Nova Explore Publications
Nova Journal of Engineering and Applied Sciences
Vol. 3(3), 2014: 1-8
PII: S229279211400011-3
www.novaexplore.com

Research Article

Unsupervised Object Category Recognition from Image Datasets

Saima Sultana, Syeda Ohishee Rahman, Md. Alamgir Hossain
Department of Computer Science and Engineering, Sylhet Engineering College, Tilagarh, Sylhet 3100 Bangladesh
**Corresponding author: Saima Sultana; E-mail : sinthi315@gmail.com*

Abstract

The variability of the object models is treated as flexible constellations of rigid parts which is represented by a joint probability density function (pdf) on the shape of the constellation and the output of part detectors. Firstly we use method an Affine Invariant Salient Region Detector to identify the distinctive parts in the training set for clustering data. Then we use Markov chain Monte Carlo expectation maximization (MCMC-EM) algorithm to learn the statistical shape model of the object and discover object categories in an unsupervised manner. In the MCMC-EM algorithm, the high-dimensional integrals required in the EM algorithm are estimated using MCMC sampling. The MCMC sampler requires simulation of sample paths from a continuous time Markov process, conditional on the beginning and ending states and the paths of the neighboring sites.

Keywords: *Object recognition, dataset, algorithm, models, detector, probability density function (pdf).*

Introduction

Great efforts have been undertaken over the last half century to build machines that can see. A great number of fundamental problems have been solved. Examples of such problems are the detection of edges in images, the computation of basic flow fields representing motion throughout a sequence of images, or the point-by-point registration of a pair of stereoscopic images. At the same time, one cannot escape the impression that the biggest problems are still far ahead of us. The focus of this thesis is on the problem of learning principled representations of objects automatically from sensory signals, and on how such representations can be used to detect objects. We consider an object to be a part of, or token in, a sensory signal. The precise representation of the object within the signal can undergo changes such as scaling, translation, or other deformations, or it can be contaminated by noise or be partially occluded. These changes give rise to an entire collection or class of signals which can all still be associated with the original object. Objects in signals often correspond to physical objects in the real world environment from which the signals have been recorded. This is the case for objects in images of natural scenes. However, objects can also be defined solely in the universe of signals. Classes of objects are collections of objects that are similar. The similarity might be based on high-level cognitive principles, e.g., in the case of the class of chairs. In this case a recognition system might have to embody or develop some understanding of these concepts in order to identify objects of the class. In a more tractable scenario, the similarity largely manifests itself at the level of the signal representation. For example, portions of the signals of two different objects from the same class could be identical.

There are two fundamentally different approaches to building recognition systems. One is to design every aspect and detail of the system before it is put to use. Under the second approach, instead of designing a complete solution, one would endow the system with some prior knowledge about the problem at hand, as well as some mechanism that will allow it to extract, from some collection of *training signals*, the necessary remaining information to finally solve the problem. This concept is referred to as learning. The amount of learning can vary from a single parameter being adjusted to a complete representation of the problem setting being acquired. The extreme case would be a system that embodies a simple method of comparing signals and a huge memory to store every signal that might possibly be encountered. It could then recognize objects by simply looking up the answer to the given problem in memory, using the input image as an index into the list of answers. Since such large amounts of memory are unavailable for most problems, and since training of such a system would require tutoring until all likely signals have been encountered at least once, this procedure is rather impractical. Especially since many simple rules, such as invariance relations with respect to changes in overall brightness in an image, can easily be incorporated into a recognition system. Learning methods have proved very successful in computer vision and many other fields. Some of the most successful methods for face detection, for example, make extensive use of learning techniques. They need to be trained on a set of a few hundred images, before they become powerful recognition tools.

