ERES International Journal of Computer Networks

Vol. 1. No. 2 2014

SVM CLASSIFIER BASED OBJECT RECOGNITION

Ms.V.Amaraselvi[#1]
Department of ECE
Regional Office of Anna University
Madurai, Tamil Nadu, India
Email:vamaraselvi@gmail.com

Abstract-Object Recognition Plays an important role in analyzing the contents of image and video. This paper is to find the tag of the input image based on the training image and perform the image retrieval by the generating tag. To tag the image first find which content found on that image. To do this work the training images is needed because this project makes comparison between the input image and the input image to find the similarities. Then put the tag based on the most similar images. The feature extraction technique is used to find the content of the image. The DWT, Gabor filter, Edge histogram descriptor, Color moment and SIFT are used for feature extraction. To find the similarities between the training and the testing images feature extraction is used. After extracting the features one classifier is needed. SVM is used in this project and this classifier finds the most similar image from the training image then the tag of the similar image is considered as the final tag. Then this final tag is applied on the input image.

I.INTRODUCTION

Wearable devices are developed for new mobile applications such as augmented reality, robot and computer vision and electronic aids for the people who are visually impaired an essential step in these applications. Object recognition plays an important role in analyzing the contents of images and videos. For augmented reality object recognition is employed to locate the position and orientation of markers to render 3-D graphics so that users are able to interact with virtual objects in a real scene.

For robot and computer vision images captured from sensors are analyzed so that machines can locate the target based on the object recognition and complete the tasks without human intervention. For electronic aids for visually impaired people the devices serve as a pair of artificial eyes for users and object recognition assists users in detecting obstacles and helps them walk around safely. Since the computations for object matching are heavy the development of techniques for wearable devices that can effectively accelerate the processing speed is a critical issue.

An object recognition system finds objects in the real world from an image of the world using object models which are known as priori. This task is surprisingly difficult. Humans perform object recognition effortlessly and instantaneously. Algorithmic description of this task for implementation on machines has been very difficult. In this methods the different steps in object recognition and introduce some techniques that have been used for object recognition and the different types of recognition tasks that a vision system may need to perform. It will analyze the complexity of the tasks and present approaches useful in different phases of the recognition task.

We discuss previous work on related topics in section II and explain about system design of proposed method in section III and section IV demonstrates simulation results of the system model and discuss the experiment results. Finally section V concludes the paper.

II.RELATED WORK

To enhance the user's perception and interaction with the real world through supplementing the real world with 3D virtual objects that appears to co-exist the same space as the real world. The state of the art in AR today is comparable to the early years of VR –many research systems have been demonstrated but few have matured beyond lab based prototypes.

We define AR systems to share the following properties: Blends real and virtual in a real environment. Real time interactive and registered in 3D.Registration refers to the accurate alignment of real and virtual objects. Without accurate registration the illusion that the virtual objects exist in the real environment is severely compromised in paper [1].

A network on chip based parallel processor is presented for bio-inspired real time object recognition with virtual attention algorithm. It contains an ARM10 compatible 32 bit main processor, 8 single instruction multiple data clusters with 8 processing elements in each cluster, a cellular neural network based visual attention engine, a matching accelerator and a DMA like external interface.

The VAE with 2-D shift register array finds salient objects on the entire image rapidly. Then the parallel processor performs further detailed image processing within only the preselected attention regions. The low latency NOC employs dual channel, adaptive switching and packet based power management. A 201.4 GOPS real time multi object recognition processor is presented with a three stage pipelined architecture. Visual perception based multi object recognition algorithm is applied to give multiple attentions to multiple objects in the input image.

For human like multi object perception a neural perception engine is proposed with biologically inspired neural networks and fuzzy logic circuits. In the proposed hardware architecture three recognition tasks are directly mapped to the neural perception engine, 16 SIMD processors including 128 processing elements and decision processor respectively and executed in pipeline to maximize throughput of the object recognition.

For efficient task pipelining proposed task/power manager balances the execution times of the three stages based on intelligent workload estimations. In addition a 118.4 GB/s multi casting network on chip is proposed for communication architecture with incorporating over all 21 IP blocks in [3]. In [4] proposes hardware architecture for object detection based on an adaboost learning algorithm with haar like features as weak classifiers.

We analyze and discuss the parallelism in this detection algorithm and propose a partially parallel execution model suitable for hardware implementation. This parallel execution model exploits the cascade structure of classifiers in which classifiers located near the beginning of the cascade are used more frequently than subsequent classifiers and assign more resources to these earlier classifiers to execute in parallel than to subsequent classifiers.

