International Journal of Engineering & Technology, 7 (2.16) (2018) 130-134

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

Comparative study of Object Recognition Algorithms

R.S.Jaiswal [1]*, M.V.Sarode[2]

[1] Department Of CSE,COET, Akola,MH-444004
[2]Department Of CSE, GCOP, Yavatmal, MH-445001
**Email: rashmiramdhanijun@gmail.com*

Abstract

The world we live in is full of enormous masses of digital visual information. This enormous amount of digital visual information motivates us to develop robust and efficient object recognition technique. Most of the work reported in this paper focuses focus light upon efficient techniques that can be used for recognition of object and its applications. Here in this paper, various techniques for object recognition in an image are discussed.

Keywords: *Comparison of Object Recognition Techniques; Feature Selection ;Feature Extraction; Object Recognition; Object Classification*

1. Introduction

There is a need of efficient Image analysis technique for analyzing and organizing the enormous volume of visual information. The content of the image determines its importance in most of the potential uses. Hence automatic analysis of semantic contents of images or videos is a need of time. One important aspect of image content analysis is objects present in the images. So we are required to develop efficient object recognition techniques[1].

Object recognition involves identification of a specific object in a digital image or video. Here input is image . Output is attributes extracted from image. Object recognition algorithms covers matching, learning. edge finding, gradient calculation, histogram analysis of oriented gradients (HOG), Haar wavelets, and linear binary patterns are used in appearance-based, feature-based approaches . These are pattern recognition algorithms . Deep learning based approaches such as convolutional neural networks are used for object recognition. Stabilization of videos applications, assistance systems for drivers and identification of various diseases uses object recognition techniques.

Section II elucidates object recognition system & various difficulties encountered in object recognition under varied circumstances. Section III presents various object recognition techniques. In Section IV applications of object recognition are discussed. We conclude in Section V.

2. Object Recognition System

In below figure , pre-processing removes the noise. Model learning/estimation performs mapping of features.
Object Recognition involves :
• Identification of separate objects

• Describing geometry and positions of objects in 3D
• Categorization/Classification of objects as being one of a known class

•Scene and context categorization/Understanding – of spatial relationships between objects.

Object recognition systems have the following two processing steps[2] :

1. In this step, the image is scanned for a set of characteristic features of the object. For example, in a car detection system a canonical gray-value template of a wheel might be cross-correlated with the input image to localize the wheels of a car. These local object features are the components of an object, parts, patches or fragments.

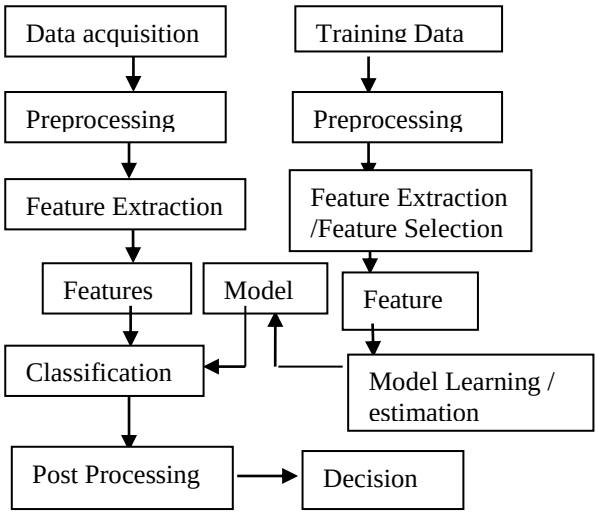

Fig.2.1 Object Recognition System Component

2. In a second step, the results of the previous stage are combined to determine whether the input image contains an object of the given class.

Another approach for object classification is to perform global approach . Global approach involves matching for the object as a

Copyright © 2016 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

whole, by computing the cross-correlation between a template of the object and the input image. Here, a single classifier takes as input a feature vector containing information about the whole object.

Problems Encountered In Object Recognition

In order to develop an efficient and robust object detection system following problems need to be conquered:
1. In-door and outdoor images for same object can have varying lightning condition. Shadows in the image can affect the image light.
2. Object position in the image may be changed. If template matching is used, the system must handle such images uniformly.
3. Image might be in rotated form.
4.The condition when object in an image is not completely visible is referred as occlusion.
5. 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 that need to be overcome by object recognition system.

