Journal of Engineering and Applied Sciences 14 (16): 5753-5761, 2019
ISSN: 1816-949X

An Efficient Approach for Visual Object Categorization based on Enhanced Generalized Gabor Filter and SVM Classifier

[1]Hayder Ayad, [2]Loay Edwar George, [3]Mamoun Jassim Mohammed, [3]Raad Ahmed Hadi and [1]Siti Norul Huda Sheikh Abdullah
[1]Department of Artificial Intelligence Technology, College of Information Systems and Technology, University of Kebangsaan, Bangi,Malaysia
[2]Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq
[3]Department of Computer Engineering, College of Engineering, Al-Iraqia University, Baghdad, Iraq
hayder.alsultany@gmail.com, 009647811111030
loayedwar57@yahoo.com, loayedwar57@cs.uobaghdad.edu.iq, 009647902234325
raad.hd@gmail.com, 009647731288613
mamoun_87@yahoo.com, mamoun.jassim@aliraqia.edu.iq, 009647702968307
mimi@ftsm.ukm.my, +60192377984

Abstract: Filter banks such as the Gabor Filter (GF) are widely used to describe objects. The main disadvantage of the Gabor filter is that it constructs redundant and incompact filters that may decrease system recognition performance. The purpose of the current study primarily is to enhance the categorization problem through generalizing the GF method (GGF). The unsupervised machine learning algorithm, denoted by the k-means clustering algorithm is proposed to implement generalization on a GF set. To assess the performance of the proposed method, the standard GF is used as a benchmark. Furthermore, the first 20 classes and the overall classes from the dataset Caltech 101 have been utalized in the performance demonstration of the newly suggested method. Based on a single classifier and combination feature (Naive approach), the proposed GGF outperforms and shows higher potential results than the standard GF in describing objects.

Key words: VOC technique, Gabor filter, Naive combination approach, SVM classifier, benchmark, unsupervised

INTRODUCTION

In new technology, image investigation in several areas including computer vision and its diverse techniques (i.e., Visual Object Categorization (VOC) and Content-Based Image Retrieval (CBIR)) make an alert to manage these techniques. The VOC technique is the main focus of current research and represents an open area that still needs to be researched. The present study is mainly dedicted to solve the categorization problem (semantic gap). Based on state-of-the-art-research, the VOC technique commonly uses the texture feature as it provides meaningful information that can achieve high-level features and improve the categorization problem (Abdullah and Wiering, 2007; Liu *et al*., 2007; Varma and Zisserman, 2003).

However, much of the research in the area has used a filter to describe the object (Choi *et al*., 2008; Cruz-Aceves *et al*., 2016; Long *et al*., 2003; Ramdman *et al*., 2014; Yavuz *et al*., 2016). Based on the literature, filters can be divided into two kinds: linear filters such as the Sobel detector and non-linear filters (filter banks) such as the Gabor Filter (GF). The former uses a single orientation and fixed mask size when constructing feature maps which may inaccurately capture the important salient features, particularly when applying orientation descriptors. Furthermore, it is sensitive to orientation; specifically to non-edge orientation. However, these edges that are oriented within a narrow range are replied to by the highly the highly discriminative edge filters (Shristi *et al*., 2016; Nadernejad *et al*., 2008).

The second one (non-linear filters) can efficiently overcome the limitations of the former in capturing the most discriminative edges and salient features rely on a different scale, orientation and filter size (Yavuz *et al*., 2016). Using an orientation descriptor based on this type of filter can provide meaningful and distinctive features for correctly describing an object (Abdullah *et al*., 2009). More advanced techniques have frequently used the GF

Corresponding Author: Hayder Ayad, Department of Accounting, College of Bussines Administration, Al-Bayan University, Baghdad, Iraq, hayder.alsultany@gmail.com, 009647811111030

to describe objects. Choi *et al.* (2008) showed the GF proved to be efficient when detecting, recognizing and tracing objects. Lin-Lin and Zhen (2009) suggested using the GF with an SVM classifier for object categorization and their experiment showed that the proposed method performed well. Many researchers also used the GF for 3-D object representation and recognition (Wang *et al.*, 2001). These studies all showed the benefit of using the GF to detect the salient feature that can improve system performance.

Generaly, many descriptors have been introduced to extract texture features. To be selected, a descriptor should have a certain ability in human vision to produce the most discriminative and important features. Constructing edge features are among the basic vision features in the human system. It is thus possible to describe objects efficiently using the edge orientation information. Orientation descriptors including the edge histogram have become highly well-known systems with respect to object categorization (Abdullah *et al.*, 2009).