In this work, we are particularly interested in learning methods that do not require any supervision by a human operator during the learning phase. We review some of the methods proposed for object recognition that are related to ours. Other widely used but less closely related approaches include recent large-scale object labeling efforts [1, 2, 3] have demonstrated the difficulties in deciding on the granularity of the categories to be labeled. A categorization or labeling of the world thought up by one person may not in fact be the most useful for training a machine how to see. There has also been success in learning robust representations of specific classes in constrained situations, notably frontal faces [4] and pedestrians in street scenes [5, 6], but models that can be trained to recognize generic object categories remain elusive. In contrast, an unlabelled training set comes virtually free; one only needs to point a camera out at the world to obtain an unlimited supply of training images. There has been recent research interest in learning from unlabelled data, including unsupervised algorithms for object categorization [7, 8] and segmentation [9, 10]. These algorithms have functioned as proofs of concept, demonstrating in some cases that models from the statistical analysis of text can be modified to apply to unsupervised analysis of images.

Applications of image dataset

Applications for recognition systems are manifold and their number is steadily growing with the rapid improvements in signal acquisition and storage capabilities. Different types of sensors are almost omnipresent and provide an abundance of data that needs to be processed. Therefore, automatic tools that understand or at least extract useful information from images are needed. Object recognition would be key. Interaction with machines is often difficult and cumbersome, if not impossible, unless these machines understand something about their environment, in particular the objects and creatures they interact with. In order to navigate in any environment, vision is the most important sensory modality. In interactions between humans, visual input is almost as important as auditory information. This suggests that vision should play a crucial role in human-machine interactions. It is of great importance that these systems be able to detect objects such as roads, pedestrians, traffic signs, and other vehicles. Another rapidly expanding area in computer vision is medical image processing. Here, machines that can accurately detect objects such as bones, organs, boundaries between different types of tissues, or tumors would be of great practical usefulness. Such biological objects lend themselves particularly well to the approach we have chosen to model object classes.

Modeling Objects as Collections of Parts

Peter, Gyuri, Cordelia, Hendrik, and Freitas [11] showed improvements over state-of-the-art local feature recognition systems, how to formulate principled models for automatic local feature selection in object class recognition when there is little supervised data, and how to formulate sensible spatial image context models using a conditional random field for integrating local features and segmentation cues (super-pixels). By adopting sparse kernel methods, Bayesian learning techniques and data association with constraints, the proposed model identifies the most relevant sets of local features for recognizing object classes, achieves performance comparable to the fully supervised setting, and obtains excellent results for image classification. In principle this is possible with our approach and we have experimented with this idea according to some change for better result. Shimon and Nathan [12] invoked philosophical observations, computational arguments, behavioral data and neurobiological findings to explain why computer vision researchers should care about unsupervised learning, statistical inference and the visual brain. We then outline a neuromorphic approach to structural primitive learning motivated by these considerations, survey a range of neurobiological findings and behavioral data consistent with it and conclude by mentioning some of the more challenging directions for future research. Kadir, Andrwe and Brady [13] introduce a novel technique for detecting salient regions in an image. The detector is a generalization to affine invariance of the method introduced by Kadir and Brady [14]. The detector deems a region salient if it exhibits unpredictability in both its attributes and its spatial scale. They compare the performance of the saliency detector to other standard detectors including an affine invariance interest point detector. It is demonstrated that the saliency detector has comparable viewpoint invariance performance, but superior insensitivity to perturbations and intra-class variation performance for images of certain object classes. We have followed this method in our experiment.

