ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 4, April 2013

Techniques for Object Recognition in Images and Multi-Object Detection

Khushboo Khurana, Reetu Awasthi

***Abstract*— The modern world is enclosed with gigantic masses of digital visual information. Increase in the images has urged for the development of robust and efficient object recognition techniques. Most work reported in the literature focuses on competent techniques for object recognition and its applications. A single object can be easily detected in an image. Multiple objects in an image can be detected by using different object detectors simultaneously. The paper discusses various techniques for object recognition and a method for multiple object detection in an image.**

***Index Terms*— Multi-object detection, Object recognition, Object recognition applications.**

I. INTRODUCTION

The modern world is enclosed with gigantic masses of digital visual information. To analyze and organize these devastating ocean of visual information image analysis techniques are major requisite. In particular useful would be methods that could automatically analyze the semantic contents of images or videos. The content of the image determines the significance in most of the potential uses. One important aspect of image content is the objects in the image. So there is a need for object recognition techniques.

Object recognition is an important task in image processing and computer vision. It is concerned with determining the identity of an object being observed in an image from a set of known tags. Humans can recognize any object in the real world easily without any efforts; on contrary machines by itself cannot recognize objects. Algorithmic descriptions of recognition task are implemented on machines; which is an intricate task. Thus object recognition techniques need to be developed which are less complex and efficient.

Many successful approaches that address the problem of general object detection use a representation of the image objects by a collection of local descriptors of the image content. Global features provide better recognition. Color and shape features can also be used. Various object recognition techniques are presented in this paper. Difficulties may arise during the process of object recognition; a range of such difficulties are discussed in this paper. The robust and efficient object recognition technique can be developed by taking into account these difficulties and overcoming them.

Rest of this paper is organized as follows. Section II elucidates various difficulties in object recognition under varied circumstances. Section III presents various object recognition techniques. In Section IV applications for object recognition are discussed. In section V we have proposed a method for multi-object detection in an image and finally, we conclude in Section VI.

II. Difficulties In Object Recognition Under Varied Circumstances

1. *Lightning:* The lightning conditions may differ during the course of the day. Also the weather conditions may affect the lighting in an image. In-door and outdoor images for same object can have varying lightning condition. Shadows in the image can affect the image light. Whatever the lightning may be the system must be able to recognize the object in any of the image. Fig.1 shows same object with varying lightning.

Fig.1 Objects with different lightning.

2. *Positioning:* Position in the image of the object can be changed. If template matching is used, the system must handle such images uniformly.

3. *Rotation:* The image can be in rotated form. The system must be capable to handle such difficulty. As shown in fig.2, the character ‘A’ can appear in any of the form. But the orientation of the letter or image must not affect the recognition of character ‘A’ or any image of object.

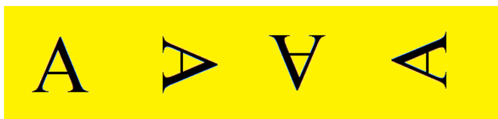

Fig.2 Different orientation of character ‘A’

4. *Mirroring:* The mirrored image of any object must be recognized by the object recognition system.

Manuscript received Feb, 2013.
***Ms.Khushboo Khurana**, Computer Science & Engg. department, S.R.O.C.E.M.., Nagpur, India.*
***Ms. Reetu Awasthi**, Computer Science department, SFS College, Nagpur, India.*

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 4, April 2013

5. *Occlusion:* The condition when object in an image is not completely visible is referred as occlusion. The image of car shown in a box in fig.3 is not completely visible. The system of object recognition must handle such type of condition and in the output result it must be recognized as a car.

   In [1], a segmentation aware object detection model is presented with occlusion handling.

Fig.3 Occluded car

6. Scale: Change in the size of the object must not affect the correctness of the object recognition system.

Above stated are some of the difficulties that may arise during object recognition. An efficient and robust object detection system can be developed by conquering the above stated difficulties.

III. Object Recognition Techniques

A. *Template matching*

Template matching is a technique for finding small parts of an image which match a template image. It is a straightforward process. In this technique template images for different objects are stored. When an image is given as input to the system, it is matched with the stored template images to determine the object in the input image.