The literature presents the main issue of using the GF and orientation descriptors including the Edge Histogram Descriptor (EHD) to describe objects. Clearly, Abdullah *et al.* (2010) used the EHD to extract the texture feature to describe objects. This descriptor used a single orientation to extract the texture feature which may not be sufficient in describing objects. Applying an advanced filter such as the GF, to construct various feature maps that rely on various scale, orientation and filter size will facilitate to the descriptor to produce excellent outputs and highly outstanding texture features. Conversely, the main drawback of the GF is that it constructs redundant and incompact filters that may affect system recognition accuracy and decrease system performance. Zhu *et al.* (2004) proposed a Principle Component Analysis (PCA) that helps decreseaing the dimentionality of GF and resulting producing an optimal filters that help detect faces. The aligorithm that is calld Fast Fourier Transform (FFT) has been invested in the dimentionality reduction of GF within the frequency domain (Yavuz *et al.*, 2016).

In this research, a new filter-implying descriptor and classifier that has a Generalized Gabor Filter (GGF) has been suggested. The latter helps enhance the categorization problem. The GF parameters are selected. GF is generalized via. the machine learning algorithm of the unsupervised type. The algorithm is marked by the technique of k-means clustering. This stage ensures the removal of redundant filters which eventually leads to a more compact and optimized filter set. We also use an orientation-based descriptor, denoted by edge histograms to extract the texture feature. A Naive approach that consists of combination features and single classifier have been used to construct feature vectors and classify them by a Support Vector Machine (SVM). To assess GGF method's peformance, VOC technique framework has been implemented. This framework is based on the suggested method that involves categorizing the objects that lie within the spatial domain and comparing the results obtained with that of the standard GF method. Furthermore, to assess the suggested method, the first twenty classes and all classes from the Caltech 101 dataset have been utlized. We extend the experiment by implementing the VOC technique framework based on the GF and proposed GGF methods in the frequency domain and compare it to the suggested VOC method within the spatial domain.

MATERIALS AND METHODS

First, we briefly review the GF method computation. We next present an orientation descriptor, denoted by an Edge Histogram Descriptor (EHD). We then discuss the Naive approach Support Vector Machine (SVM) and k-means clustering methods.

Gabor Filter (GF): The filter banks are ubiquitous in extracting feature textures. The responses of these filters are mainly based on the distribution, joint filter response distribution and the scale and orientation used to create these filters. These filter banks can be used for segmentation, classification and synthesis (Yavuz *et al.*, 2016).

Recent texture classification has widely used the filter banks. The biological plausibility and the hypothesis in various scale and orientation must extract sufficient features and perform accurate classification (Senthilkumar and Paulraj, 2016; Wang *et al.*, 2001). Examples of filter banks include the gabor, structure tensor and steerable filters, among others (Li *et al.*, 2010).

The GF has become important in several applications due to its characteristics such as providing classification schema and a multi-resolution schema for feature texture. This filter also supplies significant pieces of information through the use of various magnitude and orientation. Besides, a number of salient features can be captured using this filter, including spatial frequency characteristics, spatial locations and orientation selections. The mathematical equations below describe the GF (Zhu *et al.*, 2004):

$$\psi k(z) = \frac{\left\| k^2 \right\|}{\sigma^2} e^{-\|k|^2 z|^2 /2\sigma^2} \left[e^{ikz} - e^{\sigma^2/2} \right]$$

The parameters are defined, thus: k refers to the wavelength together with the orientation of the kernel $\psi k(z)$ in the image. This equation has two parts: the oscillation refers to the term in brackets and the DC comprises is the reset. σ denotes the standard deviation of the Gaussian function. (k) is defined thus:

$$k(\mu,\upsilon) = K_{u^{e^{i\varnothing\mu}}}$$

where, parameters μ and υ refer to the orientation and scaling of the Gabor kernel, respectively:

$$k_u = k_{max}/f\upsilon \text{ and } \varnothing\mu = \pi\mu/8$$

Here, parameter k_{max} refers to the maximum frequency whereas f refers to the space factor that is between the kernels within the frequency domain.

The extension GF convolves the image before constructing the image primitives. If I(x) represents the grey level of the image and $\psi k(x)$ the GF kernels, the following equation can yield the image primitive:

$$Ok(x) = I(x) * \psi k(x)$$

Where:
Ok(x) = The convolution result at k and parameter
* = The convolution operator

Figure 1 gives an example of the cosine part (even) of GF.

