Penerbit Akademia Baru

Journal of Advanced Research in Computing and Applications

Journal homepage: www.akademiabaru.com/arca.html

ISSN: 2462-1927

Large scale image retrieval for remote sensing images using low level features

Open Access

Khaled H Mohamed[1,*], Safa M Gasser[2], Mohamed S El-Mahallawy[2], Mohamed Waleed W Fakhr[3]

[1] Department of Electrical and Computer Engineering, college of Electrical Engineering Institute HTI Egypt, 10th of Ramadan city, Egypt
[2] Department of Electronics and communication, college of Engineering AAST, Egypt
[3] Department of Computer Engineering, college of Computer Science AAST, Egypt

ARTICLE INFO

Article history:
Received 5 June 2017
Received in revised form 10 July 2017
Accepted 15 July 2017
Available online 4 December 2017

Keywords:
Low level features, PCA, KNN, SVM

ABSTRACT

Content Based Image Retrieval (CBIR) system extracts features relevant to query image using feature extraction method. Many low level features are proposed to retrieve accurate similar image, but the problem is no method provides accurate results. In this paper, we discuss getting an accurate result for retrieving remote sensing images from (USGS) United States Geological Survey database using different low level features. The lower level features used to construct the feature vector are Discrete Cosine Transform (DCT), Karhunen-Loève transform (KLT), Wavelet transform (WT), Histogram of orientation (HOG), and Gist. Different combinations of these features are used to train two classifier (KNN) K-nearest neighborhood and (SVM) Support Vector Machine classifier. A dimensionality reduction technique (PCA) principal component analysis is used to reduce the dimensionality of the feature vectors and see the effect of PCA on the accuracy of the classifiers.

1. Introduction

Due to the growing number of images added every day on social media, police stations, hospitals, etc., these images are grouped to gather in a database. These databases could be a medical image diagnosis [1], architecture & engineering design, art collection images, criminal database, Banks finger print, geographical information & remote sensing, military, photographic archive & journalists [2-3]. So this field attracted so people and received a growing amount of attention from vision community [4-5]. The Large Scale Image Retrieval (LSIR) means we search on these databases on a certain image or similar to it.

The need for fast and intelligent retrieval systems for Remote Sensing (RS) images is needed due to the fast increase of the number of images acquired every day. Synthetic Aperture Radar (SAR) sensors approximately collect about 10: 100 Gbyte of images daily. All the recent retrieval systems depend on the content of the image more than the high semantic features due to the semantic gap between them.

* *Corresponding author.*
E-mail address: kh2my2007@gmail.com (Khaled H Mohamed)

There are two methods for LSIR; Text based image retrieval & Content Based Image Retrieval (CBIR). In the Text based image retrieval, the metadata is used in this technique and the image is retrieved using keyword, tags or description associated with the image. The disadvantages of this technique are that different users can use different keywords to search with and all the images in the data have to describe manually. In CBIR, the image is presented using image based features. Image features used in CBIR are categorized into two types, high level features and low level features. High level features where the image is described using the scene from the human perspective, e.g. house, the flower and the drawback of this type of feature is that many users see the image with different perspectives, hence the low level features give us an accurate result for describing the images based on its internal features as color histogram and texture features [6-7]. The color histogram gives the information about the image color and not the location of the color in the image. Gabor filter and Gray Level Co-occurrence Matrix (GLCM) are used in extracting texture features [8-10]. GLCM is very computationally intensive algorithm and rotation variant, while the Gabor filter is rotation invariant and it is used in extracting the Gist features.

This paper focuses on searching the best combination of low level features that can be extracted from the RS images, to achieve best accuracy using KNN and SVM classifiers. Data reduction techniques such as PCA is also applied to reduce the dimensionality of the feature vector, hence reduce the system complexity.

The remainder of this paper is organized as follows. In section 2, the LSIR methodology is introduced. Dimensionality reduction technique (PCA) is briefed in section 3. In section 4 the dataset used in this paper is described. Section 5 the results of our experiments. Section 6 concludes the paper.

2. LSIR Methodology

The methodology used in the LSIR system evaluated in this paper is introduced in this section as shown in Figure 1. The feature extractor used lower level features to represent the images. The dimensionality reduction algorithm employs PCA. These features are stored in a codebook. For image retrieval part, we use SVM and KNN as classifiers.

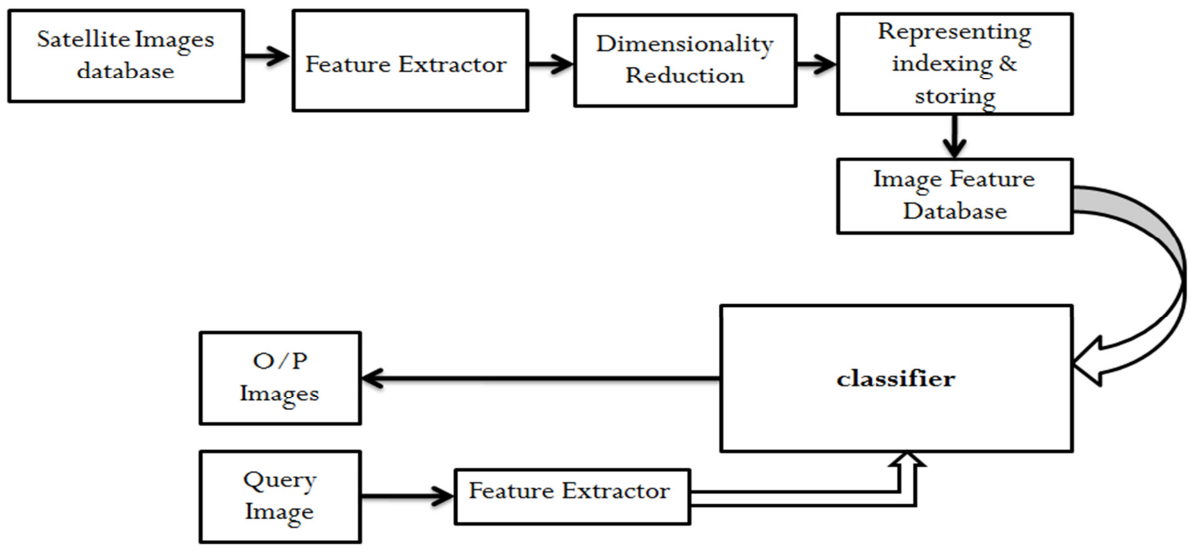

Fig. 1. LSIR Block Diagram

Penerbit Akademia Baru

2.1 Low Level Features Extraction

There are many low level features extraction algorithms for LSIR used in the literature. The lower level features used in this paper are DCT, KLT, WT, HOG & Gist features:

The DCT expresses the image in terms of sum of cosine functions with different frequencies and amplitudes. Its compress the energy of the signal in a few values [11].

The Karhunen-Loève transform is an orthonormal transform. That transforms the image into new orthonormal basis, which reveals the information distribution, Structure & magnitude of the image [12].

The Wavelet transform (WT) provide a multi-resolution approach to texture analysis, it decomposes the image into a summation of time-domain basis functions of various frequency resolutions, and it provides a progressive encoding of the image at various scales [13].

The Histogram of Gradient (HOG) features encode the local shape information of an image. HOG features group the gradient magnitudes into bins in a histogram based on its orientation [14].

In the Gist features, the image is convolved with different Gabor filters at different orientations [15].Gist features are biologically plausible and have a low computational time.

2.2 Classifiers

KNN and SVM classifiers are used in this paper to measure the accuracy of different feature vectors employed in LSIR system.

2.2.1 KNN classifier

KNN used in statistical estimation and pattern recognition as a non-parametric technique. A case is classified by a majority vote of its neighbors, with the case being assigned to the class most common amongst its K nearest neighbors measured by a distance function [16]. If K = 1, then the case is simply assigned to the class of its nearest neighbor base on the Euclidean distance as in equation (1)

$$D_{EUC\,(v,w)} = \left\|\boldsymbol{V} - \boldsymbol{W}\right\|_2 = \sqrt{\sum_{i=1}^{N} (V_{(i)} - W_{(i)})^2} \quad (1)$$

where||V||2 is the norm-2 of the vector V; N is the number of the features selected in the feature vector V and W.

2.2.2 SVM classifier

Support Vector Machine classifier is one of the most accurate classification methods used in many applications [17]. It can construct the learning model by means of less number of data sets regardless of the number of features since it is based on a strong foundation. As a solution to the two-class learning problem, SVM finds the best classification function which can clearly distinguish between members of the two contrasting classes in the training data.

For classification, SVM's function by finding a geometrical hyper-plane in the possible input space. This hyper-plane will split the two regions, one with positive examples and the other with negative examples. The split is computed such that the distance of the data points in the input space from the hyper-plane gets maximized, as: [18-19].

$$W \,.\, x_i \geq +1 \text{ For } y_i = +1 \tag{2}$$

$$W \,.\, x_i \leq -1 \text{ For } y_i = -1 \tag{3}$$

Where x_i data points, W is the normalized vector to the hyper plane (weight vectors), y_i is the class label of the point x_i and b is bias term. SVM minimizes the quadratic optimization equation (4) to achieve accurate result where C is the realization parameter and ζ is the slake variable.

$$\min \frac{1}{2} ||W||^2 + C \sum_i^N \xi_i \text{Subject to } y_i(W^T x_i + b) \geq 1 - \xi \text{ For i = 1, 2, 3, 4… N} \tag{4}$$

3. PCA Dimensionality Reduction

One of the widely used methods of statistical data analysis used in image retrieval is PCA. It is a kind of linear dimensionality reduction method where the data space is projected to a lower dimensional space using orthogonal transformation [20].

The PCA tries to find the best projection direction which represents the original data and minimizes the least mean square error.

4. Remote Sensing Dataset

The USGA is an extensive manually labelled ground truth dataset used to perform quantitative evaluation. The dataset consists of images of resolution of one foot. Larger images were downloaded from the United States Geological Survey "USGS" national map of the following US regions: Birmingham, Boston, Buffalo, Columbus, Harrisburg, Houston, Jacksonville, Las Vegas, Los Angeles, Miami, Napa, New York, Reno, San Diego, Santa Barbra, Seattle, Tampa, Tucson, and Ventura [21].

100 images of size 256 x 256 pixels were manually selected for each the following 21 classes as shown in Figure (2); agricultural, airplane, baseball diamond, beach, buildings, chaparral, dense residential, forest, freeway, golf course, harbour, intersection, medium density residential, mobile home park, overpass, parking lot, river, runway, sparse residential storage tanks, and tennis courts images.

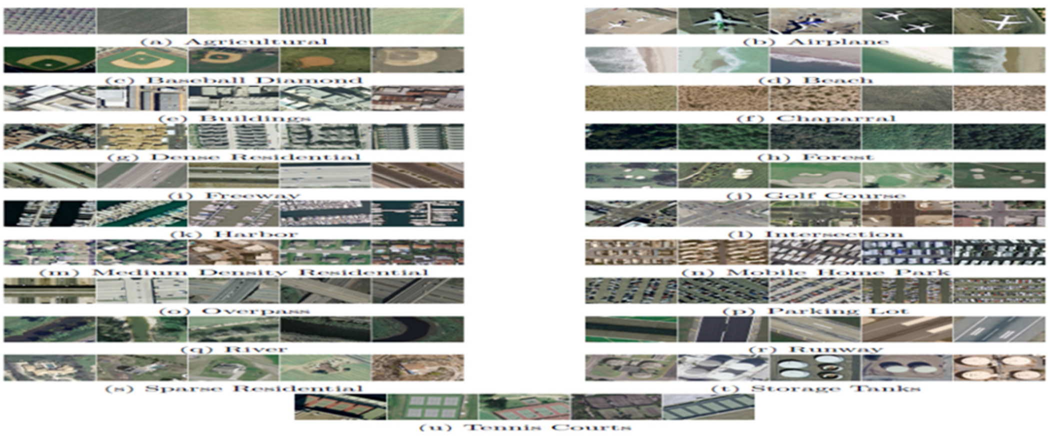

Fig. 2.USGA image database

5. Results and Discussion

In this paper, HOG, Gist, KLT-features, DCT-features, and Wavelet-features are extracted. The dimension of each feature vector is chosen as shown in the Table 1to achieve the best accuracy using KNN classifier.

