Journal of Faculty of Engineering & Technology

Journal homepage: www.pu.edu.pk/journals/index.php/jfet/index

Visual Speaker Identification Using Lip and Body Movements

M.U.G Khan, H.M. Shahzad, K.H. Asif, T. Ahmad, M. Afzal,

Dept of Computer Science and Engineering, University of Engineering and Technology, Lahore. Pakistan

Abstract

Speaker identification has been studied in many fields such as image processing, audio processing, artificial intelligence and speech recognition. Two of these areas are integrated together in order to identify the speaker. This research work is focused on two main approaches which are lip movements and body movements. These two approaches are combined together to achieve the speaker identification. Based on the proposed methodology, a speaker is identified in different scenarios, if there is a single speaker or if there is multiple speakers in the video or if the speaker's lips are not in view. For evaluation purposes, we have used precision, recall and accuracy scores applied on manually created videos dataset.

Index Terms: speech identification, video, skin detection, skin segmentation, lip detection, body movement detection

Introduction

Speaker identification has been a wide and attractive area of research. The area of speaker identification is concerned with extracting the identity of which person spoke the utterance. There are an increasing number of audio visual materials available (e.g., broadcast news, dramas, movies, sports video). Recently there have been a broad range of studies relating to identification of human objects, faces, and their actions. This study correlates two main areas extracted from the video. The first approach is based on lip movements, identifying the region of the lip of the speaker(s) and determines whether lips are open or closed, thus producing results for the close lip from the video. The second approach uses tracking of the body movements of the person(s) in the video to identify who is speaking. In particular, this study is to identify a speaker at each moment in a stream of video where multiple human objects are present. Moreover, when the lips are not in view of the camera, in this case how can we recognize the speaker?

Our aims for this research work are defined as the following: Firstly, identify whether the person in the video is speaking, to find out the best possible features for the identification of the speaker. Secondly, identify who is speaking when the lip is not in view or difficult to process the lip region. In addition, identify the speaker when there are multiple speakers in the video and also when the actual speaker is not seen in the video.

Proposed algorithm was tested against manually created dataset taken from TRECVid high level feature extraction task [1]. Precision values for individual tests based on body movements and lip detection alone were low. On the other hand, when we combined lip detection with body movements, precision values increased dramatically. The remainder of this paper discusses the research, design, and implementation work that we carried out; the results and our analysis of them; the problems that we encountered; and finally a summary of our achievements and conclusions.

Related Works

2.1 Body Movement

How can we identify who is speaking when the lip is not in view or the lip region is not clear? One solution would be to track body movements to identify a speaker. To be more specific, we will be looking at head and hand movements. The person who makes more head and hand movements is most likely to be the person speaking. This claim is backed up by Rose and Clarke [2], who conducted an experiment that supports the hypothesis that speakers move more than listeners.

Bull and Conelly [3] also showed that body movements can be linked to speech and aimed to show a significant relationship between body movement and phonemic clause structure. A phonemic clause consists of about five words [4], with changes in pitch, loudness and rhythm indicating just one main stress. The clause stops at a juncture, when the pitch, loudness and rhythm average out again. The next phonemic clause then starts after the juncture. They further investigated the relationship of body movement and vocal stress, using phonemic clauses. They concluded a strong relationship between the two. They found the tonic stresses were accompanied by body movements, to be more precise over 90% of the tonic stresses had body movements.

There are some recent works which are based on body movements for speaker detection which results in useful applications development. Bojan, et al. proposed a system for smart audio/video playback control which was based on presence detection and user localization in home environment [5]. Their work presented the design and implementation of a simple software-based home control platform used for the automatic control of audio/video devices within some specific range of 5 meters.

2.2 Lip Movement

Cootes [6] have come up with the Active Shape Models (ASM) which is capable of capturing the variability of image structures which belong to the same class. ASM maintains a training set where images are marked with set of landmark points which will exist in almost all the similar images. A Point Distribution Model (PDM) is derived out of the figures of the variations for the labelled points in the training set. The shapes are allowed to be varied within this model. A new image is identified with the model if it exists within the deformity of the set of images the model contains.

