International Journal of Engineering Sciences Paradigms and Researches (IJESPR)
(Vol. 21, Issue 01) and (Publishing Month: June 2015)
(An Indexed, Referred and Impact Factor Journal)
ISSN (Online): 2319-6564
www.ijesonline.com

A Survey on Object Recognition Techniques

Manju[1] and Dr. Satpal[2]

[1]**Research Scholar, Department of Computer Science, Baba Mastnath University, Rohtak, HR (India)**
dr.manjumalik21@gmail.com

[2]**Professor, Department of Computer Science, Baba Mastnath University, Rohtak, HR (India)**

Abstract

Object recognition is the necessity for the real advancement of the technology. Any machine can recognize the object based by using the object recognition techniques. This paper studies various object recognition techniques. The attributes based techniques seem to be promising for the higher accurate recognition. The attributes based classification is of two type; direct attribute based prediction and indirect attribute based prediction. The paper also gives the zero shot learning as training samples are not available for the various real world objects.

Keywords*: DAP, IAP, Object Recognition, Zero Shot Learning.*

I. Introduction

RECOGNITION is one of the most useful functions of our visual system [1]. We recognize materials (marble, orange peel), surface properties (rough, cold), objects (my car, a willow tree), and scenes (a thicket of trees, my kitchen) at a glance and without touching them. We recognize both individuals (my mother, my office), as well as categories (a1960s hairdo, a frog). By the time we are six years old, we recognize more than 104 categories of objects, and keep learning more throughout our life. As we learn, we organize both objects and categories into useful and informative taxonomies and relate them to language. Replicating these abilities in the machines that surround us would profoundly affect the practical aspects of our lives, mostly for the better. Certainly, this is the most exciting and difficult puzzle that faces computational vision scientists and engineers in this decade [1].

II. Object Recognition

Object recognition is one of the fundamental challenges in computer vision [2]. Object recognition concerns itself with a two-fold problem. First, what is the form of visual object representation? Second, how do observers match object percepts to visual object representations? Unfortunately, the world isn't color coded or conveniently labeled for us. Many objects look similar (think about four-legged mammals, cars, or song birds) and most contain no single feature or mark that uniquely identifies them. Even worse, objects are rarely if ever seen under identical viewing conditions: objects change their size, position, orientation, and relations between parts, viewers move about, and sources of illumination turn on and off or move. Successful object recognition requires generalizing across such changes. Thus, even if an observer has never seen a bear outside of the zoo, on a walk in the woods they can tell that the big brown furry object with teeth 20 ft in front of them is an unfriendly bear and probably best avoided or that the orange-yellow blob hanging from a tree is a tasty papaya [3].

III. Zero-Shot Learning

Zero-shot learning (ZSL) for visual classification has received increasing attentions recently . This is because although virtually unlimited images are available via social media sharing websites such as Flickr, there are still not enough annotated images for building a visual classification model for a large number of visual classes. ZSL aims to imitate human's ability to recognize a new class without even seeing any instance. A human has that ability because he/she is able to make connections between an unseen class with the seen classes based on its semantic description. Similarly a zero-shot learning method for visual classification relies on the existence of a labeled training set of seen classes and the knowledge about how each unseen class is semantically related to the seen classes. An unseen class can be related to a seen class by representing both in a semantic embedding space.

IV. Object Recognition Methods

Object recognition methods can be classified according to a number of characteristics.[4] The object recognition classification as:

a) Appearance Based Methods

Appearance-based approaches have become a popular method for object identification and pose location in intensity images, and has also been used to recognize free-form objects in range data, as noted below. This popularity is in part due to these methods' ability to handle the effects of shape, pose, reflectance, and illumination variations for large databases of general objects. An appearance-based recognition system encodes individual object views as points in one or more multidimensional spaces. The bases for these spaces are obtained from a statistical analysis of the ensemble of training images. Recognition of an unknown view is typically performed by projecting that view into the space(s) along the stored basis vectors and finding the nearest projected view of a training image [5].

