A Key Point Extraction Method of Lip Contours Based on Jumping Snake

Yuanyao Lu[a*], Song Geng[b]

College of information Engineering, North China University of Technology, Beijing, 100144, China

[a]E-mail:luyy@ncut.edu.cn [b]E-mail:gengsong2009@sina.com

Keywords: lip segmentation; key point extraction; color space; jumping snake

Abstract: Lip segmentation and key point extraction from human mouth image are essential steps in lip-reading applications. In this paper, a method is proposed to improve key point extraction. First, mouth is detected according to the structure and proportion of the face. Second, two corners of the mouth are detected and extracted. Third, we study many color spaces and find more appropriated color space and gradients for outer and inner lip contours. And at the last, we use the jumping snake algorithm to extract key points along outer and inner lip contours. The proposed method gives promising results but still has little inaccuracy with inner contours.

Introduction

Many researchers have proposed audio speech recognition (ASR) systems with high recognition rate. But their recognition rate decreases significantly in a noisy environment. Once the system is applied to the real environment, such as office, factories, airports, etc, their performance will descend rapidly. It is well known that Visual speech recognition (VSR) can significantly increase speech comprehension under noisy condition [1, 2]. Lip-reading technology can greatly improve the recognition rate of speech recognition in noisy environment.

Lip contours extraction is a very important part for lip-reading technology. The current methods for lip contours extraction are, combining edge detection and region segmentation[3], active shape models and snakes[4], parametric model[5], cooperative scheme of Bayesian segmentation and active contour[6] and gradient vector flow with parabolic template[7]. Some methods is not so robust and fails if lighting conditions are not favorable and other methods is robust to lighting conditions but is not accurate[8].

In this paper, a method based on jumping snake algorithm to extract key points of lip contours is proposed. Our method chooses more appropriate color space and gradients and it is more accurate and robust in key point extraction.

System architecture

Fig. 1 shows the schematic of our method. First step deals with mouth localization, second step concerns with the detection of corners of the mouth, and the last step is the detection of the key points in outer and inner lip contours.

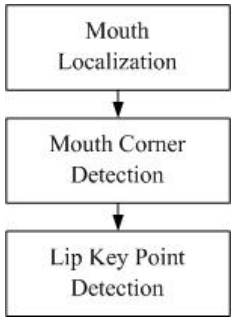


Fig. 1 Schematic diagram

Mouth localization

There have been many researchers done for mouth localization. Some researchers made the whole face image as the study area [9] and some researchers determined the lip area by means of locating the obvious key points in the face area, such as eyes, forehead, jaw, naris [10]. But these methods contained a lot of redundant information. So we can extract the mouth area according to the structure and proportion of the face [11], the mouth area is calculated as:

$$\frac{1}{4}W_{face} < W_{mouth} < \frac{3}{4}W_{face} \text{ , and } \frac{2}{3}H_{face} < H_{mouth} < \frac{9}{10}H_{face} \text{ ,} \tag{1}$$

where W_{face} and H_{face} are the width and height of the face area respectively. W_{mouth} and H_{mouth} are the width and height of the mouth area respectively. Fig. 2 gives the result using the method.

Mouth corner detection

In fact, there have been lots of researchers done and on the journey for corners of the mouth detection. Lip edge detection method used to fix the corners of the mouth was raised by Steifelhagen et al. [12], but it did not reach to higher level of accuracy. In this paper, the method for corners detection is searched by region as the corners of the mouth goal and reaches to high accuracy. The two corners is shown in Fig. 3 marked with P_1 and P_2.