Iwano [7] have used the Optical Flow Analysis Method for tracking the lip movement. By this method we can determine the object movement. The images of the lip region are extracted from the video and then it takes a pair of adjacent images for calculating the optical flow velocities. The horizontal and vertical features of the images are computed which tells you whether the mouth is moving or not. It basically helps when the mouth pauses or mouth is shut.

Robert Kaucic, Andrew Blake [8] discuss lip tracking and that most approaches utilise Kass's snake approach which track the outer lip despite the impressive performance that comes from this approach it is not distinctive enough. Another contribution was that instead of relying on prior models it could detect the nostrils, and then colour thresholds were used to identify the black area in the inner mouth region, contour was then grown around the area identified as the inner mouth.

Bhat, et al. described a possible way of tracking the lips by using dynamic contour tracking through sparse representation of the lip contours, via splines, combined with a Kalman filter utilising prior shape and motion models of deforming lips [9]. For image feature detection, edge detection is used for the lips but it can be difficult as lip colour is similar to the skin colour. Using Bayesian classification, which uses probability for whether the pixel belongs to the lips by its classification of colour. Afterwards Fishers linear discriminate analysis is used to determine the boundary between the lips and facial skin. The next step is the inner-outer lip contour tracking and inner contour tracking enables additional reasoning to be made about the presence of the tongue and teeth.

Yuille, et al. used shape templates with snakes in order to extract the lip contours from an image face [10]. The lips can be described using a parameterized shape template. The shape template models an object within the image. By adjusting the parameters the model can be made to deform to fit the object in the image. The shape of this template is based on prior knowledge of the shape of the lips. One of the most successful lip reading systems to the present date is developed by Bregler and Omohundro [11].

Approaches

3.1 Body Movement

A person can be detected by identifying skin colours. We implement the condition, proposed by Peer et al [12]; on the RGB colour map .The condition boundaries are as followed:

“(R > 95*) AND (G >* 40*) AND (B >* 20*) AND (max{R, G, B} − min{R, G, B} >* 15*) AND (|R − G| >* 15*) AND (R > G) AND (R > B)”* [4]

Where R stands for red, G for green and B for blue colour spaces in a given image. We now create an image which just shows areas considered to be skin coloured. This method is by no means perfect, it doesn’t work for all skin colours at once and it will more than likely detect background areas which are not a person. This seems to be a major problem, however we are concern with body movement or skin areas which are moving, so the problem of background being detected as skin colour is not so much of a problem if we assume the background won’t be moving.

The next task is to find which detected skin areas are moving, and likely to be a person’s body movements. We use image subtraction between the consecutive frames in the video to detect any movement. Looking at the results from the image subtraction, we discovered that the algorithm is still detecting non-body movements from the background. However it was clear which cluster of pixels were body movements and which were not. Using grouping methods we take these clusters and turn them into objects. We also store properties about these objects such as ‘Area’. We identified that the objects which were significant body movements were bigger than those objects which were not. Hence we set a threshold and only keep objects above this certain size. It was difficult to set one threshold for all videos because the person could be close or far away from the camera. In particular when a person is far away from the camera, the skin detected areas become small enough to be the same size as non-skin detected areas. At a certain point there will be a limit to how far away the person can be to the camera. Here is an example of algorithm in action:

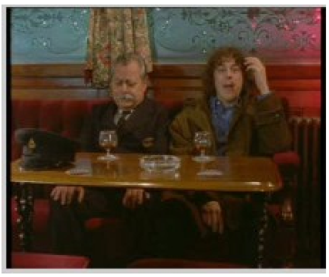

Figure 1 – current frame in video

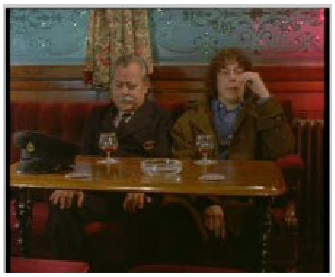

Figure 2 – next frame in the video

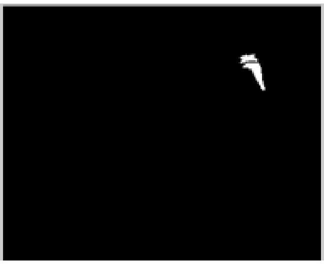

Figure 3 – The body movement detected between figure 1 and figure 2

Now we have an algorithm working for one person in a frame, we can look at two people in a frame and the possibility of multiple people in a frame. As we already have a working algorithm for one person, we ask the question, how can we split the frame with two people to get two frames with one person? Our approach was cropping the frame, it is by no means the best solution but it is effective. If the project was to be extended this section would have to be improved, particularly if we extended to more than two people.

To crop the frame we make the assumption that the positions of the people in the frame are either side of the centre of the frame. We get the size of the frame and split it in half down the middle, hence creating two frames. A left frame containing "person 1" and a right frame containing "person 2".

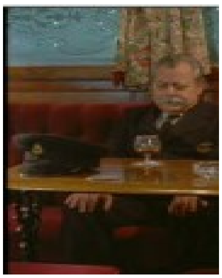

Figure 4 – left frame after crop

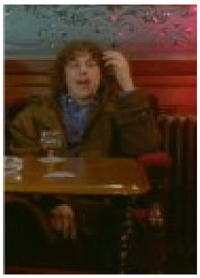

Figure 5 – right frame after crop

We create a graph representation of body movement we have detected. To implement this we still keep the idea about splitting the frame into two, left frame being "person 1" and right frame being "person 2". The algorithm is changed to say if no objects are identified in the in a frame after skin detection and image subtraction methods were applied (i.e. no movement detected) then give the frame number a score of 0. If objects were identified (i.e. significant movement detected) then give the frame number a score of 1. Here is an example of a graph representation:

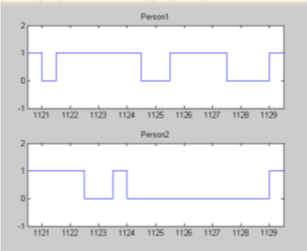

Figure 6 – output of graph showing the frames in which body movement was detected.

3.2 Lip Movement

To identify the lip of a person, initially face is detected from the given image [13]. Firstly, we divide the video into small number of frames by which we get a number of continuous static images which gives clear movement of lips. To identify the lip movement accurately we use 5 frames per second. After getting the frames, we sharpen the images by applying the image processing techniques in order to get a clear view of the image. Now that the image is ready, we need to identify the person in the image. We use the same skin detection methods, mentioned in the previous section, for detecting the face of a person in the image.

Since the results detect all skin of any body part while performing skin detection, we remove other parts except face by the width height ratio of the objects which we got from skin detection method (how to get the objects is mentioned in the body movement section). In the connected region, the width height ratio is higher for the face than other parts of the body. So now the face has been identified in the image, the next task is to detect lips out of the face. The method we use for this is cropping. Cropping the frame to the area of the skin and lip region was not our original approach that we were going to take but due to the image quality not always being perfect we use this approach. As mentioned previously, if the project was extended the cropping approach would have to be changed.

Returning the mouth region is also achieved using the crop method. This is done by taking an approximation on where the mouth region will be on the face and for most cases this crops the mouth region successfully. The next step is to use the cropped image of the mouth to identify the lips and judge whether these are open or closed which later we decide if the speaker is talking or not.

Firstly the lips have to be identified in the cropped image. We do this by using certain threshold ranges, before this we determine the structuring elements size. Next we apply rotated hue to the image, then taking two threshold ranges to the image and final a double threshold. This should leave the lip region identified as white where the non-lip region is identified as black.

