CSE 344/544 Computer Vision
Paper Critique

Kaustav Vats (2016048)
B.Tech, Computer Science and Engineering
Indraprastha Institute of Information Technology
Delhi, India
kaustav16048@iiitd.ac.in

I. Paper Information

1) **Title:** On Detection of Multiple Object Instances using Hough Transforms
2) **Authors:**
   - Olga Barinova *Moscow State University*
   - Victor Lempitsky *University of Oxford*
   - Pushmeet Kohli *Microsoft Research Cambridge*
3) **Year Published:** 2010
4) **Conference Name:** Computer Vision and Pattern Recognition
5) **Link:** https://www.microsoft.com/en-us/research/wp-content/uploads/2016/11/blk$_cvpr2010.pdf$1001[1]

II. Summary

Hough Transform is a very classical feature extraction technique. Hough transform was first suggested as a method for line detection in image. Then is was further extended to be used for low parameter object detection. It converts image to a new representation in Hough space, where each point represent as point of interest and intensity at that point represent confidence of the detection. Hough transform basically splits the input image in different voting elements. Each of these elements votes for different hypothesis. Based on the votes made by large number of elements usually centroid of the object is denoted as area of object. It's not necessarily true that centroid will give exact location of the object, it will give point which lie in area of that object.

Annotating image with location of interest and observing distribution of parameters on training data. This approach for object detection is robust to imaging noise, occlusion etc. Consider voting in hough image, for object with strong vote will also contribute to vote for other objects thus strength of those votes are not inhibited. On considering the flaw of lack in consistent probability model, some non max supression heuristic need to be added. Paper introduced new object detection methods that are similar to hough transform. Some additional important advantages over hough transform- Energy optimization in contrast to heuristic peak location followed by non max supression in vanilla Hough Transform, Probabilistic nature of model make it easy to integrate with other approaches. Downfall with these approach is that, it increases complexity of the algorithm thus making it slow.

To start with probabilistic interpretation of the hough based detection. Denoting each pixel as a voting element. h is denoted as hypothesis about presence of the object in the image. Hough based object detection consider each pixel as independent and finds out which hypothesis object was generated by which pixels. They have tried to solve this problem using binary random variables, which denotes 1 if pixel correspond to object else 0. Votes were considered as probability of occurrence of hypothesis conditioned on the pixel descriptor, which denotes geometric position of the voting element. Each of these conditional probabilities are added and peaks are considered as valid hypothesis. Rather than just summing the votes, authors have used a joint distribution distribution model using random variables in probabilistic ways. After that they determine there values via the Maximum a Posterior (MAP) configuration. Authors also used a greedy procedure for non max supression of the hough image. Then they keep picking hypothesis corresponding to highest peak value until the peak value is greater than some threshold λ. Some of the optimizations were performed to increase computation speed. Authors tested there algorithm firstly on line detection. They used YorkUrbanDB dataset. This dataset contains about 102 urban scene images. Authors have used a very limited number of images for parameter estimation with MAP. They used 20 images for estimation and rest 82 for parameter validation. Authors have derived there own mean of precision based on dataset annotations. They tried matching selected line with ground truth within 2 pixel range. Second measure they used was to count number of non-Manhattan lines which are denoted as error. Where as both these approaches still penalizes the correctly detected lines. Then they considered angle between the lines as well to overcome the problems faced above.

For each image they first mark edges of object using canny edge detector. The considered each pixel as voting element. Based on these voting element and using greedy version of their framework which is explained above, they detected line in the image. For bench marking they have used soft voting scheme followed by identifying local maxima in the hough image and performing non max supression.

They also tested their approach for detecting pedestrian. They used two video sequences of TUD-Campus and TUD-Crossing, which contains mostly profile view of crowded crossing. Since the videos were 50% annotated, they manually annotated rest of the videos and extracted frames from the two videos. To detect objects authors used Hough Forest on 16x16 patches which were extracted from the frames. Authors tested their result with and without non max supression. They observed that with non max supression it could not filter out duplicate detection without filtering out correctly detection as well. Final observation that they have made is that, using there probabilistic framework which is explained above, non max supression can be fully avoided and significant change is observed in accuracy.

III. PROS

- Improvement over a classical technique to solve fast object detection.
- They are able to identify actual peaks with the classical thresholding method. There random variables estimation by MAP has shown great result of object detection.
- They have introduced there new probability frame work for likelihood of the object in certain region. They have used most of the virtues of the Hough transform.
- They have considered each pixel liklihood of the hypothesis and have also detected object in much faster way than classical object detection with hough transform, which has multiple issues like memory bound and parameters to consider for arbitrary shapes.
- Results for both with and without non max supressions were shown. There work showed that non max supression can be avoided if calculating peaks using there probabilistic framework, thus decreasing computation time significantly.
- Results for various approach and images generated during the process are shown in the images, dataset considered actually had the problem they were trying to overcome by these approach.

IV. CONS

- Since the parameter learning is based on distribution observation, there approach won't work if an object is present in image with similar distribution.
- There procedure is very robust to occlusion, deformation, image noises etc.
- Biggest issue is that there's no consistent probability model. This make hypothesis prediction erroneous.
- Log probabilities shouldn't be considered because it gives very low range of value and since thresholding is done on those range then it can affect final detection outcome.
- They have assumed that distribution over hypothesis which is generated by votes is independent of elements.
- There approach cant be used of problem statement is related to counting number of object of interest in the image. Since the result given by there approach is multiple for a single object.
- There approach can detect multiple instance of object of interest. They have also not given any method to clear this problem.
- They have considered various random variable for occurrence of the hypothesis in the image.

V. IMPROVEMENTS

- They have used Maximum a posterior estimation to learn various parameters based on number of training images, but they have not given any theoretical proof that this estimation gives best result. They could have tried different estimation like Maximum likelihood estimation to estimate likelihood of each pixel belong to a hypothesis of object. They should have tried these above method to check various changes in their evaluation metric and could have finalized based on that. Completely discarding methods can effect research idea in a heavy way.
- To detect multiple occurrence of the same object, they could have used some heuristic approach of identifying such object. Solution basically requires to look for nearest peak points which end up detecting same object. Having some heuristic like area of detection or like recursively not counting nearest peak value as hypothesis. Based on the heuristic if distance between two peak point is less than some threshold and they belong to same hypothesis then discard the points with lower value of peak. An assumption can be made that close points belong to same object of interest. Taking a small size kernel could also help in reducing number of overlaps. Kernel for all high peak points won't affect computation much, but this depends on the number of points detected in the image.
- Authors have only shown result for one prior calculation approach. This would affect there results. It can also be tried by various probability distribution and evaluating based on the new distributions for all priors of hypothesis. There work can be further extended in image segmentation using hough transform. There approach give object of interest, which can be extracted using segmentation.

REFERENCES

[1] Multiple object detection using Hough transform followed by non max supression or mode of seeking in order to locate and distinguish peak in hough image. Paper Link: https://www.microsoft.com/en-us/research/publication/detection-multiple-object-instances-using-hough-transforms-2/?from=http