3. Object Recognition Techniques

Objects Recognition Techniques can be categorized as:

1.Feature Based

2.Appearance based

3.1. Feature Based Technique

A function of measurements of some quantifiable property of an object is nothing but a feature, and is computed such that it quantifies some significant characteristics of the object [3].

Features defines the characteristics of object. One object can not only be described by one attribute but multiple features. Colors, contour lines, geometric forms or edges (gradient of pixel intensities) are features that can be used to describe the object. The feature-based object recognition strategies rely on the following concept:

A search is conducted for finding image with specific type of feature, this feature is then compared to a database containing models of the objects in order to verify if there are recognized objects. Here, comparison of object features and image features is done to find feasible matches.

Feature Classification can be done as follows:

• General features: Color, texture, and shape are the general features and they are application independent features. These features can be further divided into:

3.1.1 Pixel-level features

Pixel-level features depend on color, location of pixels.

3.1.2 Local features

It involves subdivision of the image band and calculation of features over the subdivided image. Just small parts of the image (local feature) are enough to find features and their descriptors. Descriptors of local features are more robust and therefore performance of algorithms with local features is often better than global feature-based methods. The local feature strategy is further described with referencing to two famous algorithms: SIFT(Scale-Invariant Feature Transform) and SURF(Speeded-Up Robust Features).

-**Bag-of-words models** with features such as SURF and MSER

In natural language processing & information retrieval (IR), the **bag-of-words model** is used. It is a vector space model. Here, a text (such as a sentence or a document), ignoring grammar and word order represents bag but keeping multiplicity.

It is applied in the applications where the (frequency of) occurrence of words is used as a feature for training a classifier.

3.1.3 Global features

It involves calculation of features over the entire image or just regular sub-area of an image. Whole image is used to find the features and their descriptors . An histogram of the pixel intensity or color is simple example of global feature. A correct recognition is not possible due to changes in illumination, position (occlusion) or rotation .
A system which uses HOG-based multistage approach for object detection and object pose recognition for service robots is designed[4]. This system used features of multi-class and bi-class HOG-based detectors to form a three-stage algorithm at low computing cost. First stage involves estimation of orientation of a potential target object in the image using multi-class classifier with coarse features; in the second stage, filtering of most of false positives is done using a bi-class detector with intermediate level features; and in the third stage, fine features are used to achieve accurate detection with low rate of false positives. The multi-class and bi-class training SVMs in different levels is demonstrated with their respective features. Real-world environment experimentation shows that as the proposed method uses only multi-class detector, it is much more accurate than the detection method. The proposed method uses a bi-class detector for each possible orientation so it is much more efficient than the detection method. The approach outperforms on the scenarios where the SIFT-based detector may fail. Real-time object detection, localization, and pose recognition on a P4 2.4GHz PC can be achieved using the proposed method.

3.1.4 Domain-specific features

Human faces, their fingerprints, and conceptual features are application dependent features . Application dependent features are nothing but domain-specific features. These are

- Shape based features
- Color based features

Classification of facial expressions involves extraction of facial features based on the RGB information and morphological operations. In order to detect most facial feature points around the eyes, eyebrows and mouth , two eyes inner points was used as the base of face area from the images of FG-net database. They used 90 static images consisting of the neutral and emotional states from the FG-net database in the experiment process. Eight feature points comprising four points related to upper and lower eyelids, two points from eyebrows and two points from lips corners position were detected automatically based on the two inner corners of eyes while two points located in the upper and lower of lips were extracted manually in the proposed facial features extraction. Therefore, 8 facial feature points and for all subjects (18) with all expressions (5 included neutral) 720 feature points were detected for every expression.
The coordinate of feature points which were detected according to the proposed feature points extraction technique, has been compared with the location of points which were marked manually, in order to evaluate the accuracy of detection performance. The automatic detection of facial feature points is successful if the comparison between marking manually and automatic labeling shows less than 10 percent bias , according to the Tie (2011) study[5] .