Appearance based methods typically include two phases. In the first phase, a model is constructed from a set of reference images. The set includes the appearance of the object under different orientations, different illuminants and potentially multiple instances of a class of objects, for example faces. The images are highly correlated and can be efficiently compressed using e.g. Karhunen- Loeve transformation (also known as Principal Component Analysis - PCA) [4].

In the second phase, "recall", parts of the input image (subimages of the same size as the training images) are extracted, possibly by segmentation (by texture, colour, motion) or by exhaustive enumeration of image windows over whole image. The recognition system then compares an extracted part of the input image with the reference images (e.g. by projecting the part to the Karhunen-Loeve space). A major limitation of the appearance-based approaches is that they require isolation of the complete object of interest from the background. They are thus sensitive to occlusion and require good segmentation [4].

b) Geometry-Based Methods

In geometry- (or shape-, or model-) based methods, the information about the objects is represented explicitly. The recognition can than be interpreted as deciding whether (a part of) a given image can be a projection of the known (usually 3D) model [6] of an object. Generally, two representations are needed: one to represent object model, and another to represent the image content. To facilitate finding a match between model and image, the two representations should be closely related. In the ideal case there will be a simple relation between primitives used to describe the model and those used to describe the image [4].

c) Attribute-Based Classification Method

Attribute-based classification models object classes relative to an inventory of descriptive attributes. For a given class, each attribute can be either active or inactive, resulting in a characteristic association signature for that class. Following the probabilistic formulation of the DAP model, let $a_y = (a_1^y, \dots\dots a_m^y)$ be a vector of binary associations$a_m^y \epsilon \{0,1\}$ between attributes $a_m$and training object classes y. A classifier for attribute am, trained by labeling all images of all classes for which $a_m^y = 1$ as positive and the rest as negative training examples, can provide an estimate of the posterior $p(a_m|x)$ of that attribute being present in image x. Mutual independence yields $p(a|x) = \prod_{m=1}^{M} p(\,a_m|x)$ for multiple attributes [7].

In order to transfer attribute knowledge to an unknown class z, we take a binary vector a^z for which $p(a|z) = 1$ for a = a^z and 0 otherwise. The posterior probability of class z being present in image x is then obtained by marginalizing over all possible attribute associations a, using Bayes' rule

$$p(z|a^z) = \frac{p(a^z|z)p(z)}{p(a^z)} = \frac{p(z)}{p(a^z)}$$

$$p(z|x) = \sum_{a\epsilon\{0,1\}^M} p(z|a)p(a|x)$$

$$= \frac{p(z)}{p(a^z)} \prod_{m=1}^{M} p(\,a_m|x) a_m^z (1)$$

Assuming identical class priors p(z) and a factorial distribution for $p(a) = \prod_{m=1}^{M} p(\,a_m)$, we obtain

$$p(z|x) \propto \prod_{m=1}^{M} \left(\frac{p(a_m|x)}{p(a_m)}\right) a_m^z \;\; (2)$$

Attribute priors can be approximated by empirical means over the training classes $p(a_m) = \frac{1}{K}\sum_{k=1}^{K} a_m^{yk}$. Classifying an image x according to test classes z_L uses MAP prediction:

$$f(x) = \max_{l=1,\ldots,L} \prod_{m=1}^{M} \left(\frac{p(a_m|x)}{p(a_m)}\right) a_m^{zl} \quad (3)$$

V. Attribute Classifiers

In creating a classifier for a particular attribute, we could simply extract all types of low-level features from the whole object, and let a classifier figure out which are important for the task and which are not. Attribute classifiers C_i are built using a supervised learning approach. Training requires a set of labeled positive and negative images for each attribute [8]. The goal is to build a classifier that best classifies this training data by choosing an appropriate subset of the feature set F (I). Each classifier's performance is evaluated using cross validation. The features used in the classifier with the highest cross-validation accuracy are added to the output set. We continue adding features until the accuracy stops improving. When no training samples are available the following technique can be used.
The attribute-based classification is indeed a solution to the problem of learning with disjoint training and test classes. Two generic methods to integrate attributes into multi-class classification [9] are: Direct attribute prediction (DAP), Indirect attribute prediction (IAP).

a) Direct Attribute Prediction (DAP)

It uses an in between layer of attribute variables to decouple the images from the layer of labels. During training, the output class label of each sample induces a deterministic labeling of the attribute layer. Consequently, any supervised learning method can be used to learn per attribute parameter. At test time, these allow the prediction of attribute values for each test sample, from which the test class label are inferred. DAP is illustrated in figure 1.