Templates are frequently used for recognition of characters, numbers, objects, etc. It can be performed on either color or gray level images.

Template matching can either be pixel to pixel matching or feature based. In feature based the features of template image is compared to features of sub-images of the given input image; to determine if the template object is present in the input image.

In [2], authors have proposed a mathematical morphological template matching approach for object detection in inertial navigation systems (INS). The major focus of the paper is to detect and track the ground objects. The flying systems equipped with camera were used to capture the photos of ground; to identify the objects. Their method is independent of the altitude and orientation of the object.

In [3], an approach for measuring similarity between visual images based on matching internal self-similarities. A template image is to be compared to another image. Measuring similarity across images can be complex, the similarity within each image can be easily revealed with simple similarity measure, such as SSD (Sum of Square Differences), resulting in local self-similarity descriptors which can be matched across images. As shown in fig.4, the template image of the flower is compared with all the corresponding descriptors.

(a) (b) (c)

Fig.4. (a) Image template (b) Image against which it is compared (c) Detected object superimposed on gray-scale image (from[2])

B. *Color based*

Color provides potent information for object recognition. A simple and efficient object detection scheme is to represent and match images on the basis of color histograms.

Fahad Khan, et.al. [4] proposed the use of color attributes as an explicit color representation for object detection. The color information is extended in two existing methods for object detection, the part-based detection framework and the Efficient Subwindow Search approach . The three main criteria which should be taken into account when choosing an approach to integrating color into object detection are feature Combination, photometric invariance and compactness. The paper investigates the incorporation of color for object detection based on the above mentioned criteria and demonstrate the advantages of combining color with shape on the two most popularly used detection frameworks,namely part-based detection with deformable part models and Efficient Subwindow Search (ESS) for object localization. The resulting image representations are compact and computationally efficient and provide excellent detection performance on challenging datasets. Fig.5 provide how the extension correctly detects all Simpsons in the image; Simpsons is an American animated sitcom. The technique correctly detects challenging object classes where state-of-the-art techniques using shape information alone fail.

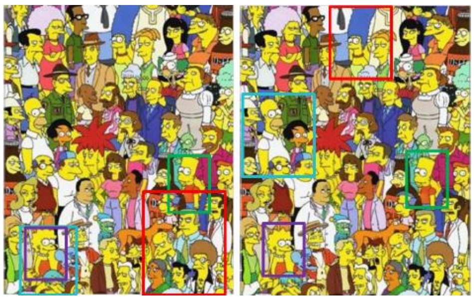

Fig.5. Find the Simpsons. On the left, the conventional part based approach fails to detect all four members of Simpsons. On the right, our extension of the part-based detection framework with color attributes can correctly classify all four Simpsons.

All Rights Reserved © 2013 IJARCET

In [5] the aim is to arrive at recognition of multicolored objects invariant to a substantial change in viewpoint, object geometry and illumination. Assuming dichromatic reßectance and white illumination, it is shown that normalized color *rgb*, saturation *S* and hue *H*, and the newly proposed color models $c1c2c3$ and $l1l2l3$ are all invariant to a change in viewing direction, object geometry and illumination. Further, it is shown that hue *H* and $l1l2l3$ are also invariant to highlights. Finally, a change in spectral power distribution of the illumination is considered to propose a new color constant color model $m1m2m3$. To evaluate the recognition accuracy di¤erentiated for the various color models, experiments have been carried out on a database consisting of 500 images taken from 3-D multicolored man-made objects. The experimental results show that highest object recognition accuracy is achieved by $l1l2l3$ and hue *H* followed by $c1c2c3$, normalized color *rgb* and $m1m2m3$ under the constraint of white illumination. Also, it is demonstrated that recognition accuracy degrades substantially for all color features other than $m1m2m3$ with a change in illumination color. In this paper the authors aim to examine and evaluate a variety of color models used for recognition of multicolored objects according to the following criteria:

1. Robustness to a change in viewing direction
2. Robustness to a change in object geometry
3. Robustness to a change in the direction of the illumination
4. Robustness to a change in the intensity of the illumination
5. Robustness to a change in the spectral power distribution (SPD) of the illumination.