Kamarainen, Hamouz, Kittler, Paalanen, Ilonen and Drobchenko [15] applied state-of-the-art approach to object detection and localisation by incorporating local descriptors and their spatial configuration into a generative probability model. They applied a supervised approach where local image features (landmarks) are annotated in a training set and therefore their appearance and spatial variation can be learnt. Their method enables working in purely probabilistic search spaces providing a MAP estimate of object location, and in contrast to the recent methods, no background class needs to be formed. Using the training set, pdfs for both spatial constellation and local feature appearance were estimated. By applying an inference bias that the largest pdf mode has probability one, they were able to combine prior information (spatial configuration of the features) and observations (image feature appearance) into posterior distribution which can be generatively sampled, e.g. using MCMC (Markov chain Monte Carlo) techniques. The MCMC methods are sensitive to initialization, but as a solution, we also propose a very efficient and accurate RANSAC-based method for finding good initial hypotheses of object poses. The complete method can robustly and accurately detect and localize objects under any homography. We have introduced some changes to the model and detection method and developed unsupervised learning algorithms for estimating the models from data.Hichem, Audibert and Keriven [16] focused in this paper on object recognition using a new type of kernel referred to as "context-dependent". Objects, seen as constellations of interest points are matched by minimizing an energy function mixing a fidelity term which measures the quality of feature matching, a neighborhood criterion which captures the object geometry and a regularization term. They also showed that the fixed-point of this energy is a context-dependent kernel (CDK) which is also positive definite. Experiments conducted on object recognition show that when plugging our kernel in SVMs, we clearly outperform SVMs with context-free kernels (CFK).

Wen Wu and Jie Yang [17] presents a semi-supervised learning (SSL) ap-proach to find similarities of images using statistics of local matches. SSL algorithms are well known for leveraging a large amount of unlabeled data as well as a small amount of labeled data to boost classification performance. Their approach proposes to formulate the problem of matching two images as an SSL based classification problem of image pairs with a minimal amount of labeled pairs. We apply a Gaus-sian random field model to represent each image pair as vertices in a weighted graph and the optimal configuration of the field is obtained by harmonic energy minimization. Their approach consist of only local unsupervised image segmentation but not spatial. Amiya, Soumajit and Arindam [18] described an evolutionary approach for unsupervised gray-scale image segmentation that segments an image into its constituent parts automatically. The aim of this algorithm is to produce precise segmentation of images using intensity information along with neighborhood relationships. In this paper, fuzzy c-means clustering helps in generating the population of Genetic algorithm which there by automatically segments the image. This technique is a powerful method for image segmentation and works for both single and multiple-feature data with spatial information. Validity index has been utilized for introducing a robust technique for finding the number of components in an image. Experimental results shown that the algorithm generates good quality segmented image. Latter Amiya and Soumajit [19] improved their experiment using Fuzzy Hopfield Neural Network (FHNN) clustering helps in generating the population of Genetic algorithm which there by automatically segments the image.

Other Deformable Models

Deformability of objects has also been modeled without decomposing objects into specific parts. Work along these lines has been done by Minshu Cho, Jungmin Lee and Kyoung Mu Lee [20]. In this approach, they present an efficient method for feature correspondence and object-based image matching, which exploits both photometric similarity and pair-wise geometric consistency from local invariant features. They formulate object-based image matching as an unsupervised multi-class clustering problem on a set of candidate feature matches, and propose a novel pair-wise dissimilarity measure and a robust linkage model in the framework of hierarchical agglomerative clustering. The algorithm handles significant amount of outliers and deformation as well as multiple clusters, thus enabling simultaneous feature matching and clustering from real-world image pairs with significant clutter and multiple deformable objects. Felzenszwalb, Girshick and McAllester [21] developed a complete learning-based system for detecting and localizing objects in images. Their system represents objects using mixtures of deformable part models. These models are trained using a discriminative method that only requires bounding boxes for the objects in an image. Isonas Kokkinos and the Project-Teams GALEN [22] used bounding-based techniques, such as Branch-and-Bound (BB) and Cascaded Detection (CD) to efficiently detect objects with Deformable Part Models (DPMs). Instead of evaluating the classifier score exhaustively over all image locations and scales, they use bounding to focus on promising image locations. The core problem is to compute bounds that accommodate part deformations; for this they adapt the Dual Trees data structure to our problem.

