

Image Based Techniques for Lip Segmentation – A Review

Sneha K S¹

M.Tech Student, Department of School of Computer Science, Mahatma Gandhi University, Kottayam, Kerala, India¹

ABSTRACT: The lips and mouth region are the main visual parts of the human speech system; these parts hold the most visual speech information. A few techniques have been recommended to perceive lip tracking with the focus on segmentation of lip regions or extraction of lip contours. Commonly used approaches for lip tracking are Image-based methods. In the Image based methods a set of characteristic points are identified with the help of spatial cues such as color and edges. These types of methods are quite sensitive and ambiguous to the initial position of feature points. This paper reviews the various components of Image based techniques used in lip tracking.

KEYWORDS: Lip segmentation, Image based methods

I. INTRODUCTION

The main complication of lip feature extraction lies in the accuracy and reliability of the system. Recent research has shown that color is a powerful tool. Unlike the gray-level approach, color image analysis increases the efficiency and robustness of locating the lips, and easily modifies to detect beards, teeth and tongue. Color-based techniques operate at pixel or neighbourhood level, and try to differentiate lip and skin pixels based on color features. Lip tracking algorithms usually start by transforming the RGB image to a suitable color plane image by applying a suited color transform. Since there is a difference between the color of lips and the color of the face region around the lips, detecting lips using color information attracted researchers because of simplicity, not being time consuming, and the use of fewer resources, e.g. low memory, allowing many promising methods for lip detection using color information to emerge. The most important information that focus on include the red and the green colors in the RGB color system, the hue of the HSV color system, and the component of the red and blue in the YCbCr color system. Image based methods are computationally less expensive but it adversely affected by illumination. This approach is bottom-up. In this the image data is examined at a low level and looking for local structures such as regions or edges, which are assembled into groups in an attempt to identify objects of interest. Without a global model of what to expect, this approach is difficult and prone to failure. Skin color in an image is affected by many factors:

- **Illumination:** A change in the light source distribution and in the illumination level such as indoor, outdoor, highlights, shadows and non-white lights produces a change in the color of the skin in the image i.e., color constancy problem. The illumination variation seriously degrades the performance.
- **Camera characteristics:** Even under the same illumination, the skin-color distribution for the same person varies from one camera to another depending on the camera sensor characteristics.
- **Ethnicity:** Skin color also varies from person to person belonging to particular ethnic groups and from persons across different regions. For example, the skin color of people belonging to Asian, African, Caucasian and Hispanic groups is different from one another and ranges from white, yellow to dark.
- **Individual characteristics:** Individual characteristics such as age, body parts and sex also affect the skin-color appearance.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor & UGC Approved Journal)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

- Other factors: Different factors such as subject appearances like makeup, hairstyle and glasses, background colors, shadows and motion also influence skin-color appearance

II. RELATED WORK

Vezhnevets et al. [1], and Kakumanu et.al. [2] describe the most widely used skin color detection techniques, their key ideas; summarize their advantages, disadvantages and characteristic features. In the table below show the best results of different methods, for the Compaq dataset. Albiol et al. [3] provide a theoretical proof to find the optimum color spaces for skin detection.

Method	TP	FP
Bayes SPM in RGB [Jones and Rehg 1999]	80% 90%	8.5% 14.2%
Bayes SPM in RGB [Brand and Mason 2000]	93.4%	19.8%
Maximum Entropy Model in RGB [Jedynak et al. 2002]	80%	8%
Gaussian Mixture models in RGB [Jones and Rehg 1999]	80% 90%	~ 9.5% ~ 15.5%
SOM in TS [Brown et al. 2001]	78%	32%
Elliptical boundary model in CIE-xy [Lee and Yoo 2002]	90%	20.9%
Single Gaussian in CbCr [Lee and Yoo 2002]	90%	33.3%
Gaussian Mixture in IQ [Lee and Yoo 2002]	90%	30.0%
Thresholding of I axis in YIQ [Brand and Mason 2000]	94.7%	30.2%

Fig 1 Performance of different skin detectors

For color image segmentation, histogram based and clustering based methods have been widely used. In [4], [5], [6] & [7] propose a fuzzy clustering algorithm in which color lip segmentation problem treated as a two-class clustering and segmentation problem. The proposed algorithm helps to exploit the spatial interactions between neighbouring pixels by using novel dissimilarity index. Spatial fuzzy clustering utilizing both luminance and chrominance features from the CIELAB and CIELUV color spaces is used to produce a probability map of the lip image. A shape function is enclosed in the dissimilarity measure of the objective function. Thus pixels having similar color information which is located in different regions can be clearly differentiated.

Panning, Axel, et al [8] suggests a method for lip segmentation based on rg-color histogram. They use the rg-color-space for the purpose of pure color based segmentation. Initially select an adaptive threshold that ensures that all pixels in that region are skin pixels and not lip pixels are selected. Use those pixels to build a gaussian model which represents the skin pixels distribution and is used to obtain a refined, excellent threshold. They do not use shape or edge information.

