iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** IV **Month of publication:** April 2018

DOI: http://doi.org/10.22214/ijraset.2018.4675

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

Review and Analysis of Various Lip Reading System Techniques

Poonam, Mridul Chawla[1]
[1]*Department of Electronics and Communication Engineering Deenbandhu Chhotu Ram University of Science and Technology Murthal, Sonepat, India*

Abstract: ***Speech is the most common and natural methodology of communication. There are many approaches available for lip reading. In this paper, various techniques for lip reading are being reviewed. The important stages are speech, lip reading, feature extraction and recognition. Many classification techniques have been surveyed in this paper like neural networks, lip extraction technique, cues of lip shape, sign language recognition, 3D tracking system, 3D-discrete cosine transform etc.***
Keywords: ***Lip reading, Feature extraction, Segmentation, Automatic speech recognition, Neural Network.***

I. INTRODUCTION

The most natural methodology of communication among people, in general, is "talking" which we call it as "speech." But sadly, this natural kind of communication for the individuals those are dumb and hearing lessened cannot use. The strategy of word recognition offered to assist hearing lessened or dumb individuals talk to the others during a normal technique. It's a visible methodology of talking method within which solely lip movements are applied established vocalized words [10]. Visual speech recognition may be a technique that acknowledges the words by movement of the lip. Visual speech recognition is the process to reading the lip. The deaf person and the hearing-impaired person can easily recognize the speech by lip movements. Since earlier times, people have had apprehended that the movement of the lip has had some speech knowledge. In various application areas like speech in the changing region, spaces where you are doing not ought to be compelled to talk and catastrophe situation (volcanic activity) visual speech is significant [6].

A. Lip reading

Recently, lip reading becomes a lot of well-linked analysis, as a result of its large spread facilities. The Automated Speech Recognition (ASR) system is used widely, for instance in Artificial Intelligence (AI), radiophone and private computer. But, within the buzzing surroundings, the performance of ASR can fall. The problem is going to be overcome by the lip-reading system. Method of Lip reading will even be used for speech interface which is silent for a laryngectomee (Person United Nation agency has no larynx). In few of the security system, lip impression is employed for person identification. Several lip segmentation ways supported gray-scale image, various ways units supported color image. Some ways used the color image directly, the choice technique used color conversion to different color space to expand the utterly totally different of the lip and thus the background [3]. Lip reading is made up of two necessary phases: the initial is an extraction of the feature, and second is identification. Feature extraction includes detection of the key purpose of lip curve. The edge-based approaches use abstraction cues like color and intensity information designed for chase lip while talking. As an example, abstraction cues such as saturation and hue are used for feature extraction. For boosting segmentation, "snakes" are employed by mistreatment and physical property. For effective segmentation, the form of the lip prior has to ensure to made lip from mistreatment current from Model and Active Contour Model, etc. However, parameter adjustment problem within the on top of models enthused researchers to accept section primarily relayed method [10].

B. Feature extraction

Visual features of speech i.e. visemes are valuable in the many fields, such as visual speech recognition audio-video synchronization. It is observed that mostly visual information correlated to speech is contained in the movements of the lip. However, it is the most difficult portion of visual speech recognition to find an efficient method for extracting visual speech features. It is very complex to take out visual features from the input image or video which is taken in different situations under varying lighting conditions, different pose and for different skin tones. A viseme extraction and lip localization many algorithms remained proposed. Some groups are working with template models created on active shape models, deformable model, and dynamic contours [8].

Lip segmentation could be an essential methodology for several uses. It is used for lip reading, speech recognition and expression analysis (example- facial expression, evaluation of spirit, recognition of pain). Every application has certain restriction relating to accuracy, toughness, and speed. The necessities used for facial expression recognition could also be really utterly completely diverse counting on the application area. Many algorithms planned to try and do action on the color footage. Initially, a color alteration is to be done to require blessings of the assorted visual property of lips starting from the skin. Here, these transformations focus on red, inexperienced color data, relating to the RGB color house. The color alteration outcome could also be one channel strength plot of the mouth portion and the mouth is highlighted. On the other hand, to boot various works area units unconcealed acting on ancient gray footage from monochrome cameras. Mainly, the segmentation techniques are divided into two parts. A primary cluster is that focus on recognition of the lip edges inside the mouth Region of Interest (ROI). These imply dynamic contour models accustomed the color reworked a region of interest or utilize deformable templates. Several works steady their active contours victimization sustain following points. The major assumption of the edge-based rule is, that lips produce distinguished edges on skin-lip crossing [1].