Face Detection & Facial expression recognition

Now-a-days the most prominent and most frequently addressed application of object recognition is detecting and recognizing human faces and also the expression of human faces in images. Recently, several systems have been presented, which are able to detect faces as well as facial expression from several viewing directions in cluttered scenes with high accuracy. Kashem, Nasim, Shamim and Mahbub [23] proposed a face recognition system for personal identification and verification using Principal Component Analysis (PCA) with Back Propagation Neural Networks (BPNN) is proposed. This system consists on three basic steps which are automatically detect human face image using BPNN, the various facial features extraction, and face recognition are performed based on Principal Component Analysis (PCA) with BPNN. The dimensionality of face image is reduced by the PCA and the recognition is done by the BPNN for efficient and robust face recognition.Thanh-Toan, Ewa [24] applied Co-occurrence Histograms of Oriented Gradient (CoHOG) in face recognition, which is an extension of HOG, on the face recognition problem. Some weighted functions for magnitude gradient are tested. They also proposed a weighted approach for CoHOG, where a weight value is set for each subregion of face image. The recognition results using CoHOG are competitive with some of the state of the art methods. This proves the effectiveness of CoHOG descriptor for face recognition.

Santhosh and Manjula [25] developed an algorithm which is capable of identifying a person's facial expression and categorize them as happiness, sadness, surprise and neutral. Their approach is based on local binary patterns for representing face images. In their project they use training sets for faces and non faces to train the machine in identifying the face images exactly. Facial expression classification is based on Principle Component Analysis. In their project, they have developed methods for face tracking and expression identification from the face image input. Applying the facial expression recognition algorithm, the developed software is capable of processing faces and recognizing the person's facial expression. The system analyses the face and determines the expression by comparing the image with the training sets in the database. They have followed PCA and neural networks in analyzing and identifying the facial expressions. Arumugam and Purushothaman [26] developed an algorithm which is used to identify the person's emotional state through facial expression such as angry, disgust, happy. This can be done with different age group of people with different situation. They used a Radial Basis Function network (RBFN) for classification and Fisher's Linear Discriminant (FLD), Singular Value Decomposition (SVD) for feature selection.

Problem of learning object models

The problem of learning object models from unlabeled and unsegmented images has been neglected in most instances. It is often assumed that the images contain a constant background, or correspondence needs to be established by intervention of a human operator, whose task might be to hand-label feature points in the images. A first step towards unsupervised learning methods is the automatic computation of correspondence between training images. This problem has been addressed recently by Amiya, Soumajit and Arindam [18]. But all FCM based algorithm minimize intra-cluster variance as well, but have the same problems as *k*-means; the minimum is a local minimum, and the results depend on the initial choice of weights. A learning problem similar to ours is addressed by Kadir, Andrwe and Brady [13]. Here the author shows the optimal choice for region selection depends ultimately on the application, however, there are three broad classes of image change under which good performance may be required: Global transformations, Local perturbations and Intra-class variations.

Object Recognition with Constellation Models

Global Geometry and Local Photometry : In order to recognize objects as members of classes, one has to deal with the, possibly large, variability of the signals across a class. We propose a principled strategy to model important sources of variability explicitly. It is based on the idea of decomposing objects into components. Before we introduce our model based on "global geometry and local photometry," we will illustrate how this approach allows us to address the different sources of variability.We will then derive the theoretically optimal detector based on this object class model. Unfortunately this detector is computationally too demanding for most practical recognition problems. Therefore, we will introduce, in the next chapter, an abstract version of our model which can be implemented more efficiently.

Most ideas included in this chapter are based on previous work by Kadir, Andrwe and Brady [13].If we study the visual world surrounding us, we discover many objects that are visually similar. A few examples are human faces, bicycles, hands, cars, coffee mugs, telephones, shoes, pens, chairs, staplers, houses, tree leaves, flowers of the same species, and bottles. All these objects could be thought of as instantiations of *object classes*. Some classes can only be defined based on high-level cognitive concepts. The class of chairs, for example, could be defined in a rather abstract way as "anything man-made for a single person to sit on." As a consequence, two chairs are often visually extremely dissimilar. Just about any particular feature one might choose, in order to devise a more concrete definition of this object class, such as an arm rest, back rest, or a leg, is not present in many chairs. There is not much hope to model such abstract classes based solely on visual appearance. However, there is an intermediate range of classes, e.g., human faces, leaves of trees, or cars, the members of which are more consistently visually similar, yet not identical. It is the problem of identifying members of those classes that we address here (see Figure 2.1).