Threshold-based segmentation strategies give a simple and effective approach to implement lip segmentation. Automatically finding the thresholds is a major challenge. [9] Propose a method called adaptive threshold optimisation (ATO) which selects the threshold by using feedback of shape information. ATO make use of shape information by constructing a lip shape model from statistical analysis of labelled images. The algorithm consists of two components: the base algorithm, which serve as a typical threshold-based segmentation algorithm, and the adaptive threshold optimisation (ATO) algorithm.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor & UGC Approved Journal)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

Lucey, Simon et al. [10] use a fuzzy based approach in which the temporal information from progressive image frames is used to calculate a stable threshold. Maximum entropy principle and fuzzy c-partition method [11], [12] are also used to select the threshold values. In this a fuzzy partition on a two-dimensional (2-D) histogram is performed. The 2-D histogram is built such that the horizontal axis represent one of the pixel color component and the vertical axis represents the local average of that color component, Then color segmentation is accomplished by performing the fuzzy partition on each of the color components and then combining the results.

Statistical-color model approach [13], [14] used to automatically find and track the face and other facial features in the image. And also propose a method to estimate the 3-D pose of a user's head by tracking a few facial features. A few years later, Feris et al. [15] present a real time system which detect the facial features and track them. They also use the same statistical color model.

III. CONCLUSION

This paper has reviewed the main components of an image based techniques used in lip segmentation. Histogram and clustering methods are widely used. Each method has its own advantages and difficulties in implementation. Comparing to other methods thresholding is the simplest and computationally fast. Image based methods are less expensive than the model based methods and the progress in these areas will allow audio-visual speech recognition to play a key role in intelligent computer interfaces.

REFERENCES

- [1] Vezhnevets, Vladimir, VassiliSazonov, and AllaAndreeva. "A survey on pixel-based skin color detection techniques." Proc. Graphicon. Vol. 3.
- [2] Kakumanu, Praveen, SokratisMakrogiannis, and NikolaosBourbakis. "A survey of skin-color modeling and detection methods." Pattern recognition 40.3 (2007): 1106-1122.
- [3] Albiol, Alberto, Luis Torres, and Edward J. Delp. "Optimum color spaces for skin detection." Image Processing, 2001. Proceedings. 2001 International Conference on. Vol. 1. IEEE, 2001
- [4] Liew, Alan Wee-Chung, Shu Hung Leung, and Wing Hong Lau. "Lip contour extraction using a deformable model." Image Processing, 2000. Proceedings. 2000 International Conference on. Vol. 2. IEEE, 2000
- [5] Leung, Shu-Hung, Shi-Lin Wang, and Wing-Hong Lau. "Lip image segmentation using fuzzy clustering incorporating an elliptic shape function." IEEE transactions on image processing 13.1 (2004): 51-62.
- [6] Liew, AW-C., Shu Hung Leung, and Wing Hong Lau. "Segmentation of color lip images by spatial fuzzy clustering." IEEE transactions on Fuzzy Systems 11.4 (2003): 542-549.
- [7] Liew, Alan Wee-Chung, et al. "Fuzzy segmentation of lip image using cluster analysis." Eurospeech. 1999.
- [8] Panning, Axel, et al. "A new adaptive approach for histogram based mouth segmentation." Proceedings of the World Academy of Science, Engineering and Technology 56 (2009): 779-784.
- [9] Gritzman, Ashley D., et al. "Automatic computation of histogram threshold for lip segmentation using feedback of shape information." Signal, Image and Video Processing 10.5 (2016): 869-876.
- [10] Panning, Axel, et al. "A new adaptive approach for histogram based mouth segmentation." Proceedings of the World Academy of Science, Engineering and Technology 56 (2009): 779-784.
- [11] Cheng, H. D., Y. H. Chen, and X. H. Jiang. "Thresholding using two-dimensional histogram and fuzzy entropy principle." IEEE Transactions on Image Processing 9.4 (2000): 732-735.
- [12] Cheng, H. D., Jim-Rong Chen, and Jiguang Li. "Threshold selection based on fuzzy c-partition entropy approach." Pattern recognition 31.7 (1998): 857-870.
- [13] Stiefelhagen, Rainer, Uwe Meier, and Jie Yang. "Real-time lip-tracking for lipreading." Eurospeech. 1997.
- [14] Stiefelhagen, Rainer, Jie Yang, and Alex Waibel. "A model-based gaze tracking system." International Journal on Artificial Intelligence Tools 6.02 (1997): 193-209.
- [15] Feris, Rogério Schmidt, TeófiloEmídio de Campos, and Roberto Marcondes Cesar Junior. "Detection and tracking of facial features in video sequences." Micai. 2000.