6. images\Rivendell-001.png : 1.8218286251445748
    7. images\Rivendell-004.png : 2.077383152425639
    8. images\Rivendell-002.png : 2.1462431938842412
    9. images\Golbin-005.png : 2.1696694308870828
    10. images\Goblin-001.png : 2.1979721168253947
```

B. Image Database

Content Based Image Retrieval systems are working with various type of image file formats and also different types of size, themes, features, information, etc. The image file formats are

- TIFF 6.0 (Tagged Image File Format)
- GIF 89a (Graphics Interchange Format)
- JPEG (Joint Photographic Expert Group)
- JFIF (JPEG File Interchange Format)
- JP2-JPX
- JPEG 2000
- ImagePac, Photo CD
- PNG 1.2 (Portable Network Graphics)
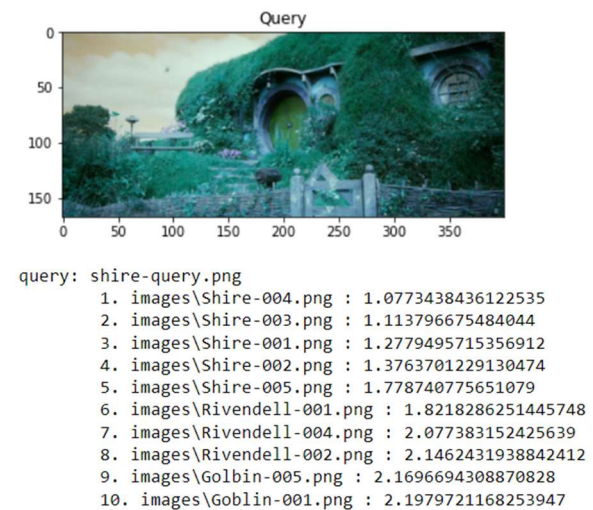- PDF 1.4 (Portable Document Format)

6.Results

Input Image :

Figure 13 : Input
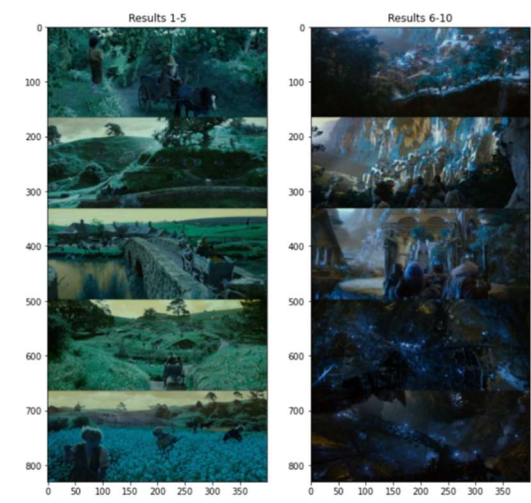
Chi Square Distances :

Output Image:

Figure 14 : Output

7. Conclusion & Future Work

Content based image Retrieval System is a process to find the similar image in image database when query image is given.

In this paper we presented CBIR system using feature classification with Modified K-Nearest Neighbor Algorithm. CBIR system based on visual feature of image that color of image

enables to classify the image.

In the future we will next extend the CBIR system used to implement interesting visual features of image and providing label with help of MKNN classification by using the user input. We will also improve the performance of CBIR system using MKNN image classification algorithm.

8. References

[1] D. Zhang, M. M. Islam, and G. Lu, "A review on automatic image annotation techniques," Pattern Recognition, vol. 45, no. 1, pp. 346–362, 2012.

[2] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," Pattern Recognition, vol. 40, no. 1, pp. 262–282, 2007.

[3] T. Khalil, M. U. Akram, H. Raja, A. Jameel, and I. Basit, "Detection of glaucoma using cup to disc ratio from spectral domain optical coherence tomography images," IEEE Access, vol. 6, pp. 4560–4576, 2018.

