

文章编号 1004-924X(2012)12-2768-05

基于主成分分析的唇部轮廓建模

王丽荣¹, 王建蕾^{1,2*}

- (1. 长春大学 电子信息工程学院, 吉林 长春 130022;
2. 长春理工大学 电子信息工程学院, 吉林 长春 130022)

摘要:研究了基于主成分分析(PCA)的唇部轮廓建模方法。首先,对 5 000 个样本唇部轮廓进行标定并对标定的坐标数据进行 Procrustes 分析,使数据归一化。然后,通过 PCA 算法寻找形变模式,在保持形变范围内最大限度地降低数据维数并利用所得到的均值和特征向量构建唇部轮廓模型。最后,利用 PCA 得到的前 16 种模式所建立的模型对 5 000 个样本原始的唇部轮廓进行重构。实验结果显示:PCA 得到的前 4 种模式分别描述了唇部角度、下唇、尺度以及唇角等的形变信息,其余模式描述了唇部更细致的形变,模型重构的唇部轮廓与相应样本原始唇部轮廓的每个特征点之间平均差异均不大于 0.6 个像素宽。结果表明所建唇部模型能满足特征定位精度要求。

关键词:唇部轮廓;特征提取;主成分分析法;建模

中图分类号:TP391 **文献标识码:**A **doi:**10.3788/OPE.20122012.2768

Lip contour modeling based on PCA

WANG Li-rong¹, WANG Jian-lei^{1,2*}

(1. School of Electronics and Information Engineering, Changchun University, Changchun 130022, China

2. School of Electronic and Information Engineering, Changchun University of Science and Technology, Changchun 130022, China)

* Corresponding author, E-mail: wangjlei118@163.com

Abstract: A modeling method of lip contours was proposed based on Principle Component Analysis (PCA). Firstly, the lip contours of 5 000 training samples were labeled, and the Procrustes analysis was performed on the coordinates gotten by labeling to normalize the data. Then the PCA was used to identify modes in data and compress the data dimension without losing the lip contour information. Furthermore, the lip contour model was constructed by using a mean value and eigenvectors gotten by PCA. Finally, the model constructed by the first 16 modes gotten by PCA was taken to reconstruct the lip contours of the original 5 000 samples. Experimental results indicate that the first 4 modes respectively describe the rotation, lower lip, scale and corners of the lip, and other modes describe more detailed lip variation. The mean difference of every feature point between the lip contour gotten by model reconstruction and the original contour is less than a width of 0.6 pixels. The model can satisfy

收稿日期:2012-07-17;修订日期:2012-09-25.

基金项目:吉林省科技厅自然科学基金资助项目(No. 20101518);教育部重点项目(No. 201049)

the precision requirements of the feature location.

Key words: lip contour; feature extraction; Principle Component Analysis(PCA); modelling

1 引言

近年来,语音识别技术的研究取得了突破性进展,已经接近实用阶段。运用语音识别技术将说话者的语音转换为文字,让听障者可以直接“看到”说话内容,这种技术具有直观、实时的优点。语音识别技术的主要瓶颈是语音识别率,语音识别率不高会影响传递信息的品质,使得语音识别系统无法达到预期效果,甚至基本无法正常使用。因此,研究提高语音识别率的方法是语音识别技术应用的关键,而如何将语音信息与唇部信息进行关联以提高识别率是该项研究的热点,其中提取唇部的相关特征^[1]并对唇动视频进行特征跟踪^[2-5]尤其受人关注。

早期唇部特征提取的研究中,描述唇部轮廓主要采用可变形模板的方法^[6],该方法由一系列手工编码的多项式即可变形模板对唇部轮廓进行定位。由于多项式的初始选择限制了可变形模板的形变范围,所以这种方法不能很好地定位唇部轮廓的细节。20世纪90年代中期,兴起了采用主动轮廓模型对唇部轮廓进行描述的方法^[7],即使用能量最小化曲线来对唇部轮廓进行定位,这种方法可以定位唇部轮廓细节,但是定位的结果与主动轮廓初始位置有关,而且对复杂形变的目标定位不是很理想。