C. Recognition

Methods to language detection are going to be separated into isolated sign detection and continuous language detection. Isolated sign recognition may be a unique case of gesture detection. Gesture recognition systems area unit usually designed to acknowledge artificial gestures. The user should learn these gestures thus on speak with the systems. Thus, on own o idea regarding the only approach needed to make a symbol language's recognition systems it becomes critical to review the previously developed systems. For this reason, the analysis above the electronic databases and reading of a small number of papers that were most related to the systems has been allotted. Associate analysis and comparison of the ways applied for the development of the similar systems would facilitate to use the acceptable method to build up the system. Several approaches in different areas are units utilized so as to vogue languages recognition system [7]. recognition, the Learning Vector Quantization (LVQ) for model detection and probabilistic neural network for an organization of action units (AUs). Actuated by talented Self -Organizing Map NN (SOMNN) for lip detection used for helping the pc lip's analysis actuated North American nation to use a similar for recognition purpose within the conferred approach [10]. Mostly the region-based technique uses applied mathematics characteristic of lip region to grasp lip pursuit. Favored regions base approach mostly includes a deformable constant example, 3-state parabolic geometric model. On high of how were proficient in producing symmetrical conic thus the lip type model didn't work accurately. A solution was given by victimization a pair of parabolas for the higher boundary that was afterward changed by quadratics. Some analysis has applied Multi Feature Active from Model (MF-ASM) to spice up convergence to the uproarious situation. This model rule programs, Associate in the nursing extension of ASM presented by Matthews's that has chemist analysis on a gray level case. The on high of how unit economical enough to figure out lip type with physically concern standardization, but the presentation deteriorated once used for durable lip pursuit. The performances of the top of ways might deteriorate once occlusions like teeth or tongue square measure gift. To enhance the performance, a localized active contour model (LACM) supported native datum was purposed. All on top of mentioned algorithms would like repetitive modeling, therefore leading to extended and complicated model fitting. (ANN) Artificial Neural Network square measure commonly applied classifiers which apply supervised also as Associate in nursing unsupervised technique to learn. ANNS square measure higher as compared to HMM owing to its unfitness to the selection of the inside parametric. Previous use ANNs embodies Multi-Layer Perceptron (MLP) neural network for the lip bioscience, a back propagation algorithmic program for a text-dependent spokesperson

II. LITERATURE REVIEW

Panning et al. (2009) described that the mouth and lips segmentation could be an elementary problem in the detection of the face. This paper projected a technique for segmentation of lip supported RG-color bar chart. Primarily an uneven adjustive threshold takes a bar chart section that assures that each one dots theirs in section area unit skin pixels. Authors made the Gaussian model that describes skin pixel allocation and it is employed for a distinguished, optimum threshold. [1].

Pana et al. (2012) discussed the various strategies used for lip segmentation. It still remains a difficult problem due to high changeable lip color and low chromatics distinction between the lip and skin. An explicit automatic lip segmentation algorithmic program supports an explicit color transformation in RGB rather than advanced color models. The comparative study with some existing lip segmentation programs has indicated the superior performance of the developed algorithms [2].

Nasuhal et al. (2013) analyzed that the lip's reading may be a widespread application, example -Audio-Visual Automatic Speech Recognition (AV-ASR), that is speech interface and personal identification. Segmentation of lip is an important part in lip's reading. Lip following may be a way of the finding lip to associated lip in successive video outline. The chan-vese model may be a section-based segmentation rule, which equally is employed as tracing methodology. This rule can sense the border of an object that is not made public by gradient, where a standard active curve can't be implied [3].