It is dramatically improves the total processing speed without a great increase in circuit area. Moreover the partially parallel execution model achieves flexible processing performance by adjusting the balance of parallel processing. To find an efficient parallel processing method we categorize the available parallelism in the detection algorithm. In [5] high performance devices can improve the life quality of visually impaired people.

It allows detection of obstacles very close to the user and requires instantaneous reaction when an obstacle is encountered. Most have used the transmission of energy wave and the reception of echoes from objects or charge coupled devices cameras. The ESSVI can be classified on the basis of the target. one category is clear path indicators which provide information about the presence and the approximation position of obstacles meaningful advances have been made in the past decades in the development of electronic sensory systems for visually impaired people.

In sensory supplementation the function of ESSVI is to provide additional information that is critical and not available otherwise. K-means is an important clustering algorithm that is widely applied to different applications including color clustering and image segmentation. To handle large cluster numbers in embedded systems to develop the systems.

To achieve low power consumption the visual vocabulary processor which is based on bag of words matching algorithm is used to advance the matching stage from the feature level to the object level and results in a 97% reduction in the required memory bandwidth compared to previous recognition systems. The matching efficiency of the VVP enables the system to support full time processing. In [8] a heterogeneous many core object recognition processor is proposed to realize robust.

It's efficient object recognition on real time video of cluttered scenes to achieve high effective GOPS/W or EGOPS/W which only counts operations carried out in a meaningful regions of input image. This is achieved by the unified visual attention model which confines complex scale invariant feature transform feature extraction to meaningful regions while rejecting meaningless background regions. The intelligent inference engine a mixed mode neuro fuzzy inference system performs the top down familiarity of the UVAM which guides attention towards pre learned objects.

III. SYSTEM DESIGN

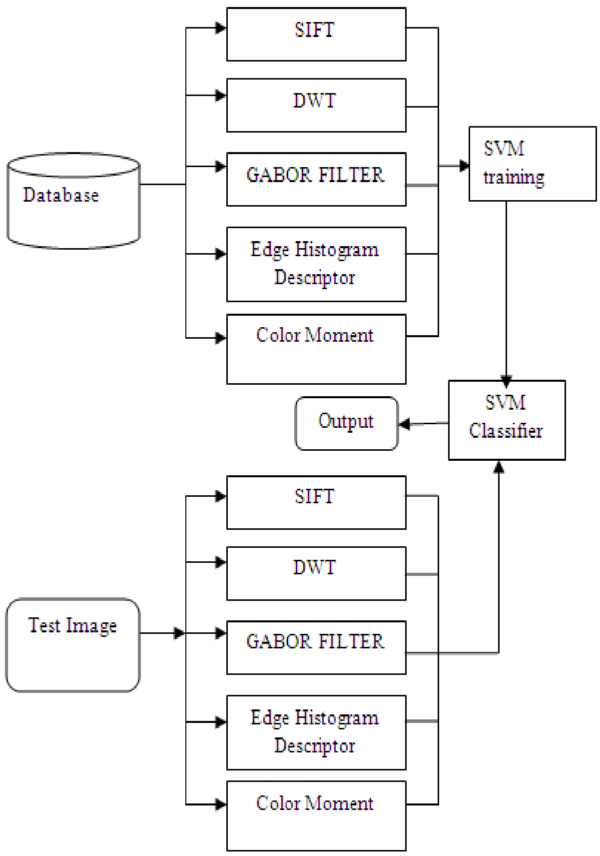

Fig.1. Proposed Method for System Design

A. DISCRETE WAVELET TRANSFORM

Discrete Wavelet transform is a multi resolution analysis. In DWT the image is filtered in both the horizontal and vertical directions using separable filters. This creates four sub bands. Sub band LL1 represents the horizontal and vertical low frequency components of the image. Sub band HH1 represents the horizontal and vertical high frequency components of the image. Sub band LH1 represents the horizontal low and vertical high frequency components.

Sub band HL1 represents the horizontal high and vertical low frequency components. In numerical analysis and functional analysis a discrete wavelet transform is any wavelet transform for which the wavelets are discretely sampled .As with other wavelet transforms advantage is that it captures both frequency and location information.

B. GABOR FILTER

In image processing a Gabor filter named after Dennis Gabor is a linear filter used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system and they have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal wave.

