International Journal of Current Engineering and Technology
ISSN 2277 - 4106

Available at http://inpressco.com/category/ijcet

Research Article

Extraction of Lip Contour from Face

Riti Kushwaha[a], Neeta Nain[a*] and Promila Jangra[a]

[a]Department of Computer Engineering Malaviya National Institute of Technology, Jaipur, Rajasthan, India

Accepted 3June 2012, Available online 8 June 2012

Abstract

Lip segmentation is an essential stage in many multimedia systems such as video-conferencing, lip reading, or low bit rate coding communication systems. It is also useful in various image / video acquisition devices encouraged the development of many computer vision applications, such as vision-based surveillance, vision based man machine interfaces, vision-based biometrics, and so on. Hence extraction of lip becomes a broad field. This paper presents an algorithm for extraction of lip contour. We have used two algorithms for this. Our first algorithm is a color based algorithm derived from a quadratic polynomial, with the assumption that color of lips should lie in the range of blue. The accuracy of the first proposed algorithm is 91% in case of normal skin people, and 95% in case of fairer skin color people. Our second algorithm is model a based algorithm which gives accuracy up to 96%.

Keyword: *Lip contour detection, Lip segmentation, Lip extraction from face.*

1. Introduction

Detection of the lip contour is a fundamental procedure of mouth feature extraction. It forms a preliminary stage of face image analysis, which is essential for numerous application areas including man-machine interaction based on the observed human behavior, video-telephony, face and person identification, bimodal speech recognition, face and visual speech synthesis, facial expressions classification etc. All of these applications require an efficient and fully automated mouth feature extraction method that can be achieved using an automatic lip contour detection technique. Observing the pixels of lips, we find that lip color of fairer skin people ranges from dark red to purple, and for normal skin color people it is in the range of blue under normal light conditions. From the perspective of human visual perception, the lips are very easy to be differentiated from the face, cause of different contrasts in colors. The outer labial contour of the mouth is having very poor color distinction when compared against its skin background; it makes extraction of lip a difficult problem. In order to improve the contrast between lip and the other face regions we need different type of lip extraction techniques. But if we are taking the assumption that lip color varies from dark red to purple then this condition is satisfied with the people who have very fair complexion. For the normal skin people we are assuming that the color of lip also varies in the range of blue. We thus used the chromatic color space (Cheng Chin *et al*, 2003), which is constructed from the RGB color space (Cheng Chin *et al*, 2003), to exclude dark pixels because chromatic color space is not suitable to distinguish between bright pixels and dark pixels. For chromatic color space, each pixel is represented by two values, denoted by (r,g). Their relationship is explained in section IV-A. Both dark pixels and bright pixels might have the same converted r and g values due to the normalization effect in brightness in chromatic color space thus we have used both the color spaces. Our second algorithm is model based algorithm (T.F.Cootes *et al*, 1995) hence we don't need any color constrain for that.

Section II describes the related work, section III describes the dataset taken of different types. Section IV and V describes the proposed algorithms for lip contour extraction techniques. In section VI result analysis of the proposed methods are obtained. Section VII contains the whole summary of the experiments.

2. Related work

A number of techniques are used for detection of lip contours from a given face image, which can be generally classified into categories, which are model based approaches like deform-able templates (L.Yuille *et al*, 1989), active shape models J.Luettin *et al*, 1996 and snakes (M.Kass *et al*, 1987). These approaches generally use a set feature point for approximating the lip contours with spline functions and give a model for the lip contour. High robustness is achieved if reliable constraints on the deformation of mouth model are included in the cost function, which is defined on the basis of heuristics or through generally supervised learning from examples. However, the minimization of the cost function involves computationally expensive algorithms, and the model

* Corresponding author: Neeta Nain

definition often employs semi-automated user assisted procedures or a priori knowledge of some characteristics of the mouth image. The second category is Colour based approach like (Yao Hongxun *et al*, 2002). Another category is level set based approach and rule based approach (Cheng Chin *et al*, 2003). We have used rule based approach. In this approach firstly colour deviation is done by some technique then rules for extracting the lip contour is applied.