The openness of the mouth is then calculated for the image by calculating the area of the holes present in the image larger than the structuring elements size. The area is then normalized to the image size and we either fill the holes in the image or extract them more. Once this is detected the points around the mouth are identified taking eight points, namely the extremes of the width and the extremes of the height. We place a line from the left most point to the right most point of the width this is then used later in gathering the results. We then need some more key information such as we calculate the distance between the centroid and the line created from earlier. The centroid is identified as the middle point in the image. An example what the image looks like after all these test is shown below.

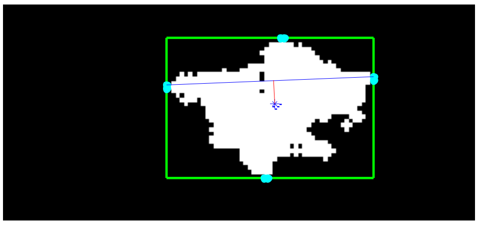

Figure 7 – An open mouth region with the eight points of the extremes

The first test is whether the mouth region is open. When applying the threshold, if it leaves an area in the middle of the mouth that is black and not white then this is classed as the mouth region open, the image needs to be of good quality to detect this, as for most images tested this is not found. After this we look at the lip height, the difference in lengths between the upper and lower lips. Based on this difference, we conclude whether the lip is open or not. When the difference is very small, we define the lips are closed, else if the lips height is very high then it is open, otherwise the lips still need to be classed through further methods. The next test is on the lip's width; we do this test because when people's mouth is open the distance between the lips is less than when closed. If the distance is very short or very large the lips are classed as open or closed respectively. If the lips are still unclassified then we continue to do more tests mixing lip height and width height to determine whether the lips are open. If the lips remain unclassified after all tests then they are classed as closed. If the lip's region is not found during the stage identifying the eight points of extremes then the lips are classed as closed by default to enable the algorithm to continue checking the next frames. In the results we have identified this by saying how many times the mouth region has not been identified, the better the results of this the more accurate that the lip tracking is.

Now that the mouth region in the frame has been identified in the video, we output whether the person is speaking or not in each frame. This is whether the speaker's mouth is open or not and later we determine whether the speaker is speaking or not. This is done by identifying if three consecutive frames are identified as mouth region closed, then the speaker is classed as not speaking. Here is an example of the output generated for two speakers. For two speakers, we use the same cropping method mentioned in the body movement section. The generated output for the lip movement also shows the manual test results (ground truths).

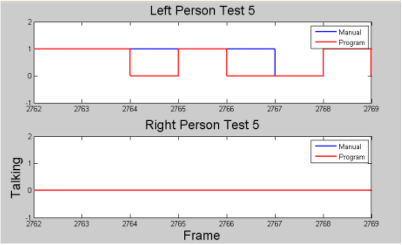

Figure 8 – The output of graph showing whether the lips are open or closed. The blue line is where the results differ from manual results

Experiments

4.1 Data Set

The data set that used in this experiment was taken from TREC Video Retrieval Evaluation (TRECVid) [14], which is an international benchmarking that provides a huge test collection, uniform scoring procedures and a forum for organization. These sets created for researchers who are interested in information retrieval and want to evaluate their works [15]. TRECVid can be applied on many applications such as automatic and manual searching, search for shots within a video corpus, automatic detection of a variety of video features.

When searching for data to test the body movement algorithm, we looked for videos that matched the following criteria; One, the video had two people in. Two, the camera view was fixed, i.e. the camera angle didn't move. Three, the body or at least hands were visible, i.e. no videos of head shots or hands covered by gloves. Four, there was a reasonable distance between the two people and no contact was made between them.
Although we have varying video conditions, we found the best results were videos which had a dark colored static background.

For the lip movements the data sets we looked for were where the speaker in the video is looking directly at the camera and the camera not moving. For this we found that news reports were the best video's to use for the data set. We tried to make sure the video background contained as small amount as possible of colors that could be classed as skin color.

The content of the videos varies. We have videos of people sitting or standing or both; we have different types of people including males and females, Caucasian and Asian. Our skin boundary condition did not extend to African skin color. We have various locations, differing from news studios to outdoors next to a road.