从2000年开始,主动形状模型(ASM)方法在图像特征提取的研究中逐渐成为热点^[8-9],这种方法通过一系列特征点来对目标区域轮廓进行描述,并对大量训练样本目标区域的轮廓坐标值进行统计,进而建立轮廓模型。与主动轮廓模型不同,这种方法运用了全局形状限制来使模型只能在合理的范围内形变,从而保证了模型的鲁棒性。目前,主动形状模型大部分应用在人脸和医学图像的研究中。本文基于主动形状模型,对样本外唇和内唇进行标定,利用主成份分析(PCA)算法对唇部轮廓进行建模,并利用所建立的模型重构样本原始轮廓,取得了较好效果。

2 唇部模型建立

2.1 建模总体过程

建模的总体过程包括样本标定,Procrustes分析,主成分分析(PCA),如图1所示。

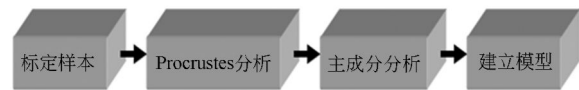


图1 建模过程示意图

Fig. 1 Diagram of lip modeling process

2.2 样本标定

唇部分为外唇和内唇两部分^[10],外唇轮廓可表示唇部的大部分形变,内唇轮廓可表示唇部张开程度,本文采用同时标定外唇和内唇的方案^[8,10]。Tim Cootes在他的主页中描述了如何采用68个点对人脸进行标定^[11],其中唇部采用19个点,这19个点不仅清晰地描述了唇部外唇和内唇轮廓,而且相对于较密集的标定方案,该方法更有利于提高后续特征定位的实时性。本文也采用19个点的唇部标定方案,如图2所示,特征点的坐标值是建模关键。



图2 唇部标定方案

Fig. 2 Scheme of lip labeling

2.3 Procrustes分析

唇部可能出现在样本不同位置,如果直接对

坐标数据进行分析,那么建立的模型没有意义。另外,轮廓模型只和形状变化有关,和样本唇部轮廓的尺度以及角度等信息无关,采用 Procrustes 分析方法可以使坐标数据归一化。Procrustes 分析方法^[12]在这里不做详细介绍,处理结果使所有轮廓数据均与第一个样本轮廓数据对齐。

2.4 主成分分析

主成分分析(PCA)是一种识别数据模式并将数据表现为一种突出它们相似性和差异性的方法。它可以在高维数据中寻找模式,并在保持信息量的前提下最大限度降低数据维数^[13]。

第一步 进行 Procrustes 分析,将得到的每个样本的轮廓数据表示为

$$\mathbf{x} = (x_1, y_1, x_2, y_2, \dots, x_{18}, y_{18}, x_{19}, y_{19})^T. \quad (1)$$

第二步 计算样本向量的均值,以第 1 个特征点为例

$$\bar{x} = \frac{1}{5\,000} \sum_{i=1}^{5\,000} x_i, \quad (2)$$

其中: x_i 为第 i 个样本的第 1 个特征点。

第三步 计算每个样本的每个特征点与均值的差值,以第一个样本第 j 个特征点为例

$$d_{x_{1,j}} = x_{1,j} - \bar{x}_{1,j}, \quad (3)$$

其中: $x_{1,j}$ 为第一个样本的第 j 个特征点, $\bar{x}_{1,j}$ 为其对应的均值。

第四步 构建协方差矩阵

$$\mathbf{a} = \begin{bmatrix} d_{x_{1,1}} & \cdots & \cdots & \cdots & d_{x_{5000,1}} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ d_{x_{1,38}} & \cdot & \cdot & \cdot & d_{x_{5000,38}} \end{bmatrix}, \quad (4)$$

$$\text{cov} = \mathbf{a} \times \mathbf{a}^T, \quad (5)$$

其中: $d_{x_{1,1}}$ 为第一个样本第 1 个特征点的 x 坐标与其均值的差值, $d_{x_{1,38}}$ 为第一个样本第 19 个特征点的 y 坐标与其均值的差值。