Its impulse response is defined by a sinusoidal wave multiplied by a Gaussian function. A set of Gabor filters with different frequencies and orientations may be helpful for extracting useful features from an image.

C. EDGE HISTOGRAM DESCRIPTOR

An image histogram is a type of histogram that acts as a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each tonal value. By looking at the histogram for a specific image a viewer will be able to judge the entire tonal distribution. The histogram is a graphical representation of the distribution of data.

It is an estimate of the probability distribution of a continuous variable. A histogram is that the height of the rectangle is also equal to the frequency density of the interval. The total area of the histogram is equal to the number of data. A histogram may also be normalized displaying relative frequencies.

D. COLOR MOMENT

Color moments are measures that characterize color distribution in an image in the same way that central moments uniquely describe a probability distribution. Color moments are mainly used for color indexing purposes as features in image retrieval applications in order to compare how similar two images are based on color. Usually one image is compared to a database of digital images with Recomputed features in order to find and retrieve a similar image. Color moments can be computed for any color model. Color moments are scaling and rotation invariant. It is usually the case that only the first three moments are used as features in image retrieval applications as most of the color distribution information is contained in the low order moments. Since color moments are encoded both shape and color information they are a good feature to use under changing lighting conditions, but they are a good feature to use under changing conditions, but they are cannot handle occlusion very successfully.

Color moments can be computed for any color model. Three color moments are computed per channel. Computing color moments is done in the same way as computing moments of a probability distribution.

E. SCALE INVARIANT FEATURE TRANSFORM

Scale invariant feature transform is an algorithm in computer vision to detect and describe local features in images. Applications include object recognition, robotic mapping and navigation, image stitching, 3D modeling, gesture recognition, video tracking and individual identification of wildlife and match moving. For any object in an image interesting points on the object can be extracted to provide a feature description of the object.

F. SUPPORT VECTOR MACHINES

Support vector networks are supervised learning models with associated learning algorithms that analyze data and recognize patterns used for classification and regression analysis. The basic SVM takes a set of input data and predicts for each given input which of two possible classes from the output making it a non-probabilistic binary linear classifier. Given a set of training examples each marked as belonging to one of two categories an SVM training algorithm builds a model that assigns into one category or the other.

An SVM model is a representation of the examples as points in space mapped so that the examples of the separate categories are divided by a clear gap that is wide as possible.In addition to performing linear classification SVMs can efficiently perform a nonlinear classification using what is called the kernel trick implicitly mapping their inputs into high dimensional feature spaces.

A Classifier that separates a set of objects into their respective groups with a line. Most classification tasks however and not simple and often more complex structures are needed in order to make an optimal separation. I.e. correctly classify new objects of the test cases on the basis of train cases that are available. Compared to the previous one it is clear that a full separation of the green and red objects would require a curve. Classification tasks based on drawing separating lines to distinguish between objects of different class memberships are known as hyper plane classifiers.

IV. SIMULATION RESULTS AND DISCUSSIONS

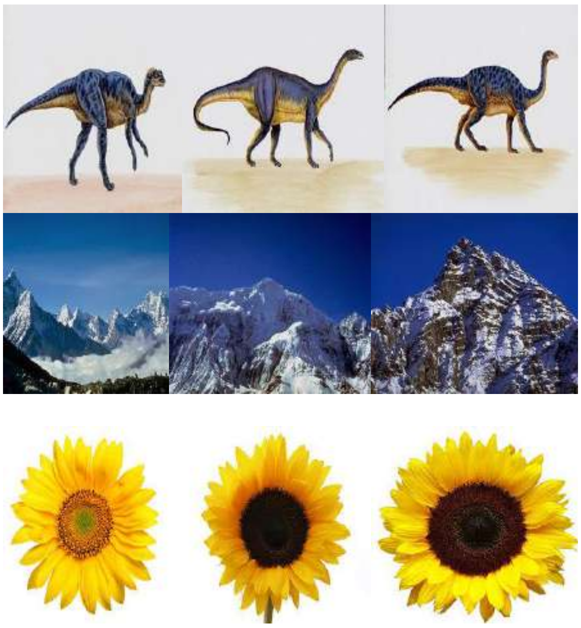

Fig.2. Experimental Images

The performance of the algorithm is evaluated on several real images.These pictures are the most widely used standard test images used for the algorithm in this method. The image contains a nice mixture of detail, flat regions, shading and texture that do a good job of testing various image processing algorithms. These are till in the industry standard for tests. It is a good test images.These images are used for many image processing researches.

Feature extraction techniques are given below and the features are extracted by using the above explained features.