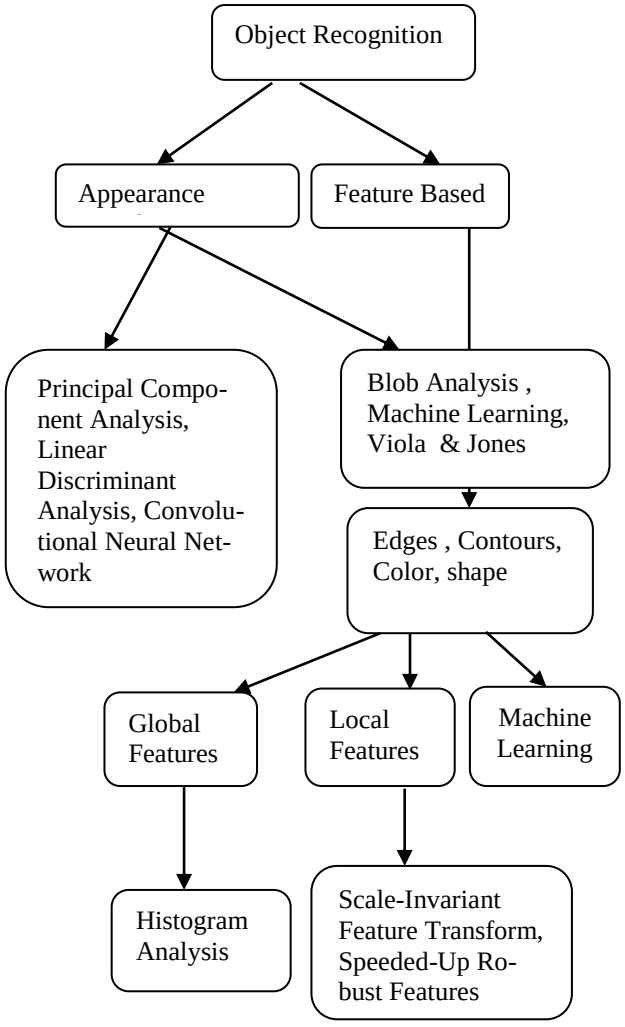

Fig.3.1 Object Recognition Techniques

3.2. Appearance-based technique

Face or handwriting recognition applications uses these methods. A set of reference training images, which are highly correlated, is needed for this strategy. Dimensionality reduction techniques compresses dataset to obtain a lower dimension subspace, also called eigen space. Correspondence is examined by projecting parts of the new input images on the eigenspace. **E**xample images (called templates or exemplars)are used in appearance-based technique to perform object recognition.
Change in light or color, change in viewing direction, change in size / shape changes the look of object. A single exemplar is unlikely to succeed reliably. However, all appearances of an object cannot be represented. Appearance-based methods include Convolutional Neural Networks, PCA and LDA.

3.2.1. Image segmentation and blob analysis

A region of a digital image differing in properties, such as brightness or color compared to surrounding regions is detected in **blob detection** methods can be used in computer vision. A region of an image in which some properties are constant or approximately constant is Blob ; the points which can be considered similar to each other in some sense are present in Blob . Most common method for blob detection is Convolution .

3.2.2. In order to recognize a variety of objects, including faces and upper bodies, the **Viola-Jones** algorithm can be used.

3.2.3. Machine Learning : can be feature based or appearance based.

To classify unknown feature vectors given a large database of feature vectors whose classifications are known is the job of a machine learning algorithm.