Huang and Lin (2013) described the recommended rapid mouth recognition theme. In this mouth recognition, the unit correlates technique that couldn't accurately realize the correct mouth location, specifically outline face or light-weight influence. Foremost procedure in this technique contains face removal, mouth features withdrawal, mouth detection. In face extraction, the authors applied Viola and Jones detector for taking out the face areas. Results demonstrate that the methodology may acquire an elevated correctness and fast dead time with the face taking out and the mouth features localization [4].

Mardiyantol and Sardjono (2015) discussed the main points of lip shape and offers necessary signals of lip form tracing that were applied for speech detection, lip analysis and plenty of transmission application. A special acceptable threshold segmentation was given for 6 key points lip's feature abstraction rules. Color transformation in the Red-Green-Blue house and adaptive threshold were applied for the lip segmentation. Curve of the segmental lip was being outlined and also stuffed along with bounding color. Lastly, the 6 main points that are corners of right or left, minor purpose, 3 points of the cupid on the bow of the lips are observed. Presentation of projected techniques was appropriated and related to offer strategies to achieve a necessary enhancement in the correctness [5].

Morad and Patnaik (2015) analyzed the lip's reading and the choice of options showed significant part. In lip analysis application info is video, thus three-Dimensional transformations are suitable to take out lip action data. The state of skill in the lip's analysis predicates about outline standardization and edgewise features withdrawal. Additionally, all the frames can't be thought-about evenly as they allow variable motion data. Defined in that paper 3D remodel technique is planned for features extraction. That option is the key to Genetic Algorithmic program (GA) model for detailed study. The genetic Algorithmic program is employed for spatiality decrease and to boost the presentation of the classifiers at a small value calculation. Each test and coaching period for the classifier condenses by the compact features size. For the testing of the digit utterance CUAVE and Tulips, database is being used [6].

Dour and Sharma (2015) described that sign language recognition has attracted respectably in computer vision. Symbol languages may be a way of conveying the words by using hand, arm, body, and face images to express ideas and solutions. Such as sign languages, spoken languages came out and developed naturally among the deaf individuals. Sign languages aren't general. There is no internationally customary and standardized linguistic communication used for all deaf individuals [7].

Dave (2015) discussed lip localization and tracking. They are helpful in lots of applications- facial animation lip's reading, visual speech recognition, lip synchronization, visual speech recognition etc. To synchronize lip movements with input audio, author firstly segmented lip region from the input image or video frame. This paper presents an imaginative color-based method for localization of lips, which is the initial stage for tracing lips in real time. A phoneme is a basic unit of speech and a viseme is an image expression of phonemes or outline of the mouth while a note of an exact phoneme. The main goal of this work is to implement a system for matching lips with the input speech. To extract visual features i.e. visemes from input video frame or image, HSV and YCbCr color model along with many morphological procedures are used. The author has developed the algorithm to work with normal lighting conditions and natural facial images of the male and female [8].

Oztel and Kazan (2015) despite the benefits of online searching in terms of amount and variety, its drawbacks includes the fact that the users can't try products. The application was made for lipstick trial. The users will see the applied lipstick in front of screen. Authors tend to apply look segmentation for identification lips space, color segmentation for extracting lips and color area alteration for finer effect [9].

Rathee (2016) analyzed recognition of the speech supported the outline of lip actions while talking. Audio speech detection system was the standard for decades and has attained some achievement, however recently the visual's speech detection created curiosity in the minds of researchers for lip's reading. Lip's reading has a spare benefit of the elevated correctness and the sound freedom. Author gave associate degree algorithmic program for repeated lip's reading. The algorithmic program consisted of 2 major steps: features extraction and classification of words recognition. Lip's info is derived using lip's geometric and lip's appearance. Correctness attained from the projected method was ninety- seven percent [10].

Zhang (2016) proposed an improved version of the word segmentation algorithm that is used to find out the section which represents the pronunciation piece from the video outline sequence inside the computer lip-reading systems. In the computer lip-reading system, after extracting lip's features, deviation of lip's feature is calculated. The result showed that this method can only be

used to determine the approximate pronunciation section and that it cannot accurately determine the location inside the beginning and the ending in the pronunciation position. So, an improved method was presented in which the location of beginning and ending of pronunciation section was determined according to the distribution of the gradient of the average deviation [11].