The color models have High discriminative power; robustness to object occlusion and cluttering; robustness to noise in the images.

C. Active and Passive

Object detection in passive manner does not involve local image samples extracted during scanning. Two main object-detection approaches that employ passive scanning:

1. The *window-sliding* approach: It uses passive scanning to check if the object is present or not at all locations of an evenly spaced grid. This approach extracts a local sample at each grid point and classifies it either as an object or as a part of the background [6].
2. The *part-based* approach: It uses passive scanning to determine interest points in an image. This approach calculates an interest value for local samples at all points of an evenly spaced grid. At the interest points, the approach extracts new local samples that are evaluated as belonging to the object or the background [7].

Some methods try to bound the region of the image in which passive scanning is applied. It is a computationally expensive and inefficient scanning method. In this method at each sampling point costly feature extraction is performed, while the probability of detecting an object or suitable interest point can be squat.

In active scanning local samples are used to guide the scanning process. At the current scanning position a local image sample is extracted and mapped to a shifting vector indicating the next scanning position. The method takes successive samples towards the expected object location, while skipping regions unlikely to contain the object. The goal of active scanning is to save computational effort, while retaining a good detection performance [8].

The active object-detection method (AOD-method) scans the image for multiple discrete time steps in order to find an object. In the AOD-method this process consists of three phases:

(i) Scanning for likely object locations on a coarse scale
(ii) Refining the scanning position on a fine scale
(iii)Verifying object presence at the last scanning position with a standard object detector.

D. Shape based

Recently, shape features have been extensively explored to detect objects in real-world images. The shape features are more striking as compared to local features like SIFT because most object categories are better described by their shape then texture, such as cows, horses and cups and also for wiry objects like bikes, chair or ladders, local features unavoidably contain large amount of background mess. Thus shape features are often used as a replacement or complement to local features.

A. Berg, et.al. [9], have proposed a new algorithm to find correspondences between feature points for object recognition in the framework of deformable shape matching. The basic subroutine in deformable shape matching takes as input an image with an unknown object (shape) and compares it to a model by solving the correspondence problem between the model and the object. Then it performs aligning transformation and computes a similarity based on both the aligning transform and the residual after applying the aligning transformation. The Authors have considered various reasons like Intra-category variation, Occlusion and clutter, 3D pose changes that makes correspondence problems more difficult. Three kinds of constraints to solve the correspondence problem between shapes are Corresponding points on the two shapes should have similar local descriptors, Minimizing geometric distortion, Smoothness of the transformation from one shape to the other.

In [10], a new shape-based object detection scheme of extraction and clustering of edges in images using Gradient vector Girding (GVG) method is proposed that results a directed graph of detected edges. The algorithm used contains a sequential pixel-level scan, and a much smaller second and third pass on the results to determine the connectiveness. The graph is built on cell basis and the image is overlaid with a grid formed of equal sized cells. Multiple graph nodes are computed for individual cells and then connected corresponding to the connectivity in the 8-neighbourhood of each cell. Finally, the maximum curvature of the result paths is adjusted. The Authors have also proposed several techniques to increase the performance of the method and

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 4, April 2013

improve the quality of the result. The method is particularly proposed for a RoboCup scenario but can also be applied to any other images if the prerequisites are met.

In [11] K.Schindler, D. Suter presented a method for object class detection based on global shape based on elastic matching of contours. An edge map with closed edge chains is produced by segmentation into super-pixels. A probabilistic measure for the similarity between two contours is derived and then combined with an optimisation scheme to find closed contours in the image, which have high similarity with a template image. The method only requires a single object template. The basic idea used is to define a distance measures between shapes and then try to find minimum distance. The Authors discussed different deformable template matching techniques listed chamfer matching and spline-based shape matching and compares to elastic matching method and finally proposed a search strategy which relieves the problem of false local minima by re-evaluating the shape distance at each search step.

