Three-dimensional Object Recognition Based Intelligence System for Identification

Chiranji Lal Chowdhary
School of Information Technology and Engineering, VIT University
Vellore (TN) - India
c.l.chowdhary@gmail.com

Kanguana Muatjitjeja
Polytechnic of Namibia
Namibia
kmuatjitjeja2010@gmail.com

Dharm Singh Jat
Polytechnic of Namibia
Namibia
dsingh@polytechnic.edu.na

***Abstract*—If we compare the object recognition abilities of human and computer-based system, it is much complex task for a machine. Human brain can recognize an object quickly but for a computer system accuracy depends on the level of algorithms, software and tools used for recognition. Image processing, pattern recognition and compute vision are being challenging but becomes a crucial component for developing such a computer system in the modern digital world. Last four decades many researchers offered many algorithms to recognize objects from an image. If knowledge is acquired, how objects are appearing and an image to be identified from those objects. The intelligent system must be able to recognize different objects from the scenes or images. In this paper, we present a three-dimensional object recognition system based intelligence system to identify individuals.**

Keywords—surveillance camera; scale invariant feature transform; pre-processing; three dimensional object recognition; histogram of oriented gradients.

I. INTRODUCTION

When there is to recognize a face image from dataset by using any image processing technique, it is required to extract different common features from the images and later go for classification. For face recognition, selecting feature is more easily compared with an image of different objects or an image of a single object or a video sequence. Information retrieval from the images is being main task by object recognition techniques. The different objects are a part of information. It is easy for a human to recognize different objects in images, even in conditions when pose of are at some arbitrary angles, of image of any size or image may be at any scale i.e. translated or at different rotation. Human can recognize a partially obstructed object. Researchers do good work for computer vision and image processing systems but still remain a challenging task. For three-dimensional object recognition, pose of the object is also of interest for recognizing and determining a photograph or range scan. Appearance-based method and geometric-based method are two approaches in three dimensional object recognition areas.

Object recognition may also be known as pattern recognition too. One pattern is an arrangement of descriptors. Descriptors may have more forms, but they are primarily vectors and strings. If there is availability of different poses of given three dimensional object than by object recognition observers can recognize depending on varying viewing conditions. The idea of a finite set of classes is quite limiting. Corresponds with industries the use of object recognition is very much of application specific.

The computer object recognition techniques lack some abilities which are simple for humans. Techniques thus far only classify objects based on their shape, color, texture, etc. These are only representative of the light reflected by an object. Humans classify objects many ways, including an object's function. We classify a ring of rocks with a fire inside as a fire pit. We classify a board as a joist once it is installed as support for the floor. We classify our computer as a paperweight once it is more than five years old.

This paper is organized as follows: Section II covers literature survey. Section III introduced different input data formats in recognition task. Section IV provides the different detectors and descriptors used for proposed intelligence systems. Proposed System uses of three dimensional object recognition techniques for intelligent system are given in Section V. At last, Section IV is with conclusion.

II. LITERATUR SURVEY

Earlier three-dimensional object recognition was in light of feature-based geometry-based object recognition which is currently concentrated on appearance-based systems [19]. These appearance-based strategies are having propelled feature descriptors with new pattern recognition calculations. The Eigen object techniques are productive in numerical reckoning with more exactness. The eigenvectors are figured from a situated of vectors where one object image is being a raster scanned vector of grey-scale picture element values. The each of the eigenvector to catch certain variance among all the vectors, and a little arrangement of eigenvectors catches all the appearance variation of object images in the training set. By considering a text image spoke to as a vector of grey-sale pixel values; its character is dictated by discovering the nearest neighbour of this vector in the wake of being anticipated onto a subspace spanned by an arrangement of eigenvectors. In this way, every object image can be spoken to by a direct mix of Eigen object with least blunder, and this straight blend constitutes a smaller re-introduction. The Eigen object methodology has been received in recognition bland objects crosswise over distinctive viewpoints and modelling illumination variation. Neural networks, independent component analysis, nearest neighbour, Fisher linear discriminate, sport vector machines, principal component analysis, sparse network winnows, point clods and boosting

978-1-4799-7707-9/15/$31.00 ©2015 IEEE

algorithms help for recognition of a three-dimensional object by use of two-dimensional images. The appearance-based systems give great results for object recognition under viewpoint and illumination changes contrast with occlusion.