Below is a breakdown of the length of the videos in the data set and the amount of time in the video any speaker is speaking for.

Table 1- Statistics for data sets

Test Case	Number of Tests	Average Length of Video	Average Time Speaking in Video	Max video length	Min video length	Max speaking time	Min Speaking time
Lip movement	5	14.5s	11.4s	20.2s	12.2s	18.2s	6.8s
Body movement	14	18s	13s	43.5s	8s	35.5s	0.5s

4.2 Extraction of frames from video

To extract the frames from the video we used ffmpeg, which is a piece of software used to manipulate audio and video files. Ffmpeg can be used to extract frames at a specific frame per second speed [16]. The default speed is set at 25 frames per second, which was too quick to detect significant body movement and lip movement with our algorithms. For the body movement algorithm we used 2 frames per second and the lip movement used 5 frames per second. For running the tests on both approaches we used 2 frames per second. The frames are then stored in a predefined directory with the file names "frame'*i*'.jpeg" where *i* is the frame number.

4.3 Evaluation Strategy

We used precision and recall test strategy for the algorithms. The graphs that are generated by the algorithm are then compared to the ground truths. These ground truths are the same type of graphs created manually by hand based on human perception of speaker identification.

For the lip movement used the following conditions were used to create tables listing the number of true positives (when a mouth is classed as open and manually recorded as open), false positives (when a mouth is classed as open and manually recorded as closed), true negatives (when a mouth is classed as closed and manually recorded as closed), false negatives (when a mouth is classed as closed and is manually recorded as open).

The conditions for the body movement were as follows, true positives (when a person is detected moving and manually recorded as speaking), false positives (when a person is detected moving and manually recorded as not speaking), true negatives (when a person is not detected moving and manually recorded as not speaking), false negatives (when a person is not detected moving and is manually recorded as speaking). We then create tables for the accuracy and positive rates.

We then create tables for the accuracy and positive rates for each test and summarized the all the tests of different types in one table.

4.4 Body Movement

The definitions for true positives, false positives, true negatives and false negatives for each case are explain in the previous paragraphs.

Table 2 – Body movement case results

Test case description	True Positives	False Positives	True Negatives	False Negatives	Total frames
Body movement tested on videos with 2 people	162	220	441	177	1000

A. Table 3- Body movement individual test accuracy, recall and precision

Test number	accuracy	Recall	Precision
1	55%	51%	54%
2	74%	52%	92%
3	56%	39%	58%
4	50%	100%	53%
5	79%	78%	74%
6	47%	33%	50%
7	45%	42%	40%
8	52%	100%	14%
9	58 %	20%	44%
10	73%	33%	57%

11	72%	100%	9%
12	48%	71%	33%
13	60%	73%	58%
14	64%	52%	62%

The "body movement tested on videos with two people" test case is based upon results for 10 video clips, ranging from 16 to 87 frames per video clip. All 10 video clips contain two people, some clips contain one person speaking and others contain both speaking, there is a mixture of people sitting and standing and the videos are taken in various locations.

4.5 Lip Movement

Table 4 - Lip movement case results

Test case description	True Positives	False Positives	True Negatives	False Negatives	Total frames
Lip Movement tested on videos with 1 person	190	105	32	34	362
Lip Movement tested on videos with 2 people	76	46	58	50	230

Table 5 - lip movement with 1 person Individual test accuracy

Test number	Accuracy	Recall	Precision
1	52%	79%	54 %
2	77%	88%	82 %
3	52%	88%	53 %
4	59%	73%	66 %
5	65%	92%	67 %

Table 6 - Lip movement with 2 people Individual test accuracy

Test number	Percent of true frames	Recall	Precision
1	59%	63%	61%
2	52%	55%	65%
3	81%	100%	57%

The "lip movement tested on videos with one person" test case is based upon results for 5 video clips. All video clips contain one person facing the camera. These clips are mostly taken from news interviews and only show the top half of the body.

The "lip movement tested on videos with two people" test case is based upon results for 3 video clips. All video clips contain two people facing the camera where only the top half of the body is shown.

