Identification of Precise Object among Various Objects using Sparse Coding

Giby Jose
Electronics & Communication Engineering Department,
Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India.

P. Manimegalai
Electronics & Communication Engineering Department,
Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India.

ABSTRACT

In order to identify the exact coconut object form the image, the following methodology is proposed. The input image is preprocessed. Its quality is enhanced using histogram equalization to produce a better result for region-based feature extraction. The edges are then detected in the image segmentation process, as this information is most essential for the classifier algorithm. After detecting the edges the CHT algorithm is applied to identify the coconut. The performance measures viz. Recognition rate, Precision, Recall and simulation time are computed. To enhance the performance, the proposed algorithm is applied and the performance parameters are tabulated. A comparative study is made between the CHT and proposed algorithm to validate the superiority of the proposed algorithm. Sparse Representation-Based Classification (SRC) is face recognition innovation in current years, which has effectively addressed the recognition problem. It recognizes an object based on the training images made available in the gallery. In this paper, SRC has been intended for coconut object identification.

Keywords
Coconut Identification, Sparse Representation-Based Classification (SRC), Recognition Rate.

1. INTRODUCTION

The coconut harvesting machines, being designed now are available with camera sensors. When complete automation of these machineries is needed, then the images captured by these sensor plays a vital role. Based on the images captured, the coconut and its stem shall be identified. After identifying the coconut, the coconut plucking mechanism shall be activated. This process can be done completely through a remote control if exact coconut objects are identified. Despite the various technological advances, the reliability of identifying the coconut from the image is tedious task. It depends on various factors which include, the quality and position of the sensor, the angle at which the image is captured and the height and diameter of the coconut tree. In spite of these parameters, if an image is obtained, then identifying the coconut objects from such images is a real challenge. This paper addresses this issue with the objective of identifying the coconut from an image irrespective of the input nature of the image [1].

2. RELATED WORKS

The regular human computer interaction (HCI) technique, in which one particular user faces a computer and interacts by using it via the mouse or perhaps a keyboard, were designed to point out the indication of explicit messages though ignoring implicit info on the consumer, such because the user's changes within the affective states [2]. Human activity

analysis is usually a significant component in image and online video understanding and also the large visual variance together with semantic ambiguity underlying this topic causes it to be a struggle [11]. Human perception is incredibly effective with identifying human being subjects via facial images, most modern-day face identification systems have never achieve similarly good identification accuracy and there may be numerous variables affecting the system performance [5].

Object detection in addition to recognition with noisy in addition to cluttered images is complicated problem with computer vision [8]. From the context involving holistic strategies, where every one of the pixels of a face image rather than a part of it could be seen since candidate functions, abundant work has been devoted to research various strategies projecting the actual high dimensional makeup image straight into lower dimensional function spaces. Eigen faces, Fisher faces and Laplacian faces are the most used examples [7]. Sparse representations are already recently exploited in several pattern identification applications. These approaches provide the assumption which a test approximately is based on a lower dimensional subspace spanned by the training data thereby can end up being compactly represented by the few training samples [4].

Sparse representation face recognition (SRC) is actually modelled based on the image subspace assumption, which signifies that various training images could form an impartial partition for that image subspace, and virtually any test image could be expressed by the linear combination of the very same face training image [10]. The sparse representation has been used in a number of use-cases [9]. Sparse representation, also referred to as compressed sensing, has been applied not too long ago to image based recognition and with sparse representation we can easily represent each and every face by a set of features, which sufficiently characterize each individual [6]. With sparse representation based face Recognition, usually all of us assume how the face images are in-line. Recently, sparse representation has been extended to fix the misalignment or pose adjust [3].

