Decision level data fusion in Speech and Image Recognition Systems

Gaurav Sanjay Newalkar (12EC43),
Nithin Rao Koluguri(12EC51),
Nikhil Lunavath(12EC68)

Project Guide: Dr.Ashvini Chaturvedi
Dept of E & C Engg

Jan - May 2016

Objectives

- To recognize 4 different words through automatic speech recognition.
- BASED ON THE RECOGNIZED WORD, A IMAGE RECOGNITION SYSTEM TO BE DEVELOPED TO RECOGNIZE THE OBJECT PERTAINING TO THE COMMAND GIVEN.

Confusion Matrix

Words/Digits	1	2	3	4
1	37	2	0	1
2	1	39	0	0
3	2	0	37	1
4	2	0	1	37

- We have achieved a total of 98.33 percentage accuracy on training data and 93.75 percentage accuracy on test data samples.
  Edit: We changed the test data with different speakers.

Objective 2- Approaches

- Correlation Method
- Supervised Learning Method

Correlation Method

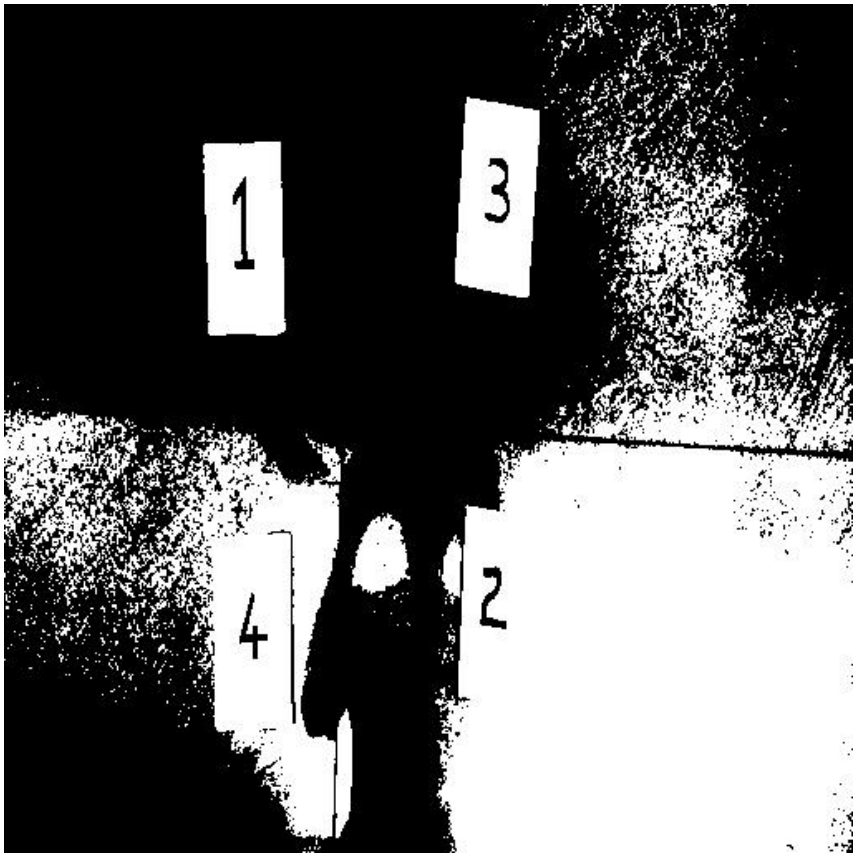

Figure: Image taken in room

Correlation Method Disadvantages

- Prototype image has to traverse through all the pixels and has to do cross correlation at each step.
- Time consuming.

