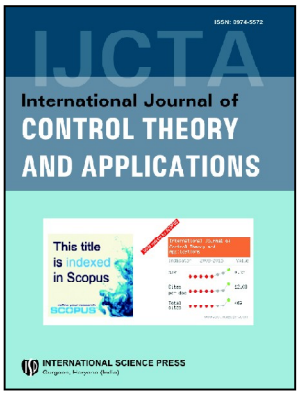

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 10 • 2017

Bi-Feature Descriptor for the Large Datasets based with Probabilistic SVM Classification

Komal Preet Kaur[1] and Mandeep Singh[2]

[1] *M. Tech Scholar, Department of computer science engineering, Chandigarh University , Mohali, India*
E-mail: Komalpreet1990@gmail.com
[2] *Associate Professor, Department of Computer Science Engineering, Chandigarh University , Mohali, India*
E-mail: mandeeptinna@gmail.com

Abstract: The image retrieval systems are used to retrieve the image out of the stored image dataset by analyzing and comparing the query image against the training dataset. The image retrieval algorithms use the various kinds of the feature descriptors based upon the texture, color or patterns using the feature descriptors such as histogram of gradients, scale invariant feature transform, speeded up robust features, laplacian of Gaussians, etc. In this paper, the proposed model has been designed with the bi-feature descriptor basd upon the color and low-level feature descriptors, which has been realized using the histogram of gradients (HoG) along with the laplacian of Gaussians (LoG). The bi-feature scheme in the proposed model has been applied over the support vector machine based classification to achieve the high robustness during the classificationprocess. The laplacian of Gaussian has been found suffering from the null features, which causes the mis-similarity between the two features, which has been solved in the proposed model. The proposed model accuracy has been measured in the various performance parameters such as specificity, sensitivity, positive predictive value and prevalence for its evaluation against the existing models. The experimental results have justified the performance of the proposed model against the existing model by minimum 10 percent on each parameter evaluated under the performance evaluation.

Keywords: Bi-features, Laplacian of Gaussian, LoG, efficient information retrieval.

INTRODUCTION

Image classification is associated with the semantic indexing of the input images for the retrieval of the desired image out of the input data. It also becomes important to describe the features, classify the image and to rank the image to run the CBIR efficiently. From application point of view, image classification methods become very useful in-content based image retrieval. [1-2, 8, 11] The accurate CBIR image classification method helps in better organization for the image retrieval application. [10] Image classification is highly valuable in automatic navigation to the new dataset while searching the query image on the CBIR system. [12]

In general any image contains useful and unwanted data. The system must differentiate between the each. By considering the below image wherever the person reading a book is that the helpful information and the

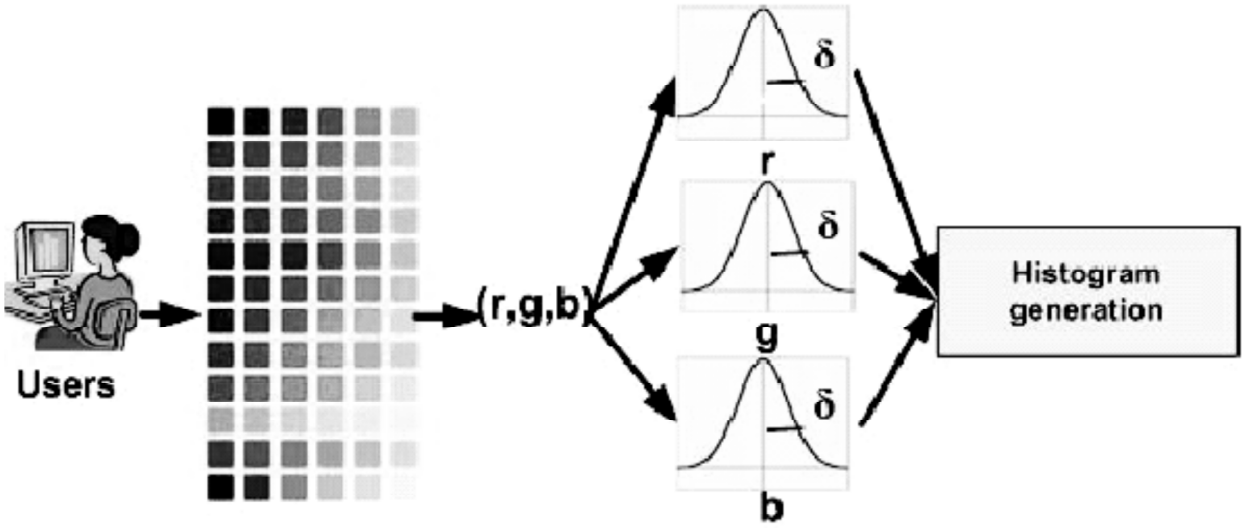

Figure 4.1: Color Histograms

background, where people and the market is the unwanted information. The system must group along the continual pattern to spot the objects within the image. as an example below is given the array for the a part of the shirt and this pattern is continual again.

The low-level image features of the image have been used for the evaluation of the images against the query image. The color, intensity, texture and pattern features are popular among feature descriptors for the images. [9] The images can be classified in the various categories and the query image can be classified prior to the image matching and ranking evaluation. [7] In the case of BIR, we can define multi-class data to classify an image in, low-level features are sufficient. [5] The studies on CBIR have shown that when number of classes increase in case of image categorization, these low-level features show promising results. But image reorganization in the CBIR has still remained a challenging task in computer vision due to severe intra-class similarities and inter-class variations. [5-6]

Image retrieval is cultured with many low level image features as the matching objects. So image retreival techniques simply based on color, texture and intensity are not very effective to classify indoor scenes. Pioneering works used SIFT [5], GIST [18] etc in combination with supervised learning. But these techniques fail to distinguish many types of image in the CBIR data. Bridging the semantic gaps becomes very important for the image classification and image representation for the purpose of accurate image recognition for the CBIR operations. [18]

One way to bridge semantic gap between image representation and image recognition is to make use of more and more sophisticated models, but good learning and inference is extremely difficult task for such models. [14] Alternatively semantic gap between low-level features like color, intensity, texture etc. and high-level category label can be reduced by introducing object-based representation as intermediate representation. [7, 13] As the performance of image recognition is heavily dependent on feature representation, this object-based intermediate representation proves to be useful in enhancing classification results. [17] Recently objects-based techniques for image classification have proven to be showing promising performance over other state-of-art techniques. [14]

In this work, we will review the recent and significant techniques that have been used for image retrieval and recognition. [15-16, 20] Besides we will identify the key approaches being used in image classification for

CBIR. The major contributions made by each significant work and the challenges posed to efficient classification will also be discussed. [19]

LITERATURE REVIEW

Wang *et al.* [23] represented that a solution for hierarchal single-keyword search relating to to certain relevancy score. For the first time this paper outline and solve the matter of effective however secure hierarchal keyword search over encrypted cloud information. hierarchal search greatly enhances system usability by returning the matching files in a hierarchal order relating to to certain relevancy criteria.

Ren *et al.* [4] suggested that similar secure per-file index, wherever an index including trapdoors of all unique words is built for every file. Here, the author's define many important securities challenges and inspire more investigation of security resolution for a trustworthy public cloud atmosphere. Penatti, Otávio AB et. [15] al. has performed a comparative study universal color and texture descriptors for internet image retrieval. This paper presents a comparative study of color and texture descriptors considering the online as the atmosphere of use. The authors have taken under consideration the range and large-scale aspects of the online considering a huge variety of descriptors (24 color and 28 texture descriptors, as well as each standard and recently proposed ones). The analysis is formed on two levels: a theoretical analysis in terms of algorithms complexities and an experimental comparison considering efficiency and effectiveness aspects. ibrahim *et al.* [11] concluded after making some experiments that to guard the privacy, users got to encrypt their sensitive information before outsourcing it to the cloud. However, the standard encoding schemes are inadequate since they create the appliance of categorization and looking operations tougher tasks. consequently, searchable coding systems are developed to conduct search operations over a group of encrypted knowledge. sadly, these systems solely enable their shoppers to perform a particular search however not approximate search, a vital need for all the present information retrieval systems. Xia *et al.* [27] represented that the results may come back not solely the precisely matched files, however also the files together with the terms semantically associated with the question keyword. within the proposed model, a corresponding file information is built for every file. Then both the encrypted information set and file collection are uploaded to the cloud server. With the information set, the cloud server builds the inverted index and constructs semantic relationship library (SRL) for the keywords set. when receiving a question request, the cloud server 1st finds out the keywords that are semantically associated with the question keyword according to SRL. Cao *et al.* and yang *et al.* [28] proposed a algorithm that scheme for multi-keyword stratified search, wherever "Inner product similarity" is employed for result ranking. This paper for the first time defines and solves the difficult issue of privacy conserving multi-keyword ranked search over encrypted cloud knowledge. Ahmad *et al.* [1] described the detailed view of texture, color and shape descriptors for CBIR. A comparison of different challenges in this field is discussed. There is an exponential growth of images around and the information which they carry needs to be effectively utilized. So the authors have extracted the low level features from the image such as texture, shape and color.

EXPERIMENTAL DESIGN

Feature Detection: The process can be started with keypoint detection. For this, a Difference of Gaussians pyramid can be constructed by subtracted Gauss-filtered images where the standard deviation ó differs by factor k:

$$D(x, y) = L(x, y, k\sigma) - = L(x, y, \sigma)$$

Next, extrema can be detected inDoG images which is compared each pixel to 8 neighboring pixels and nine pixels in the scales above and below. If a pixel value is larger or smaller than all of the perfect neighbors, the feature rich key point candidate is determined as the preliminary candidate.

Algorithm 1: Laplacian of Gaussian based Scale Invariant Feature Transform (LoG-SIFT)

1. Map the input image matrix by analyzing the matrix data in the omni-directional blocking and return the matrix I
2. Then perform the scale & space representation practice using the hybrid matrix representation over the input matrix I and return RI
3. Compute the laplacian of the given block
4. Compute the local externa factor by utilizing the spherical and super block LoG features
5. Run the iteration for each image block
   a. Compute the Local Spin Density (LSD) correlation over the current block
   b. Apply the exchange function for gradient-correction
   c. Return the evaluated feature Ei

The workflow of spherical SIFT algorithm can be summarized in Algorithm 1. Each one of the steps is defined in details in following sections. In this paper, spherical image can be defined in a (θ, ϕ)-grid where columns are points of constant longitude, $\phi \in [0, 2\pi)$,and rows are points of constant latitude, $\theta \in [0, \pi]$. The detector window is tiled with a grid of overlapping blocks in which Laplacian of Oriented Gradient (LoG) feature vectors are extracted.

The linear SVM based been utilized for the purpose of the CBIR classification in this research. The non-overlapping sliding window function scans the image data by analyzing all of the image areas in order to understand the image data scales and suppression based non-minimized factors for the overall computation of the output pyramid for the detection of the objects in the versatile instances. In this paper, we have focused upon the unique LoG features to add the additional robustness to the proposed model design. The proposed model classification works on the basis of the following algorithm flow:

Algorithm 2: SVM based classification

1. Load database in the Mat lab workspace
2. Resize the image according to the smaller sized image
3. Convert image from RGB to Gray
4. Normalize the gray image for fixed mean
5. Generate the histogram of RGB
6. Find entropy, standard deviation and local range of Gray
7. Combine the image feature
8. Load the test image
9. Apply the procedure 2-7 to find combine feature of test image
10. Determine the normalized Euclidean distance of test image with stored image of database
11. Sort the normalized Euclidean distance values to perform indexing.
12. Display the desired number of results on Results Window using indexing array

RESULT ANALYSIS

Assuming the positive evaluation set is a trustworthy representation of real world images in the content based image retrieval dataset and arranged in the form of different sized images which includes the grayscale, colored or binarized images. The image retrieval dataset has been obtained from Professor Wang's official blog and contains multiple category and multi-color scale images classified in the different categories of beach, monuments, urban transport, digital and tribal images, whereas different numbers of images have been kept under the image

different categories to form the final training image set. The proposed multi-stage hierarchical CBIR image dataset has been defined with the sensitive image data to test the real-time accuracy of the CBIR system. The proposed model has been found accurate with the accuracy level of more than 85% in comparison with the existing models using the different feature descriptors and classifiers.

Table 1
The table of properties calculated from the hypothesis of statistical type-1 and type-2 errors on the Category-A testing dataset

Property Name	*Proposed Model*
Positive	113 (TP 106, FP 7)
Negative	18 (TN 5, FN 13)
Sensitivity	89.08%
Specificity	41.67%
Positive Likelihood Ratio	1.53
Negative Likelihood Ratio	0.26
Prevalence	90.84%
Positive Predictive Value	93.81%
Negative Predictive Value	27.78%

131 images from various categories among the test set has been obtained for the testing of the CBIR system from the given training dataset of almost 1000 images of different categories. The training data has been used for testing the fully automated proposed image recognition and retrieval system. Some of the image taken under the consideration are the clean and noise free images, whereas some of the images with the different noise levels has been shortlisted for the testing of the proposed CBIR system

Table 2
Statistical evaluation of the proposed model against the other model over standard dataset

Property Name	*SASI*	*SIFT*	*Existing*	*Proposed*
Sensitivity	73.78%	81.22%	82.45%	89.08%
Specificity	40.12%	35.01%	40.73%	41.67%
Positive Likelihood Ratio	0.64	1.13	1.27	1.53
Negative Likelihood Ratio	0.56	0.45	0.35	0.26
Prevalence	69.12%	79.77%	86.14%	90.84%
Positive Predictive Value	73.29%	83.70%	87.58%	93.81%
Negative Predictive Value	45.33%	31.10%	29.84%	27.78%

CONCLUSION

The proposed model has been tested with the high variance image database for the testing of the proposed model based on content based image retrieval (CBIR).The proposed model performance has been evaluated in the form of statistical measures. The image database of Professor Wang has been utilized, which contains the 1000 images and know as the similarity dataset. The Professor Wang's dataset contains the images from the several categories such as people, urban, natural scenes, digital, etc. The bi-featured content based image retrieval system has been proposed in this research model, which utilizes the Laplacian of Gaussian (LoG) and the histogram of gradients (HoG) for the feature extraction from the testing and training image sets. The support vector based classification

has been proposed over the bi-feature scheme in the proposed model. The proposed model has been recorded with the sensitivity of nearly 89% against the existing schemes based upon SIFT, SASI and Early Elimination based HOG. All of these feature descriptors (including proposed) have been implemented with the SVM classifier. The proposed model specificity has been also recorded higher at the value of 41%, which is better than 40, 35 and 40.73 percent offered by the SASI, SIFT and HoG based Early Pruning model.

REFERENCES

[1] Ballard, L., Kamara, S., & Monrose, F. (2005). Achieving efficient conjunctive keyword searches over encrypted data. In Information and Communications Security (pp. 414-426). Springer Berlin Heidelberg.

[2] Bellare, M., Boldyreva, A., & O'Neill, A. (2007). Deterministic and efficiently searchable encryption. In Advances in Cryptology-CRYPTO 2007 (pp. 535-552). Springer Berlin Heidelberg.

[3] Boneh, D., Di Crescenzo, G., Ostrovsky, R., & Persiano, G. (2004, January). Public key encryption with keyword search. In Advances in Cryptology-Eurocrypt 2004 (pp. 506-522). Springer Berlin Heidelberg.

[4] Boneh, D., Di Crescenzo, G., Ostrovsky, R., & Persiano, G. (2004, January). Public key encryption with keyword search. In Advances in Cryptology-Eurocrypt 2004 (pp. 506-522). Springer Berlin Heidelberg.

[5] Bringer, J., & Chabanne, H. (2012). Embedding edit distance to enable private keyword search. Human-centric Computing and Information Sciences, 2(1), 1-12

[6] Cao, N., Wang, C., Li, M., Ren, K., & Lou, W. (2014). Privacy-preserving multi-keyword ranked search over encrypted cloud data. Parallel and Distributed Systems, IEEE Transactions on, 25(1), 222-233.

[7] Chang, Y. C., & Mitzenmacher, M. (2005, January). Privacy preserving keyword searches on remote encrypted data. In Applied Cryptography and Network Security (pp. 442-455). Springer Berlin Heidelberg.

[8] Chuah, M., & Hu, W. (2011, June). Privacy-aware bedtree based solution for fuzzy multi-keyword search over encrypted data. In Distributed Computing Systems Workshops (ICDCSW), 2011 31st International Conference on (pp. 273-281). IEEE.

[9] Chuah, M., and W. Hu. "Privacy-aware bedtree based solution for fuzzy multi-keyword search over encrypted data." In Distributed Computing Systems Workshops (ICDCSW), 2011 31st International Conference on, pp. 273-281. IEEE, 2011.

[10] Curtmola, R., Garay, J., Kamara, S., & Ostrovsky, R. (2006, October). Searchable symmetric encryption: improved definitions and efficient constructions. In Proceedings of the 13th ACM conference on Computer and communications security (pp. 79-88). ACM.

[11] Curtmola, R., Garay, J., Kamara, S., & Ostrovsky, R. (2006, October). Searchable symmetric encryption: improved definitions and efficient constructions. In Proceedings of the 13th ACM conference on Computer and communications security (pp. 79-88). ACM.

[12] Goh E-J Secure indexes. Cryptology ePrint Archive, Report 2003/216

[13] Golle, P., Staddon, J., & Waters, B. (2004, January). Secure conjunctive keyword search over encrypted data. In Applied Cryptography and Network Security (pp. 31-45). Springer Berlin Heidelberg.

[14] Hwang, Y. H., & Lee, P. J. (2007). Public key encryption with conjunctive keyword search and its extension to a multi-user system. In Pairing-Based Cryptography–Pairing 2007 (pp. 2-22). Springer Berlin Heidelberg.

[15] Li, J., Wang, Q., Wang, C., Cao, N., Ren, K., & Lou, W. (2010, March). Fuzzy keyword search over encrypted data in cloud computing. In INFOCOM, 2010 Proceedings IEEE (pp. 1-5). IEEE.

[16] Liu, C., Zhu, L., Li, L., & Tan, Y. (2011, September). Fuzzy keyword search on encrypted cloud storage data with small index. In Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference on (pp. 269-273). IEEE.

[17] Ren, K., Wang, C., & Wang, Q. (2012). Security challenges for the public cloud. IEEE Internet Computing, (1), 69-73.

[18] Song, D. X., Wagner, D., & Perrig, A. (2000). Practical techniques for searches on encrypted data. In Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on (pp. 44-55). IEEE.

[19] Song, D. X., Wagner, D., & Perrig, A. (2000). Practical techniques for searches on encrypted data. In Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on (pp. 44-55). IEEE.

[20] Song, Dawn Xiaodong, David Wagner, and Adrian Perrig. "Practical techniques for searches on encrypted data." In Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on, pp. 44-55. IEEE, 2000..

[21] Stefanov, E., Papamanthou, C., & Shi, E. (2013). Practical Dynamic Searchable Encryption with Small Leakage. IACR Cryptology ePrint Archive, 2013, 832.

[22] Wang, C., Cao, N., Li, J., Ren, K., & Lou, W. (2010, June). Secure ranked keyword search over encrypted cloud data. In Distributed Computing Systems (ICDCS), 2010 IEEE 30th International Conference on (pp. 253-262). IEEE.

[23] Wang, C., Cao, N., Ren, K., & Lou, W. (2012). Enabling secure and efficient ranked keyword search over outsourced cloud data. Parallel and Distributed Systems, IEEE Transactions on, 23(8), 1467-1479.

[24] Wang, C., Ren, K., Yu, S., & Urs, K. M. R. (2012, March). Achieving usable and privacy-assured similarity search over outsourced cloud data. In INFOCOM, 2012 Proceedings IEEE (pp. 451-459). IEEE.

[25] Witten, I. H., Moffat, A., & Bell, T. C. (1999). Managing gigabytes: compressing and indexing documents and images. Morgan Kaufmann.

[26] Yang, C., Zhang, W., Xu, J., Xu, J., & Yu, N. (2012, November). A fast privacy-preserving multi-keyword search scheme on cloud data. In Cloud and Service Computing (CSC), 2012 International Conference on (pp. 104-110). IEEE.

[27] Zerr, S., Demidova, E., Olmedilla, D., Nejdl, W., Winslett, M., & Mitra, S. (2008, March). Zerber: r-confidential indexing for distributed documents. In Proceedings of the 11th international conference on Extending database technology: Advances in database technology (pp. 287-298). ACM.

[28] Zerr, S., Olmedilla, D., Nejdl, W., & Siberski, W. (2009, March). Zerber+ r: Top-k retrieval from a confidential index. In Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology (pp. 439-449). ACM.