IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 2, April 2014.

www.ijiset.com

ISSN 2348 - 7968

Frontal Face Clustering In Video Surveillance

S.G.Kartheeswari[1],Ms.B.Vijaya Lakshmi[2] M.E(Ph.D),
[1]Final Year Research Student, [2]Associate Professor,Department of ECE, K.L.N.College of Information Technology, Madurai,Tamilnadu, India.

Abstract - **In the video sequence human faces have unlimited orientations and positions so face detection and clustering is very important.Clustering results are used for security, criminal records or identification, verification of person, etc.In this paper,we proposed a method to cluster human faces from the video sequence based on hidden markov method.First we used a face detector to localize faces in all frames of the video and extract various features of the detections. Feature extraction is a useful technique for recognizing faces through systems. Face detection is done by using viola jones detector and feature extraction is done by means of HOG algorithm.Optical flow estimation is used to compute the movement between any of two frames.In optical flow estimation pyramidal version of lucas kanade algorithm is used. This estimation is used to compute a dissimilarity matrix such as appearance dissimilarity,space time dissimilarity. Clustering tracks the similar faces from the video sequence faces may have,different image orientation or it may occluded by objects. Finally, an optimization method involving clustering of faces this is used for face recognition.**

*Index Terms:**Face Detection,Optical Flow,Clustering,***

Tracking,Dissimilarity.

I.INTRODUCTION

Face detection is required as the first step of the automatic face analysis system.This work has been widely investigated in recent years because it overlies many areas of application: face recognition, man-machine interaction systems, visual communication systems, video-surveillance, etc. However face detection is a challenging task due to variation in illumination, variability in scale, location, orientation and pose,Facial expression, occlusion and lighting conditions also change the overall appearance of face. In this paper face is detected by means of viola-jones algorithm and then feature extraction is done which is used to cluster the faces.Optical flow estimation is calculated between any of two frames which has been extracted from the video.In optical flow method pyramidal version of the Lucas–Kanade algorithm to represent both small and large displacements.From the above result output dissimilarity matrix is calculated and then Clustering is performed based on the feature extraction.

II.RELATED WORK

Many novel methods have been proposed to resolve Face detection and clustering. For example, the template-matching methods are used for face localization and detection by computing the correlation of an input image to a standard face pattern. The feature invariant approaches are used for feature detection of eyes, mouth, ears, nose, etc. The appearance-based methods are used for face detection with eigenface, neural network , and information theoretical approach.Nevertheless, implementing the methods altogether is still a great challenge. First, the all the faces are vertical and have frontal view, second, they are under almost the same illuminate condition. Face pattern detection discriminates and localizes the face within the identified body parts.Faces and bodies of users are tracked over several temporal scales: short-term (user stays within the field of view), medium-term (user exits/reenters within minutes), and long term (user returns after hours or days). Short-term tracking is performed using simple region position and size correspondences,while medium and long-term tracking are based on statistics of user appearance.Various algorithms are for face detection.I proposed a method to cluster-specific object detections of a video sequence, which we applied to face detections. Due to the number of detections extracted from video, an automatic clustering of face detections is interesting for visual surveillance applications. For archive browsing or for face tagging on videos, it is easier to investigate with an album of faces than with a set of all the detected faces.Our effort focus on real visual surveillance constraints: cluttered scenes, uncontrolled, and containing multiple small faces.