Method-2

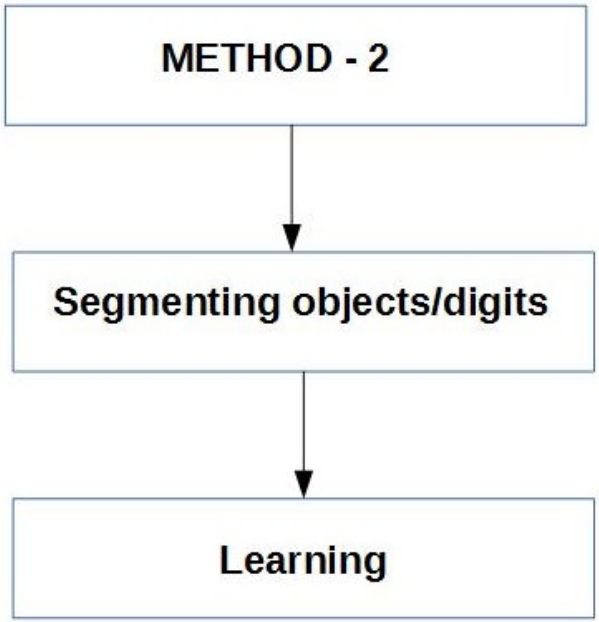

Figure: Flow of Method-2

Segmenting objects

- Separate out each object/digit.
- Attach/tag the coordinates of each segmented digit.(test images)

Learning

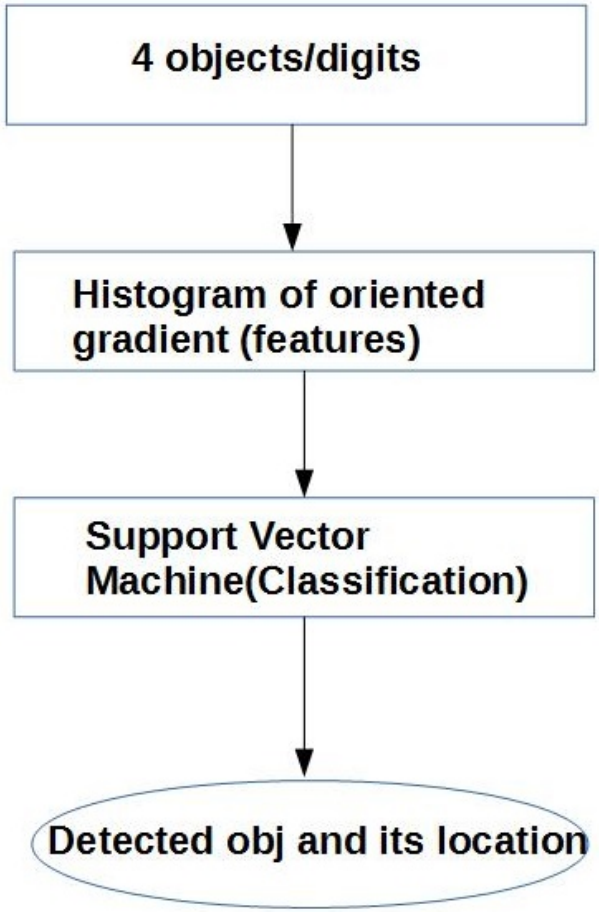

Figure: Flow of Learning

Why HOG and SVM?

- One significant advantage of SVMs over neural nets is the relative ease and speed during training
- The oriented histogram based features significantly outperform raw pixel features when the number of training examples are small.
- we get an approach which is the best in terms of all three criteria:accuracy,computation time at training and computation time at testing.
  Source : Fast and Accurate Digit Classification by Maji,UCB.

References

N. Dalal and B. Triggs, *" Histograms of Oriented Gradients for Human Detection",Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 886-893*,2005

LeCun, Y., Bottou, L., Bengio, Y., and Haffner. P *"Gradient-based learning applied to document recognition." Proceedings of the IEEE, 86, 2278-2324,* 1998

Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, A.Y. Ng, *"Reading Digits in Natural Images with Unsupervised Feature Learning"*, NIPS Workshop on Deep Learning and Unsupervised Feature Learning 2011.

Subransu Maji,Jithendra Malik , *"Fast and Accurate Digit Classification"*, University of California at Berkeley.

Thank You !!!