3. Database acquisition

For testing both the methods we have chosen different databases. If a person is having very fair complexion then we have applied dark red to purple color on his lips, if he is having normal skin color then we choose blue color. A dataset of 150 persons for fairer skin color people with red color on lips is taken from internet for testing results. For normal skin color people we construct the database by applying blue color on lips, we are having 120 person's face. Two types of pictures are taken in both the cases which are enlisted below.

- Images when a part of face with lips is available
- Images when whole face is available

We have tested all the listed type of images for first algorithm. And the second algorithm is tested on the images when a part of face with lips is available.

4. Proposed Method

The overview of the first proposed method for extracting the lip contour from face is shown in Figure 1

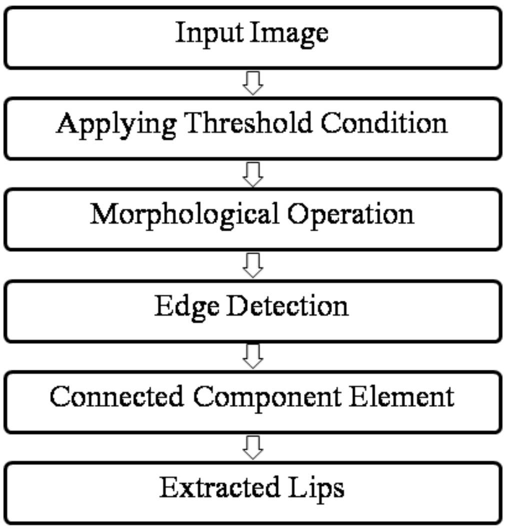

Fig.1 flow chart for color based approach

a) Rules for extraction

This algorithm adopts the chromatic color coordinate for color representation. For chromatic color space, each pixel is represented by two values, denoted by (r, g). The conversion from conventional RGB color space to chromatic color space is defined as follows:

$$r = R \big/ R+G+B \tag{1}$$

$$g = G \big/ R+G+B \tag{2}$$

Where R, G, B denotes the intensities of pixels in red, green and blue channels respectively. According to the above transformation, it is easy to see that the brightness variation in images can be normalized in chromatic color space. The color representation in this two-dimensional chromatic color space enables us to visualize and analyze very easily when building the color model for skin pixels.

$$lip(r) = -.776r^2 + .5601r + .2123 \tag{3}$$

$$f(r) = -.776r^2 + .5601r + .1766 \tag{4}$$

As the objective of our algorithm is to detect lips from face in real time, we adopt a simple technique which uses two quadratic polynomials to find out the upper and lower contour of lip region using equation 1 and 2. This extracts the darker region of face with a O(2) computation cost. After applying this discriminate function the pixels from darker part of face like eyes, eyebrows, nose holes etc., are also included in the extracted region which are removed by applying threshold as explained in the section 4.2.

b) Applying threshold

As we know that the threshold varies for different skin color people hence we calculate two threshold conditions. One for those with fairer complexion and another for those who are having normal complexion as given below:

1) For (FSC) Fairer Skin Color

Threshold condition for fairer skin color people is not having lots of variation in their ranges, hence the intensity value of red green and blue pixels lie in approximately same range. The threshold is calculated as:

$$F(N) = \begin{cases} 1 & if \quad f(r) \le g \le lip(r) \;\&\; R \ge 20 \;\&\; G \ge 20 \;\&\; B \ge 20 \\ 0 & \quad otherwise \end{cases} \tag{6}$$