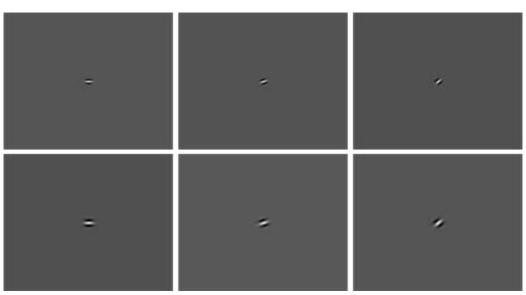

Fig. 1: The GF with parameters values ($\upsilon = 2$, $\mu = 3$, $\sigma = 2.5$ and filter size = 128×128)

Edge histogram: The edge feature is an important feature for describing image content, especially considering the object shape and texture (Mignotte, 2012). Thus, an edge histogram is used to refer to the edge distribution of an image. At the same time, the edge histogram refers to the frequency and directionality of an edge in that image. It also captures the spatial distribution of the edge (4 directional edges and 1 non-directional edge), describes and extracts the non-homogenous texture (Manjunath *et al*., 2002). Figure 2 illustrates sub-images and blocks inside the sub-images.

Extracting the edge histogram first requires partitioning a specified image into blocks of 4×4 non-overlapping sub-images. Each of the latter has a group of image-blocks. The ED method then applies five different edge detectors to perform texture extraction in different directions: horizontal, vertical, diagonal of both 45 and 135 and non-directional (Manjunath *et al*., 2002). Figure 3 illustrate the ED filters.

In each sub-image, the image blocks are convolved with the filters coefficient which represents various edge detectors as in Fig. 3. Only the highest value among these edge strengths is compared with a specified edge threshold. Accordingly, if the value is bigger than of the threshold, the block of the image is considered as having the corresponding edges. Finally, the edge histogram produces 80 bins of edge features that represent the image (Manjunath *et al*., 2002). The following equation shows how to obtain the edge strength among different 5 filters:

$$\text{edg-stg}(i,j) = \left|\sum_{k=0}^{n} A_k(i,j) * \text{edge-filter}(k)\right|$$

Here, n refers to the sub-block number in each of the image blocks, $A_k(i, j)$ refers to an image block and edg-stg(i, j) indicates any of the five edge detectors. However, Fig. 4 shows the spatial distribution of the edge histogram.

Naive approach: Combining features outperforms using a single descriptor when representing an image in a Visual

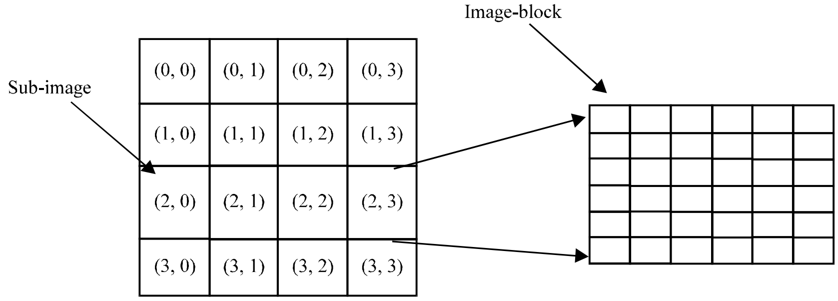

Fig. 2: An image divided into sub-images and image-blocks

Object Categorization (VOC) (Alemu *et al*., 2009). The Naive approach considers a combination method where it combines a number of feature vectors in one feature vector to be later used in feeding the classifier. Similarly, it places several feature vectors produced by various sources directly into an input vector and enters it into the classifier for learning and testing. However, the feature vector produced by the Naive approach leads to large dimensional feature vectors. Nonetheless, the Naive approach outperforms a single feature vector; however, using the Naive approach with combined features increases dimensional feature vectors, produces over-fitting and affects performance (Abdullah, 2010).

Support Vector Machine (SVM): Recent research has shown how the SVM algorithm outperforms these techniques in different applications. The basic part of the algorithm in question is the optimal hyper-plane that isolates the generated data into two types of classes or categories {+1, -1} where the topest margin is in the higher dimensional feature space. SVM helps in design a model that invests the generated training data to be able to predict the target values of the test data attributes (Cortes and Vapnik, 1995).

For the classification schema, if X is the input value, SVM will classify this input into two classes y∈{-1, +1} according to the following equation:

(a)

1	-1
1	-1

(b)

1	1
-1	-1

(c)

$\sqrt{2}$	0
0	$-\sqrt{2}$

(d)

0	$\sqrt{2}$
$-\sqrt{2}$	0

(e)