If we are asked to give a definition of, say, the class of human faces, it comes naturally to do so by describing a decomposition of a prototypical face in terms of distinctive features or parts. A face is composed of two eyes, a nose, etc. Similar decompositions suggest themselves for cars (in terms of wheels, body, headlights, bumper) or hands (five fingers). This idea of decomposing objects into components was introduced to computer vision research almost three decades ago by Fischler and Elschlager [FE73]. Although it is difficult to give a formal definition, a "part" shall be understood as a subset of an object satisfying four key requirements: (1) The parts of one object correspond to parts in other objects belonging to the same class. (2) Corresponding parts are similar; more so than objects belonging to the same class are similar. (3) Detectors can be devised that can detect the part with some degree of success. (4) A part has geometrical properties, such as a spatial position, often a scale, an orientation, a spatial extent, a velocity.

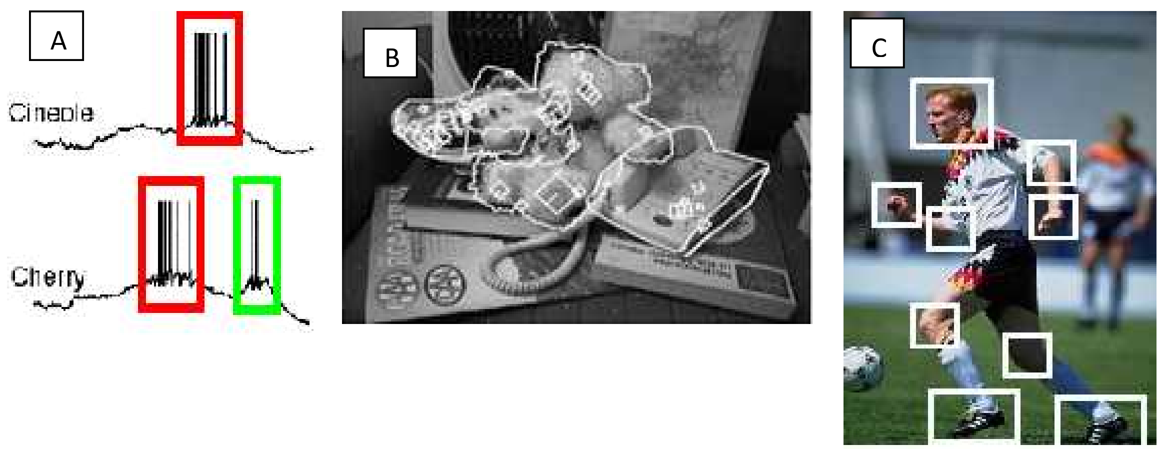

Figure 2.1: Different data may be represented as 'parts' and 'shape'. (a) Sequences of spikes produced by neurons in the mushroom body of the locust in response to different odors. Either individual action potentials, or bursts of action potentials could represent the parts. The mutual distance between such parts is the 'shape'. (b) A human face may be thought of as a collection of textured patches in a specific mutual position. (c) The human body in motion may be thought of as a collection of textured patches, each having a position and velocity.

Since it is often possible to model and detect parts based on the underlying photometric pattern, we refer to this portion of the model as *local photometry*. A complete description of an object cannot be given by simply listing all its parts. It is also important how the parts are arranged spatially. We refer to this information as the *global geometry* of an arrangements of parts. From this global perspective, parts are reduced to a set of attributes, such as type, position, orientation, etc.