Fig. 2 Extracted mouth area

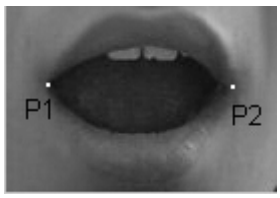


Fig. 3 Detected corners of mouth

Lip contour extraction

Color spaces and gradients

In this paper, we use jumping snake algorithm to extract lip contours. The jumping snake, introduced by Eveno et al. [13], is an active contour that converges in a succession of jumps and

growth phases. A jumping snake need gradient information to fit contours accurately and the difficulty is to define strong information able to accentuate the difference between lip and skin for outer contours and between lip and oral cavity for inner contours[14]. In this paper, C2 and C3 component in Discrete Hartley Transform are used. So, four different mixed gradients used for jumping snake convergence are deduced from color information, and they are computed as follows:

$$G_1 = \nabla[H_1 - L], \tag{2}$$

$$G_2 = \nabla[H_1 + L], \tag{3}$$

$$G_3 = \nabla[R - u - L - H_2], \tag{4}$$

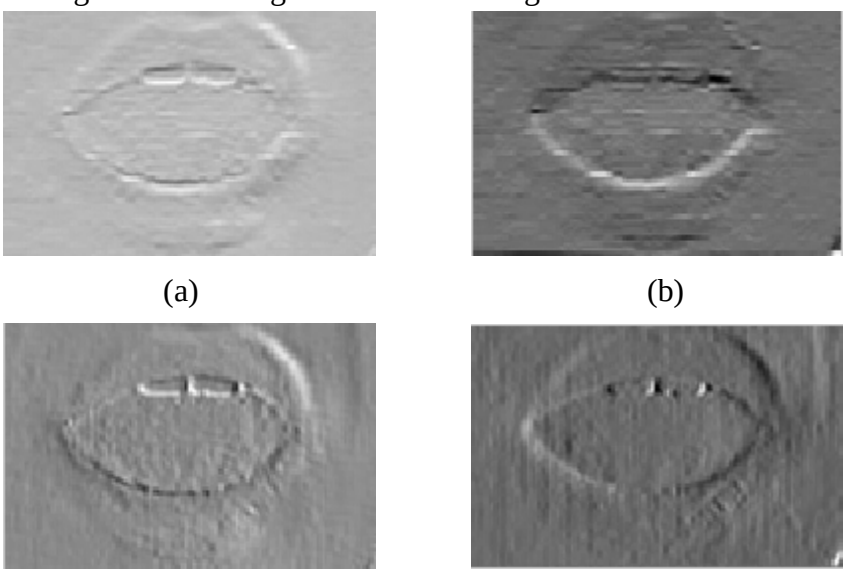
$$G_4 = \nabla[u \times L], \tag{5}$$

where × represents a dot to dot product. H_1 and H_2 used in [14] are two other components called pseudo-hues. They are computed as follows:

$$H_1 = U + C2 + C3 \text{ , and } H_2 = \begin{cases} \dfrac{R}{R+G} & if\ R > G, \\ 0 & otherwise. \end{cases} \tag{6}$$

All components are normalized between 0 and 1. U is U component in LUV.

For the outer contours, G_1 is used to highlight the outer upper and G_2 the outer lower lip contour. For the inner contours, G_3 is used to highlight the inner upper and G_4 the inner lower lip contour [14]. The four gradients for Fig. 2 are shown in Fig. 4.



(a) (b)

(c) (d)

Fig. 4 Four gradients for mouth ((a) for G_1, (b) for G_2, (c) for G_3, (d) for G_4)

Key point extraction of the outer contour

Key point extraction of the outer upper contour

In this paper, key point extraction of the outer upper contour uses jumping snake algorithm. In the growth phase, the location of the initial seed S0 is set manually and depended on the two corners of the mouth. The coordinate values are computed as follows:

$$S0(x) = \frac{P_2(x) - P_1(x)}{2}, \tag{7}$$

$$S0(y) = 1, \tag{8}$$

where $S0(x), S0(y)$ are coordinate values of the initial seed S0. $P_1(x)$, $P_2(x)$ are coordinate values of the left and right corner in Fig. 3.

The upper snake in the outer upper contour which is shown in Fig. 5 convergences with the maximization of the gradient flow G_1, and the number of the points of the snake is 9. We choose 3 points as the key points. They are shown in Fig. 6 marked with P_3, P_4, and P_5.