In [12], boundary structure segmentation (BoSS) is presented. A ground segmentation method for extractions of image regions that resembles the global properties of a model boundary structure and are perceptually prominent. The global boundary-based shape representation is called Chordiogram, which is defined as the geometric relationships between pairs of boundary edges. The perceptual prominence cue used favour coherent regions different from the background. The method relates the object detection based on novel holistic shape descriptor to ground segmentation and perform them simultaneously. Similarity in shape and perceptual prominence are the two major properties of foreground regions selected by BoSS while matching an image to an object model. The results provide a unified framework that integrates both object segmentation and object detection.

E. *Local and global features*

The most common approach to generic object detection is to slide a window across the image and to classify each such local window as containing the target or background. This approach has been successfully used to detect rigid objects such as faces and cars in [13] and [6].

In [14], a method of object recognition and segmentation using Scale-Invariant Feature Transform (SIFT) and Graph Cuts is presented. SIFT feature is invariant for rotations, scale changes, and illumination changes. By combing SIFT and Graph Cuts, the existence of objects is recognized first by vote processing of SIFT keypoints. Then the object region is cut out by Graph Cuts using SIFT keypoints as seeds. Both recognition and segmentation are performed automatically under cluttered backgrounds including occlusion.

Authors in [15], present a method for object recognition with full boundary detection by combining affine scale invariant feature transform (ASIFT) and a region merging algorithm. The algorithm is invariant to six affine parameters namely translation (2 parameters), zoom, rotation and two camera axis orientations. The features give strong keypoints that can be used for matching between different images of an object. They trained an object in several images with different aspects for finding best keypoints of it. Then, a robust region merging algorithm is used to recognize and detect the object with full boundary in the other images based on ASIFT keypoints and a similarity measure for merging regions in the image. Fig. 6 shows the trained image for an object (left) and detected image of the object in a image (right).

Fig.6. Training image and detected object in the image. (from [15])

In [16], Histogram of Gradients (HOG) based multistage approach for object detection and object pose recognition for service robots is used. It makes use of the merits of both multi-class and bi-class HOG-based detectors to form a three-stage algorithm at low computing cost. In the first stage, the multi-class classifier with coarse features is employed to estimate the orientation of a potential target object in the image; in the second stage, a bi-class detector corresponding to the detected orientation with intermediate level features is used to filter out most of false positives; and in the third stage, a bi-class detector corresponding to the detected orientation using fine features is used to achieve accurate detection with low rate of false positives. The training of multi-class and bi-class support vector machine (SVM) with their respective features in different levels is described.

Antonio Monroy, Angela Eigenstetter and Bjorn Ommer [17], have presented an approach that directly uses curvature cues in a discriminative way to perform object recognition. Integrating curvature information substantially improves detection results over descriptors that solely rely upon histograms of orientated gradients (HoG). The joint descriptor is refered as HoGC. Because of the histogram-nature of the feature vectors, SVM with histogram intersection kernel is used as a classifier.

An innate extension of these local approaches is to use sliding window to detect object parts, and then assemble the parts into a whole object. Problem with local features is that recognition may fail because of insufficient local information. This can be solved by using the context of the image as a whole i.e., global features.

In [18], the gist of an image is computed. First a steerable pyramid transformation is applied, using 4 orientations and 2 scales; then the image is divided into a 4x4 grid. Object presence detection determines if one or more instances of an object class are present. They have combined local features and global features- GIST for object recognition.

All Rights Reserved © 2013 IJARCET

IV. Application Of Object Recognition

1. *Biometric recognition*: Biometric technology uses human physical or behavioral traits to recognize any individual for security and authentication [19]. Biometrics is the identification of an individual based on distinguished biological features such as finger prints, hand geometry, retina and iris patterns, DNA, etc. For biometric analysis, object recognition techniques such as template matching can be used.

2. *Surveillance*: Objects can be recognized and tracked for various video surveillance systems. Object recognition is required so that the suspected person or vehicle for example be tracked.

3. *Industrial inspection*: Parts of machinery can be recognized using object recognition and can be monitored for malfunctioning or damage.