Addressing Signal Variability Explicitly

Describing objects as a collection of parts is not only a natural approach, but it also allows to account for signal variability in a principled way. We will soon show how exactly one could represent the variability in the geometry of an arrangement of parts and in the appearance of parts themselves using probability models. The selection of a set of image regions forms the first step in many computer vision algorithms, e.g object recognition or computing image correspondences. The optimal choice for region selection depends ultimately on the application, however, there are three broad classes of image change under which good performance may be required:

Global transformations : Features should be repeatable across the expected class of global image transformations. These include both geometric and photometric transformations that arise due to changes in the imaging conditions. For example, we often require that feature segmentations commute with viewpoint change.

Figure 2.2: Detected regions, denoted by the green ellipses, sould commute with viewpoint change.

Local perturbations: Features should be insensitive to classes of semi-local image disturbances. For example, a feature responding to the eye of a human face should be unaffected by any motion of the mouth. A second class of disturbance is where a region neighbours an object boundary. The detector can be required to detect the region corresponding to the object despite changes in the background.

Figure 2.3: Features should be insensitive to background changes

Intra-class variations: Features should capture corresponding object parts under intra-class variations in objects. For example, the headlight of a car for different brands of car (imaged from the same viewpoint). Stated another way, features should be *relevant* to the scene entities to which they correspond.

Figure 2.4: Similar objects should exhibit similar features. For example, if a feature fires on the eye of the top face it should do so on those of the other too.

<u>Variability due to Absence of Features:</u> Not all objects of an object class need to be composed of the exact same set of parts. For example, not all motorcars have the same number of wheels, headlights or doors, not all human faces are decorated with beards, hair, or eyeglasses, photographs of hands do not always show the same number of fingers. If we design a model which explicitly handles the event of part absence, be it due to occlusion or other causes, then we can limit our representation to the visible portion of the object and have gained a significant advantage over many methods which need to represent such information implicitly.

<u>Variability due to Deformations:</u> When we compare a non-rigid object before and after it has undergone a physical deformation or two geometrically different instances of a class of rigid objects, we find that the *local* differences in appearance are small compared to the *global* change. It might therefore be possible to model the local changes with a simple model, or even to neglect them. The deformation can then be described explicitly by indicating the positions of the local patterns in each image. We would expect that there is an optimal size for the local patterns, since patterns which are very small are not very distinctive and are thus easily confused with other patterns. On the other hand, the assumption that a part is a constant and, therefore, rigid pattern will become less justifiable with increasing part size.

<u>Variability Due to Pose Changes:</u> Deformations of a signal due to changes in pose of the underlying objects can be considered a special case of the geometric deformations just mentioned. If pose changes in three dimensions are over a limited range, it would still be possible to use a simple model to explain the resulting changes in local photometry. The global geometry model might in this case be based on explicit knowledge about the three-dimensional structure of objects as well as allowed transformations, e.g., under a certain camera model.

<u>Variability Due to Background Clutter:</u> We can decompose an entire signal comprised of an object of interest and background clutter into a collection of components. We can then assign individual components to either foreground or background and model their actual appearance independently.

<u>Variability Due to Lighting:</u> When an entire scene undergoes changes in lighting conditions, the brightness patterns across an object often change in a complex fashion. By concentrating on a local neighborhood on the object's surface, we might find changes in lighting conditions that can be accounted for by a much simpler model than would be needed to represent the global change. It might even be possible to eliminate those local changes through simple normalization steps.

Similarity invariant saliency : The key principle underlying the method consists of three steps: (i) Calculation of Shannon entropy of local image attributes (e.g. intensity or colour) over a range of scales-$H_D(s)$; (ii) Select scales at which the entropy over scale function exhibits a peak – s_p; (iii) Calculate the magnitude change of the PDF as a function of scale at each peak – $W_D(s)$. The final saliency is the product of $H_D(s)$ and $W_D(s)$ at each peak. The histogram of pixel values within a circular window of radius s, is used as an estimate of the local PDF. Steps i and iii measure the feature-space and the inter-scale predictability respectively, while step ii selects optimal scales.

The block diagram of unsupervised object recognition is given below:

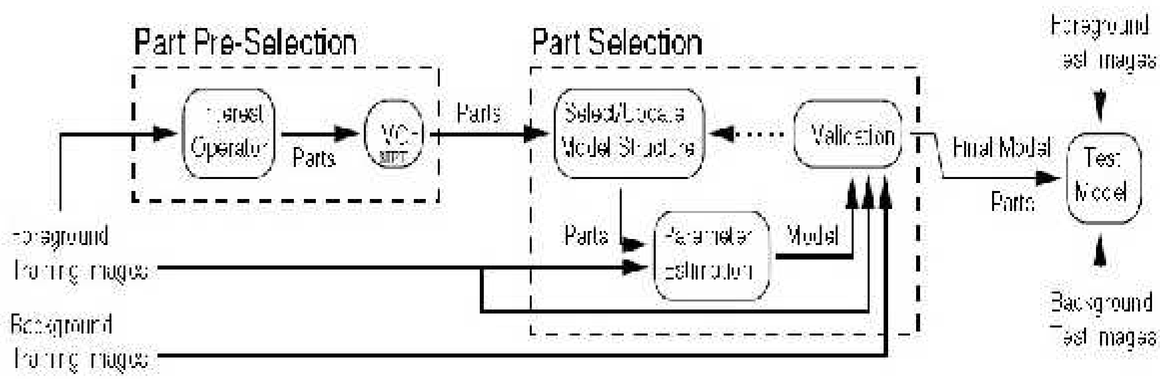

Figure 3: Block diagram of unsupervised object recognition.

Conclusion

In this study we proposed a probabilistic model for simultaneously segmenting and recognizing consistent object or objects parts without the use of human supervision. Our system differs from previous work, which either cascaded or interleaved segmentation and recognition instead of integrating them into a single process. We have presented a novel model for object categories. Our model allows efficient unsupervised learning, bringing the learning time to a few hours for full models and to minutes for variational approximations. The significant reduction in complexity allows to handle many more parts and features than comparable algorithms. The detection performance of our approach compares favorably to the state of the art even when compared to purely discriminative approaches. Also our model is capable of learning the spatial extent of the objects without supervision with good outcomes. Our model is divided into two parts. In the first part, we discuss object models, particularly the *constellation model*, while the second part is dedicated to learning these models. This combination of fast learning and ability to localize is required to tackle challenging problems in computer vision. Among the most interesting applications we see unsupervised segmentation, learning, detection and localization of multiple object categories, deformable objects and objects with varying aspects.

References

[1] http://www.pascal-network.org/challenges/VOC/.

[2] Russell BC, Torralba A, Murphy KP, Freeman WT. LabelMe: a database and web-based tool for image annotation. IJCV. 2008; 77(1–3):157–173.

[3] Winn J, Criminisi A, Minka T. Object categorization by learned universal visualdictionary. In ICCV. 2005; I:756-763.

[4] Viola P, Jones MJ. Robust real-time face detection. Int. J. of Comp. Vision. 2004; 57(2):137–154.

[5] Leibe B, Seemann E, Schiele B. Pedestrian detection in crowded scenes. In Proceedings of the IEEE Conference on Computer Vision and PatternRecognition, 2005.

[6] Mikolajczyk K, Schmid C, Zisserman A. Human detection based on a probabilistic assembly of robust part detectors. In Proceedings of theEuropean Conference on Computer Vision, 2004.

[7] Grauman K, Darrell T. Unsupervised learning of categories from sets of partiallymatching image features. In CVPR, 2006.

[8] Sivic J, Russell BC, Efros AA, Zisserman A, Freeman WT. Discoveringobjects and their location in images. In ICCV, 2005; 370–377.

[9] Russell B, Efros AA, Sivic J, Freeman WT, Zisserman A. Using multiple segmentations to discover objects and their extent in image collections. In CVPR, 2006.

[10] Todorovic S, Ahuja N. Extracting subimages of an unknown category from a set of images. In CVPR, 2006.

[11] Carbonetto P, Dork´o G, Schmid C, Kuck H, de Freitas H. Learning to recognize objects with little supervision. International Journal of Computer Vision, 2008; 77 : 219-237.

[12] Edelman S, Intrator N. Unsupervised statistical learning in vision:computational principles, biological evidence. Miscellaneous Papers, Available from kybele.psych.cornell.edu, 23 April, 2004.

[13] Kadir T, Zisserman A, Brady M. An affine invariant salient regiondetector. Proceedings of the 8th European Conference on Computer Vision, Prague, Czech Republic. 2004.

[14] Kadir T, Brady JM. Scale, saliency and image description. Intl. J. of Computer Vision, 2001; 5(2):83-105.

[15] Kamarainen JK, Hamouz M, Kittler J, Paalanen P, Ilonen J, Drobchenko A. ObjectLocalisation Using Generative Probability Model for Spatial Constellation and Local Image Features. ICCV, pp.1-8, 2007 and IEEE 11th International Conference on Computer Vision, 2007.

[16] Sahbi H, Audibert J, Keriven R. Context-Dependent Kernels for Object Classification. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2011; 33(4) : 699- 708.

[17] Wu W, Yang J. Semi-Supervised Learning of Object Categories from Paired Local Features. ACM International Conference on Image and Video Retrieval, CIVR 2008, Niagra Falls, Canada, 07. July 2008.

[18] Halder A, Pramanik S, Kar A. Dynamic Image Segmentation using Fuzzy C-Means based Genetic Algorithm. International Journal of Computer Applications. 2011; 28(6) : (0975- 8887.

[19] Halder A, Pramanik S. An Unsupervised Dynamic Image Segmentationusing Fuzzy Hopfield Neural Network based Genetic Algorithm. International Journal of Computer Science Issues. 2012; 9(2).

[20] Cho M, Lee J, Lee KM. Feature correspondence and deformable object matching via agglomerative correspondence clustering. Computer vision, 2009 IEEE 12th International Conference on, date of conference: Sept. 29 2009- Oct. 2 2009: 1280-1287.

[21] Felzenszwalb PF, Girshick, McAllester. Discriminatively Trained Deformable Part Models. This project has been supported by the National Science Foundation under grant no. 0534820, 0746569 and 0811340. http://people.cs.uchicago.edu/~pff/latent-release4/.

[22] Kokkinos I. the Project-Teams GALEN.Rapid Deformable Object Detection using Bounding-based Techniques. Informatique/Vision par ordinateur et reconnaissance de formes, RR-7940, ISSN 0249-6399 ISRN INRIA/RR--7940-FR+ENG, research Report no 7940. 2012: 19.

[23] Kashem MA, Akhter MN, Ahmed S, MahbubAlam M. Face Recognition System Based on Principal Component Analysis (PCA) withBack Propagation Neural Networks (BPNN). Canadian Journal on Image Processing and Computer Vision. 2011; 2 (4).

[24] Do T, Kijak E. Face Recognition Using Co-occurrence Histograms of Oiented Gradients. IEEE International Congerence on Acoustics, Speech and Signal Processing, Kyoto International Conference Center, Koyoto, Japan, paper: IVMSP-P9.2,session: Feature Extraction and Analysis, location: poster Area E, presentation time, March 29, 2012, presentation: poster.

[25] Santhosh S, Manjula VS. Face Emotion Analysis Using Gabor FeaturesIn Image Database for Crime Investigation. International Journal of Data Engineering. 2011; 2 (2).

[26] Arumugam D, Purushothaman S. Emotion Classification Using Facial Expression. International Journal of Advanced Computer Science andApplications. 2011; 2(7).