2) For (NSC) Normal Skin Color

As we all know there are lots of variation in the normal skin color that implies variation in the intensity value of the pixels. Thus R, G and B values have variations in their range. The threshold is calculated as shown below. Where L (N) =1 indicates a lip region pixel. And f(r) and lip(r) is taken from the equation 3 and 4. We have used standard morphological operation for image enhancement and noise removal, followed by edge detection technique for finding the edge detection technique for finding the edges of extracted component. Connected component labeling is also done for the better region extraction. The algorithm is discussed below:

$$L(N)=\begin{cases}1 & if \quad f(r)\le g\le lip(r)\ \&\ R\ge 30\ \&\ 30\le G\le 70\ \&\ 55\le B\le 122\\ 0 & otherwise\end{cases} \tag{7}$$

Step1. Morphological Operation: After applying threshold stray pixels as side effects also called salt noise is left in the image (I). This is removed by morphological Erosion operation using a square (size varying from 3 - 15) as the StrEl1 (structuring element) as shown in equation 8.

$$I = I\Theta StrEl1 \tag{8}$$

These small components from the face up to the size of 15 pixels. The Erosion operation also removes some part from the lip component which is restored by Dilating I (the lips) as shown in equation 9, using Disc as a StrEl2.

$$I = I \oplus SrtEl2 \tag{9}$$

This restores the original lips and also connects both upper and lower lip portions which sometimes break due to Erosion. Both the StrEl's are shown in the table 1.

Step2. Edge Detection: To extract the upper and lower contour of lips which are mostly diagonal edges we apply edge detection technique using Sobel edge detector as it smoothes the image in the direction perpendicular to edge detection which makes it a good edge detector for diagonal edges.
Step3. Connected Component: As we know that the largest connected component in the face will be lips, we have applied the 8-connected component algorithm to identify lips. Figure 3 shows an instance of the above steps for normal skin color people and Figure 2 shows the result for fairer skin people.

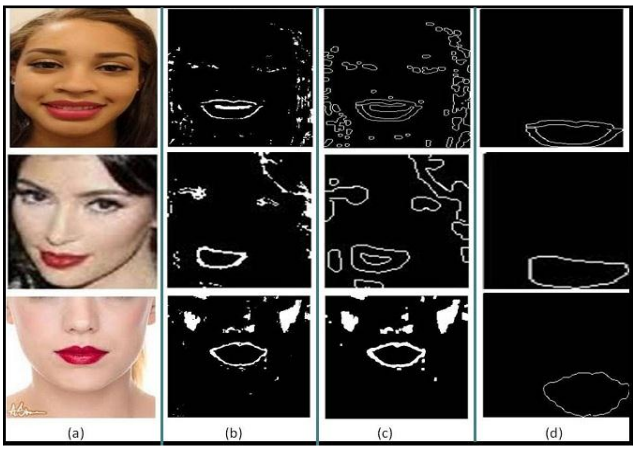

Fig. 2 Result on fairer skin color people (a) Input image (b) Applying threshold conditions (c) Morphological operation and edge detection (d) Extracted lips.

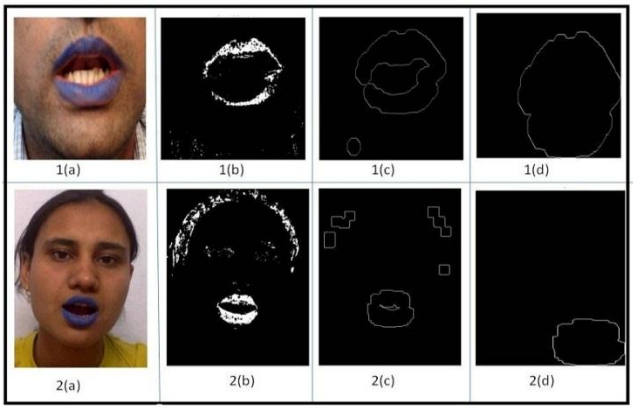

Fig. 3 Result on Normal skin color people (a) Input image (b) Applying threshold conditions (c) Morphological operation and edge detection (d) Extracted lips.