Recognition of object and face is constantly an explored point. Great segmentation is the best approach for best object recognition. In the event that one object is introduced in the image than a large portion of the genuine range images contain mistaken regions resampling shadows and self-occlusion. Thus, present object recognition frameworks for range images must be robust to occlusions. The primary exemplary object recognition methodologies are utilizing Eigen images or Eigen spheres as worldwide systems or they are relying upon their beginning segmentation. Such systems may discover great results on clean images; however they may be vulnerable to occlusions because of reliance on worldwide properties. Some local feature methodologies like shading histograms, feature vectors, gradient histogram, surface shape histogram, and curve segments was proposed and utilized by a few analysts as a part of last a few years. [2, 8]

Target recognition and identification need three-dimensional object retrieval [3]. Point clouds are a standout amongst the most primitive and a major representation of three-dimensional objects, and working straightforwardly with such representation is discriminating and testing in the meantime. The shapes can be exceptionally recognized with a high probability by their dissemination of Euclidean distances between sets of focus [3]. The dissemination separation is equivalent to the Euclidean distance in an embedding space, which makes this discord about dispersals correlated to dispersion removes in the implanting space as well [3]. Graph-based object recognition strategy was useful for matching based local shape patches. The graph-based matching strategies [3, 23] with the inborn distance distributions utilized the presented structure begins fabricating in this bearing of fractional matching. By such prior speculative and computational results, histogram-based signatures for three-dimensional shape recognition frameworks was made for overall and adjacent connection between shapes addresses by point clouds [2, 3, 5].

Object detection [20-23] is a dynamic examination subject in computer vision and there are a few methods created to achieve the object detection assignment. The object detection system, by and large, can be ordered as either worldwide or local approach. Worldwide approaches were cantered around the detection of a full object in which the item is regularly encoded by a feature vector. The appearance of objects was talked about with utilization of some picture peculiarities, for example, histogram of oriented gradients and features. Contrasted and worldwide discovery, nearby location has playing point of having the capacity to detect objects with high explanation and adapt to the issue of occlusion [4]. Object recognition and classification [20-24] have been one of the objectives of computer vision. The identification of object in different poses is in novel review conditions were made a huge advance as of late. With special cases, a little advance has been made for the general assignments of multi-class object classification. Numerous recognition routines have not been tried on multi-class categorization, so little is thought about their separate abilities to sum up past known and seen objects [5].

The idea of programmed face recognition is more invaluable over other biometric innovations. This is because of the characteristic and high throughput properties in face data acquisition. From numerous years, numerous scientists were working more than two dimensional face / object recognition utilizing intensity images as input data. Face recognition is still an awesome test because of varieties in pose, illumination, and expressions. Three-dimensional face/ object data can be caught all the more rapidly and precisely on the grounds that it speaks to faces as three dimensional point sets or range data. This was so as a result of the quick improvement and dropping expense of three-dimensional digital acquisition devices. As Three-dimensional face/ object data was containing unequivocal Three-dimensional geometry, more pieces of information can be utilized to handle the varieties of face pose, expression and illumination. Accordingly, the difficulties confronted by two dimensional objects / face recognition can be determined by three-dimensional face/ object recognition approaches. So centre should be given on three-dimensional face/ object recognition shapes for development of framework execution [1, 7, 9].

Some authors suggested invariants as a local object description. Building up an institutionalized representation of local surface appearance and that can be used to pick ensuring matches in three-dimensional object modelling and recognition errands by joining the possibility of invariant neighbourhood shape decryption with the relative regions [10, 27]. Multi-view geometric constraints were utilized to speak to the bigger three-dimensional surface structure, hold gatherings of steady matches, and reject incorrect ones. A few examinations were demonstrating the guarantee of a way to three-dimensional object recognition [11, 27] limitation of this methodology is its dependence on composition: Some objects are basically texture less, yet effortlessly recognizable. On the other hand, numerous objects are vigorously textured, yet the relating patterns may be more distracting than trademark. Taking care of such objects will oblige new image descriptors that better pass on shape data, yet catch a suitable level of viewpoint invariance [10, 11].

Testing issues object recognition assorted views of an object, different light conditions, surface reflections, and noise achieved by picture sensors, are overseen in optical object recognition. Calculations of scale- invariant feature transform or speeded-up robust features are invariant towards scaling and rotation [12-15, 20-24]. These calculations were computationally mind boggling and require intense equipment to work progressively [16].