4. *Content-based image retrieval (CBIR)*: When the retrieval is based on the image content it is referred as CBIR. A supervised learning system, called OntoPic, which provides an automated keyword annotation for images and content–based image retrieval is presented in [20].

5. *Robotic:* The research of autonomous robots is one of the most important issues in recent years. The humanoid robot soccer competition is very popular. The robot soccer players rely on their vision systems very heavily when they are in the unpredictable and dynamic environments. The vision system can help the robot to collect various environment information as the terminal data to finish the functions of robot localization, robot tactic, barrier avoiding, etc. It can decrease the computing efforts, to recognize the critical objects in the contest field by object features which can be obtained easily by object recognition techniques [21].

6. *Medical analysis:* Tumour detection in MRI images, skin cancer detection can be some examples of medical imaging for object recognition.

7. *Optical character/digit/document recognition:* Characters in scanned documents can be recognized by recognition techniques.

8. *Human computer interaction:* Human gestures can be stored in the system, which can be used for recognition in the real-time environment by computer to do interaction with humans. The system can be any application on mobile phone, interactive games, etc.

9. *Intelligent vehicle systems*: Intelligent vehicle systems are needed for traffic sign detection and recognition, especially for vehicle detection and tracking. In [18], such a system is developed. In detection phase, a color-based segmentation method is used to scan the scene in order to quickly establish regions of interest (ROI). Sign candidates within ROIs are detected by a set of Haar wavelet features obtained from AdaBoost training. Then, the Speeded Up Robust Features (SURF) is applied for the sign recognition. SURF finds local invariant features in a candidate sign and matches these features to the features of template images that exist in data set. The recognition is performed by finding out the template image that gives the maximum number of matches.

V. Method For Multi-Object Detection In An Image

A single image may consist of single or multiple objects. If all the objects in an image need to be detected the method shown in fig.7 can be used.

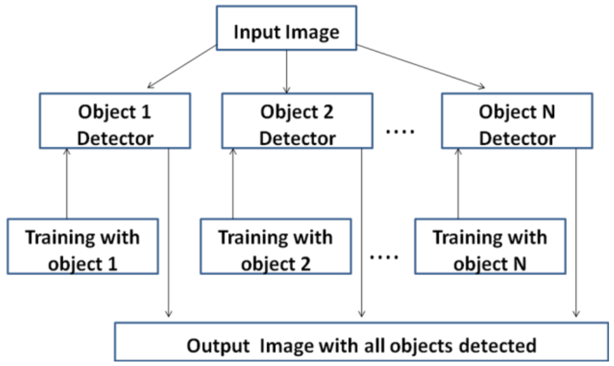

Fig.7. Method for multi-object detection in an image

The method trains different object detectors with individual objects, as shown in fig.7. there are N object detectors which are trained to detect N different objects. Any of the above mentioned object recognition techniques can be used depending upon the application area. An image is provided as input to the system. The same image is given as input to all object detectors. Each detector will determine if the object is present or not. We propose to use object detector along with boundary detector. If the object is present, the detector will find its boundary and tag the object name in the image.
So, after the image has passed via all the detectors all objects will be detected along with object boundary and its tag displayed in the output image.

Also, when the output image is displayed, we can move the cursor over the image. The tag shown for an object inside the complete boundary of the object remains same. Such multi-object detection in the image can greatly improve the performance of the content based image retrieval systems.

The performance can further be improved by letting the object detectors run in parallel.

VI. Conclusion

In this paper, we have discussed various object detection techniques. The template matching technique requires large database of image templates for correct object recognition. Hence it must be used only when limited objects are to be detected. Global features and shape based method can give

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 4, April 2013

better result and are efficient as compared to local features. These techniques help in easy access of the images. They also find their application in fields such as biometric recognition, medical analysis, surveillance, etc. A method for multiple object detection is also presented.

REFERENCES

[1] T. Gao, B. Packer, D. Koller, "A Segmentation-aware Object Detection Model with Occlusion Handling", In IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), June 2011.

[2] W. Hu, A.M.Gharuib, A.Hafez, "Template Match Object Detection for Inertial Navigation Systems," Scientific research (SCIRP), pp.78-83, May 2011.

