LIPS DETECTION IN LOW RESOLUTION IMAGES

Ing. Jaroslav KADLEC, Doctoral Degree Programme (1)
Dept. of Computer Graphics, FIT, BUT
E-mail: kadlecj@fit.vutbr.cz

Supervised by: Dr. Pavel Zemčík

ABSTRACT

This paper presents a novel method for lips detection in low resolution images of human faces. The presented technique uses edge detection and color filtering for noise reduction and enhancement of the desired recognition of lips. The paper also discusses advantages of this method, its use, and future development.

1 INTRODUCTION

Although in present days there are lots of high resolution cameras and scanners, there still can remain pictures (e.g. portions of a high resolution images) with small dimensions which are important for image analysis (e.g. faces in high resolution image). This problem is usually very hard to solve. This paper describes a method for lips recognition in such low resolution images. Recognized lips would help to improve accuracy of speech recognition algorithms and person identification.

2 IMAGE PROCESSING

Any processed image is very important to filter before it is possible to find desired features. Noises should be removed by image smoothing and sometimes it is important to convert the image to other color space from RGB (e.g. grayscale). Finally, image enhancing is applied and such picture is ready for further analysis or recognition.

2.1 NON-RELEVANT PARTS REMOVAL

As there is only need for lip recognition, all other parts of the face and picture are not important and could cause misclassification. The advantage of low resolution images is the fact, that almost every face feature, including mouth, is edge or consists of very few edges. That's why edge detection can be used for image filtering. SUSAN (Smallest Univalue Segment Assimilating Nucleus) edge operator [3] has been used for edge detection. It is fast and has better results then other edge detection operators. Only a part of the SUSAN edge detection operator was used to get the edges enhanced as in this time edge directions and other edge information is not used.

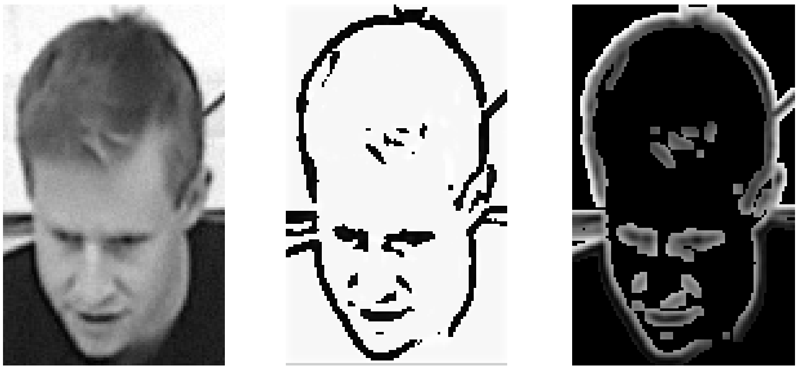

Fig. 1: *Detected edges and image subtraction*

The edge image is considered to be a mask. The original image is masked and only edge parts of the image are considered. The image contains only a few color pixels on potential edges, which could be desired lips or mouth. From now lips and mouth detection can be performed.

2.2 COLOR EXTRACTION

RGB color scheme of the image is not suitable for immediate processing as it contains a lot of mixed information about lightness etc. Another color scheme should be used. Very convenient color scheme is YCbCr as it separates luminance, blue chrominance and red chrominance. These colors are very convenient as the mouth is considered to contain high red and low blue components in comparison with other face regions. The YCbCr colors can be computed from the RGB as follows:

$$Y = 0.229 * R + 0.587 * G + 0.144 * B$$
$$C_B = 0.168 * R - 0.3313 * G + 0.5 * B + 128$$
$$C_R = 0.5 * R - 0.4187 * G - 0.0813 * B + 128$$

Fig. 2: *Computation of luminance, blue and red chrominance*

With this color scheme it is possible to start lip detection. To be more precise the pixels of the lips can be described by the probability, that current pixel is mouth pixel. Because lips pixels contain high red and low blue components (see text above) the lips detection can be correlated to the red chrominance (Cr):

$$isLipColor(x, y) = C_R(x, y)^2 * (C_R(x, y)^2 - \eta * \frac{C_R}{C_B})^2$$

Fig. 3: *isLipColor returns probability of lip color*

$$\eta = 0.95 * \frac{\sum_{(x,y)} C_R(x,y)^2}{\sum_{(x,y)} \frac{C_R(x,y)}{C_B(x,y)}}$$

Fig. 4: *Constant η fits final value in range 0..255*

isLipColor function returns values in range of <0, 255> where 0 represents values, which are not lip color similar and 255 for colors very similar to lip color. In next picture, each pixel of the filtered and edge masked picture 2 is computed by isLipColor function.

Fig. 5: *Picture generated by isLipColor function and its thresholded version*

2.3 LIP DETECTION

Light regions in left image of 5 represent very high probability of lip pixels of the processed image. After thresholding, all pixels with high probability level are presented in the right image of 5. Largest region in the picture represents mouth. This can be said thanks to edge filtering, so there will always be only few large regions among which lips would be the largest one. The largest region is recognized by image erosion. After several applications of simple “image erosion” filter, only few pixels remain. These remaining pixels are certainly from the largest region and can be used as a seed for color filling. Filled area is certainly the largest area, which is lip region and the detection is almost done. The center of the area is computed, its height, width and also deformation coefficient describing average distance of the pixels from the center, which can be used to determine mouth shape. Image 6 shows filled image and final output of the detected lips to the processed image.

Fig. 6: *Filled lips and the detected lip position*

3 FUTURE WORK

The algorithm works well on low resolution images. Future work will be focused on better region identification as there can be regions of similar size, which lead in lip misidentification. Also novel approaches using multiple masks based on more features such as hue color of the face, lip corners etc. should be used.

4 CONCLUSION

The presented method is very useful for detecting mouth position and size from the low resolution images. It is also independent on head position. Algorithm is very fast and can be used in real-time applications. The gained information can be used for identification of speakers in the video streams and may improve accuracy in speech recognition algorithms.

REFERENCES

[1] Hsu, Rein-Lien, Abdel-Mottaleb, Mohamed, Jain, Anil K., Face Detection in Color Images, IEEE ICIP 1999, pp 622-626

[2] Lai-Kan-Thon, Olivier, Lips Localization, Brno 2003

[3] Smith, S. M., Brady, J. M., SUSAN – a new approach to low level image processing, International Journal of Computer Vision, 23(1):45-78, May 1997

[4] Ahlberg, J.: A system for face localization and facial feature extraction, Linkoping University, Tech.Rep. LiTH-ISY-R-2172

[5] Albiol, A., Torres, L., Delp, E. J.: Optimum color spaces for skin detection, In Proceeding of the International Conference on Image Processing 2001, vol. 1, 122-124