For arbitrary poses of distinctive object of character texts, with a few cases of cartoon objects, are rested for recognition with three dimensional angles [18]. Data registration, object recognition and object tracking got to be fascinating by utilizing three dimensional point clouds. To attain to the previously stated undertakings, tackle the correspondence issue by matching features that two contrasting images have in like manner. This can be separated in three stages to the feature matching methodology: detection, description and

matching. Each of the three stages is discriminating. The descriptions of features must be sufficiently different to channel out wrong matches while staying strong to changes that happen between images, including rotation, translation, and noise. A mixed bag of synthetic and real data was demonstrated by created descriptor performs with respect to a generally utilized descriptor, spin images. Descriptor is more different than spin images while remaining turn and interpretation invariant. The change in execution is most clear when an object has attributes that are mirror pictures of each other in view of symmetry [17].

III. Different Input Data Formats in Recognition Task

A. Eigenvector:

Eigenvector is important in face and object recognition algorithms. The eigenvectors are formed by multiplication of eigenvalues and this is just linear transformation. The eigenvector doesn't change or reverse the direction.

B. Euclidean distance:

Euclidean distance is measured by finding distance between two pixel points. If P and Q are two pixels points as $P = (p_1, p_2, p_3, \dots p_n)$ and $Q = (q_1, q_2, q_3, \dots, q_n)$, than the Euclidean distance will be defined as:

$$\sqrt{(p_1-q_1)^2+(p_2-q_2)^2+\dots+(p_n-q_n)^2}=\sqrt{\sum_{i=1}^{n}(p_i-q_i)^2}$$

C. Interest point detection:

By concepts of computer vision there is an important terminology found as the interest point detection for subsequent processing. An interest point in the image can be characterized where it has a mathematical, clear and well-found definition which is a well-defined position in image space.

D. Match point detection:

Match point detection is refers to the detection of the matching of the interest points between the given two images.

E. Haar features:

Haar features are most intuitive as Haar wavelets. Haar features are used in object recognition of computer vision as digital image features. Haar features were mainly started as real-time face detector.

F. The Hessian Detector:

The Hessian detector scans for object areas that display solid subsidiaries in two orthogonal headings. It is in view of the network of second derivatives, the alleged Hessian.

G. The Harris Detector:

The famous Harris detector was expressly intended for geometric soundness. It characterizes key focuses to be "focuses that have provincially maximal organizing toward oneself exactness under translational slightest squares format coordinating". By and by, these key guides frequently relate toward corner-like structures. The Harris identifier returns via scanning for focuses x where the second-minute grid C around x has two huge eigenvalues.

IV. Different Detectors and Descriptors Used for Intelligence Sytsems

Descriptors can be used for a variety of tasks, from registration to comparison to analysis and retrieval. There have been a variety of detectors and descriptors developed for these tasks, which have been evaluated in isolation or in pairs, but not against each other. This paper proposed a novel intelligence system using object recognition with different detectors and descriptors.

Scale-invariant feature extraction strategy proposed by Lowe [15] is being a direction change in object recognition (Fig. 1). In such descriptor-based object recognition system it is obliged to assess and enhance execution. They assessed the repeatability rate of the detectors for sets of pictures, i.e., the rate of recognized investment districts which show "sufficient" spatial cover between the two images of an image pair. The repeatability rate is resolved for various types of picture changes, e.g., JPEG compression artefacts, viewpoint or scale changes, and so forth and diverse scene sorts.

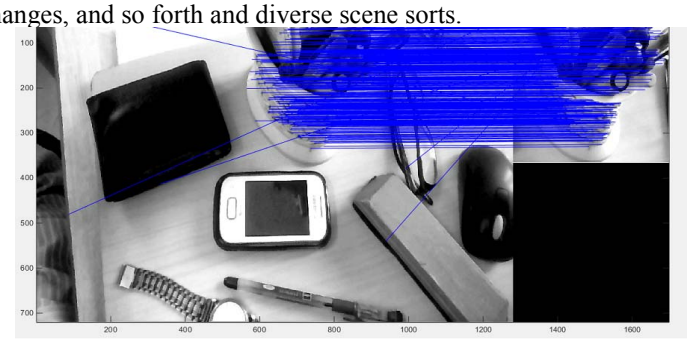

Fig. 1: Object Identification [15]

Therefore, there was no detector that obviously boated all others for all varieties or scene sorts. Much of the time, however by a wide margin not all, the maximally stable extremal regions detector attained to best results, took after by the Hessian-affine detector. There were significant contrasts between diverse detectors the extent that the quantity of detected regions and also their detected size is concerned. Besides, distinctive detectors react to diverse district sorts.