III.PROPOSED WORK

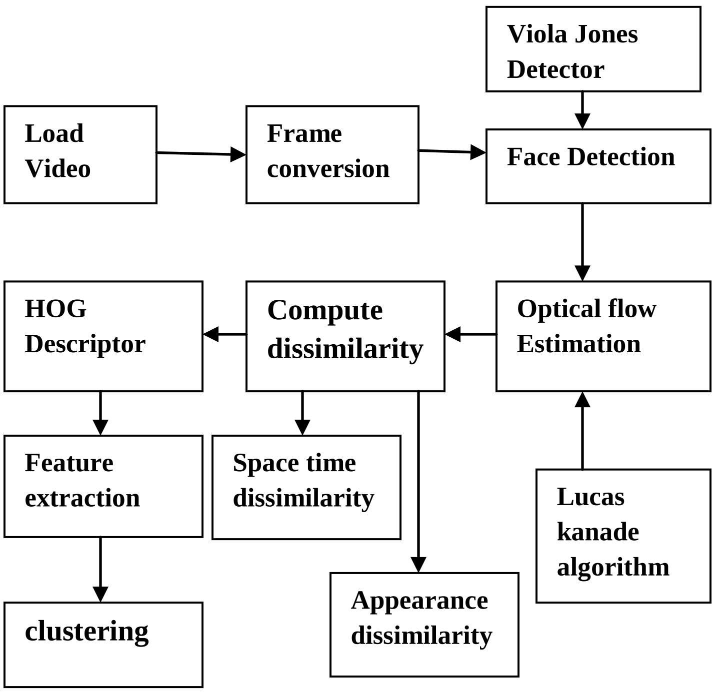


Figure.1.Block diagram of proposed system

The above block diagram shows the flow of the process what we did in this project and the process are explained in the following modules.

1.Load Video

Load the video using the avi reader method.Read Audio/Video Interleaved (AVI) file.Video is taken from the database.
Database:www.eecs.qmul.ac.uk/andrea/avss2007_d.html.

2. Frame Conversion

In this module, Loaded video is convert into frames.we have extracted upto 130 frames and we have shown first fifteen frames out of 130 frames.Extracted frames are stored in workspace as .jpg image format.

3.Face Detection: Mutiple Faces present in the video sequence is detected by means of viola-jones algorithm.

3.1 Viola-Jones's Detector: This framework, designed for rapid object detection, is based on the idea of a boosted cascade of weak classifiers [1] but extends the original feature set and provides different boosting variants for learning [2].The cascade learning algorithm is similar to decision tree learning. Essentially, a classifier cascade can be seen as a degenerated decision tree. For each stage in the cascade a separate subclassifier is trained to detect almost all target objects while rejecting a certain fraction of the non-object patterns. The resulting detection rate, *D*, and the false positive rate, *F*, of the cascade is given by the combination of each single stage classifier rates:

$$D = \prod_{i=1}^{K} d_i \qquad F = \prod_{i=1}^{K} f_i \qquad (1)$$

Under this approach, given a 20 stage detector designed for refusing at each stage 50% of the non-object patterns (target false positive rate) while falsely eliminating only 0.1% of the object patterns (target detection rate), its expected overall detection rate is 0.99920 ≈ 0.98 with a false positive rate of 0.520 ≈ 0.9* 10^{-6}. In a single stage classifier one would normally accept false negatives in order to reduce the false positive rate classifiers are combined. The cascaded classifier is composed of stages each containing a strong classifier.

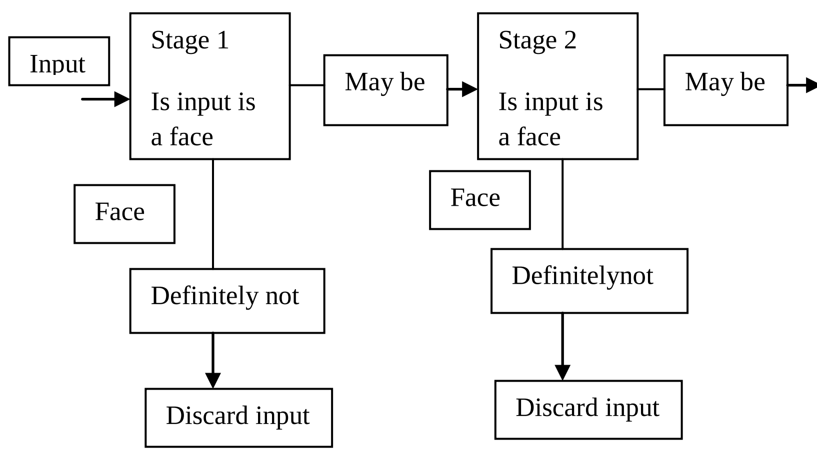


Figure.2.Cascade classifier

4.Feature extraction:

Feature extraction from the video is done by using HOG (Histogram of Oriented Gradient) algorithm. This technique counts occurrences of gradient orientation in localized portions of an image. The HOG descriptor maintains a few key advantages over other descriptor methods.The HOG descriptor operates on localized cells, the method upholds invariance to geometric and photometric transformations,except for object orientation. Such changes would only appear in larger spatial regions.The implementation of these descriptors is achieved by dividing the image into small connected regions called cells, and for each cell compiling a histogram of gradient directions or edge orientations for the pixels within the cell. The combination of these histograms represents the descriptor.For improved accuracy, the local histograms can be contrast-normalized by calculating a measure of the intensity across a larger region of the image called a block, and then using this value to normalize all cells within the block. This normalization results in better invariance to changes in illumination or shadowing. This method filters the color or intensity data of the image with the following filter kernels [-1 0 1] $[-1\ 0\ 1]^T$ .Then the normalization factor f is as follows, Let υ be the block to be normalized and e be a small constant.

$$f = \frac{v}{\|v\| + e}$$

In feature extraction 380 x 81(rows x col) features are exracted.