Weihong Deng et al. [12] have extended SRC to applications where there are very couple of or obviously any good single, training images for each subject. They have got assumed how the intra class variations of just one subject could be approximated by the sparse linear combination of those involving other subjects, Extended Sparse Representation-Based Classifier (ESRC) received applied a auxiliary intra class variant book to represent the possible variation between training in addition to testing images. The dictionary atoms routinely have represented intra class sample distinctions computed via either the actual gallery faces themselves or the

generic faces which can be outside the actual gallery. Experimental results on the AR in addition to FERET database have showed that ESRC offers better generalization capability than SRC for just sampled encounter recognition within variable movement, illuminations, hide, and a long time. The superior connection between ESRC include suggested that when the dictionary was appropriately constructed, SRC algorithms could generalize well on the large-scale encounter recognition problem, even having a single training image for each class.

John Wright et al. [13] have provided a fresh framework for two crucial issues in encounter recognition: function extraction in addition to robustness for you to occlusion. Regarding feature extraction, they include showed that when sparsity within the recognition problem was appropriately harnessed, the selection of functions was no more critical. Their framework can handle errors as a result of occlusion in addition to corruption evenly by exploiting the point that these errors can be sparse with regards to the standard (pixel) basis. The concept of sparse representation had assisted to predict simply how much occlusion the actual recognition algorithm can handle and how to choose the training images to maximize robustness for you to occlusion. They include conducted extensive experiments in order to verify the actual efficacy with the proposed formula work.

Jian Lai and Xudong Jiang [14] have proposed the novel construction of robust face recognition based on the sparse representation. Image was divided straight into modules in addition to each module was prepared separately to discover its reliability. A reconstructed image through the modules weighted through their consistency was formed for that robust identification. They include proposed make use of the modular sparsity in addition to residual jointly to determine the modular reliability. The recommended framework received advanced both the modular in addition to global sparse representation approaches, especially in working with disguise, huge illumination modifications and expression changes. Weighed against the related state-of-the-art strategies, experimental effects on standard face database had tested the advancement with the proposed process.

Yongkang Wong et al. [15] have proposed a different approach for SR-based face verification, where SR development was carried out on nearby image patches as opposed to the entire face. The received sparse signals were pooled via averaging to create multiple region descriptors, which then formed a general face descriptor. Because of the deliberate decrease in spatial contact within each and every region (caused through averaging), the actual resulting descriptor had been robust for you to misalignment in addition to various image deformations. Within their proposed work, they received evaluated a number of SR development techniques: $l^{(1)}$-minimisation, Sparse Auto encoder Neural Network (SANN) in addition to an implicit probabilistic technique based on Gaussian mixture models. Detailed experiments about AR, FERET, Yale W, BANCA in addition to Choke Point datasets include showed how the local SR strategy obtains drastically better plus much more robust overall performance than a number of previous talk about of-the-art of utilizing holistic SR strategies, on both the traditional closed-set identification task and also the more pertinent face verification task. The experiments also provide showed in which $l^{(1)}$-minimisation-based encoding incorporates a considerably better computational cost when compared with SANN-based in addition to probabilistic development, but had cause higher recognition rates.

3. PROPOSED METHODOLOGY

The main objective of the proposed work is to develop an algorithm for identifying the coconut objects in the images. Manual screening may not be effective in the semi or fully automated coconut harvesting systems. To make the system to be more reliable and effective through automation, processing the images captured is needed. This demand has led to the development of several techniques to identify the coconuts in the images. The research objectives are as follows:

- Extensively review the available image classification and identification schemes. In addition, the literatures relevant to image preprocessing techniques viz. image enhancement, edge detection, histogram equalization and feature extraction shall be also reviewed. These terms are needed as the image captured has to be pre-processed through all these techniques.

- Obtaining a preprocessed image by applying histogram equalization and edge detection algorithms.

- Extract the feature of the image through down sampling method

- Identify the coconut image from the feature extracted though Circular Hough Transformation. Analyze the performance of the CHT technique.

- Develop the simulation model for proposed Sparse Representation Based Classifier (SRC) algorithm to identify the coconut object and analyze its performance.

- Compare the performance of the CHT and proposed SRC algorithms.