5. Proposed method 2

This algorithm is designed for specific set of frontal face images with open mouth, from the whole database. This model based algorithm uses Mean Shape Model (MSM). The database used for training is hand designed. The algorithm consists of four major steps: (a) Training the data set, (b) calculate MSM, (c) normalize the MSM by aligning it with the min angle to principal axis, (d) MSM deformation.

5.1 Training Set

The training set contains any set of 20 open mouth images to make MSM. Open mouthed images are selected to cater for the purpose of maximum outer lip contour segmentation of any input image in any size database. Thus for this experiment, the training set contains only open mouth images to make the large shape model which covers up all other type of movement of lips by just inner deformation of the model.

5.2 Mean Shape Model

A SV (Shape Vector) is defined manually for all images of the training set. The SV consists of 66 landmark points which specify the outer lip contour as shown in Fig. 4 (b). These landmark points are chosen with the constraint that these points exactly define the lip contour shape. The ith SV is defined as equation 10.

$$SV_i = \{(x_{i1}, y_{i1}), (x_{i2}, y_{i2}), \ldots, (x_{i66}, y_{i66})\} \tag{10}$$

After evaluating the SV of all training set images, the MSM is calculated using equation 11.

$$\overline{SV} = \frac{1}{N}\sum_{i=1}^{N} SV_i \tag{11}$$

Here, N is the number of landmark points for any training set image. One instance of MSM is shown in Figure 4 (c).

5.3 Alignment of MSM

The MSM calculated in the above step is aligned with minimum angle with the principal axis as shown in Figure 4(d). To achieve the required alignment we have to perform scaling, rotation and translation so that image and MSM corresponds as closely as possible.

5.4 Mean Shape Model Deformation

After proper normalization of MSM, we start the deformation of MSM to extract the lip contour of the input image. Before performing any deformation, we threshold the RGB image to a binary image (B) using equation 12.

$$B = \begin{cases} 1 & if \quad R \le 120 \,\&\&G \le 120 \,\&\&B \le 170 \\ 0 & otherwise \end{cases} \tag{12}$$

The deformation can be done as:

1) The MSM is divided into two parts- upper lip model and lower lip model. The upper lip model contains 35 landmarks points whereas; lower lip model contains 31 landmarks points of the MSM.
2) MC (model curve) for these both parts is extracted using curve fitting algorithm (www.mathworks.in). We find the polynomial of degree n that fits the data points x_i to. y_i Let suppose the mean shape points of upper lip contour are:

$$X_u = \overline{x_1}, \overline{x_2}, ..., \overline{x_{35}} \tag{13}$$

$$Y_u = \overline{y_1}, \overline{y_2}, ..., \overline{y_{35}} \tag{14}$$

$$p(x) = p_1.x^n + p_2.x^{n-1} + ... + p_n.x + p_{n+1} \tag{15}$$

Here, $p_1, p_2 ...$ *and* p_{n+1} are the polynomial coefficient of a polynomial equation. The degree n of the curve is chosen such that all data points related to * lies on the polynomial equation of the curve

Where, * is

$(\overline{x_i}, \overline{y_i})$

3) Construct the tangent on the MSV curve. For example, if $y=p(x)$ then slope of tangent is dy/dx . Then the tangent line at

$(\overline{x_i}, \overline{y_i})$
is defined as:

$$y - \overline{y_i} = \frac{dy}{dx}.(x - \overline{x_i}) \tag{16}$$

These tangents drawn on the model curve are equal to the number of landmark points which define the MSM.

4) Find tangent normal for each pixel which occurred in between the MSM and the lip contour of input image (this is nothing but the intensity change of the binary image). Normal line is the line perpendicular to the tangent line to the curve. The slope of normal line is $-1/(dy/dx)$
and thus the equation of normal line is