Fig.3. Preprocessed Images

In preprocessed images the contrasts of the training images are enhanced.The contrast enhanced images are used for feature extraction.

Fig.4. Feature Extraction for Training Images

In the feature extraction the features of the images are extracted. The image features are DWT,SIFT,color moment,gabor,edge histogram.

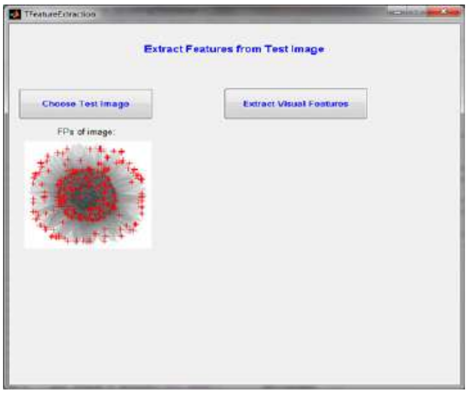

Fig.5. Extract features from test image

In Fig.5 To find the similarities between the training and the testing images feature extraction is used.After extracting the features one classifier is used.

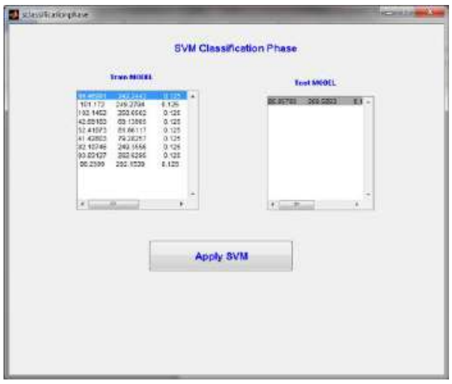

Fig.6. SVM classification

In Fig.6 SVM Classifier is used and this classifier finds the most similar image from the training images.

Fig.7. Output Tag

In Fig.7 the tag of the most similar image is considered as the final tag and then this final tag is applied on the input image.

V.CONCLUSION

This project retrieves image based on the content of the image and specifies a suitable tag for the given image automatically using the training datasets. This project gets image as the input and the closest related tag is given to the new image from the training images. The feature extraction technique is used to find the content of the image. The DWT, Gabor filter, SIFT, Edge histogram descriptor, color moment said to be the image features which are used for feature extraction and SVM is used to classify the image and the tag is specified to the given new image.For the experimental purposes several real world images are used.From the experimental results it is shown that the proposed method performs well than the existing methods.

REFERENCES

[1] B. Ando (2003) ‘Electronic sensory systems for the visually impaired’, IEEE Instrum.Meas. Mag., vol.6, no. 2,pp. 62-67,Jun.

[2] R. Azuma,Y. Baillot, R. Behringer,S. Fiener,S. Julier, and B.Mac-Intyre (2001) ‘Recent advances in augmented reality’,IEEE Comput. Graph.Appl., vol. 21, no. 6,pp. 34-47, Nov.

[3] T.-W. Chen and S.-Y. Chien, (2011) ‘Flexible hardware architecture of hierarchical K-means clustering for large cluster number’, IEEE Trans. Very Large Scale Integr.(VLSI) Syst.,vol. 19, no.8, pp.1336-1345, Aug.

[4] M. Hiromoto, H. Sugano, and R. Miyamoto (2009) ‘Partially parallel architecture for adaboost-based detection with haar-like features’, IEEE Trans. Circuits Syst. For Video Technol.,vol. 19,no. 1, pp.41-52, Jan.

[5] J.-Y. Kim,M. Kim, S. Lee, J. Oh, K. Kim, S. Oh, J.-H. Woo, D. Kim,and H.-J. Yoo (2009) ‘A 201.4 GOPS 496 mw real time multi-object recognition processor with bio-inspired perception engine,’ in ISSCC Dig. Techn. Papers. pp.150-151.

[6] K. Kim,S. Lee,J.-Y. Kim,M. Kim,D. Kim,J.-H. Woo, and H.-J.Yoo, ‘A 125 GOPS 583 mw network-on-chip based parallelprocessor with bio-inspired visual-attention engine,” in ISSCCDig. Techn. Papers, 2008,pp.308-309.

[7] S. Lee,J. Oh, M. Kim, J. Park, J. Kwon, and H.-J. Yoo (2010)‘A 345 mw heterogeneous many-core processor with an intelligent inference engine for robust object recognition’, inISSCC Dig. Techn. Papers, pp. 332-333.