Demonstrate the superiority of the proposed SRC algorithm.

4. SPARSE DICTIONARY

Sparse dictionary learning is a demonstration learning method. The objective of this learning method is to find a sparse representation of the input data, in the form of a linear combination of basic elements. These elements are called as atoms and they compose a dictionary. The atoms may not be orthogonal and so, the dimension of the representation space may be higher than the input space. The above two properties lead to having seemingly redundant atoms that allow multiple reconstruction ways but also provide an improvement in sparsity and flexibility of the representation. The basic principle of dictionary learning is that the dictionary can be inferred from the input data. Earlier, signal processing has been effected with predefined dictionaries such as Fourier or Wavelet transforms. But the evolution of sparse dictionary, which is inferred from the input data, makes the data significantly improve the sparsity. This enhances the results of signal processing. The applications of sparse dictionary learning includes image denoising, classification, video and audio processing.

Sparse representations are used in several pattern identification applications. The SRC identification method of object detection consists of three main stages, Sparse Dictionary Design, Feature Extraction and Object Detection. Primarily, a sparse dictionary is designed along with large variety of different training images. The images comprise of various items which include target image too. The next process is feature extraction. If sparsity inside the recognition is correctly harnessed, the choice of features will not be an

issue in feature extraction. The problem associated with feature extraction is the volume of features, which can sufficiently substantial and compute the sparse representation in the right way. In order to avoid such problems, down sampling method of feature extraction is attempted. The unconventional features such as down sampled images along with random projections perform equally well as typical features viz. Eigen and Laplacian faces. This is because the dimension in the feature surpasses certain threshold predicted from the theory of sparse representation in face recognition. Finally, in the assessment process, the sparse dictionary characteristically fixes the object to be identified. The sparse dictionary method focuses on providing a sparse representation by constructing a dictionary enforced by the Equation 1.

$$Z = XY \qquad (1)$$

Where,

Z is sparse approximation

Y is a sparse matrix

X is pre-defined analytical dictionary

4.1. Sparse Dictionary Design

Dictionaries are designed to fit the proposed model by selecting from one of a pre-specified set of linear transforms or adapting the dictionary to a set of training signals. Selection of dictionary is an important phenomenon in SRC. Large numbers of training samples are to be collected and maintained for building a dictionary. Increasing the training samples, will improve the representation ability and help in solving complex sparse problems. The training samples need to be decided with utmost care as the dictionary atoms will cover the sub-space made up of various samples. For instance, samples with different gestures and expressions, under various illumination conditions and with no manual occlusion are selected to make a dictionary. Dictionary learning leads to creation of dictionaries that are sparser compared to the given instruction sample arranged or the built in sample. This ensures that fewer atoms are able to represent the newest sample from the same type, and sparser and much more accurate sparse solutions will probably be obtained.

The dictionary design is over, when the databases are collected. Then the training phase starts. Throughout the training phase, the training samples usually are processed by simply multi-resolution blocking, and non-lap sub-blocks of each and every training image with distinct resolutions along with space positions are acquired. When sparse dictionary is applied for object identification, the input data set has to be matched with the atoms in the dictionary. It arise the necessity for linear combinations of input variables, which explains the object. It is done by using special techniques viz. sparse PCA and LDA.

4.1.1. Sparse PCA

In the training phase, the multivariate data sets (input data sets of the image) are treated with a technique named Sparse Principal Component Analysis (Sparse PCA). It is a technique used in statistical analysis, for the reduction of dimensionality of data. The reduction is done by adding sparsity constraint on the input variables. This creates new variables. The number of new variables created will be lower than the input variables in the original dataset. But all these variables still explain most of the variance present in the original data. Sparse PCA is mostly applied technique because it finds the linear combinations that contain just a few input variables.