A capability enabling Artificial Intelligence (AI) systems to learn from data is provided by Machine Learning. A good definition for what learning involves is the following: “Suppose E represents experience, T is class of tasks ,P represents performance measure . A computer program is said to learn from E with respect to T with performance measure P if its performance at tasks in T, as measured by P, improves with E”. Machine Learning algorithms can be classified using nature of this experience E into the following three categories: supervised, unsupervised, and reinforcement learning:(i) Algorithms are presented with a dataset containing a collection of features in supervised learning. Each sample is provided with labels or target values. The knowledge is encoded in mapping of features to labels of target values. To find the mapping from the features of unseen samples to their correct labels or target values after learning is the job of algorithm.(ii) meaningful representations are extracted in unsupervised learning and key features explanation of the data is provided. No labels or target values are necessary in this case in order to learn from the data.(iii) An AI agent interacts with a real or simulated environment in reinforcement learning algorithms. The interaction between the learning system and the interaction experience provides feedback which is useful to improve performance in the task being learned.
Viola and Jones face detector is briefly described in this paper. A hybrid face detection system which is combination of Skin color classification and Viola Jones has also been discussed and proposed. Increase in face detection speed and decrease in false positive rate is identified using proposed method[6].
In Asia, Rice utilized as a staple food hence is an important crop. The process of rice growing involves intensive human labor. Intelligent and robotic platforms can be used to automate the difficult labor of rice production . To automate the process of distinguishing between rice and weeds an intelligent agent with sensors can be used, so that a robot can cultivate fields, is proposed here. Automatic identification and distinguishment of weeds from rice plants is described in the paper. To find the points of interests such as the tips of leaf and the rice ear a Harris Corner Detection algorithm is firstly applied, multiple features for each points surrounding area are extracted to feed into a machine learning algorithm to discriminate weed from rice, multiple features for each points surrounding area are extracted to feed into a machine learning algorithm, a clustering algorithm is used for noise removal based on the points position and density as a third step. Internet downloaded images yielded very promising classification result[7].

3.2.4. Deep learning models such as CNNs

Machine learning algorithms are a superset of Deep Learning. Algorithms involving learning representations at different hierarchy levels to enable building complex concepts out of simpler ones are deep learning.

To accept 2-dimensional input data, such as images or time series data , Convolutional Neural Networks (CNNs) are conceived. The mathematical linear operation of convolution is always present in at least one of the layers of the network hence these models are known as CNN. 2D convolution of a 2-dimensional image I with a 2-dimensional kernel K is the most typical convolution operation used in deep learning.

3.3. Template matching

Pattern matching methods are sometimes called template matching methods, are often used because of their simplicity. Finding small parts of an image which matches a template image is job of Template matching. Traffic sign recognition, small parts of the input image are tried to be matched with a database full of different images of traffic signs uses template matching . This approach has occlusion, rotation, scaling, different illuminations problems .

4. Object Recognition Application

1 .**Image panoramas**: to capture images with horizontally elongated fields of view Panoramic photography a technique of photography, is used. It is sometimes known as wide format photography. Here, a photograph that is cropped to a relatively wide aspect ratio, like the familiar letterbox format in widescreen video.

2. **Face detection**: A specific case of object-class detection is face detection. Detection of frontal human faces is done in Face-detection algorithms. It is analogous to image detection in which the image of a person is matched bit by bit. Image matches with the image stores in database. Any change in facial feature in the database will invalidate the matching process.

3. Optical Character Recognition: In this, images of typed, handwritten or printed text are converted into machine-encoded text either the mechanically or electronically .Text may be from a scanned document, a photo of a document or it may be from subtitle text superimposed on an image.

4. Manufacturing Quality Control : In manufacturing, a process that ensures customers receive products free from defects and meet their needs is quality control .

5. **Content-Based Image Indexing** :Retrieval from text-based indexing of images that may employ keywords, subject headings, captions, or natural language text is known as Concept-based image indexing, also variably named as "description-based" or "text-based" image indexing/retrieval. It is opposed to Content-based image retrieval.