$$(x - \overline{x_i}) + \frac{dy}{dx}(y - \overline{y_i}) = 0 \tag{17}$$

The results of model deformation are shown in Figure 4(e). Length of the normal vector of each tangent is basically decided by the point on the model curve and end point of deformation where intensity change occurred between the lip and the skin. Here, we check the intensity value at each single pixel increment of normal vector length. So, the task of pixel by pixel checking becomes little bit time consuming. To reduce this time, we can use the binary search approach instead of pixel by pixel approach. In this approach, we can set the maximum value of length for normal vector that is feasible under the model deformation. For every iteration reduce the length by half and again make a test of intensity value at the point. If that point lies on the lip then again reduce the length by half otherwise increase the length up to half of previous normal vector length. Repeat these till we get a single point. This is the end point of the normal vector and which define the lip contour. A new SV is extracted that contains the landmark points which define the shape of our input image. An instance of the final output of the deformation of MSM is shown in Figure 4(f).

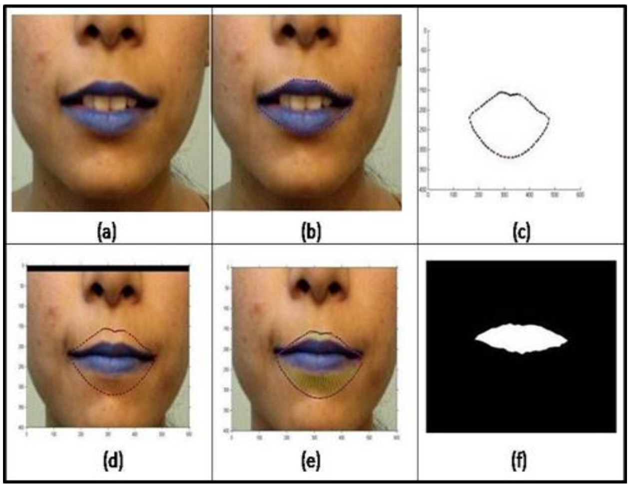

Fig. 1 Lip contour extraction processing: (a) Training set Image, (b) Landmark point represented in blue color (c) MSM, red color define the mean point (d) Alignment of MSM on input/test image (e) Deformation of MSM: green line represents the normal vector (f) Output after deformation

6. Result analysis

Algorithm 1 is tested on 270 images. 150 images of fairer skin color people with dark red to purple color lips, and 120 images of normal skin color people where, the test images contains both full face and part of face. The accuracy of the proposed algorithm is 91% for the normal skin, and 95% for the fairer skin people, where the test images contains full face, as 142 images among 150 gives correct result for fairer skin color people. And 110 among 120 gives correct result for normal skin color people.

Algorithm 2 is tested on 50 images of normal skin people database containing frontal face images having mouth only. Out of 50 images, 48 images successfully fulfilled the criteria for the acceptance of lip contour. Figure 5 shows the results in detail. Both the algorithms work well even for low-resolution images, but performance degrades with poor illumination.

Conclusions

Proposed algorithm 1 gives good results for both skin color people, and works in different well-constrained conditions of database. It has an accuracy of 91% for normal skin color and 95% for fair skin color. The accuracy is consistent across databases provided the lip color is red and image quality is good.