第五步 求出协方差矩阵的特征值 cov 和特征向量,本文保留了特征值从大到小排列的前 16 个值对应的特征向量作为唇部变化的主要模式。

第六步 建立模型

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{p}\mathbf{b}, \quad (6)$$

其中: $\mathbf{x} = (x_1, y_1, x_2, y_2, \dots, x_{19}, y_{19})^T$ 为代表轮廓

的 19 个特征点, $\bar{\mathbf{x}}$ 为样本特征点的均值向量, $\mathbf{p} = [p_1 \ p_2 \ \cdots \ p_{16}]$ 为形变的 16 种主要模式构成的矩阵, \mathbf{b} 为其对应的权向量。

唇部轮廓模型建立完成后,利用 PCA 去掉了 22 种不明显的变化模式,然后利用保留的 16 种模式对 5 000 个样本的唇部轮廓进行重构,以便求得模型重构的样本唇部轮廓与样本原始轮廓之间的差异。

$$FD = \mathbf{p}^T \times (d_{x_1}, \dots, d_{x_{38}})^T, \quad (7)$$

$$\mathbf{x}_{\text{rec}} = \mathbf{p} \times FD + \bar{\mathbf{x}}, \quad (8)$$

其中: d_{x_1} 为样本第 1 个特征点的 x 坐标值与其均值的差值, $d_{x_{38}}$ 为样本第 19 个特征点的 y 坐标值与其均值的差值, \mathbf{x}_{rec} 为重构后的相应样本轮廓。

求得误差利用

$$\text{err} = \sqrt{\frac{1}{38} \sum_{i=1}^{38} (x_i - x_{\text{rec},i})^2}, \quad (9)$$

其中: x_i 为某一样本原始轮廓的第 i 个特征点的坐标值, $x_{\text{rec},i}$ 为重构后的相应特征点的坐标值。

3 实验结果与分析

3.1 实验结果

本文采用的样本数据库是 FG-NET talking face video^[11],该数据库的 5 000 帧图像涵盖基本唇部形变范围,所以可保证建立的模型具有泛化性。样本总体大部分都位于均值的 ± 3 倍标准差之间^[8],图 3~图 7 分别反映了均值唇和在前 4 种模式下模型的变化情况。

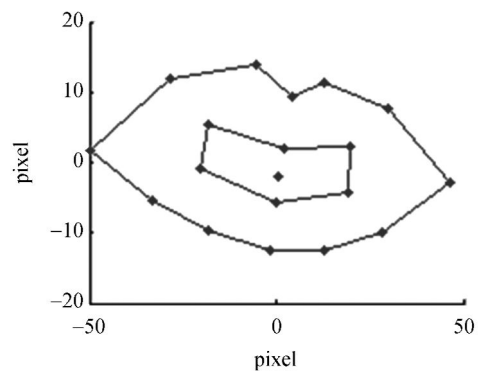


图 3 通过 PCA 获得的均值唇

Fig. 3 Mean lip gotten by PCA

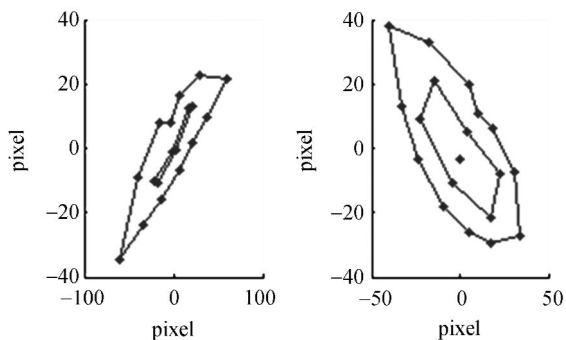


图 4 在 3 倍标准差(左)和负 3 倍标准差(右)的模式一下的唇部模型变化情况

Fig. 4 Mode 1 of lip contour variation about three times of the standard deviation(left) and negative three times of the standard deviation (right).