4.6 Lip and Body Movement

Table 7- Both algorithms tested on same videos case results

Test case description	True Positives	False Positives	True Negatives	False Negatives	Total frames
Lip Movement tested on videos used by both algorithms	51	2	8	15	76
Body Movement tested on videos used by both algorithms	14	18	21	23	76

Table 8- accuracy of lip and body movement of same videos

Test number	Accuracy	Recall	Precision
6 (lip algorithm)	71%	63%	100%
7 (lip algorithm)	83%	87%	89%
6 (body algorithm)	47%	33%	50%
7 (body algorithm)	45%	42%	40%

The test cases for which both algorithms are tested consist of two video clips, both clips are of news readers sat down at a desk facing towards the camera.

Discussion

5.1 Body Movement Detection

The results for the body movements were good for certain videos, in particular the videos with dark backgrounds, hence making the skin detection a higher percentage and therefore higher percentage of speaker identification. In other cases the accuracy was just below 50% which shows the algorithm can have a negative effect. However we should not be as surprised by these results as the algorithm detects body movements and the results are based upon a comparison against whether the person is speaking. Hence we have made an assumption that body movement is directly linked to speaking.

There have been many studies, such as Bull and Conelly [3] to show a link between body movement and speaking, but how strong that link is we can't be sure. Rose and Clarke [2] produced a study which shows that emotions can affect the link between body movement and speaking. They have shown on a small sample that when a speaker is in an emotional state such as anger or joy, it is easier to identify the speaker. Whereas emotional states such as fear or disgust are hard to identify a speaker. In this experiment there was a 50% chance of identifying a speaker by guessing, so in the case of 'fear', body movements can have a negative effect on identifying the speaker, according to this study.

Other aspects which affected the results were the content and quality of the videos. By content of the video we mean, the colour of the background, moving objects in the background (cars etc.), the position the people in relation to the camera and the location the video is set (outside or inside). Even aspects of lighting can have an effect on the skin detection.

Other aspects which affected the results were the content and quality of the videos. By content of the video we mean, the color of the background, moving objects in the background (cars etc.), the position the people in relation to the camera and the location the video is set (outside or inside). Even aspects of lighting can have an effect on the skin detection.

5.2 Lip Movement Detection

Overall the results for when the mouth was open turned out quite high accuracy results. Whereas results for when the mouth is closed did not produce the results we had hoped for. We feel this is mainly due to the fact for when the lip region is not correctly found, the default setting for this is to say the mouth region is closed and because for some of the tests this occurred in a few files, it can alter the results accuracy. To try and combat this we tried altering the threshold set but the higher or lower the threshold the larger or smaller the lip region became ending up making some of the tests void because all the values were too small or large. Some of the tests involved the speaker turning to the side or covering the mouth region with the hand, this sometimes added to the fact that the lips could not be detected, or the hand may get confused with the face, but that was rarely the problem.

For some of the results we found that the algorithm had identified the speaker to be not speaking when they are. But for the majority of the time we correctly identified when the speaker was not speaking, this is shown in the graphs produced

For the two speakers in video tests we found that it produced better results for when the speaker was not speaking than in the results gathered for one speaker. This may be down to the fact that there was less background area on the images so the face region was more clearly identified, due to the fact we used a split screen to separate the images.

We found from the results that the better quality of video the better the accuracy was, we discovered that we needed to use videos where the speaker was closer to the screen, this was due to the fact that after cropping the lip region and after enhancing it as much as possible it sometimes did not identify the region.

The results of the lip movement generated better results, in particular the recall of the case of one person in the video. One reason for this is that there is a much stronger link between lip movement and speech than body movement.