This offers confirmation to the case that distinctive detectors ought to be utilized as a part of parallel to accomplish best execution of the general object recognition plan: reciprocal properties of diverse detectors build the suitability for distinctive object sorts. An alternate viewpoint is invariance: a portion of the detectors are invariant to a larger number of sorts of changes than others. Case in point, the maximally stable extremal regions detector is invariant to

affine transformations. Contrasted with that, the fast detector is just rotation invariant.

While the enhanced invariance of maximally stable extremal regions offers focal points in circumstances where the objects to be detected really have been experienced relative projection, it is frequently not prudent to utilize detectors featuring more invariance than really required. A straightforward methodology for describing a region is to describe it by its crude intensity values. Coordinating adds up to the figuring of the cross-correlation between two descriptors. Nonetheless, this undertaking experiences its computational intricacy and also the way that it doesn't give much invariance descriptor outline goes for discovering a harmony between dimensionality reductions and keeping up discriminative force. Moreover, it ought to concentrate on changing over the data of the image region such that it gets to be invariant or if nothing else robust to common varieties, e.g., non-linear illumination changes or affine transformations because of viewpoint change.

Fundamentally, large portions of the descriptors found in writing have a place with one of the accompanying two classes: Distribution-based descriptors got from the conveyance or something to that effect of data accessible in the area, e.g., inclination introduction in scale invariant feature transform descriptors. Generally, the appropriation is depicted by a histogram or something to that effect of "normal" data. Channel based descriptors computed with the assistance or something to that effect of filtering. All the more definitely, a bank of filters is connected to the region content. The descriptor comprises of the reactions of all channels. Every filter is intended to be delicate to a particular sort of data. Regularly utilized filter sorts separate properties as a part of the frequency domain or are taking into account subordinates.

V. Conceptually Proposed System with Uses of 3D Object Recognition Techniques for Intelligent System

As we entering some public places or even in hotels, we may find a warning “You are under surveillance, CCTV cameras are ON”. These all are common now days and required also for security and safety reasons. Now surveillance cameras are also having object recognition techniques. Government agencies are fixing video surveillance cameras at places like main and crowded markets, educational institutions -schools / colleges, malls, railway stations, in parking lots, ATMs, banks, in parks and stadiums, in malls, bogies of trains, and airports. For these cameras many experts are needed to observe but it will more challenging to find the suspicious cases. These cameras will acquire a lot of information (surveillance data) and if we can develop an automatic system to recognise the cases of importance than intelligence agencies can save time and money to catch such culprits or things. In present days there are many cases of militant attacks at different places. Militant groups send their agents on such places for reconnaissance mission to discover the status. Object recognition can help us to improve security surveillance systems. Improved security surveillance systems may help us to find out such militant mission, if that can recognize the strange activities by such people at particular places. Means system should be able to find known militants, strange activities like mixing unwanted things, strange behavior, taking unexpected photos of the places which are important for security and strategic point of view and security agencies continuously receive such information with alert messages.

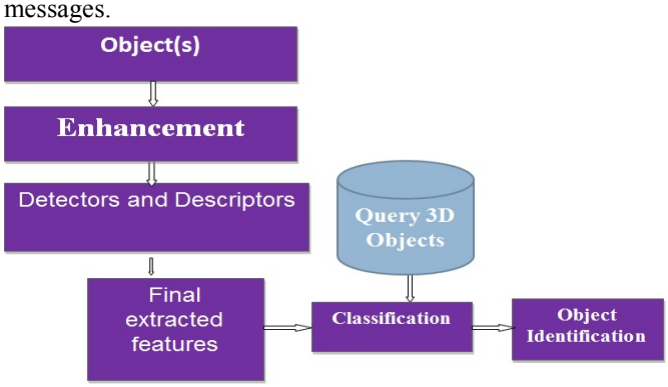

Fig 2. Conceptually three-dimensional object recognition based Intelligence System

Conceptually proposed three-dimensional Object Recognition based Intelligence System is presented in Fig. 2. Initially an object is acquired from the data set or it may be taken a camera captured image. If this image is blurred or noise affected that it will be enhanced. Next step is about proper used of descriptors and detector according to type of images in the dataset. Last step will help us to extract features and our last step will be classification stage/ Three Dimensional object recognition based wireless surveillance systems need to provide information with alerts that unwanted and suspicious person or vehicle is entering in a region with sufficient concerns that what person or vehicle is having with him. In commercial regions and for industrial inspection applications such systems are required the safety of their employees and buildings to make sure their security and safety. Medical research in breast cancer and other area may be improved by using such type of 3D object recognition based wireless systems.

VI. Discussion and Conclusion

This paper is survey of different algorithms of descriptors and detectors for 3D object recognition techniques and connecting such algorithm with 3D object recognition system based intelligence system to identify individuals. Wireless intelligent systems are very important for safety and security issues in present scenarios. Many surveillance systems are available around us but those can be improved by connecting such systems with pattern recognition techniques so that valuable information can be extracted from a bulk of surveillance data.

References

[1] Wang, Y., Liu, J., Tang, X., “Robust 3D Face Recognition by Local Shape Difference Boosting”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, pp 1858 - 1870 .

[2] Hetzel, G., Leibe, B., Levi, P., Schiele, B., “3D object recognition from range images using local feature histograms”, IEEE Computer Society

Conference on Computer Vision and Pattern Recognition, 2001, pp 394 -399.

[3] Mahmoudi, M., Sapiro, G., "Three-dimensional point cloud recognition via distributions of geometric distances", IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008, 1 – 8.

[4] Nguyen, D.T., Zong, Z, Ogunbona, P., Li, W.; "Object detection using Non-Redundant Local Binary Patterns", 17th IEEE International Conference on Image Processing (ICIP), 2010, pp. 4609 – 461.

[5] Leibe, B., Schiele, B., "Analyzing appearance and contour based methods for object categorization", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003, pp 409-415.

[6] Belongie, S., Malik, J., Puzicha, J., "Shape Matching and Object Recognition using Shape Contexts", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 509-522.

[7] Huber, D., Kapuria, A., Donamukkala, R., Hebert, M., "Parts-based 3D object classification", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004.

[8] Wu, Z., Wang, Y., Pan, G., "3D Face Recognition using local shape map", IEEE Proceedings, 2004, pp 2003-2006.

[9] Li, X., Godil, A., Wangan, A., "3D Part Identification Based on Local Shape Descriptors", 2008.

[10] Mikolajczyk, K., Schmid C., "An affine invariant interest point detector". In: European Conference on Computer Vision, 2002, pp. 128–142.

[11] Rothganger F., Lazebnik S., Schmid C., Ponce J., "3D Object Modeling and Recognition Using Local Affine-Invariant Image Descriptors and Multi-View Spatial Constraints", March 2006, pp. 231 – 259.

[12] Amato, G., Falchi F., Bolettieri P., "Recognizing landmarks using automated classification techniques: an evaluation of various visual features", Proceedings of the 2nd International Conferences on Advances in Multimedia, 2012, pp. 78–83.

[13] Bay H., Ess A., Tuytelaars T., Van Gool L., "Speeded-up robust features (SURF), Computer Vision and Image Understanding", 2008, pp. 346–359.

[14] Harris C., Stevens M., "A Combined corner and edge detector", Proceedings of the 4th Alvey Vision Conference, 1988, pp. 147–151.

[15] Lowe D., "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, 2004, pp. 91–110.

[16] Schaeferling M., Hornung U., Kiefer G., "Object Recognition and Pose Estimation on Embedded Hardware: SURF-Based System Designs Accelerated by FPGA Logic", International Journal of Reconfigurable Computing, 2012, Vol. 2012, pp. 1-16.

[17] Skelly L. J., Sclaroff S., "Improved feature descriptors for 3-D surface matching", 2007.

[18] Rivers A., Igarashi T., Durand F., "2.5D Cartoon Models", 2010.

[19] C.L.Chowdhary, "Linear feature extraction techniques for object recognition: study of PCA and ICA", Journal of the Serbian Society for Computational Mechanics, 2011, pp 19-26 .

[20] M. Brown and D. G. Lowe, "Unsupervised 3D Object Recognition and Reconstruction in Unordered Datasets", 2005, pp 56 – 63.

[21] SK Zhou, R Chellappa, B Moghaddam, "Visual tracking and recognition using appearance-adaptive models in particle filters"., IEEE Transactions on Image Processing, 2004, pp 1491-1506.

[22] "Interest Point Detection and Region Descriptors", An Introduction to Object Recognition, Springer London, 2010, pp 145-182.

[23] John S. Zelek. "Region detection and description for Object Category Recognition", Fourth Canadian Conference on Computer and Robot Vision , 2007, pp 321-328.

[24] Dominik Maximilián Ramík, Christophe Sabourin, Ramon Moreno, and Kurosh Madani, "A machine learning based intelligent vision system for autonomous object detection and recognition", Journal of Applied Intelligence, 2014, pp 358-375.