Morade and Patnail (2006) analyzed that the human uses visual information while making an attempt to find out speech, particularly in less audible and noisy conditions. Lip movement reading is the method of understanding the elemental speech via a process on the actions of lips. The detection of lip activity could be a tough job when the Region of Interest is non-linear and noisy. In this projected methodology lip's analysis has 2 stages feature withdrawal model that is precisely described and calculated economically. The primary phase 3D Discrete Cosine Transform or 3D (3D-DCT) was employed with Sensitive Discriminant Analysis for minifying the feature proportions. These options create a completely unique lip analysis system with small feature vector extent. [12].

Table-1 Depicting techniques and findings from different author

Sr. No.	Author Name	Technique/Scheme	Findings
1	Panning et al. (2009)	Lip reading, underneath speech recognition or expression study	Color and edge information.
2	Pana et al. (2012)	Lip extraction technique	Color transformation is used to develop the lip segmentation and reduces the computational complexity.
3	Mardiyantol et al. (2015)	Lip provide important cues of the lip form	Lip's feature extraction be constantly improving the lip segmentation quality.
4	Morade and Patnaik (2015)	Feature Selection	Genetic algorithm can be fine-tuned to generate improved outcome all the features selected and formed better classification accuracy.
5	Rathee(2016)	Neural Network	NN is applied for wording identification. Its characteristics are potential of huge data handling, high pace and false acceptance.
6	Jiao and Zhang (2016)	Enhanced speech segmentation algorithm	Lip's-analysis system to attain automatic segmentation of image sequence of recognition content.
7	Morade and Patnaik (2016)	3D-DCT and 3D-DWT and LSDA	Evaluate to the feature vector from other transform techniques.