5.3 Implementation Issues

This work, as many other researches, has faced some problems during the work course. For the lip movement the problems we faced were determining the skin region from the other regions on the screen, because most videos had similar background colors to those of the face. To overcome this we had to set all skin regions over a certain size which therefore means that for faces that are not within a certain distance from the camera will be ignored. This was not too much of an issue not being able to test for distant images because we found that when the speaker is far away after cropping and enhancing the image the lip region was of poor quality and gave back very poor results. Another issue was with side view videos of the person speaking because we were now using cropping which approximated the region that the mouth was in, it was not easy to detect the lip region from this. However we found again issues with the quality of the image was the main issue, so hence we concentrated on gathering results on one and two speakers in the video.

The body movement implementation faced similar problems as the lip movement. Mainly, setting the correct threshold for all videos where the people are different positions from the camera. In some of the test we had major issues with background areas being detected as skin, and in other cases problems of background movement. We believe that some of the cases of background movement were due to shadow and poor quality videos.

Conclusions

In this paper we have discussed our implementation of the lip and body movement algorithms. Based on these algorithms we have devised an approach for identification of speaking person in a video stream. Results concluded that speaker identification can be successfully performed based on lip and body movements.

Future work would include introducing an audio algorithm capable of finding stop sounds within speech and compare the audio with the lips being closed from the current lip algorithm, improving the skin detection parts of the lip and body movement algorithms and creating a new method identifying people within the screen, rather than cropping the images. If we were able to achieve these improvements we would be able to perform speaker identification on multiple people (more than two), gain a greater accuracy for speaker identification, and have the possibility to identifying a speaker not seen in the video.

References

[1] Smeaton A. F., Paul O. and Wessel K., 2009. High-level feature detection from video in TRECVid: a 5-year retrospective of achievements, Multimedia content analysis. Springer US, pp. 1-24.

[2] Rose D., Clarke, T., 2009. Look who's talking: Visual detection of speech from whole-body biological motion cues during emotive interpersonal conversation, University of Surrey.

[3] Bull, P., Conelly, G., 1985. Body Movement and Emphasis in Speech, Human Science Press.

[4] Arnon, I., Neal S., 2010. More than words: Frequency effects for multi-word phrases, Journal of Memory and Language, 62(1) pp. 67-82.

[5] Mrazovac, Bojan, et al., 2011. Smart audio/video playback control based on presence detection and user localization in home environment, 2nd Eastern European Regional Conference Engineering of Computer Based Systems (ECBS-EERC).

[6] Cootes, T., Cooper, C. D., Graham, J., 1995. Active shape models-their training and application, University of Manchester.

[7] Iwano, K., Tamura, S., Furui, S., 2001. Bimodal speech recognition using lip movement measured by optical-flow analysis, Workshop on Hands-Free Speech Communication.

[8] Kaucic, R., Blake, R., 1998. Accurate, Real-time, Unadorned Lip Tracking, University of Oxford.

[9] Bhat, Kiran S., et al., 2013. High fidelity facial animation capture and retargeting with contours, Proceedings of the 12th ACM SIGGRAPH/Eurographics Symposium on Computer Animation.

[10] Yuille, A., Hallinan, P., Cohen, D., 1992. Feature extraction from faces using deformable templates, Harvard University.

[11] Bregler, C., Omohundro, S., 1995. Nonlinear manifold learning for visual speech recognition, University of California.

[12] Peer, P., Kovac, J. and Solina, F., 2003. Human Skin Colour Clustering for Face Detection, University of Ljubljana.

[13] Erdem, C. E., et al., 2011. Combining Haar feature and skin color based classifiers for face detection, Acoustics, Speech and Signal Processing (ICASSP).

[14] Paul O. et al., 2011. TRECVID 2011 – An overview of the goals, tasks, data, evaluation mechanisms and metrics, TRECVID 2011-TREC Video Retrieval Evaluation Online.

[15] Smeaton A. F., Paul O., Kraaij W., 2006. Evaluation campaigns and TRECVid, ACM International Workshop on Multimedia Information Retrieval, Santa Barbara, USA, pp. 321–330.

[16] Zeng, H., Yuan F., 2013. "Implementation of Video Transcoding Client Based on FFMPEG." Advanced Materials Research 756: pp.1748-1752.