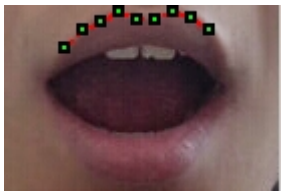


Fig. 5 Upper snake convergence

Fig. 6 Key points in upper contour

Key point extraction of the outer lower contour

As we know, the outer lower contour is a quadric curve, so we choose only 1 key point shown in Fig. 7 marked with P_6. It is chosen with the maximization of the gradient flow G_2, and P_4 and P_6 have the same x coordinate value.

Fig. 7 Key point in lower contour

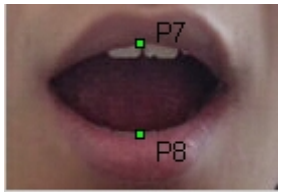


Fig. 8 Key points in inner contour

Key point extraction of the inner contour

Canny edge detector [15] is used to extract inner lip contour. In fact, there are some limitations [15-17] of edge detection so there is a little difficulty might be faced while extracting inner contour than outer one. We find 2 key points in the inner contour marked with P_7 and P_8 in Fig. 8. P_7 is chosen with the maximization of the gradient flow G_3 and P_8 is chosen with the maximization of the gradient flow G_4. P_7, P_8 and P_6 have the same x coordinate value.

Comparative study with manual results

In this paper, we extract 8 key points both in outer and inner lip contour, P1 to P8. Fig. 9 gives the final resultant image with 8 key points. In order to show the advantage of our method, the 8 key points are marked manually once again. Fig. 10 gives the manual results by mouse click.

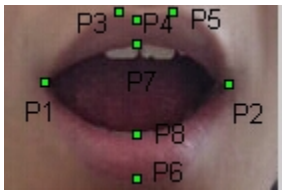


Fig. 9 All key points (P1 to P8)

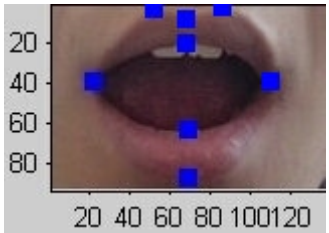


Fig. 10 Manual results by mouse click

Table 1 Manual results and results by our method

Key Point	Manual Results (x, y)[cm]	Results By our Method(x, y)[cm]	Error [cm]
P1	(21,40)	(21,39)	1.00
P2	(114,40)	(113,40)	1.00
P3	(59,3)	(58,4)	1.41
P4	(65,8)	(67,8)	2.00
P5	(87,4)	(85,4)	2.00
P6	(65,86)	(67,87)	2.24
P7	(65,21)	(67,20)	2.24
P8	(65,65)	(67,65)	2.00

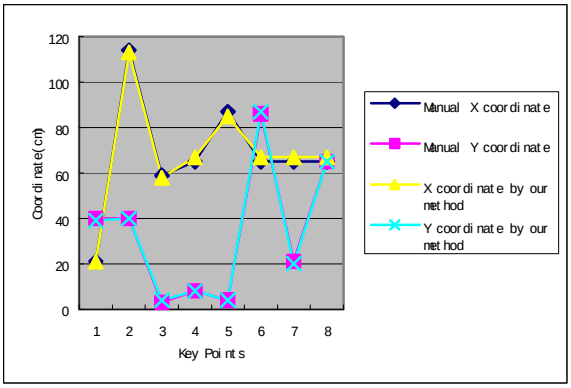


Fig. 11 Line chart for comparative study

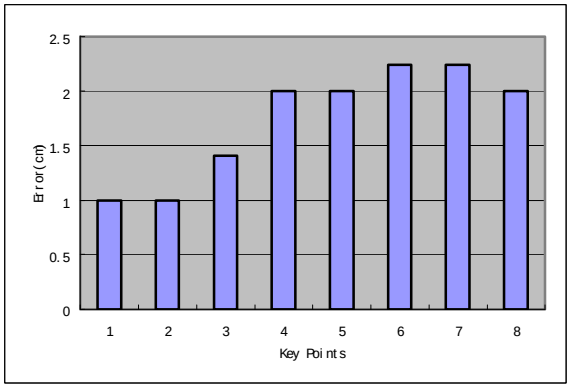


Fig. 12 Bar chart of error term

Table 1 shows the x and y coordinate values for key point computed manually (using mouse click) and by our method. Last column of this table gives an error from the desired results (manual results) for finding key points P1 to P8. This error term can be calculated using:

$$E = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}\ , \tag{9}$$

where E is an error term, (x_1, y_1) is manual key point while (x_2, y_2) is an key point using our method.