[4] S. Yang, L. Li, S. Wang, W. Zhang, Q. Huang, and Q. Tian, "SkeletonNet: a hybrid network with a skeleton-embedding process for multi-view image representation learning," IEEE Transactions on Multimedia, vol. 1, no. 1, 2019.

[5] W. Zhao, L. Yan, and Y. Zhang, "Geometric-constrained multi-view image matching method based on semi-global optimization," Geo-Spatial Information Science, vol. 21, no. 2, pp. 115–126, 2018.

[6] W. Zhou, H. Li, and Q. Tian, "Recent advance in contentbased image retrieval: a literature survey," 2017, https://arxiv. org/abs/1706.06064.

[7] A. Amelio, "A new axiomatic methodology for the image similarity," Applied Soft Computing, vol. 81, p. 105474, 2019.

[8] C. Celik and H. S. Bilge, "Content based image retrieval with sparse representations and local feature descriptors: a comparative study," Pattern Recognition, vol. 68, pp. 1–13, 2017.

[9] T. Khalil, M. Usman Akram, S. Khalid, and A. Jameel, "Improved automated detection of glaucoma from fundus image using hybrid structural and textural features," IET Image Processing, vol. 11, no 9, pp. 693–700, 2017.

[10] L. Amelio, R. Jankovi´c, and A. Amelio, "A new dissimilarity measure for clustering with application to dermoscopic images," in Proceedings of the 2018 9th International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 1–8, IEEE, Zakynthos, Greece, July 2018.

[11] S. Susan, P. Agrawal, M. Mittal, and S. Bansal, "New shape descriptor in the context of edge continuity," CAAI Transactions on Intelligence Technology, vol. 4, no. 2, pp. 101–109, 2019.

[12] L. Piras and G. Giacinto, "Information fusion in content based image retrieval: a comprehensive overview," Information Fusion, vol. 37, pp. 50–60, 2017.

[13] L. Amelio and A. Amelio, "Classification methods in image analysis with a special focus on medical analytics," in Figure 11: Example of top-5 to top-25 image retrieval results on the basis of search by query image. 18 Mathematical Problems in Engineering Machine Learning Paradigms, pp. 31–69, Springer, Basel, Switzerland, 2019.

[14] D. Ping Tian, "A review on image feature extraction and representation techniques," International Journal of Multimedia and Ubiquitous Engineering, vol. 8, no. 4, pp. 385–396, 2013.

[15] D. Zhang and G. Lu, "Review of shape representation and description techniques," Pattern Recognition, vol. 37, no. 1, pp. 1–19, 2004.

[16] R. Datta, J. Li, and J. Z. Wang, "Content-based image retrieval: approaches and trends of the new age," in Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval, pp. 253–262, ACM, Singapore, November 2005.

[17] Z. Yu and W. Wang, "Learning DALTS for cross-modal retrieval," CAAI Transactions on Intelligence Technology, vol. 4, no. 1, pp. 9–16, 2019.

[18] N. Ali, D. A. Mazhar, Z. Iqbal, R. Ashraf, J. Ahmed, and F. Zeeshan, "Content-based image retrieval based on late fusion of binary and local descriptors," International Journal of Computer Science and Information Security (IJCSIS), vol. 14, no. 11, 2016.

[19] N. Ali, Image Retrieval Using Visual Image Features and Automatic Image Annotation, University of Engineering and Technology, Taxila, Pakistan, 2016.

[20] B. Zafar, R. Ashraf, N. Ali et al., "Intelligent image classification-based on spatial weighted histograms of concentric circles," Computer Science and Information Systems, vol. 15, no. 3, pp. 615–633, 2018.

[21] W. Zhao, L. Yan, and Y. Zhang, "Geometric-constrained multi-view image matching method based on semi-global optimization," Geo-Spatial Information Science, vol. 21, no. 2, pp. 115–126, 2018.

[22] W. Zhou, H. Li, and Q. Tian, "Recent advance in contentbased image retrieval: a literature survey," 2017, https://arxiv.org/abs/1706.06064.