5.optical Flow

There are many methods to compute an optical flow between two frames, one of the main issues is to represent a large scale of displacements.In this paper, we used a pyramidal version of the Lucas– Kanade algorithm[4] to represent both small and large displacements. In each frame having a detection, we compute the optical flow from the previous to the current frame and from the current to the next frame, then these two optical flows are averaged. The resulting speed vector of detection is obtained by taking the most representative flow vector in the detection area.

6.compute dissimilarities

In this paper I calculated appearance based dissimilarity and space time based dissimilarity.Finally output dissimilarity is calculated by using above two dissimilarity.

Appearance dissimilarity: Detection appearance is represented by an HS-V histogram.This histogram is the concatenation of a 2D HS(Hue Saturation) histogram and a 1D V(Value) histogram of image pixels,where H, S, and V represent hue, saturation, and value of a color, respectively). If the S and V values are large enough for a pixel, the pixel is counted in the HS histogram, or else it is counted in the V histogram.To measure the dissimilarity between two HS-V histograms, we used the Bhattacharyya coefficient.

Space–time dissimilarity: Space time dissimilarity is almost similar to optical flow estimation which measure the trackelet time difference between the previous frame and in frame.

7.clustering

Face clustering is based on face recognition or individual identification.Clustering is performed based on feature extraction.This method is used to cluster the faces among multiple faces presented in the video.Faces may have different position and orientation this method effectively clusters the faces.Feature extraction output act as a database then based on the test data it cluster the similar faces among multiple faces.Individual face is recognized by comparing between their feature data and the ones on a database.In this project we used Hidden Markov Random Process clustering method.

Hidden Markov Random Fields: Hidden Markov Random Fields is a commonly used generative model.It is defined based on two assumptions:(a) given the latent variables Y , the observed variables X are independent, i.e. $P(X|Y) = \prod_{i=1}^{n} p(\frac{xi}{yi})$ (b) given the observed variables X, Y constitute a Markov network. The correlations between Y are embedded by a neighborhood system.In this paper, we adopted an approximate method called simulated field algorithm. Its main idea is: when treating a particular latent variable y_i, ignoring the fluctuations of its neighbors, through fixing the states of the neighbors as a result, the overall computation reduces to deal with independent variables.

IV.EXPERIMENTAL RESULTS

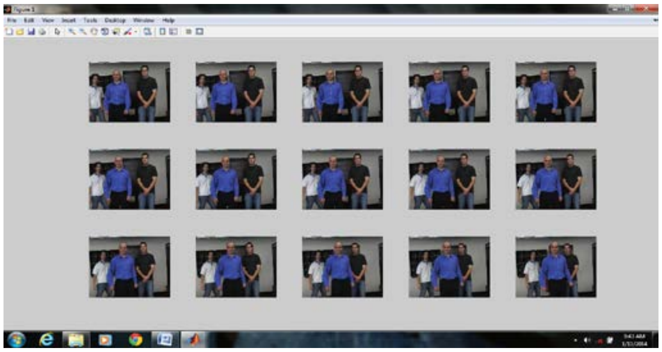


Figure.3. Extracted frame from video sequence

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 2, April 2014.
www.ijiset.com
ISSN 2348 - 7968

From the video sequence frames are extracted and we shown the first fiftenn frames in the above figure.

Figure.4.Detected faces

From the video sequence multiple faces are detected by means of viola jones detector and the Face region is indicated by means of Bounding box.

Figure.5.Frames in HSV color space

The above result shows that frames in HSV(Hues,Saturation,Value) color space.This result also used to calculate Appearance dissimilarity.