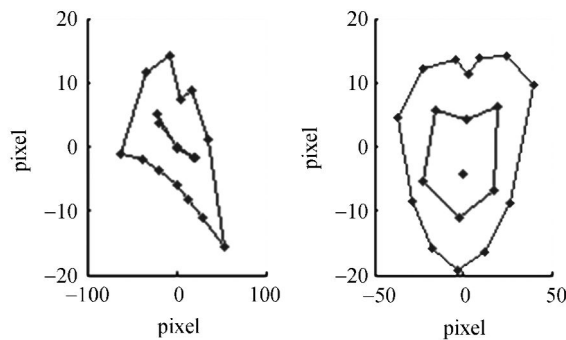


图 5 在 3 倍标准差(左)和负 3 倍标准差(右)的模式二下的唇部模型变化情况

Fig. 5 Mode 2 of lip contour variation about three times of the standard deviation(left) and negative three times of the standard deviation (right).

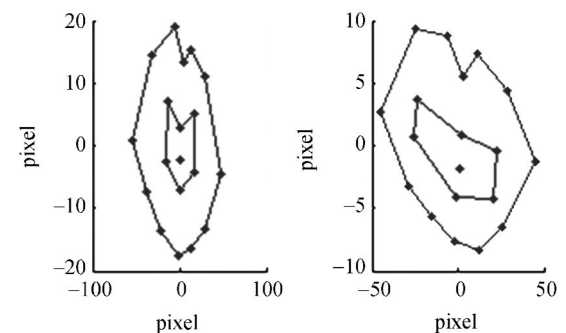


图 6 在 3 倍标准差(左)和负 3 倍标准差(右)的模式三下的唇部模型变化情况

Fig. 6 Mode 3 of lip contour variation about three times of the standard deviation(left) and negative three times of the standard deviation (right).

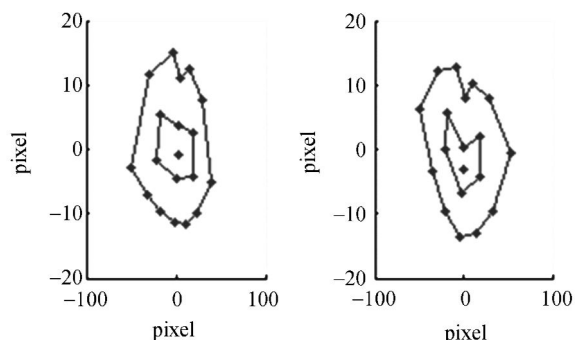


图 7 在 3 倍标准差(左)和负 3 倍标准差(右)的模式四下的唇部模型变化情况

Fig. 7 Mode 4 of lip contour variation about three times of the standard deviation(left) and negative three times of the standard deviation (right).

从上述图中可以看到,模式一主要描述了唇部角度的变化情况,模式二描述了下唇的变化情况,模式三描述了唇部尺度的变化,模式四描述了唇角的变化,其余模式描述了唇部更细致的形变。

3.2 误差分析

PCA 舍去了 22 种唇部变化模式,利用保留的 16 种模式,即 PCA 得到的前 16 个特征值对应的特征向量重构样本的唇部轮廓。得到重构轮廓与原始轮廓每个特征点之间的平均差异如公式(9)所求。如图 8 所示,在第 4 000 个样本前后特征点的平均差异是在 0.45 到 0.6 个像素宽之间,其余都在 0.2 以下。

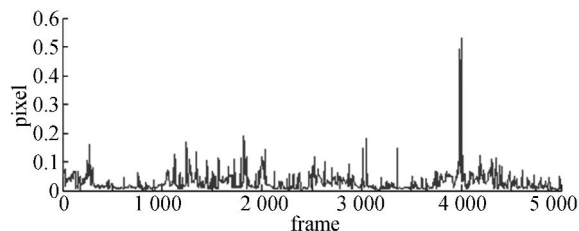


图 8 模型重构的唇部轮廓与相应样本原始唇部轮廓的每个特征点之间平均差异的分布

Fig. 8 Distribution of mean difference of every feature point between lip contours gotten by model reconstruction and original contours.