4.1.2. Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is utilized along with sparse PCA to explain the data set of the object. LDA is technique used to find linear combination of features, which characterizes or separates two or more classes of objects. The resulting combination may be used as a linear classifier for dimensionality reduction and classification. The input data set of the image is treated through sparse PCA and LDA. This results in a characteristic transformation matrix, which relive the dimensions of sub-blocks.

4.1.3. Sparse Dictionary Training Process

Figure 1 depicts the sparse dictionary training process. The probe image is the one, which is to be investigated. From the probe image, the data sets are collected and the sparse code is generated. Then the data sets are compared with the sparse dictionary which is already formed. Based on the results of matching the objects are identified and classified.
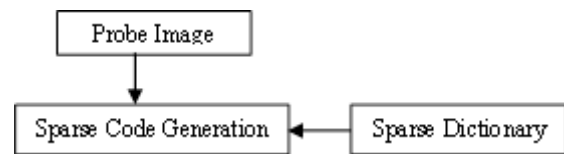

Figure1: Sparse dictionary training process

5. FEATURE EXTRACTION

Feature extraction is a special form of dimensionality reduction. It starts from an original set of measured data and builds the desired values envisioned to be useful and non-redundant. These values facilitate the learning process and assists in better human interpretations. Large number of data which are redundant can be transformed into a reduced set of features. This is termed as feature selection and the selected feature contains the relevant information from the input data. With the selected features, the desired task can be performed. The commonly used feature extraction techniques include Histogram of Oriented Gradients (HOG), Speeded Up Robust Features (SURF), Down Sampling, Local Binary Patterns (LBP), Haar Wavelets and color histograms.

5.1. Down Sampling

Down Sampling is one of the feature extraction processes. In this process the resolution of an image is changed. Changing the resolution decreases the size of an image. Down sampling improves the signal to noise ratio. This is the reason for the use of down sampling method of feature extraction.

6. IMPLEMENTATION

In the coconut detection process, the capture image is pre-processed. The input image used for the study is same as that of used in CHT. Contrast of the image is enhanced through histogram equalization. This result in large number of data sets. Based on the data sets, a sparse dictionary is created. Then the testing image is separated into sixteen blocks by down-sampling. Then the layers are sub divided into sub-blocks and the first layer is obtained. Down-sampling is done to be able to evenly partition the sub-blocks straight into four blocks to discover the second layer of sub-blocks; and continued down-sampling to discover the third layer of sub-blocks. Finally, down-sampling is used to get reduce the images into various sub-blocks of the training and testing data sets.

7. RESULTS AND DISCUSSION

The existing object detection system uses the separability filter proposed by Fukui and Yamaguchi (Trans. IEICE Japan

J80-D-II. 8, 2170-2177, 1997) to obtain the best object candidates and next, the system uses the circular Hough transform (CHT) to detect the presence of circular shape. The main contribution of the work consists of using together two different techniques in order to take advantages from the peculiarity of each of them. As the results of the experiments, the object detection rate of the system was 90% for 25 images by moving the circle template every 20 pixels to right and down.

Fig 2: Sparse Dictionary generated for a testing image

Figure 2 shows the learned sparse dictionary for a coconut testing image and Figure 3 shows the sample input images of the proposed work. Figure 4 show examples of the images for which the proposed system gave successful results.

In the proposed work 100 images were trained for the sparse dictionary. The testing process is done using 10 images taken from the training images.

Fig 3: Sample coconut images used in the proposed work

The proposed work is analysed by three metrics namely Precision, recall and Accuracy. The demonstration of these assessment metrics are specified in equations.

Precision: Precision is defined as the fraction of number of images recognized to the total no of images.