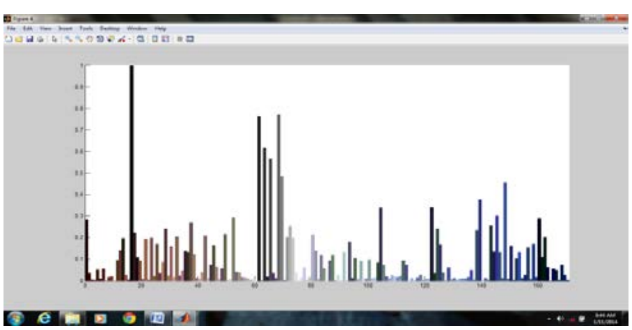
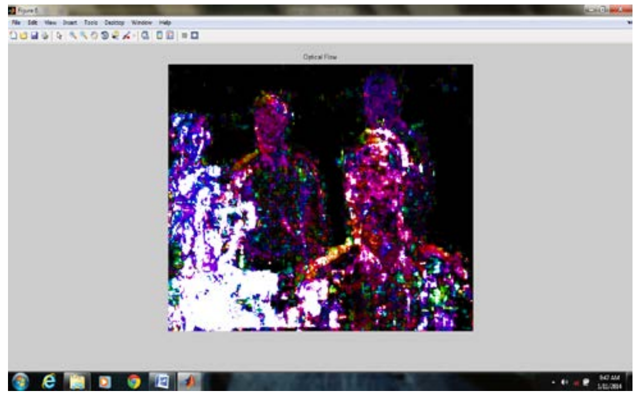


Figure.6.Color histogram

Color histogram result shows the distribution of color



density of faces.

Figure.7.optical flow

Optical flow results correlate two frames and we can recognize five faces this shows that movement dissimilarity of frames.

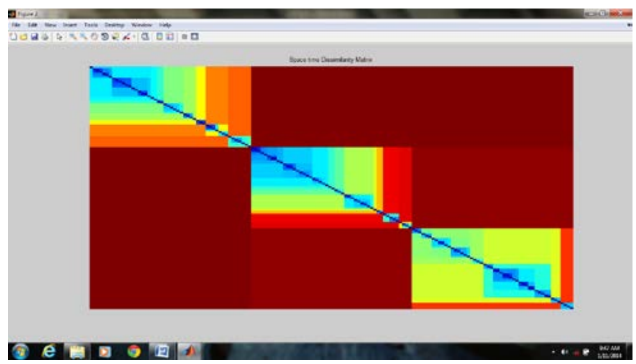
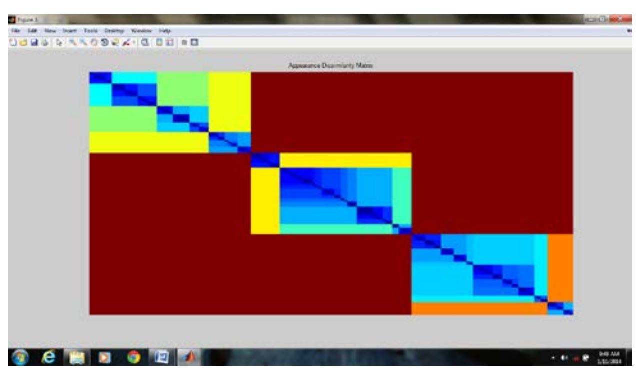


Figure.8.Space Time dissimilarity

Space time dissimilarity indicates how far the persons are located in the video.



ISSN 2348 - 7968

Figure.9.Appearance dissimilarity

Appearance dissimilarity result shows that how far the appearance of faces differs in the video.

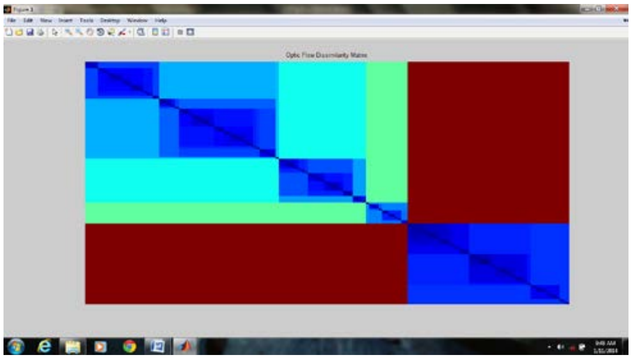


Figure.10.Optical flow dissimilarity matrix

Optical flow results are averaged and represented by means of dissimilarity matrix.

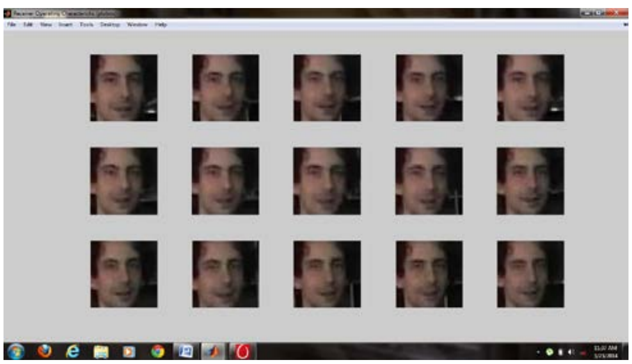


Figure.11.Clustered faces

Above result shows that person1 is recognized among other two persons from the indexed label values we shown the first fifteen frames and remaining clustered faces are stores in a work space.

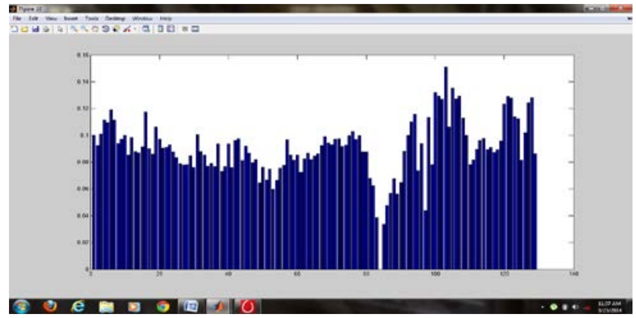


Figure.12.comparison graph

This result shows that how the predicated face differ from the training data set.

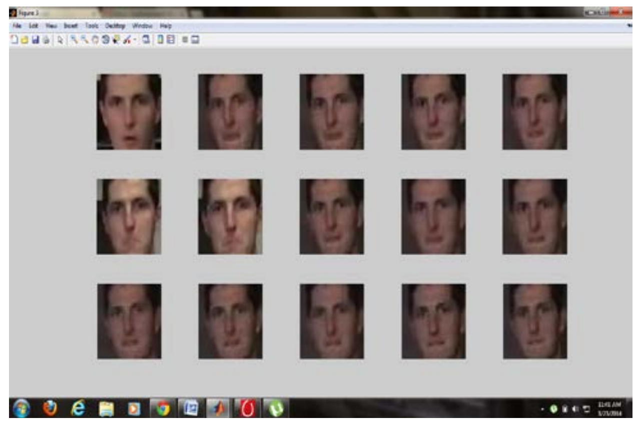


Figure.13.Clustered faces

Above result shows that person2 is recognized among other two persons.