Fig. 11 shows the line chart for the comparative study of key points P1 to P8 calculated manually and by our method. The error term is shown in Fig. 12 using bar chart. For Fig. 2, we get better results in findings of key points.

Conclusions

Our method is introduced and implemented to extract the key points of lip contours. This proposed method gives the promising results for outer and inner lip contour from the given human mouth colored images. It is obvious that accuracy of extraction of key points in inner contours mainly depends on color spaces and gradients. In the future, the important work is to make our method robust along with accurate inner lip contour extraction.

Acknowledgements

This paper is supported by the Project of Beijing Natural Science Foundation of China (4143061) and the National Natural Science Foundation of China (61371142).

References

[1] K.K. Neely, Effect of visual factors on the intelligibility of speech. J. Acoust. Soc. Am. 28(1956) 1275-1277.
[2] W.H. Sumby, I. Pollack, Visual contribution to speech intelligibility in noise. J. Acoust. Soc. Am. 26(1954) 212-215.
[3] U. Saeed, J.L. Dugelay, Combining edge detection and region segmentation for lip contour extraction, Springer Berlin Heidelberg, (2010)11-20.
[4] K.S. Jang, Lip Contour Extraction based on Active Shape Model and Snakes, International journal of Computer Science and Network Security. 10(2007)148-153.
[5] N. Eveno, A. Caplier and P.Y. Coulon, A parametric model for realistic lip segmentation, IEEE 7th International Conference on Control, Automation, Robotics and Vision, 3(2002)1426-1431.
[6] M. Lievin, P. Delmas, P.Y. Coulon, F. Luthon and V. Fristot, Automatic lip tracking: Bayesian segmentation and active contours in a cooperative scheme, IEEE International Conference on Multimedia Computing and Systems, 1(1999) 691-696.
[7] Z. Wu, P.S. Aleksic and A.K Katsaggelos, Lip tracking for MPEG-4 facial animation, ICMI'02 Proceedings of the 4th IEEE International Conference on Multimodal Interfaces. (2002)293-298.
[8] G.I. Prajapati, N.M. Patel, DToLIP: Detection and tracking of lip contours from human facial images using Snake's method, International Conference on Image Information Processing (2011)1-6.
[9] Y. Gong, Speech recognition in noisy environments: A survey, Speech Communication, 16(1995) 261-291.
[10] J. LI, C. Cheng, T. Jiang and S.Grzybowski, Wavelet de-noising of partial discharge signals based on genetic adaptive threshold estimation, IEEE Transactions on Dielectrics and Electrical Insulation, 19(2012)543-549.
[11] R. Li, C. Bao, B. Xia and M. Jia, Speech enhancement using the combination of adaptive wavelet threshold and spectral subtraction based on wavelet packet decomposition, IEEE International Conference on Signal Processing, 1(2012)481-484.
[12] T. Wark, S. Sridharan, V. Chandran, An approach to statistical lip modelling for speaker identification via chromatic feature extraction, ICPR '98 Proceedings of the 14th International Conference on Pattern Recognition, 1(1998)123-125.
[13] N. Eveno, A. Caplier, P.Y. Coulon, Automatic and accurate lip tracking, IEEE Trans. Circuits

Syst. Video Technol. 14(2004)706-715.
[14] S. Sébastien, G. Vincent and C. Alice, Lip contour segmentation and tracking compliant with lip-reading application constraints, Machine Vision and Applications 24(2013)1-18.
[15] M. Raman, H. Aggarwal, Study and comparison of various image edge detection techniques, International journal of Image Processing, 3(2009)1-11.
[16] J.F. Canny, A computational approach to edge detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 6(1986)679-697.
[17] W. Frei and C. Chen, Fast Boundary Detection: A Generalization and a new algorithm, IEEE Trans. Comput. 10(1977)988-998.