Recall: Recall is defined as the fraction of images recognized to the total number of relevant images in the dictionary.
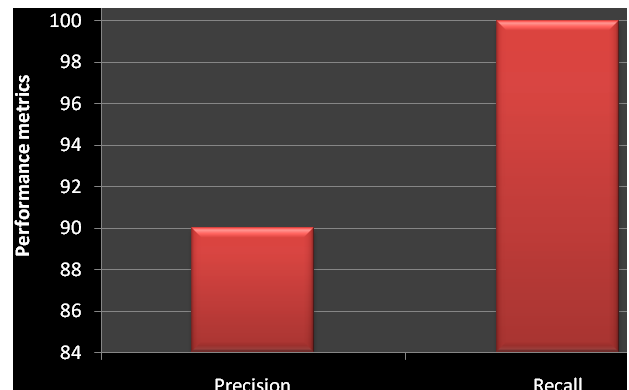

Fig 4: Detection of coconut object

Fig 5: Performance measures

Figure 5 shows the performance of the proposed system is compared with the existing circular Hugh transform [8] method in terms of recognition rate and success rate. The recognition rate obtained by the proposed and existing system is given in table 1 below.

Table 1: Comparison of Recognition Rate

Recognition Method	Recognition rate
Circular Hough transform [7]	90
Sparse representation	92

In this comparison while testing 90 percentages obtained for recognition by existing Circular Hough transform and 92 percentages for the proposed system. The graphical representation of the comparison is shown in fig 8, where the success rate is plotted to show the variation of proposed system from the existing system.
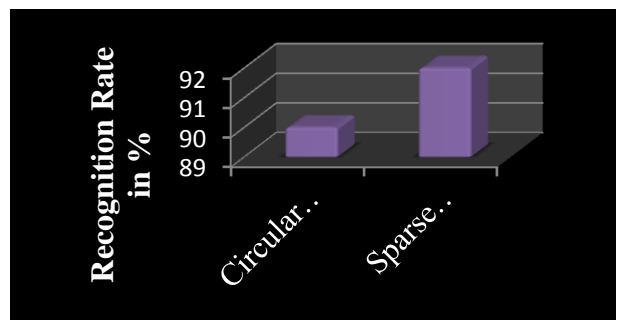

Fig 6: Comparison of Recognition rate

Figure 6 shows the comparison of recognition rate with existing techniques.Fig.6 is a chart to exhibit that, our object

detection system is with high accuracy of 92% in terms of the recognition accuracy in comparison with other existing system. In table 2 the average success rate of proposed and existing system is given.
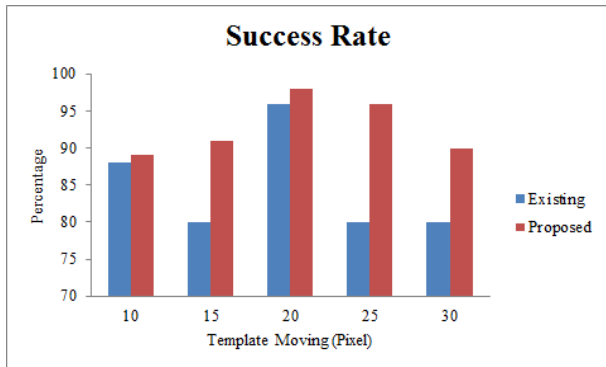

Fig 7: Comparison of Success rate

The success rate obtained by proposed and existing system in different template moving pixel level is shown in the fig 7. The maximum success ration obtained by the existing system is 96%, but by employing the proposed system the maximum success rate is 98%. The minimum success rate obtained by the existing system is 80% and proposed system the minimum success rate is 89%. The maximum success rate of both existing and proposed system is obtained for template moving value 20, but the minimum success rate is obtained many *times* in existing system i.e., for template moving value 15, 25 and 30 and only once by in proposed system at template moving value 30.This performance analysis clearly shows the improved performances of proposed system over the existing system.

8. CONCLUSION

In this research the coconut location can be detected but some constraint had been exposed. The sparse representation classifier tends to detect the possible coordinates of the coconut. We made experiments using 100 training images and 10 testing images to evaluate the performance of the proposed system. As the results of the experiments, the highest recognition rate of the proposed system to detect object was 92%. The problem arises if the coconut image is captured near and the image tends to be small. Besides that, some of the coconuts are in a bunch and overlaps between each other. These make the detection method difficult in carrying out the task. This work can be extended to detection of tender coconuts among various others by doing thermal analysis along with image processing.