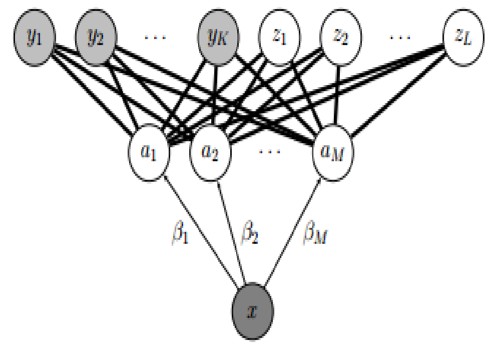

Figure 1: Direct attribute prediction (DAP) [9]

b) Indirect Attribute Prediction (IAP)

It also uses the attributes to transfer knowledge between classes, but the attributes form a connecting layer between two layers of labels, one for classes that are known at training time and one for classes that are not.

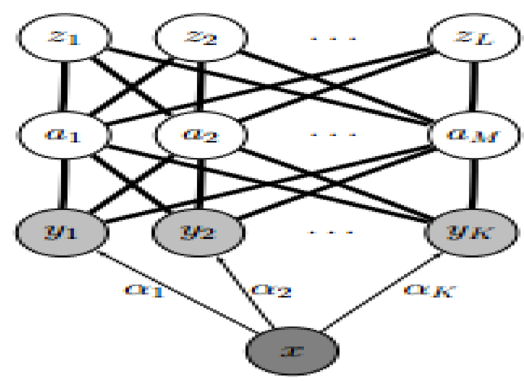

Figure 2: Indirect attribute prediction (IAP) [9]

The training phase of IAP is ordinary multi-class classification. At test time, the predictions for all training classes induce a labeling of the attribute layer, from which a labeling over the test classes can be inferred. IAP is illustrated in figure 2.

VI. Conclusion

The paper discusses the classification of object recognition techniques. The direct attributes based prediction and indirect attribute based prediction techniques have been focused. The DAP and IAP techniques are popular and promising due to zero shot learning used in the techniques. In future these techniques can be extended to enhance the accuracy of recognition.

References

[1] Fei-Fei, L., Fergus, R., & Perona, P. (2006). One-shot learning of object categories. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 28(4), 594-611.

[2] Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2010). Object detection with discriminatively trained part-based models. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32(9), 1627-1645.

[3] Tarr, M. J., & Vuong, Q. C. (2002). Visual object recognition. Stevens' handbook of experimental psychology.

[4] Matas, J., & Obdrzalek, S. (2004). Object recognition methods based on transformation covariant features. na.

[5] Campbell, R. J., & Flynn, P. J. (2001). A survey of free-form objects representation and recognition techniques. Computer Vision and Image Understanding, 81(2), 166-210.

[6] Pope, A. R. (1994). Model-Based Object Recognition. A Survey of Recent Tecniques, Technical Report.

[7] Zhao, Z., Xie, J., & Chen, X. (2014, May). Attribute based object recognition by human language. In Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems (pp. 1603-1604). International Foundation for Autonomous Agents and Multiagent Systems

[8] N. Kumar, A. Berg, P. Belhumeur, and S. Nayar ,(2011) Describable visual attributes for face verification and image search, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI),vol. 33, no. 10, pp. 1962–1977.

[9] C. H. Lampert, H. Nickisch, and S. Harmeling, (2009) Learning to detect unseen object classes by between-class attribute transfer, in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[10] Fu, Z. Y., Xiang, T., & Gong, S. (2014). Semantic Graph for Zero-Shot Learning. arXiv preprint arXiv:1406.4112.