

SPEAKER IDENTIFICATION

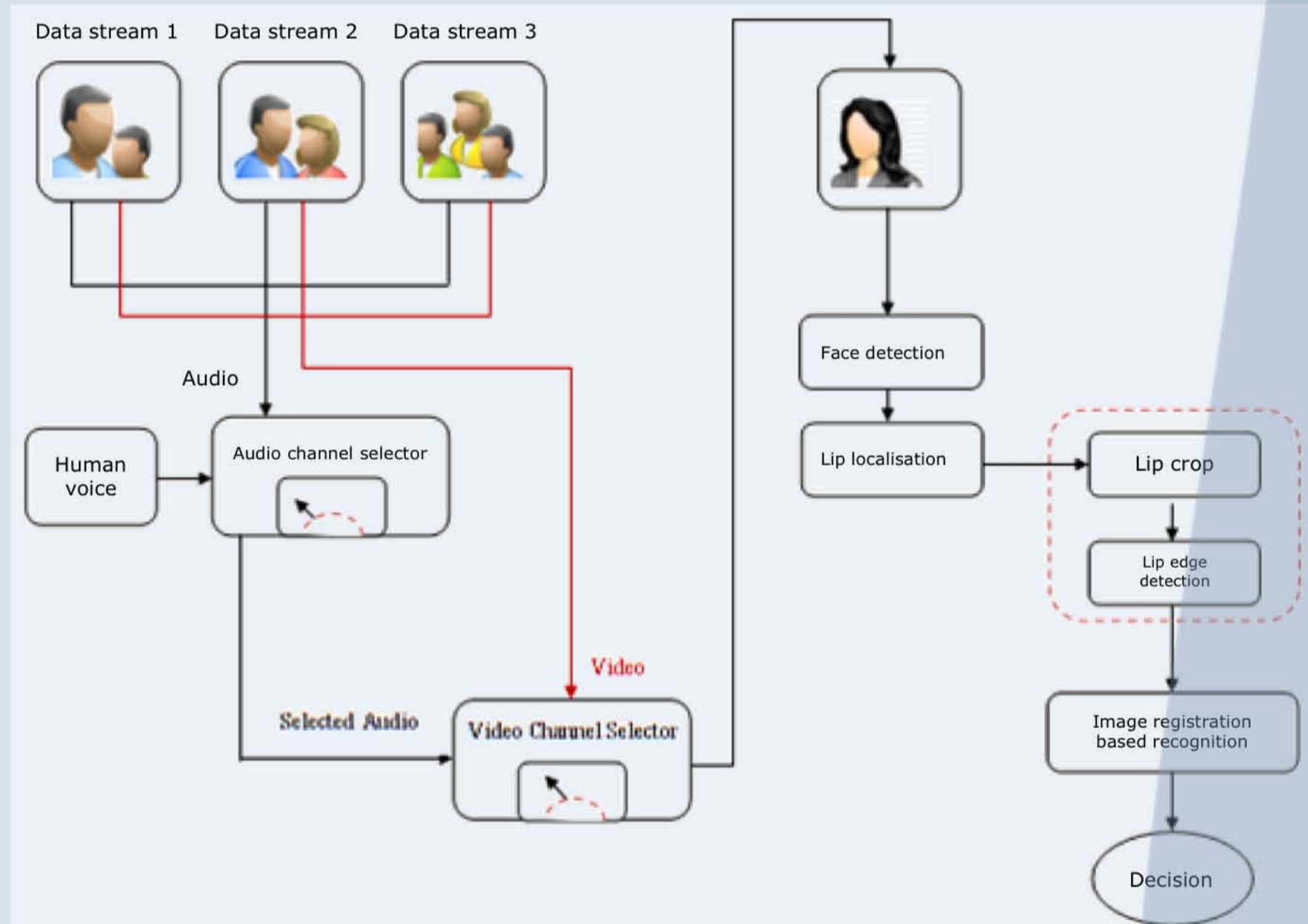
FOR VIDEO CONFERENCE

S.Saravi, I.Zafar, E.A.Edirisinghe



INTRODUCTION

Automatic speaker identification in videoconferencing will allow conference attendees to focus on the conference rather than manually identifying the active channel and the speaker. In this work we present a real-time, audio-video based approach to address this problem, but focus more on the video analysis side. The initial stage consists of a face detector, followed by a lip-localization algorithm. A novel approach for lip movement detection based on *image registration* and using the *Coherent Point Drift (CPD)* algorithm is proposed.