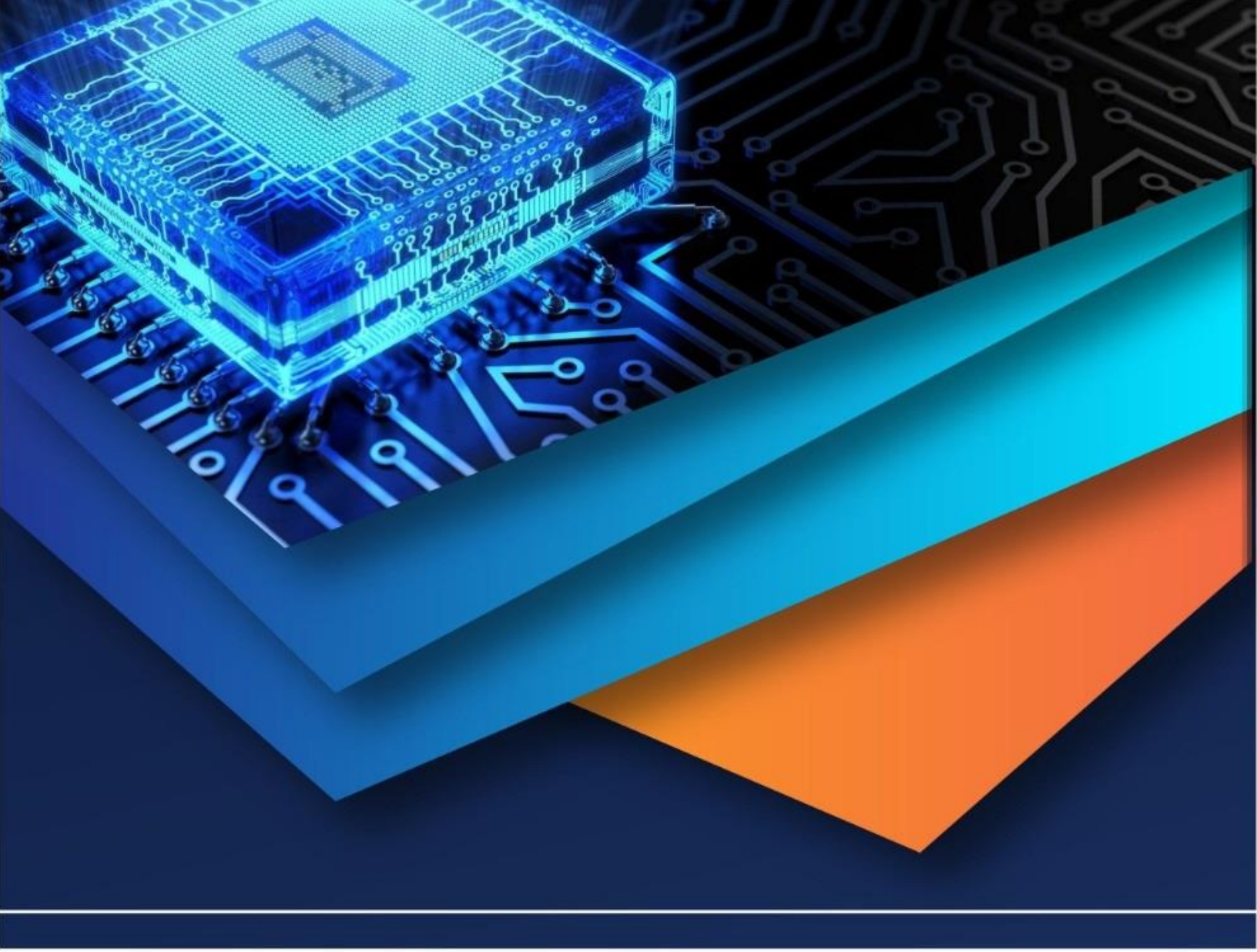
III. CONCLUSIONS

Various papers have been studied for lip's reading system techniques. From these papers, we have concluded that different techniques have different ways to recognize the lip's reading techniques. The color transformation is used for the advancement in the quality of the lip segmentation and reduces computational complexity. The procedure of lip's feature extraction was continuously enhanced in the lip segmentation quality. The neural network of the word identification ensures the better presentation owing towards its essential potential of the large data handling, high pace, and false tolerance.

REFERENCES

[1] Axel Panning, Robert Niese, Ayoub Al-Hamadi, and Bernd Michaelis, "An Adaptive Approach for Histogram based Mouth Segmentation", Vol:3, No:8, 2009
[2] Jing Pana, Yepeng Guana,b,∗, Shuangcheng Wanga, "A New Color Transformation Based Fast Outer Lip Contour Extraction?", September 1, 201

[3] Aris Nasuha1,2 Tri Arief Sardjono2, Mauridhi Hery Purnomo, "Lip Segmentation and Tracking Based on Chan-Vase Model", 2013 IEE
[4] Yu Huang and Yan-Ching Lin, "An Efficient Mouth Detection Based on Face Localization and Edge Projection", June 201
[5] Hadid Tunas Bangsawan Ronny Mardiyanto1, Tri Arief Sardjono2, "Six Key Points Lip's Feature Extraction Using Adaptive Threshold Segmentation", 2015 IEE
[6] Sunil SudamMorade, SupravaPatnaik," A Genetic Algorithm-Based 3D Feature Selection for Lip Reading", 2015 IEE
[7] Shweta Dour, M.M. Sharma," Review of Literature for the Development of Indian Sign Language Recognition System", December 201
[8] Namrata Dave1," A LIP LOCALIZATION BASED VISUAL FEATURE EXTRACTION METHOD", December 201
[9] GozdeYolcu Oztel1, Serap Kazan2," Virtual Makeup Application Using Image Processing Methods", August 2015
[10] NeeruRathee, "A novel approach for lip Reading based on neural network", 2016 IEE
[11] Ma, Xiaohui Jiao, and Hongjun Zhang," An Improved Word Segmentation Algorithm for Lip-reading", 2016 IEE
[12] Sunil S. Morade, Suprava Patnaik," Visual Lip Reading using 3D-DCT and 3D-DWT and LSDA", February 2016
[13]



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY