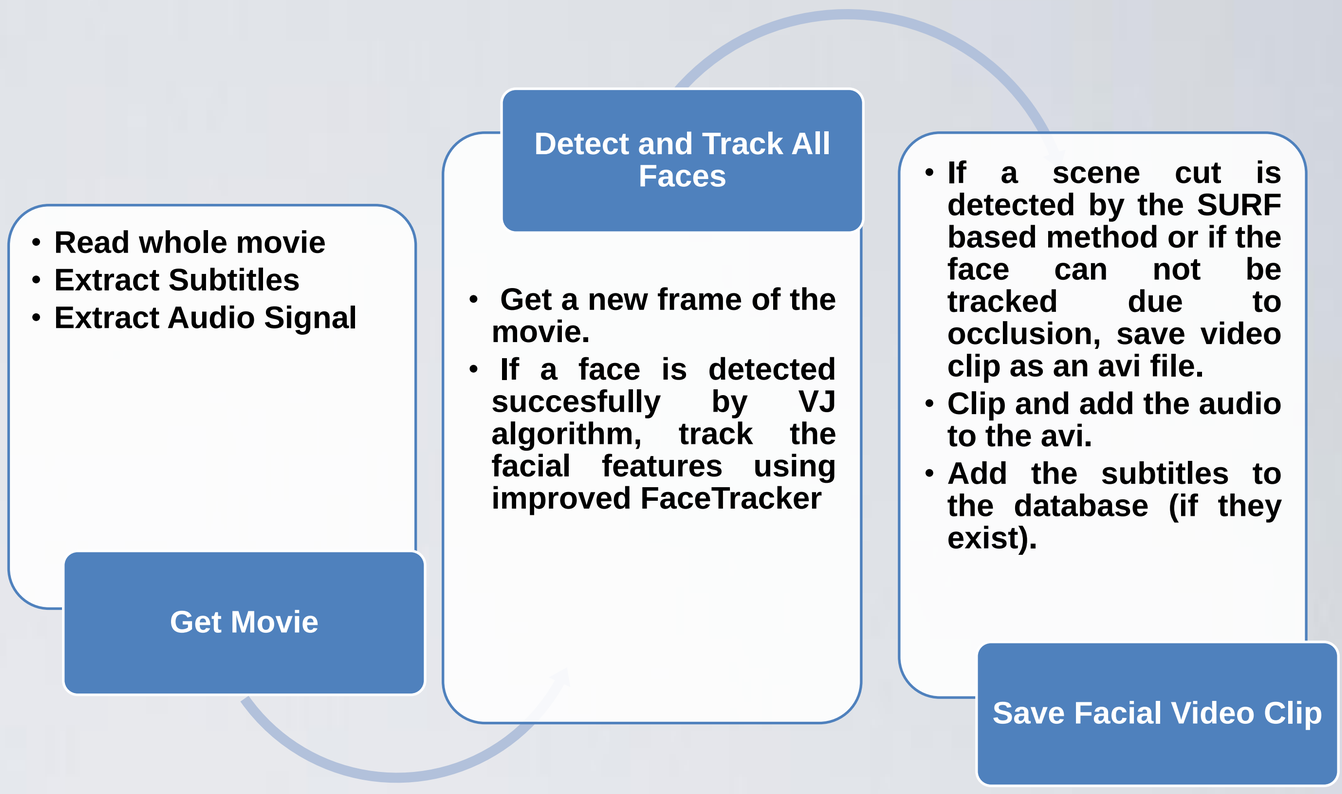
Automatic Extraction of Affective Face Videos

Can Kansın, Çiğdem Eroğlu Erdem,
Department of Electrical and Electronics Engineering, BAHÇEŞEHİR University
{cigdem.eroglu, can.kansin}@bahcesehir.edu.tr

Abstract

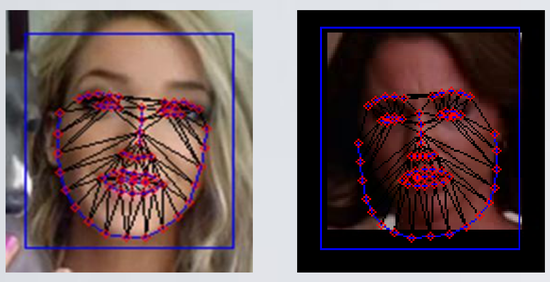
Access to naturalistic and high quality databases is essential to conduct research on affective computing including facial expression and mental state recognition from audio-visual data. Most databases that are available today have been recorded under controlled laboratory environments, and are mostly acted. We present an automatic database creation tool that can extract video clips from DVDs and TV programs containing the facial expression of actors in semi-naturalistic situations.

Method

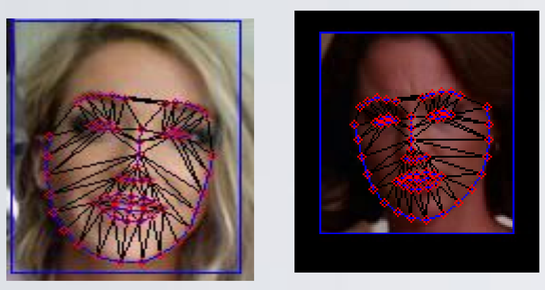


Face Detection and Face Tracking

Landmarking in first frame is incorrect if face detector uses rectangle with wrong size

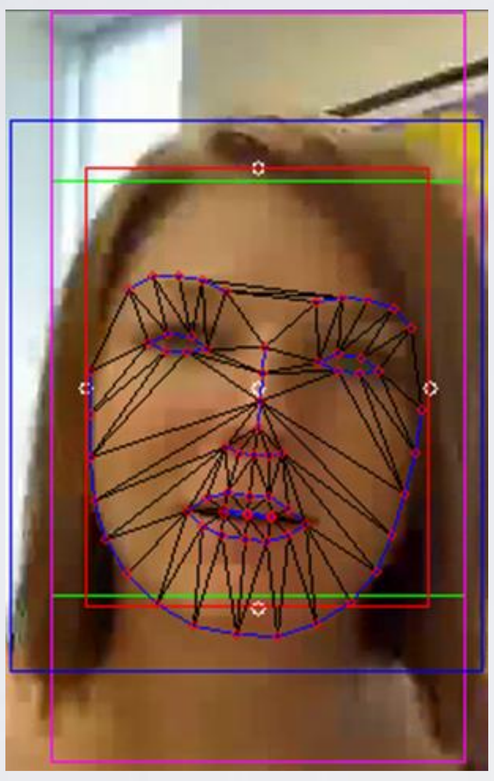


Landmarking problem solved by Enlarging the face detection window using skin color filtering.



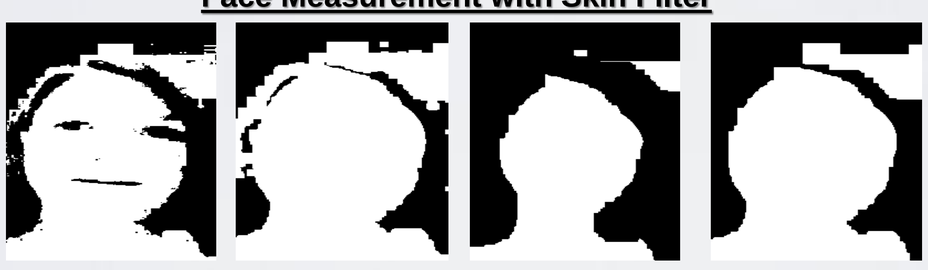
Skin color filter

$$\frac{\sum_{i=0}^{I_n} C(I_{i_{RGB}})}{\sum_{i=0}^{I_n} 1} \geq r$$



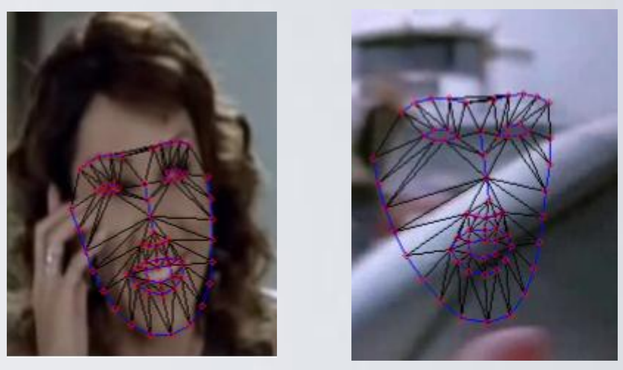
The green rectangle shows the face detection result of the Viola-Jones algorithm. Since it does not include the chin, it will problematic for the FaceTracker. The green window is extended. We enlarge it to the pink rectangle with configurable parameters. After we obtained pink rectangle we reduce it to red rectangle using skin color information. We do these operations to avoid FaceTracker's errors such as not finding eyebrows and chin.

Face Measurement with Skin Filter



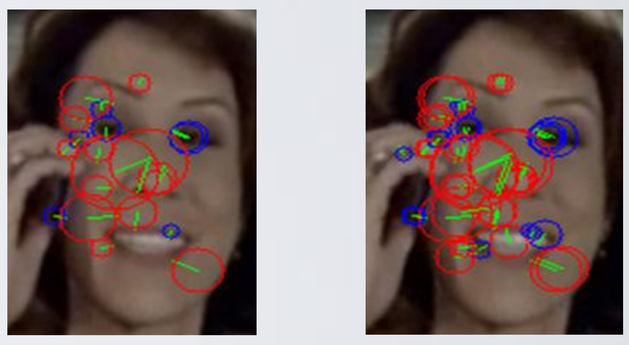
Skin Filtered | Dilate(3) | Erode(13) | Dilate(11)

Face Tracking problem because of the FaceTracker continues to track although scene change occurs



Frame 326 | Frame 327

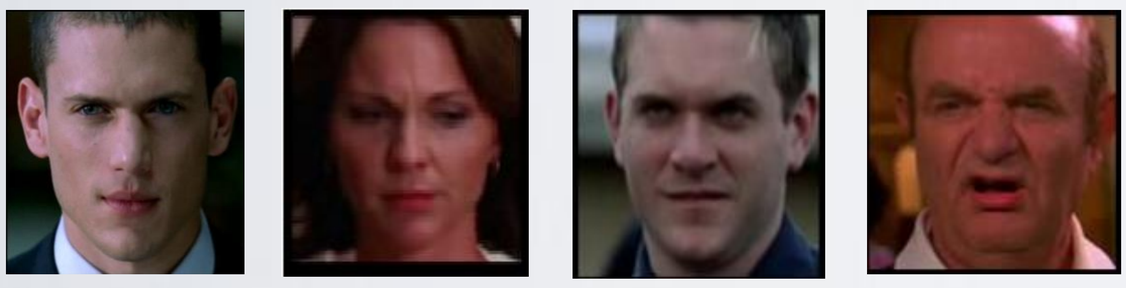
We accept the tracking result in the current frame if Most of the SURF features in that frame and the previous frame matches



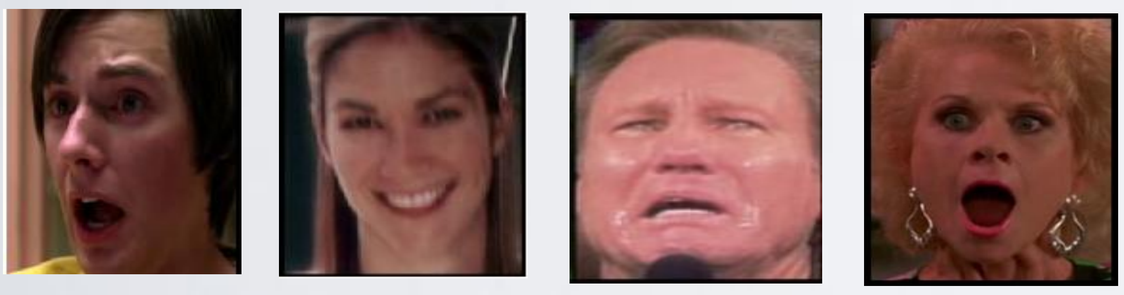
Features from previous frame | (Un)Matched features with current frame

Results and Conclusion

Examples of the accepted video clips containing the eight emotions



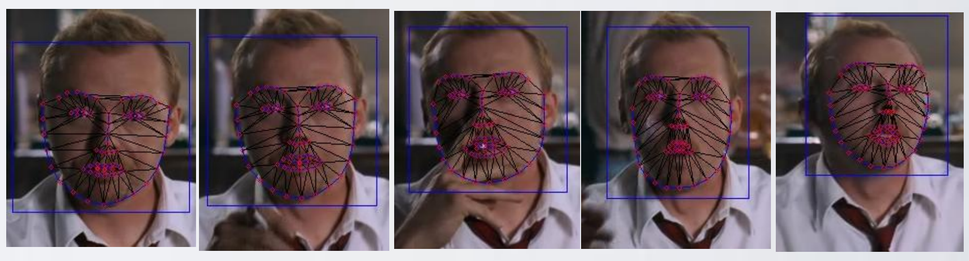
Neutral | Anger | Contempt | Disgust



Fear | Happiness | Sadness | Surprise

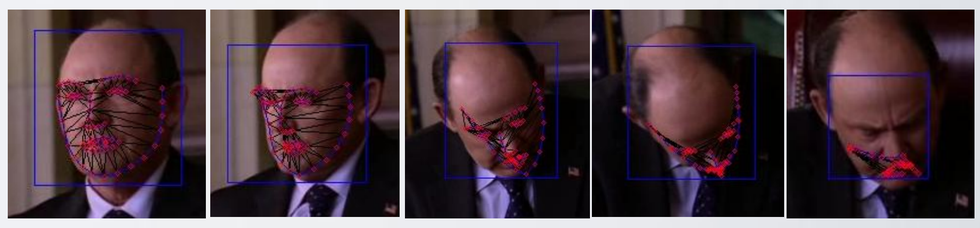
- ❖ 4850 videos generated from 24 movies
- ❖ 270 videos manually selected
- ❖ Eleminated videos if FaceTracker fails to find correct landmark points or we cannot annotate that video file with one emotion

Eleminated a video clip if there is more than one emotion



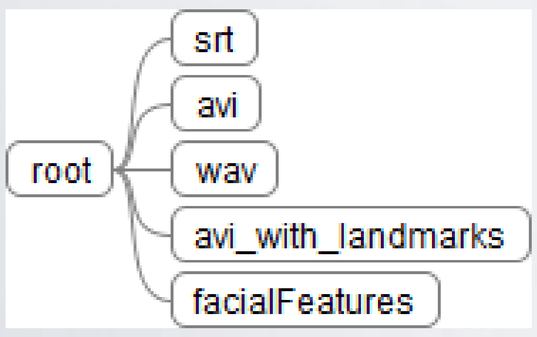
Frame 5 | Frame 44 | Frame 63 | Frame 114 | Frame 160

Eleminated a video clip if FaceTracker fails to find landmark points



Frame 5 | Frame 90 | Frame 190 | Frame 210 | Frame 290

Database Folder Organization



- ❖ srt : Extracted subtitle for the facial video as a srt file
- ❖ avi: Extracted video clip with audio as an avi file
- ❖ wav: Extracted audio as a wav file
- ❖ avi_with_landmarks: Video with drawn landmarks on it
- ❖ facialFeatures: Found landmark points recorded

This work has been supported by the Turkish Scientific and Technical research Council under Project 110E056

J. M. Saragih, S. Lucay and J. F. Cohn, “Deformable Model Fitting by Regularized Landmark Mean-Shifts”, International Journal of Computer Vision, vol. 91, pp.210-215, 2011