2	-2
-2	2

Fig. 3: Edge histogram masks for texture feature extraction; a) ver_edge_filter; b) hor_edge_filter; c) dia45_edge_filter; d) dia135_edge_filter and e) nond_edge_filter

$$y = \text{sign}(f(x))$$

Function f(x) comes from combining the Kernel function between input X and each training data point thus:

$$f(x) = \sum \alpha_m y_m K(X_m, X) + b$$

where, X represents the input vector, X_m represents the training data and the Kernel is $K(X_m, X)$. In Eq. 6, the function of decision is calculated using the training labels $y_m \in \{+1, -1\}$ and coefficient $\alpha_m \geq 0$ whereas the parameter b is set throughout the process of training that is done on the labeled dataset. When a training set of instance label pairs (x_i, y_i) is given, i = 1, ..., n, where $X_i \in R$ and the SVMs include the solution for the optimization problem (Cortes and Vapnik, 1995):

$$\min_{w,b,\xi} \frac{1}{2} w^T + C \sum_{i=1}^{I} \xi_i X_1, \ldots, X_n$$

Subjected to:

$$y_i \left(w^T \varnothing(x_i) + b \right) \geq 1 - \xi_i \, \xi_i \geq 0$$

To find the optimal hyper-plane, function∅(.) projects the input pattern Xi onto the higher dimensional space. Among several hyperplanes, SVM finds the optimal one with the highest margin for the binary problem in this space. To get a distinguished hyper-plane, a weight vector W needs to exist in the form of a training pattern subset that lies on the margin. The following equation shows how to construct W:

$$W = \sum_{i=1}^{n} \alpha_i y_i \varnothing(X_i) + b$$

where, n represents the number of support vectors whereas X_i denotes the support vector i and y_i means the label {-1, 1} of X_i. This study has adopted LIBSVM with the RBF kernel to classify the data.

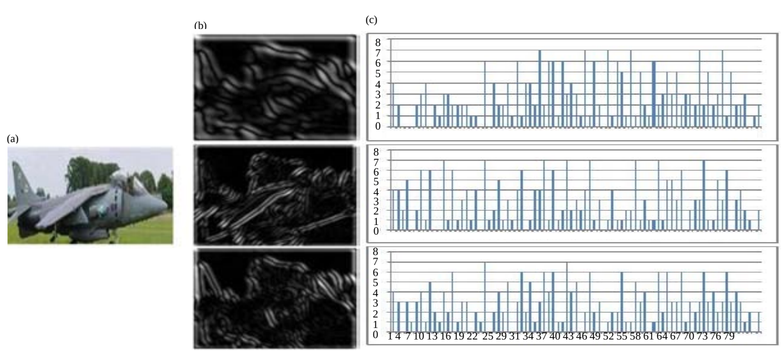

Fig. 4: The edge histogram applied to real-world images; a) Original airplane class image from the Caltech 101 dataset; b) Image primitives (feature maps) constructed by convolving the original image with the GF and c) Spatial distribution of the edge histogram

Clustering: The clustering technique is often used to cluster data into groups where each group reflects similar members whereas each group differs from one another. In the last decade, researchers have proposed different clustering approaches, namely partitioning, hierarchical, density, grid and model-based clustering. Partition clustering which includes the k-means clustering technique and hierarchical clustering is the highly famous technique invested in various applications. The basic aim behind grouping is to enhance the similarity within the intra-cluster and reduce it within the inter-cluster (Wu and Bhanu, 1997; Yang *et al*., 2010).

k-means clustering: k-means clustering is one of the highly famous grouping techniques that are used in many applications. This algorithm performs the act of clustering by allotting data in each of the various clusters based on the least distance between the cluster centroid and the data. Each cluster is then marked by its centroid, i.e., the mean of the cluster members (Jain *et al*., 1999). There are a number of merits that make this technique famous and frequently invested; these include the following: first, this technique is applicable to any standard norm; it is characterized by being fast and easy to implement. Second, it has clear parallelization and is not sensitive to data ordering. Conversely, some limitations appear in this clustering technique as the result basically depends on the initial cluster centroid. The number of the cluster should be specify in advance and the k-cluster needs, therefore, to be optimized for each case (Jain *et al*., 1999).

To explain the k-means clustering technique, assume that we have these observations $\{xi: i = 1, \ldots, L\}$. Besides, one needs to calculate the distance between these data points through using order p. k-means aims to divide the observations into different k groups where each group is denoted by its own centroid (mean) $\{\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_k\}$:

$$KCL = \sum_{i=1}^{L} \min_{1\leq j\leq k} (x_i - \hat{x}_j)^p$$

The k-means technique should follow these steps: claculate the distance between the centroid and observation data points. Identify the minimum space between the centroid and points and assign a point to the respective cluster. The k-means technique converges until the error between the new and old cluster centroids is below a fixed threshold. Figure 5 clarifies the k-means clustering technique.