6. Automated vehicle parking systems

7. **Android Eyes - Object Recognition**

5. Object Recognition Techniques Comparison

Table 5.1: Performance evaluation in terms of Robustness

Feature	SIFT	SURF	PCA	LDA	CNN
rotation	++	0	+	+	++
illumination	0	++	-	-	++
occlusion	++	+	-	-	++
speed	+	++	0	0	-

Variations of the images in the training dataset affects CNNs, but its complex architecture provides very good results even when parts of the object are occluded or even human eyes cannot recognize the object. It has comparable low speed.
SURF outperforms SIFT, because it uses approximations to be faster and adjusts parts of the steps to wipe out the problems of SIFT. In comparison with the other algorithms, SIFT and SURF are faster, because only small parts of the image or features have to be analysed, instead of CNNs using the whole image. PCA and LDA have their problem with occlusion. The object to be detected has to be fully visible and has to be easily separated from the background, because the algorithms depend on the complete appearance of the object.
Secondly, computational load and memory usage for the algorithms are compared .Number of images used in the training dataset affect computational load and the memory usage. PCA, for example, is popular for face recognition, therefore for this application just training images of faces and non-faces are needed. Whereas CNNs useful in identification of many more different objects. All objects have to be represented in the training dataset in order to achieve good results with CNNs.
This fact is more or less independent of the chosen algorithm, but there are also significant differences during runtime after the learning phase.
The memory usage and computational load is rather low for the appearance-based algorithms since they rely on relatively easy matrix equations. It is enough to have one representation for object 1, object 2 and so on in PCA and LDA to compare features. Complexity of SURF is lower than SIFT as it mostly uses approximations of the steps of SIFT. SIFT and SURF need a database with models to match the features, which results in a higher memory usage when compared to the appearance-based methods. CNNs need a lot of resources as up to one billion parameters (weights of neurons) were utilized. CNNs are very powerful.
In the following Tab. below, the gained knowledge is summarized and the results are rated from -- to ++.

Table 5.2: Performance evaluation in terms of Complexity

Feature	SIFT	SURF	PCA	LDA	CNN
computational load	+	+	+	0	--
Memory usage	-	0	+	+	--
overall complexity	0	+	++	+	--

Good recognition rate is the main strength of SIFT, whereas its processing speed is low. However, SURF has low recognition rate but outperforms the other algorithms in terms of speed and complexity. PCA does not achieve a good accuracy, but in some special cases it outperforms other algorithms. PCA reduces dimensionality and is therefore often used in combination with other approaches is its main advantage. LDA has good results in memory usage and computational load and is very popular in two-class applications. It is the second appearance based method.

6. Conclusion

Terminologies used in object recognition have been addressed, in this paper. These include object recognition methods, feature selection and object classification. Detailed explanation of most commonly used and well recognized methods for these phases have been explained. Performance of various object recognition methods is compared using parameters such as robustness, complexity. Convolutional Neural Network method provide very good results even when the parts of objects are occluded

References

[1] K.Khurana and R.Awasthi ,“ Techniques for Object Recognition in Images and Multi-Object Detection” , International Journal of Advanced Research in Computer Engineering & Technology(IJARCET) Vol.2, No. 4, 2013, pp.1383-1388.

[2] Bernd Heisele, Ivaylo Riskov and Christian R. Morgenstern ,Toward Category-Level Object Recognition, **Ponce**, J. et al.,2006,pp.225-237.

[3] Ryszard S. Choras, “ Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems” , International Journal Of Biology and Biomedical Engineering, Vol. 1,No.1,2007, pp.6-16.

[4] L. Dong, X. Yu ,L. Li, J. Kah Eng Hoe, “HOG based multi-stage object detection and pose recognition for service robot,” Proceedings of Control Automation Robotics & Vision (ICARCV), 2010, pp. 2495 – 2500.

[5] Amir. Jamshidnezhad, and MohamadMehdi. Lotfinejad, “A Morphological based Technique for Features Extraction from the Face Color Images” , International Journal of Computer Science and Electronics Engineering (IJCSEE), Vol 2, No. 3, 2014 ,pp.158-162.

[6] Zahra Sadri Tabatabaie, Rahmita Wirza Rahmat, Nur Izura Binti Udzir and Esmaeil Kheirkhah , “A Hybrid Face Detection System using combination of Appearance-based and Feature-based methods” , International Journal of Computer Science and Network Security, Vol.9 ,No.5, 2009,pp.181-184.

[7] Beibei Cheng , Eric T. Matson , “A Feature-Based Machine Learning Agent for Automatic Rice and Weed Discrimination ”, Proceedings of International Conference, ICAISC,Vol.9119, No.1 , 2015, pp 517-527.