[3] E.Shectman, M.Irani, "Matching Local Self-Similarities across Images and Videos," In IEEE International Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2007.

[4] F. Khan, R. Muhammad , et.al., "Color Attributes for Object Detection," In IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3306 – 3313, 2012.

[5] T. Gevers, A. Smeulders, "Color-based object recognition," Pattern Recognition, 1999.

[6] P. Viola and M. Jones, "Robust real-time object detection," International. Journal of Computer Vision, 57(2), pp.137–154, 2004.

[7] R. Fergus, P. Perona, A. Zisserman, "Weakly supervised scale-invariant learning of models for visual recognition," International Journal of Computer Vision, 2006.

[8] G. de Croon, "Active Object Detection," In 2nd International conference on computer vision theory and applications (VISAPP 2007), Barcelona, Institute for Systems and Technologies of Information, Control and Communication (INSTICC), pp. 97–103, 2007.

[9] A. Berg, T.Berg , J. Malik, "Shape Matching and Object Recognition using Low Distortion Correspondences," In IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pp. 26 – 33, 2005.

[10] H. Moballegh, N. Schmude, and R. Rojas, "Gradient Vector Griding: An Approach to Shape-based Object Detection in RoboCup Scenarios," from: www.ais.uni-bonn.de/**robocup**.de/papers/RS11_Moballegh.pdf

[11] K.Schindler, D. Suter, "Object Detection by Global Countour Shape," Pattern Recognition, 41(12), pp.3736–3748, 2008.

[12] F. Khan, R. Muhammad , et.al., "Color Attributes for Object Detection," In IEEE International Conference on Computer Vision and Pattern Recognition, pp. 3306 – 3313, 2012.

[13] C. Papageorgiou and T. Poggio, "A trainable system for object detection," International. Journal of Computer Vision, 38(1), pp.15–33, 2000.

[14] A. Suga, K. Fukuda, T. Takiguchi, Y.Ariki, "Object Recognition and Segmentation Using SIFT and Graph Cuts," In 19th International Conference on Pattern Recognition ,pp. 1-4, 2008 .

[15] R. Oji, "An Automatic Algorithm for Object Recognition and Detection Based on ASIFT Keypoints," Signal & Image Processing: An International Journal (SIPIJ) Vol.3, No.5, pp.29-39, October 2012.

[16] L. Dong, X. Yu ,L. Li, J. Kah Eng Hoe, "HOG based multi-stage object detection and pose recognition for service robot," Control Automation Robotics & Vision (ICARCV), 11th International Conference, pp. 2495 – 2500, Dec. 2010.

[17] A. Monroy, A. Eigenstetter, B. Ommer, "Beyond Straight Lines - Object Detection Using Curvature," In 18th IEEE International conference on Image Processing (ICIP), pp. 3561 - 3564 , 2011.

[18] K. Murphy, A. Torralba , D. Eaton and W. Freeman, "Object detection and localization using local and global features," Towards Category-Level Object Recognition, 2005.

[19] V. Bjorn, "One Finger at a Time: Best Practices for Biometric Security," Banking Information Source (Document ID: 1697301411), April, 2009.

[20] J. Schober, T. Hermes, O. Herzog, "Content-based Image Retrieval by Ontology-based Object Recognition," In KI Workshop on Applications of Description Logics , 2004.

[21] W. Chang, C. Hsia. Y. Tai, et.al, "An efficient object recognition system for humanoid robot vision," Pervasive Computing (JCPC), IEEE, December, 2009.

Khushboo Khurana received her B.E. degree in computer science and engineering from RTM Nagpur University, Nagpur, in 2010. She is currently pursuing her M.Tech. (CSE) from Shri Ramdeobaba College of Engineering and Management (Autonomous), Nagpur. Her interests include image and video processing.

Reetu Awasthi received her M.Sc. (Computer Science) degree from RTM Nagpur University, Nagpur, in 2012. She is currently working as a lecturer in SFS College, Nagpur. Her interests include biometric security and image processing.

All Rights Reserved © 2013 IJARCET