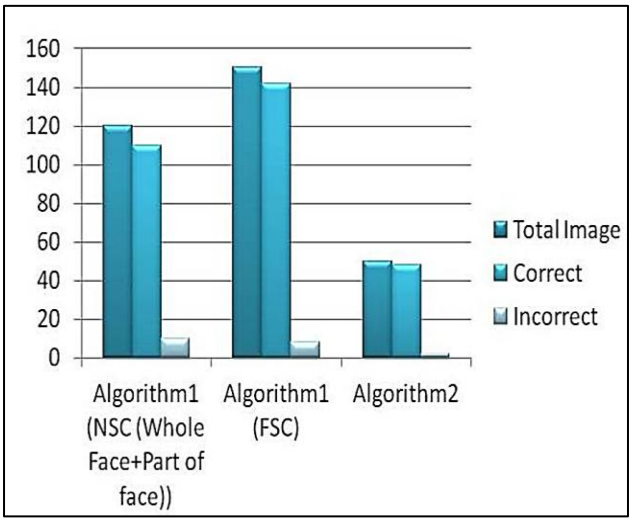

Fig. 2 Histrogram depicts the correctly classified images

The performance degrades if the lip color is neutral and any other face feature is similar to lip color. This is proposed as future enhancement of the algorithm. In algorithm 2 the lip contour segmentation is performed by using MSM deformation approach which gives good results for outer landmarks. This work can be further extended for inner landmark points of the mouth, which defines the position of mouth. This could also identify how much lips are open which is highly required in the visual speech recognition systems. Both the algorithms work well for good quality images. Small degradation is observed in poorly illuminated images.

References

L.Yuille, D.S.Kohen, and P.W.Hallinan (1989), Feature Extraction from Faces Using Deformable Templates, In: *Proceedinds of IEEE International Conference Computer Vision and Pattern Recog*., pp. 104-109, (San Diego,CA.

J.Luettin, N.A.Thacker, and S.W.Beet (1996), Visual Speech Recognition Using Active Shape Models and Hidden Markov Models, In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal.* Vol. 2, pp. 817820, Atlanta, GA.

M.Kass, A.Witkin, and D.Terzopulos (1987), Snakes: Active Contour Models. In *Intl. J. Computer Vision*, Vol. 1, No. 4, pp. 321331.

Yao Hongxun, L Yajuan, GaoWen (2002), Lip-Movement Features Extraction and Recognition Based on Chroma Analysis. In *Acta Electronica Sinica*, pp. 168172.

H.Mehrotra, G.Agrawal, M.C.Srivastava (2009), Automatic Lip Contour Tracking and Visual Character Recognition for Computerized Lip Reading, *International Journal of Computer Science*,pp.6271.

Cheng Chin Chiang, Wen Kai Tai, Mau Tsuen Yang, Yi Ting Huang, Chi Jaung Huang (2003), Novel method for detecting lips, eyes and faces in real time (*Elsevier Ltd*).

Abu Sayeed, Md.Sohail and Prabir Bhattacharya (2007), Automated Lip Contour Detection Using the Level Set Segmentation Method. In *Proceedings 14th International Conference on Image Analysis and Processing*.

M.Sadeghi, J.Kittler and K.Messer (2002),. Modelling and segmentation of lip area in face images. In *Proceeding of online confrence on Image signal processing*, Vol. 149, No.3, June.

Fahimeh Salimi, Mohammad T Sadeghi (2009), Decision Level Fusion of Colour Histogram Based Classifiers for Clustering of Mouth Area Images. In *Proceedings of International Conference on Digital Image Processing*.

Pham The Bao and Huynh Nguyen Duy Nhan (2009), A New Approach to Mouth Detection Using Neural Network In *Proceedings International Conference on Control, Automation and Systems Engineering*.

Jagdish Lal Raheja, Radhey Shyam, Jatin Gupta, Umesh Kumar, P Bhanu Prasad (2010), Facial Gesture Identification using Lip Contours, *In Proceeding of Second International Conference on Machine Learning and Computing*.

Qiang Wang, Haizhou Ai, Guangyou Xu (2002), A Probabilistic Dynamic Contour Model for Accurate and Robust Lip Tracking, *Proceedings of the Fourth IEEE International Conference on Multimodal Interfaces (ICMI02)*.

Y.P.Guan (2008), Automatic extraction of lips based on multi-scale wavelet edge detection, In *Proceedings of IET Comput. Vis*, Vol. 2, No. 1, pp. 23 33

Lain.Matthews, T.F.Cootes, J.Andrew Bangham, Stephen Cox, Richard Harvey (2002), Extraction of Visual Features for Lipreading In *Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 2

T.F.Cootes, C.J.Taylor, D.H.Cooper, J.Graham (1995), Active Shape Models- Their Training and Application In *Proceedings of Computer Vision And Image Understanding*, Vol. 61, No. 1, pp. 38-59

http://www.mathworks.in/help/techdoc/ref/polyfit.htm