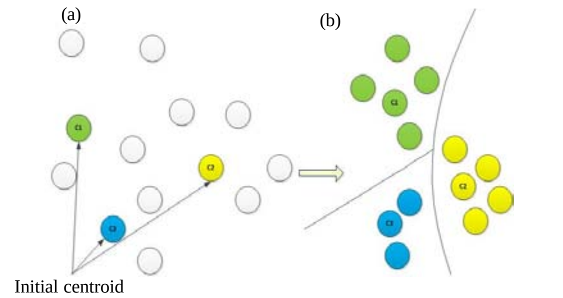

Fig. 5: The k-means clustering technique; a) data before clustering with an initial centroid and b) data after clustering, represented by their own centroids

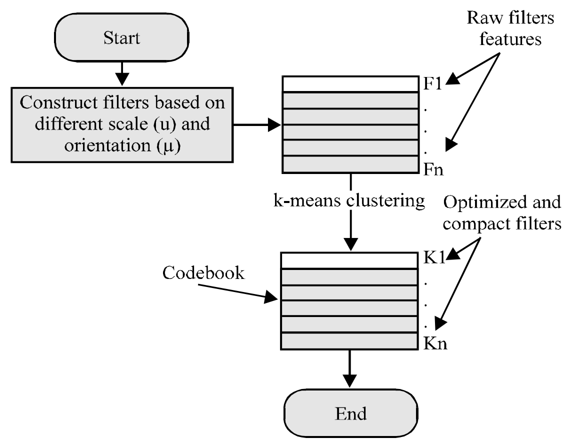

Fig. 6: Flowchart of the proposed GGF method, F1, .., Fn represents the raw filters features and K1, …, Kn represents the cluster centroid (Generalized filters)

A Novel Generalization of the Gabor Filter (GGF) method: Over the last decade, researchers have focused on using filters to describe objects. Among several non-linear filters, the GF gives impressive results in different places (Fig. 6).

The literature also indicates that the main demerit of the GF is that it constructs redundant and incompact filters which may decrease system accuracy. An effective method thus requires a generalized and compact GF that can efficiently describe an object. By investing a vector quantization method that involves a clustering algorithm, the compact and informative filters can be met. Such an algorithm is considered active in compact and generalized filters by merging the similar patterns into clusters or groups.

The current study uses the unsupervised machine learning algorithm denoted by the k-means clustering algorithm, to generalize the GF. This algorithm aims at allotting a point to every cluster centroid taking into account the minimum distance. The basic issue as far as this algorithm is concerned is the number of the clusters.

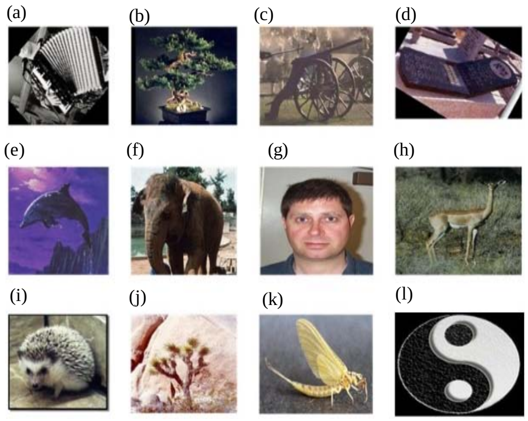

Fig. 7: a-l) Random examples from the Caltech101 dataset categories: (starting from top left) accordion, bonsai, cannon, cell phone, dolphin, elephant, Faces _easy, gerenuk, hedgehog, Joshua _tree, mayfly and yin _yang

This parameter needs to be chosen in a way that helps assure higher accuracy. However, the main goal of data clustering is to provide a precise characterization of unseen samples generated from the same probability distribution. The flowchart below shows the steps used to generalize the GF:

The proposed (GGF) method ensures the removal of redundant filters which eventually leads to a more compact and optimized filter set.

Experimental setup: This study invests the first twenty categories of Caltech 101 dataset that has 600 images in both training and testing processes. All classes of the Caltech 101 dataset including the 3030 images have been utalized in VOC performance test in both domains, the spatial and frequency.

Caltech 101 dataset: The Caltech 101 dataset involves 101 categories with 9146 images where each category contains images ranging from 40-800 ones on average. The images within this dataset have an equal size and resolution. The medium resolution of each of these images is about 300×200 pixels and all images are in the JPEG format. Moreover, objects in the images have low occlusion and clutter levels. This research uses the first twenty classes of Caltech 101 dataset to explicate the system accuracy. Furthermore, each category involves thirty images and their total number is 600. To train and test the system, each category contains 15 randomly chosen images each for training and testing (Abdullah *et al.*, 2010). We extend the experiment by using all dataset classes which provides 3030 images. Figure 7 gives an example of some classes in the Caltech 101 dataset.

SVM classifier: We used LIBSVM with the RBF kernel together with the one vs. one approach to conduct the process of classification. The normalization of the feature vector ranges between {+1, -1}; this helps avoid numerical complexities during the calculation. It further ensures that the larger values do not control smaller ones. Thus, the normalization equation is:

$$x = \frac{2(x\text{-}min)}{(max\text{-}min)}$$

The best two parameters, C and γ of RBF kernel are responsible for producing meticulous classification. The LIBSVM grid-search has directly been used on the training data to identify the best C and γ parameters. The search spaces used in this study are $\{2^{-5}, 2^{-3}, \ldots, 2^{15}\}$ and $\{2^{-15}, 2^{-13}, \ldots, 2^{3}\}$ for C and γ, respectively. Finally, when K is equal to 10, the K-fold cross-validation will be utilized in learning the classifier.

Selecting the best parameters for the Gabor Filter (GF) and k-means clustering techniques: The GF has four parameters: scale, orientation, sigma and filter size. The number of the scale's values together with the number of the clusters invloved in the k-means clustering technique differ among applications. These parameters have thus been selected to improve VOC technique performance.

Best Gabor Filter (GF) parameters and k-means clustering: Based on different runs, the experiment shows that the best parameters entered into the Gabor filter for object categorization are (υ = 5, μ = 8, σ = 2.5 and filter size = 128×128). The k-means clustering technique, however, comprises an important part of performing the proposed method. During the experiment, we found a significant difference in accuracy when performing the suggested technique in both domians, the spatial and frequency ones. The best number of clusters in the former domain, thus, differs from that of the latter.

With the spatial domain, the experiment started with 4 clusters and ran until reaching 15 clusters in step 1. It stopped at 15 clusters because there is no significant improvement in accuracy in clusters 14 and 15. The 11 clusters provide higher classification accuracy with a 67.3333% average rate based on the Naive approach. Generalization using the Gabor filter is thus considered the best choice.

For the frequency domain, we start with 4 clusters and ran until reaching 20 clusters in step 1. We stop at 20 clusters because there is no significant improvement in accuracy at cluster 20. The best k-cluster is 18 as it gives higher accuracy than the other clusters, i.e., a 65% average rate based on using combined features with the Naive approach. One advantage of using the frequency domain is that it speeds up the system while some basic information is lost from the raw data when converting it from the spatial to the frequency domains.

RESULTS AND DISCUSSION

In this part, we briefly present VOC recognition accuracy based on the standard GF and proposed GGF in both domains, the spatial and frequency ones. We, also compare performance at the end of this study.

VOC technique results within the spatial and frequency domains: This study explains in a comprehensive way the results of VOC technique depending on the GF and proposed GGF methods and by utilizing the first twenty classes from Caltech 101 dataset. Table 1 displays the results of VOC technique in the spatial domain, replying on the Naive approach and single classifier. Table 1 Categorization result based on the standard GF and proposed GGF methods in the spatial and frequency domains.

After different 10 runs, the the proposed GGF-based VOC technique will outperform the standard GF method in average rate and standard deviation, with values of 67.3% and ±0.69299, respectively. The time was reduced by 25%, using 11 generalized filters instead of 40. Table 2 shows the accuracy average of the technique being used in the frequency domain with and without the proposed GGF method.

Table 2 shows the VOC technique with and without the suggested GGF method in the frequency domain. The GGF performs worse than the VOC technique in the spatial domain because some raw data is lost when transformation the data from the spatial to the frequency domain using the FFT method. Using the FFT algorithm, however has the benefit of speeding up the system and reducing time.

Performance comparison: It is interesting to compare the classification rates of the suggested method within two different spaces. This study uses the t-test approach to calculate the important difference of the suggested method. The p-value (probability-value) of the t-test is 0.05 and the t-test result is compared to this value: if the t-test value is <0.05, the proposed algorithm is considered significant. The t-test result of the proposed method decreases as the algorithm is considered more significant.

Based on the t-test result, the proposed GGF method-based Naive approach is considered statistically significant in the spatial domain and outperforms the standard GF method-based Naive approach ($p<0.05$, $p = 0.0001$). The proposed GGF method-based single classifier is highly important from the statistical point of view and outperforms the standard GF method that relies on a single classifier (with $p = 0.00082$). The time has been reduced by 25% using 11 generalized filters instead of 40. The results of the t-test of VOC technique together with the suggested GFF method within the frequency domain is not quite statistically significant and performs a bit worse based the Naive approach and single classifier, when compared to the VOC technique-based standard GF method. This result occurs because the number of combined filters is not enough to describe the object compared to the original 40 filters and raw information losses.

Based on the t-test results, the proposed method based Naive approach and single classifier is highly important statistically within the spatial domain. It further excutes the suggested method within the frequency domain (with $p = 0.0001$). The results of the t-test of the standard Gabor filter within the spatial domain show that such a method is statistically important. It also excutes the standard gabor filter within the frequency domain (with $p = 0.0402$ for the Naive approach and $p = 0.0486$ for the single classifier). In summary, the VOC technique-based standard GF and proposed GGF methods in the spatial domain outperform those in the domain of frequency. The VOC technique in the domain of frequency performs a bit worse because it loses some basic information when transforming the raw data from the spatial to the domain of frequency.

Performing VOC technique with all dataset categories: From the results obtained using the GGF method in the spatial domain and taking into account the twenty categories from Caltech 101 dataset, it is interesting to perform the VOC technique that depends on the suggested GGF with the entire Caltech 101 dataset to see how the proposed method performs. Table 3 shows the accuracy rate of the VOC technique that works with a mono-classifier and has a simple combination approach. Such a technique that is based on the suggested GGF method in the spatial domain performs well even with all Caltech 101 dataset classes. The standard deviation shows the proposed method is stable with diverse objects.

Table 1: Categorization result based on the standard GF and proposed GGF methods in the spatial and frequency domains

	Standard GF		Proposed GGF	
Variables	Single classifier	Naive approach	Single classifier	Naive approach
Accuracy average (%)	47.345	65.1667	52.9061	67.3
SD	±1.13245	±2.16168	±0.91206	±0.69299

Table 2: Classification results of the VOC technique in the frequency domain

	Standard GF		Proposed GGF	
Variables	Single classifier	Naive approach	Single classifier	Naive approach
Accuracy average (%)	41.9958	63.2667	44.5056	62
SD	±1.08954	±1.9539	±0.86007	±1.75119

Table 3: Classification accuracy of all Caltech 101 categories based on the proposed GGF method in the spatial domain

Variables	Single classifier	Naive approach
Average accuracy (%)	40.42070	57.0784
SD	0.47436	0.98742

CONCLUSION

This study has presented the visual object categorization technique that has a filter-based descriptor and a classifier marked by the standard Gabor filter and proposed GGF techniques. The results verify the refinedness of the suggested method in describing an object, as it provides different and distinctive feature maps that are differenty scaled and oriented for each of its objects.

The VOC technique, based on both the standard GF and proposed GGF methods has been tested in two different spaces: the spatial and frequency ones. Based on the results of the t-test, it has been found that VOC technique within the spatial domain generated better results than implementing it in the frequency domain for both the standard GF and proposed GGF methods. This occurs for several reasons. One reason is that transforming raw data from the spatial domain to that of the frequency is inefficient and may lose some important information. In other words, the representation of the FFT algorithm to the data is not as efficient as the actual raw data.

Conversely, the VOC technique, based on the proposed GGF method, gives a higher accuracy rate and outperforms the standard GF method. The proposed descriptor and classifier, that has filter and is marked by the GGF method has significantly improved the object categorization problem and increased system performance.

ACKNOWLEDGEMENTS

The researchers of the present work are much indebted to the Faculty of Information Science and Technology, University Kebangsaan Malaysia, for the facilities and financial support. Furthermore, a special thank is dedicated to the Ministry of Higher Education and Scientific Research (MOHESE), Iraq, for its important contributions to this research.

REFERENCES

Abdullah, A., 2010. Supervised learning algorithms for visual object categorization. Ph.D Thesis, Institute of Information and Computing Sciences, Utrecht University, Utrecht, Netherlands.

Abdullah, A., R.C. Veltkamp and M.A. Wiering, 2009. Spatial pyramids and two-layer stacking SVM classifiers for image categorization: A comparative study. Proceedings of the International Joint Conference on Neural Networks (IJCNN 2009), June 14-19, 2009, IEEE, Atlanta, Georgia, USA., ISBN:978-1-4244-3548-7, pp: 5-12.

Abdullah, A., R.C. Veltkamp and M.A. Wiering, 2010. Fixed partitioning and salient points with MPEG-7 cluster correlograms for image categorization. Pattern Recogn., 43: 650-662.

Abdullah, A.B. and M.A. Wiering, 2007. CIREC: Cluster correlogram image retrieval and categorization using MPEG-7 descriptors. Proceedings of the International IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP 2007), April 1-5, 2007, IEEE, Honolulu, Hawaii, USA., pp: 431-437.

Alemu, Y., J.B. Koh, M. Ikram and D.K. Kim, 2009. Image retrieval in multimedia databases: A survey. Proceedings of the 2009 5th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, September 12-14, 2009, IEEE, Kyoto, Japan, ISBN:978-0-7695-3762-7, pp: 681-689.

Choi, W.P., S.H. Tse, K.W. Wong and K.M. Lam, 2008. Simplified Gabor wavelets for human face recognition. Pattern Recognit., 41: 1186-1199.

Cortes, C. and V. Vapnik, 1995. Support-vector networks. Mach. Learn., 20: 273-297.

Cruz-Aceves, I., F. Oloumi, R.M. Rangayyan, J.G. Avina-Cervantes and A. Hernandez-Aguirre, 2016. Automatic segmentation of coronary arteries using Gabor filters and thresholding based on multiobjective optimization. Biomed. Sign. Proc. Contr., 25: 76-85.

Jain, A.K., M.N. Murty and P.J. Flynn, 1999. Data clustering: A review. ACM Comput. Surv., 31: 264-323.

Li, W., K. Mao, H. Zhang and T. Chai, 2010. Designing compact Gabor filter banks for efficient texture feature extraction. Proceedings of the 2010 11th International Conference on Control Automation Robotics and Vision (ICARCV), December 7-10, 2010, IEEE, Singapore, ISBN:978-1-4244-7814-9, pp: 1193-1197.

Lin-Lin, S. and J.I. Zhen, 2009. Gabor wavelet selection and SVM classification for object recognition. Acta Autom. Sin., 35: 350-355.

Liu, Y., D. Zhang, G. Lu and W.Y. Ma, 2007. A survey of content-based image retrieval with high-level semantics. Pattern Recogn., 40: 262-282.

Long, F., H. Zhang and D.D. Feng, 2003. Fundamentals of Content-Based Image Retrieval. In: Multimedia Information Retrieval and Management: Technological Fundamentals and Applications, Feng, D., W.C. Siu and H.G. Zhang, (Eds.). Springer, Berlin, Gernany, ISBN:9783540002444, pp: 1-26.

Manjunath, B.S., P. Salembier, and T. Sikora, 2002. Introduction to MPEG-7: Multimedia Content Description Interface. 1st Edn., John Wiley and Sons, New York, ISBN: 0-47148-678-7.

Mignotte, M., 2012. An energy-based model for the image edge-histogram specification problem. IEEE. Trans. Image Proc., 21: 379-386.

Nadernejad, E., S. Sharifzadeh and H. Hassanpour, 2008. Edge detection techniques: Evaluations and comparisons. Appl. Math. Sci., 2: 1507-1520.

Radman, A., K. Jumari and N. Zainal, 2014. Iris segmentation in visible wavelength images using circular gabor filters and optimization. Arab. J. Sci. Eng., 39: 3039-3049.

Senthilkumar, D. and S. Paulraj, 2016. Ensemble deep learning for multi label classification in the design of clinical decision support. Asian J. Inf. Technol., 15: 2632-2637.

Shristi, J., B.M. Sindhu and S. Venkatesan, 2016. A narrative methodology to recognize iris patterns by extracting features using gabor filters and wavelets. Proceedings of the International Seminar on Applied Physics, Optoelectronics and Photonics (APOP 2016) Vol. 61, June 28, 2016, EDP Sciences, Les Ulis, France, pp: 1-5.

Varma, M. and A. Zisserman, 2003. Texture classification: Are filter banks necessary?. Proceedings of the 2003 IEEE International Conference on Computer Society Computer vision and pattern recognition Vol. 2, June 18-20, 2003, IEEE, Madison, Wisconsin, USA., pp: 1-8.

Wang, J.Z., J. Li and G. Wiederhold, 2001. SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. IEEE Trans. Pattern Anal. Machine Intell., 23: 947-963.

Wu, X. and B. Bhanu, 1997. Gabor wavelet representation for 3-D object recognition. IEEE. Trans. Image Proc., 6: 47-64.

Yang, Y., D. Xu, F. Nie, S. Yan and Y. Zhuang, 2010. Image clustering using local discriminant models and global integration. IEEE Trans. Image Process., 19: 2761-2773.

Yavuz, H.S., H. Cevikalp and R. Edizkan, 2016. A comprehensive comparison of features and embedding methods for face recognition. Turk. J. Electr. Eng. Comput. Sci., 24: 313-340.

Zhu, J., M.I. Vai and P.U. Mak, 2004. A new Enhanced Nearest Feature Space (ENFS) Classifier for Gabor Wavelets Features-Based Face Recognition. In: Biometric Authentication, Zhang, D. and A.K. Jain (Eds.). Springer, Berlin, Gernany, ISBN:978-3-540-22146-3, pp: 124-130.