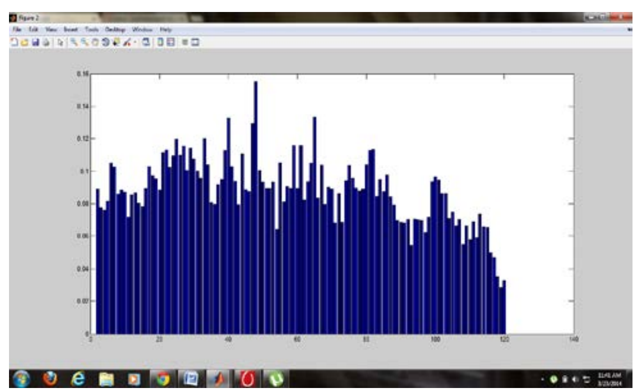


Figure.14.comparison graph

ISSN 2348 - 7968

Figure.15.Clustered faces

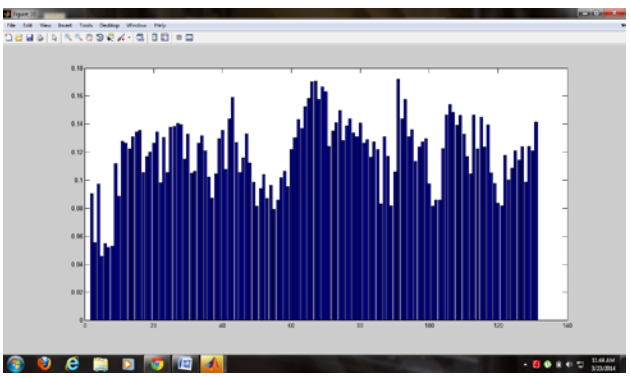


Figure.16.comparison graph

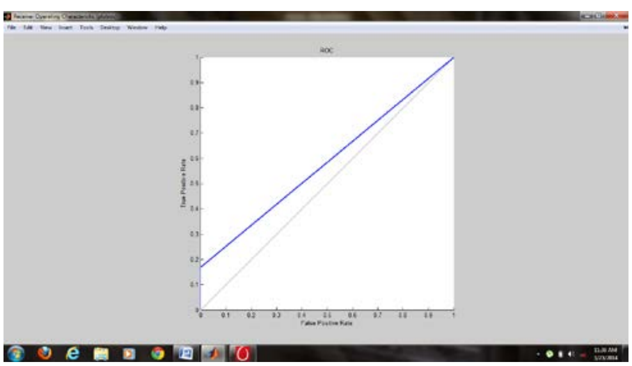


Figure.17.Performance evaluation test result.

The above result is known as a Receiver Operating Characteristic curve.It is a plot of the true positive rate(sensitivity)against the false positive rate(specificity).Prediction point in the upper left corner or coordinate (0,1) of the ROC space, representing 100% sensitivity (no false negatives) and 100% specificity (no false positives). The (0,1) point is also called a perfect classification. The diagonal divides the ROC space.Points above the diagonal represent good classification results,points below the line poor results.

V.CONCLUSION

We proposed a method to cluster faces in video sequences.Face detection is done by using viola-jones detector then feature extraction is done by using HOG descriptor.Optical flow estimation is calculated by using pyramidal lucas-kanade algorithm.Optical flow estimation,appearance dissimilarity,time dissimilarity is used to compute the difference between the frames. Feature extraction is used to extract the faces from the video sequence from that result clustering is performed.Faces are clustered by means of hideen markov clustering method.

REFERENCES:

[1] Paul Viola, Michael J. Jones, and Daniel Snow.**Detecting pedestrians using patterns of motion and appearance**..In *Proc. of the International Conference on Computer Vision*, October 2003.

[2] Rainer Lienhart, Alexander Kuranov, and Vadim Pisarevsky,**Empirical analysis of detection cascades of boosted classifiers for rapid object detection**.In *DAGM'03*, Magdeburg, Germany, September 2003.

[3] M Kim, S Kumar, V Pavlovic, H Rowley, ***Face tracking and recognition with visual*** **constraints in real-world videos**.In *IEEE Conference on Computer Vision and Pattern Recognition* (Anchorage AK, 23-28 June 2008)

[4] J Marzat, Y Dumortier, A Ducrot, ***Real-time dense and accurate parallel optical*** **flow using cuda**.In *Proceedings of The 17th InternationalConference in Central Europe on Computer Graphics, Visualization andComputer Vision (WSCG).* (WSCG, 2009). [http://www.wscg.eu/]

[5] S Dubuisson, J Fabrizio**, Optimal recursive clustering of likelihood functions for multiple object tracking**,In *Elsevier Pattern Recognition Letters , vol.30* (Elsevier, 2009), pp. 606–614. [http://www.elsevier.com]

[6] SimeonSchwab,Thierry Chateau, Christophe Blanc and Laurent Trassoudaine, Schwab, **A multi-cue spatio-temporal framework for automatic frontal face clustering in video Sequences**, EURASIP Journal on Image and Video Processing 2013, 2013:10

[7] B Benfold, I Reid**, Stable multi-target tracking in real-time surveillance video**. in *IEEE Conference on*

Computer Vision and Pattern Recognition.(Oxford, 20-25 June 2011),pp. 3457–3464

[8] L Zhang, Y Li, R Nevatia, **Global data association for multi-object tracking using network flows,** in *IEEE Conference on Computer Vision and Pattern Recognition*. (Los Angeles, 23-28 June 2008)

[9] Z Lu, MA Carreira-Perpin**,Constrained spectral clustering through affinity propagation** in *IEEE International Conference on Computer Vision*.(Portland, 23-28 June 2008

[10] T Brox, C Bregler, J Malik**, Large displacement optical flow.** in *IEEE Conference on Computer Vision and Pattern Recognition*. (Berkeley, 20-25June2009),pp. 41–48