PROPOSED SYSTEM

Step 1: Face detection

For this purpose the OpenCV face detector is used

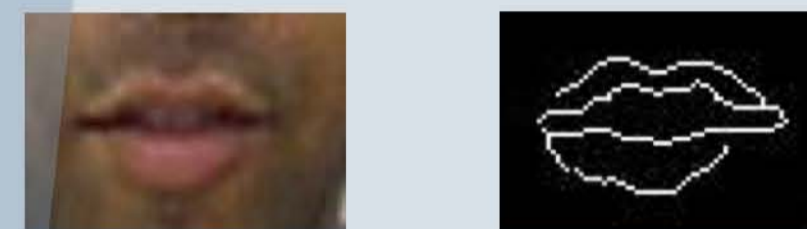
- The problem is false and undesirable object detections.
- As a solution, first step is to get the position of the biggest rectangle



(not necessarily the face) and next step is to compare its position with that of the rectangle from the previous frame; set to the previous rectangle's position if necessary.

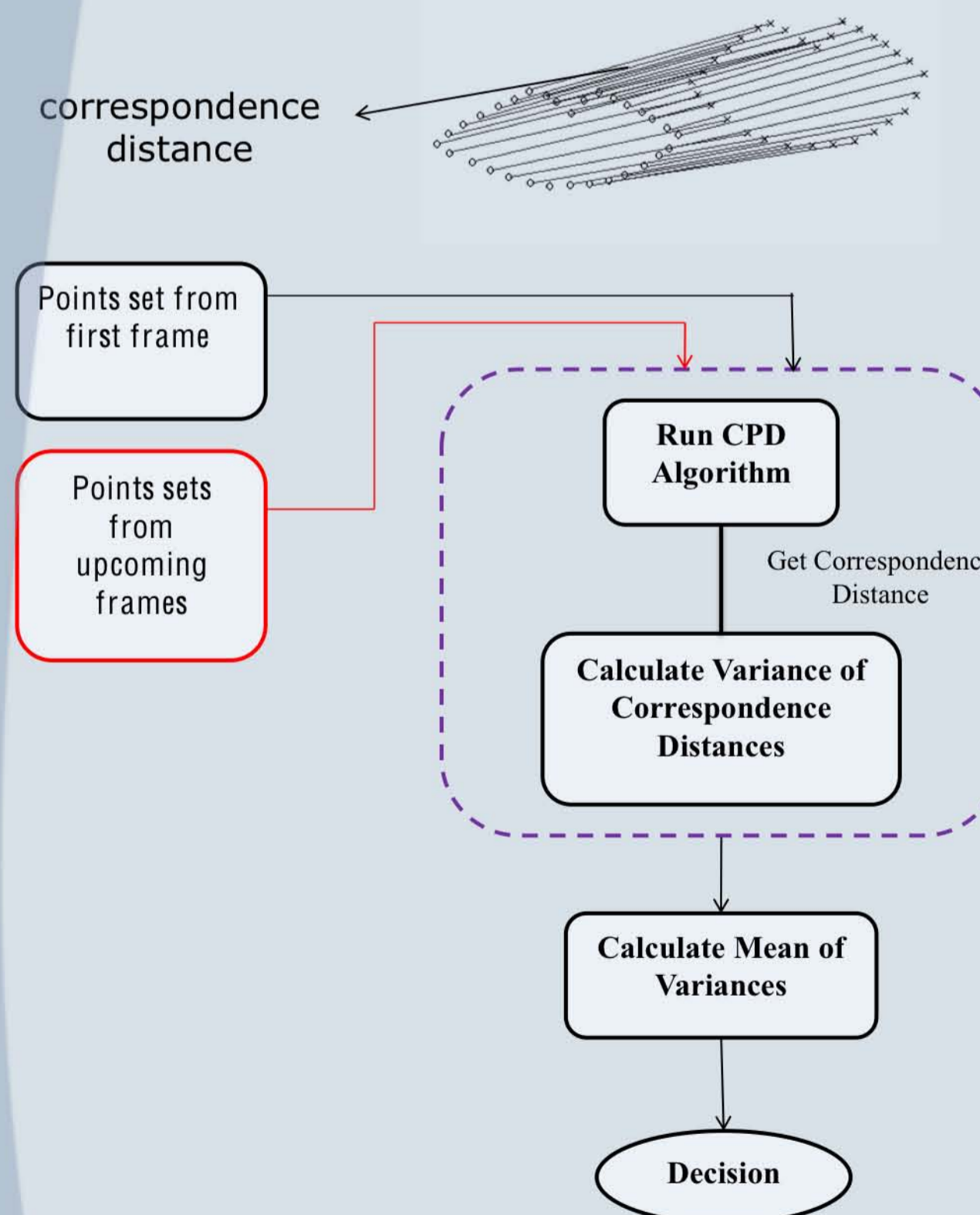
Step 2: Lip Localisation

A closer crop of the lips results in easier edge detection. Lips are located in the bottom third and middle two quarters of face. Then Canny edge detection is used to detect lip corners point by point. These points are saved as data sets for further use.



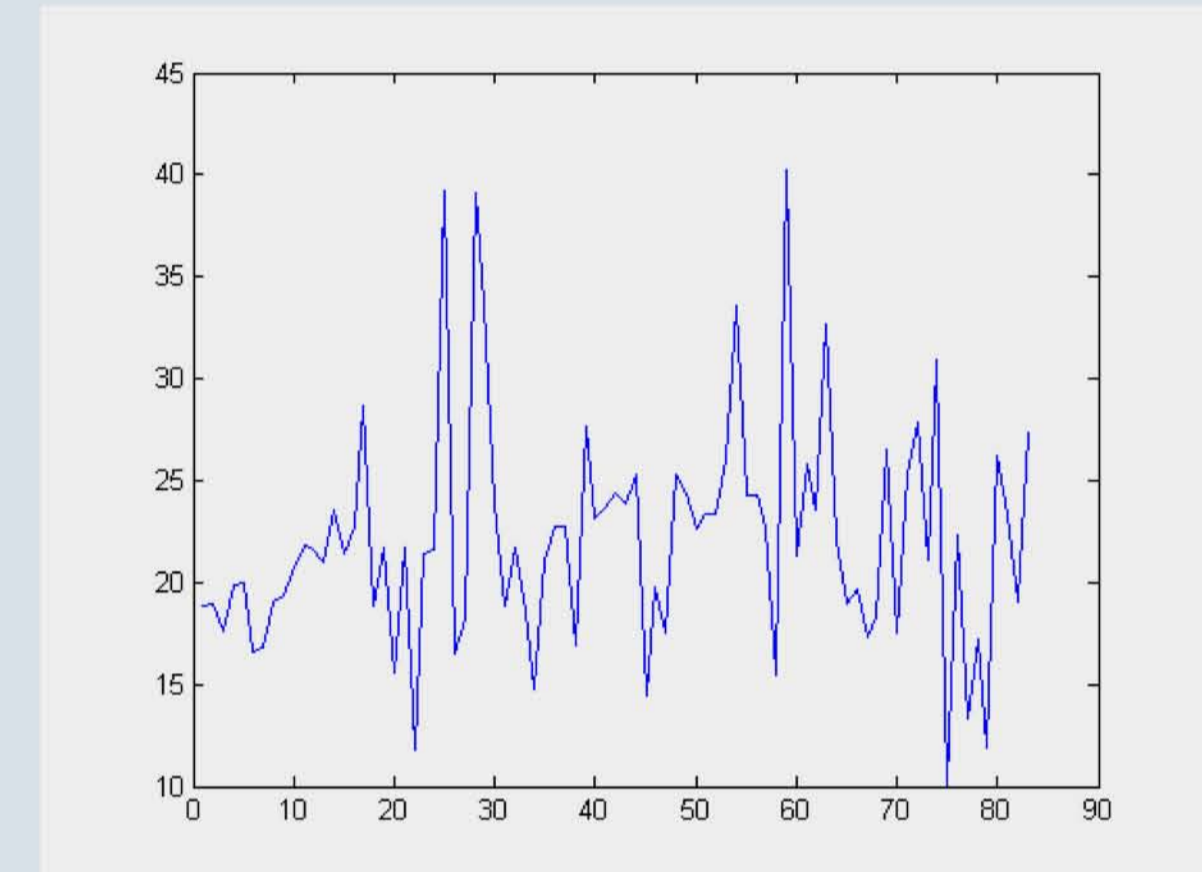
Step 3: Lip movement detection

Proposed approach for lip movement detection is based on image registration. Usage of this method is in evaluation or combination of the data acquired from different measurements. Coherent Point Drift (CPD) is a new technique for registration of point sets. The idea is to move one point set coherently to align with the other set and measure the change of shape using the variance of correspondence distance.



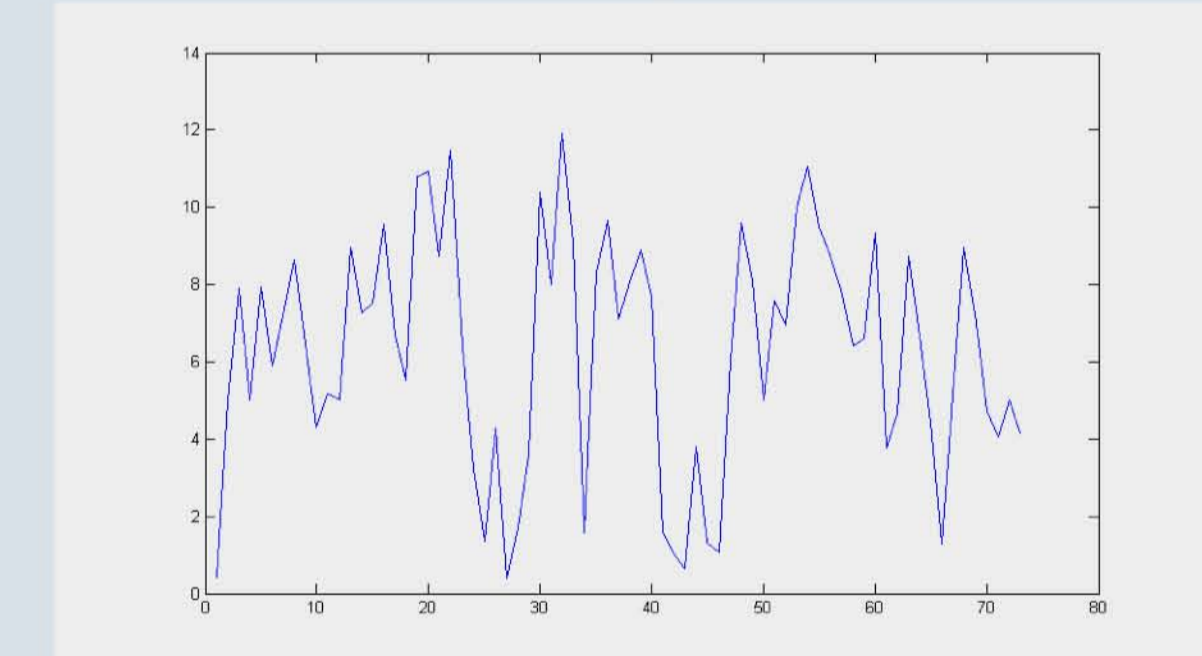
By calculating mean of variances it is obvious that average below 15 belongs to non talking and above 15 belongs to talking face.

EXPERIMENTAL RESULTS



Results of a talking human video

Mean of Variance of Corresponding Distance: 22.1025



Results of a not talking human video

Mean of Variance of Corresponding Distance: 6.2881

CONCLUSION

1. Improvement of the OpenCV face detector
2. Design and Implementation of an algorithm for lip localisation
3. Modification and use of CPD algorithm in lip movement detection

REFERENCES

1. Andriy Myronenko. 2006. Coherent Point Drift (CPD) [Online] Available at: <http://www.bme.ggi.edu/~myron/matlab/cpd/>
2. A.Myronenko, X. Song, M. A. Carreira-Perpiñan: Non-rigid point set registration: Coherent Point Drift, NIPS'19.
3. Andriy Myronenko, Xubo Song, and Miguel Carreira-Perpinan. Non-rigid point set registration: Coherent point drift. In B. Schölkopf, J. Platt, and T. Ho, editors, Advances in Neural Information Processing Systems (NIPS) 19.
4. OpenCVWiki. 2009. FullOpenCVWiki. [Online] Available at: <http://opencv.willowgarage.com/wiki/FullOpenCVWiki>