4 结 论

本文根据主动形状模型算法原理,对大量样本的唇部进行了标定,并对标定的坐标数据进行 Procrustes 分析。采用 PCA 算法构建唇部轮廓模型,利用所建立的模型对原始的 5 000 个样本

唇部轮廓进行重构。实验结果表明,模型重构的唇部轮廓与相应样本原始唇部轮廓的每个特征点之间的平均差异不大于 0.6 个像素宽,基本集中在 0.2 个像素宽以下。所建模型满足唇部轮廓特征定位的精度要求,为视觉特征提取及双模态语音识别奠定了基础。

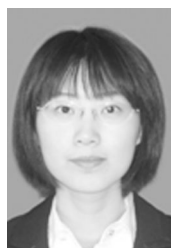
参考文献:

- [1] RONG L W, GUANG X Y, LEI J W, *et al.*. Facial expression recognition based on local texture features[C]. 2011 IEEE 14th International Conference on Computational Science and Engineering, 2011:543-546.
- [2] 吴本涛,吴敏渊,曾霖. 自适应搜索的快速分块跟踪[J]. 光学 精密工程, 2011, 19(3):703-708.
WU B T, WU M Y, ZENG L. Fast fragment based tracking using adaptive search[J]. *Opt. Precision Eng.*, 2011, 19(3):703-708. (in Chinese)
- [3] 赵燕燕,基于视频图像的唇部检测与跟踪方法研究[D]. 长春:长春理工大学,2008.
ZHAO Y Y. *Research on the methods of lip detecting and tracking based on video image* [D]. Changchun: Changchun University of Science and Technology, 2008. (in Chinese)
- [4] 王国良,刘金国. 基于粒子滤波的多自由度运动目标跟踪[J]. 光学 精密工程, 2011, 19(4):864-869.
WANG G L, LIU J G. Moving object tracking with multi-degree-of-freedom based on particle filters[J]. *Opt. Precision Eng.*, 2011, 19(4):864-869. (in Chinese)
- [5] 颜佳,吴敏渊. 遮挡环境下采用在线 Boosting 的目标跟踪[J]. 光学 精密工程, 2012, 20(2):439-446.
YAN J, WU M Y. On-line boosting based target tracking under occlusion[J]. *Opt. Precision Eng.*, 2012, 20(2):439-446. (in Chinese)
- [6] YUILLE A L, HALLINAN P, COLEN D S. Feature extraction from faces using deformable templates[J]. *International Journal of Computer Vision*, 1992, 8(2):99-111.
- [7] DAVATZIKOS C A, PRINCE J L. An active contour model for mapping the cortex[J]. *IEEE Transactions on Medical Imaging*, 1995, 14(1):65-80.
- [8] COOTES T F, TAYLOR C J, COOPER D H, *et al.*. Active shape model-their training and application[J]. *Computer Vision and Image Understanding*, 1995, 61(1):38-59.
- [9] CHEN S Y, ZHANG J. Detection and amendment of shape distortions based on moment invariants for active shape models[J]. *Image Processing, IET*, 2011, 5(3):273-285.
- [10] LUETTIN J. *Visual speech and speaker recognition*[D]. University of Sheffield, 1997:41-70.
- [11] http://www.isbe.man.ac.uk/~bim/data/talking_face/talking_face.html[OL].
- [12] GOWER J C. Generalized procrustes analysis[J]. *Psychometrika*, 1975, 40(1):33-51.
- [13] JOLLIFFER I T. *Principal Component Analysis: A Beginner's Guide*[M]. Wiley Online Library, 2012.

作者简介:



王丽荣(1966—),女,辽宁沈阳人,教授,硕士生导师,1989年于吉林工业大学获得学士学位,2002年于长春理工大学获硕士学位,2006年于吉林大学获博士学位,主要从事信息处理、模式识别等方面的研究。E-mail: wlr-661001@163.com



王建蕾(1987—),女,吉林集安人,硕士研究生,2010年于黑龙江大学获得学士学位,主要从事模式识别方面的研究。E-mail: wangjilei118@163.com