9. REFERENCES

[1] Giby Jose and P. Manimegalai, "Identification of Precise Object Among Various Objects Using Sparse Coding - A Review", Pak. J. Biotechnol. Vol. 13 (special issue on Innovations in information Embedded and communication Systems) Pp. 519- 522 (2016)

[2] Shiqing Zhang, Xiaoming Zhao and Bicheng Lei, "Robust Facial Expression Recognition via Compressive Sensing", Sensors, Vol. 12, pp. 3747-3761, 2012

[3] Lei Zhanga, Meng Yanga, and Xiangchu Fengb, "Sparse Representation or Collaborative Representation: Which Helps Face Recognition?",IEEE International Conference on Computer Vision, pp. 1-8, 2011.

[4] Yi Chen, Thong T. Do, and Trac D. Tran, "Robust Face Recognition using Locally Adaptive Sparse Representation", In Proceedings of IEEE 17th International Conference on Image Processing, pp. 1657-1660, 2010.

[5] Minakshi S. Nagmote and Dr.Milind M. Mushrif, "Review: Sparse Representation for Face Recognition Application", International Journal of Engineering Trends and Technology (IJETT), Vol. 4, pp. 1772-1775, 2013.

[6] Rania S alah El-Sayed, Prof. Dr. Mohamed Youssri El-Nahas and Prof. Dr. Ahmed El Kholy, "Sparse Representation Approach for Variation-RobustFace Recognition Using Discrete Wavelet Transform", IJCSI International Journal of Computer Science Issues, Vol. 9, No. 3, pp. 275-280, 2012.

[7] Hanxi Li, Peng Wang andChunhua Shen, "A Robust Face Recognition System via Accurate Face Alignment and SparseRepresentation", In proceedings of International Conference on Digital Image Computing: Techniques and Applications, pp. 262 – 269, 2010.

[8] Mohamed Rizon, HanizaYazid, Puteh Saad, Ali Yeon MdShakaff, Abdul Rahman Saad, Masanori Sugisaka, Sazali Yaacob, M.RozailanMamat and M.Karthigayan, "Object Detection using Circular Hough Transform", American Journal of Applied Sciences, Vol. 2, No. 12, pp. 1606-1609, 2005.

[9] Gaurav Aggarwal, Soma Biswas, Patrick J. Flynn and Kevin W. Bowyer, "A Sparse Representation Approach to Face Matching across Plastic Surgery", In Proceedings of IEEE Workshop on Applications of Computer Vision, pp. 1-7, 2012.

[10] Deng Nan, Zhengguang Xu and ShengQinBian, "Face Recognition Based on Multi-classifierWeighted Optimization and Sparse Representation", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol.6, No.5, pp.423-436, 2013.

[11] Chenyang Zhang and Yingli Tian, "Subject Adaptive Affection Recognition via Sparse Reconstruction", In proceedings of CVPR workshop, pp. 351-358, 2014.

[12] Weihong Deng, Jiani Hu, and Jun Guo, "Extended SRC:Undersampled Face Recognitionvia Intraclass Variant Dictionary", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, No. 9, pp. 1864-1870, 2012.

[13] John Wright, Allen Y. Yang, Arvind Ganesh, S. Shankar SastryandYi Ma, " Robust Face Recognition via SparseRepresentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 2, pp. 1-18, 2009.

[14] Jian Lai and Xudong Jiang, "Modular Weighted Global Sparse Representationfor Robust Face Recognition", IEEE Signal Processing Letters, Vol. 19, No. 9, pp. 571-574, 2012.

[15] Yongkang Wong, Mehrtash T. Harandi and Conrad Sanderson, "On robust face recognition via sparse coding: thegood, the bad and the ugly", IET Biometrics, pp. 1–14, 2013.