Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

Signal Processing 86 (2006) 399–413

SIGNAL PROCESSING

www.elsevier.com/locate/sigpro

Parametric models for facial features segmentation

Z. Hammal, N. Eveno, A. Caplier*, Py. Coulon

LIS, INPG, 46 Avenue Félix Viallet, 38031 Grenoble Cedex, France

Received 21 January 2005; received in revised form 2 June 2005; accepted 2 June 2005
Available online 25 July 2005

Abstract

In this paper, we are dealing with the problem of facial features segmentation (mouth, eyes and eyebrows). A specific parametric model is defined for each deformable feature, each model being able to take into account all the possible deformations. In order to initialize each model, some characteristic points are extracted on each image to be processed (for example, eyes corners, mouth corners and brows corners). In order to fit the model with the contours to be extracted, a gradient flow (of luminance or chrominance) through the estimated contour is maximized because at each point of the searched contour, the gradient (of luminance or chrominance) is normal. The definition of a model associated to each feature offers the possibility to introduce a regularization constraint. However, the chosen models are flexible enough to produce realistic contours for the mouth, the eyes and the eyebrows. This facial features segmentation is the first step of a set of multi-media applications.

Keywords: Parametric models; Deformable model; Facial features; Segmentation

1. Introduction

Facial features deformations contribute to the communication between humans: for example, lip reading allows to improve the understanding of a noisy vocal message and it is also a support of communication with hard of hearing people.

The aim of our work is the extraction of the contours of permanent facial features (mouth, eyes and eyebrows) with enough accuracy to be able to improve the human-to-human communication through a machine.

Several applications are under consideration:

- The outer contour of lips can be used in a mobile phone application in order to improve the vocal message quality in case of noisy transmission.
- This contour is going to be used in a project of phone device for hard of hearing people. This requires speech synthesis from lip shape and motion.
- The detected iris contour is going to be used to evaluate the vigilance or interest level of a user

*Corresponding author. Tel.: +33 476 574 363; fax: +33 476 577 90.
E-mail address: alice.caplier@lis.inpg.fr (A. Caplier).

0165-1684/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2005.06.006

by the analysis of his frequency blinking. It will also be used to estimate the gaze direction of a PC user.
- All the detected contours (eyes, brows and lips) are used in an automatic system of facial emotion recognition based on video data.

All the considered applications require very accurate contours and real-time processing.

In this paper, we propose new parametric models well suited for eyes, mouth and brows segmentation. For the initialisation of each chosen model, some characteristic points (eyes and brows corners for example) are extracted. Initial models are then deformed in order to maximize a gradient flow (of luminance and/or chrominance). The originality of our approach is the definition of parametric models able to take into account all the possible deformations of each considered feature. This yields to realistic and accurate segmentation.

Many algorithms have been proposed to solve the problem of lip segmentation. Some methods use low-level spatial information such as color [1]. No smoothing constraint or shape constraint is considered so that the extracted contours are often of coarse quality. In [2], a *linear discriminant analysis* (LDA) is used to segment lip pixels from skin pixels. This yields to lip contours. Though the discriminant analysis is followed by a smoothing step, the segmentation remains noisy. *Snakes* [3] have been widely used for lip segmentation ([4,5]) because snakes can take into account in a same framework smoothing and elasticity constraints. Snake-based methods yield to interesting results but the main drawback is the tuning of several parameters. Moreover, the quality of the segmentation is dependent of initialisation. A priori shape models can be used to obtain smoother contours: the extracted contour belongs to the possible space of the lip shape. For example, Active Shape Models (ASM) has been used [6]. But this approach requires a very huge learning database in order to be able to take into account all the possible deformations of a speaking mouth. And each image of the learning database has to be calibrated with very high accuracy: face orientation and illumination conditions have to be controlled. One possible solution to avoid learning stage is the use of parametric models [7].

Segmentation of eyes and brows is simpler because such features are less deformable than the mouth. Two kinds of approach exist. First of all a coarse localization of these features based on luminance information is extracted (valley images for example) [8,9]. These approaches yield to a very few accurate contours detection for eyes and brows. The second approach introduces models to be related to the searched contours [10].

In our approach, parametric models are considered. In Section 2, chosen parametric models for mouth, eyes and brows are described and justified. In the pre-processing step described in Section 3, face illumination variations are removed and iris circular contour is extracted. Section 4 presents the algorithms for the automatic extraction of facial characteristic points used for the initialisation of the parametric models. Section 5 describes the deformation model process according to gradient information to make the model coincide with the contours of the processed image. Section 6 gives qualitative and quantitative results in order to evaluated the pertinence of the chosen parametric models and the quality of the segmentation. The accuracy of the extracted contours is compliant with an application of facial emotion recognition.

2. Model choice

The analysis of face images coming from different databases shows that the models that have been proposed until now for mouth, eyes and brows are too rigid to obtain realistic contours.

2.1. Mouth model

Several parametric models have been proposed to model the lip contour. Tian [11] uses a model made of parabolas. This is very simple but the precision is limited (see Fig. 1a). Others authors propose to model the upper lip contour with 2 parabolas [12] or to use quartics [13]. It improves accuracy, but the model is still limited by its