Table 1
KNN Results On all Features

Feature vector name	Accuracy (%)	Test time (sec)	Dimension
HOG	59.62	0.0388	81
Gist	51	0.2437	512
KL-feature	34	0.0333	63
DCT feature	33.62	0.1323	192
Wavelet feature	28.19	0.3862	768

For The Gist feature, 32 different Gabor filter at 4 different scales and 8 orientations with block size equal to 4 x 4 are used so that the total feature vector for the each image became 512 features. For the HOG features, the block size used is 3 x 3 and the number of bins per block equal to 9 so the output feature vector per image is 81. For the DCT feature, the block size used is 8 x 8 for each image channel, hence the total output feature vector for the each RGB image is 192. For the KL feature, the largest 21 Eigen values were chosen for each image channel and for the RGB image the total feature vector is equal to 63. For the wavelet features, the fourth level WT using block size 16 x 16 per channel is used with to produce a feature vector of length 768 for each image.

Table 1 shows that the best accuracy is achieved using the KNN classifier using the HOG features. The HOG feature achieves the best accuracy with 81 features only. The Table 1 shows that the HOG feature is one of the fastest and accurate features, among the rest of the features compared.

Different feature vectors are combined and evaluated using the KNN classifier in Table 2. Also, the PCA algorithm is employed to reduce the system complexity. Table 2 summarizes the accuracy of different combined features evaluated with KNN classifier with and without using PCA.

Table 2
KNN result on combined feature vectors with and without PCA

Feature vector name	Accuracy (%)	Train time (sec)	Test time (sec)	Dimensionality
HOG-Gist	61.87	0.0214	0.2736	593
HOG-Gist-PCA	58.63	0.1639	0.037	21
DCT-KL- Wavelet	41.99	0.0214	0.1618	1023
DCT-KL- Wavelet-PCA	43.39	0.0203	0.0145	21

Table 2 shows that the accuracy of the system using HOG-Gist feature vector decreases by 3.24 % compared to the system implemented using HOG features only and the dimensionality of the feature vector is decreased by 96.46 % when using PCA. The test time is decreased by a large value. Also, the SVM classifier is used to measure the accuracy of the combined feature vectors and compared to that using KNN classifier. Table 3 shows that the system performance using SVM is increased compared to KNN when we used the HOG-Gist feature vector concerning the accuracy but on the expense of testing time.

Table 3
SVM result on combined feature vectors with and without PCA

Feature vector name	Accuracy (%)	Test time (sec)	Dimensionality
HOG-Gist	72.29	0.4473	593
HOG-Gist-PCA	66.86	0.4229	21
DCT-KL- Wavelet	50.81	42.4846	1023
DCT-KL- Wavelet-PCA	55.10	0.4473	21

6. Conclusion

In this paper, LSIR system is implemented and evaluated using different combinations of low level features for remote sensing images. The chosen combined feature vector achieves better accuracy compared to the other feature vectors tested by both classifiers; KNN and SVM. The PCA managed to conserve the accuracy of both classifiers and effected the testing time and the dimensionality of the feature vectors obviously.

References

[1] Müller, Henning, Nicolas Michoux, David Bandon, and Antoine Geissbuhler. "A review of content-based image retrieval systems in medical applications—clinical benefits and future directions." *International journal of medical informatics* 73, no. 1 (2004): 1-23.

[2] Markkula, Marjo, and Eero Sormunen. "Searching for photos-journalists' practices in pictorial IR." In *The Challenge of Image Retrieval, A Workshop and Symposium on Image Retrieval, Electronic Workshops in Computing, Newcastle upon Tyne*, pp. 5-6. 1998.

[3] Armitage, Linda H., and Peter GB Enser. "Analysis of user need in image archives." *Journal of information science* 23, no. 4 (1997): 287-299.

[4] Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. "Imagenet: A large-scale hierarchical image database." In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 248-255. IEEE, 2009.

[5] Sánchez, Jorge, and Florent Perronnin. "High-dimensional signature compression for large-scale image classification." In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 1665-1672. IEEE, 2011.

[6] Arjunan, R. Vijaya, and V. Vijaya Kumar. "Image Classification in CBIR systems with color histogram features." In *Advances in Recent Technologies in Communication and Computing, 2009. ARTCom'09. International Conference on*, pp. 593-595. IEEE, 2009.

[7] X. Chen, X. Yang, R. Zhang, A. Liu and S. Zheng, "Edge region color autocorrelogram: A new low-level feature applied in CBIR." In IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB),(2010). 1-4.

[8] Mohammed, Nabeel, and David McG Squire. "Improved texture features for CBIR using response scaling and locally normalised convolution." In *Content-Based Multimedia Indexing (CBMI), 2013 11th International Workshop on*, pp. 143-148. IEEE, 2013.

[9] Kaur, Manpreet, and Neelofar Sohi. "A novel technique for content based image retrieval using color, texture and edge features." In *Communication and Electronics Systems (ICCES), International Conference on*, pp. 1-7. IEEE, 2016.

[10] Rasli, Ruziana Mohamad, T. Zalizam T. Muda, Yuhanis Yusof, and Juhaida Abu Bakar. "Comparative analysis of content based image retrieval techniques using color histogram: a case study of GLCM and K-means clustering." In *Intelligent Systems, Modelling and Simulation (ISMS), 2012 Third International Conference on*, pp. 283-286. IEEE, 2012.

[11] Kekre, H. B., and Kavita Sonawane. "Row, column and fused row-col R, G, B plane's feature vector generation using DCT, DST and Kekre wavelet for CBIR." In *Communication, Information & Computing Technology (ICCICT), 2012 International Conference on*, pp. 1-6. IEEE, 2012.

[12] Liu, Lei, Jingwen Yan, Xianwei Zheng, Hong Peng, Di Guo, and Xiaobo Qu. "Karhunen-Loève transform for compressive sampling hyperspectral images." *Optical Engineering* 54, no. 1 (2015): 014106-014106.

[13] Demirel, Hasan, and Gholamreza Anbarjafari. "Discrete wavelet transform-based satellite image resolution enhancement." *IEEE transactions on geoscience and remote sensing* 49, no. 6 (2011): 1997-2004.

[14] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, pp. 886-893. IEEE, 2005.

[15] Oliva, Aude, and Antonio Torralba. "Modeling the shape of the scene: A holistic representation of the spatial envelope." *International journal of computer vision* 42, no. 3 (2001): 145-175.

[16] Kalbkhani, Hashem, Arghavan Salimi, and Mahrokh G. Shayesteh. "Classification of brain MRI using multi-cluster feature selection and KNN classifier." In *Electrical Engineering (ICEE), 2015 23rd Iranian Conference on*, pp. 93-98. IEEE, 2015.

[17] Wu, Xindong, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan et al. "Top 10 algorithms in data mining." *Knowledge and information systems* 14, no. 1 (2008): 1-37.

[18] Hmeidi, Ismail, Bilal Hawashin, and Eyas El-Qawasmeh. "Performance of KNN and SVM classifiers on full word Arabic articles." *Advanced Engineering Informatics* 22, no. 1 (2008): 106-111.

[19] Shao, Yang, and Ross S. Lunetta. "Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points." *ISPRS Journal of Photogrammetry and Remote Sensing* 70 (2012): 78-87.

[20] Cheng, Bo, Li Zhuo, and Jing Zhang. "Comparative study on dimensionality reduction in large-scale image retrieval." In *Multimedia (ISM), 2013 IEEE International Symposium on*, pp. 445-450. IEEE, 2013.

[21] Yang, Yi, and Shawn Newsam. "Geographic image retrieval using local invariant features." *IEEE Transactions on Geoscience and Remote Sensing* 51, no. 2 (2013): 818-832.