

# 基于粗略到精细分类的面部表情识别方法

崔洁, 冯晓毅

(西北工业大学电子信息学院, 西安 710072)

**摘要:** 为了更准确地识别人的表情, 在识别人脸7种基本表情(愤怒、厌恶、恐惧、高兴、无表情、悲伤和惊讶)时, 采用了局域二值模式技术提取面部特征, 进行由粗略到精细的表情分类。在粗略分类阶段, 7种基本表情中的2种表情被选为初步分类结果(候选表情)。在精细分类阶段, 选用计算加权卡方值确定最终分类结果。采用日本的Jaffe表情数据库来验证算法性能, 对陌生人表情的识别率为77.9%, 其结果优于采用同样数据库的其他方法, 且易于实现。

**关键词:** 面部表情; 局域二值模式; 表情分类; 加权卡方值

## Method of Facial Expression Recognition Based on Coarse-to-fine Classification

CUI Jie, FENG Xiaoyi

(School of Electronic and Information, Northwestern Polytechnical University, Xi'an 710072)

**【Abstract】** This paper presents an effective method for facial expression recognition. It analyzes seven basic expressions: angry, disgust, fear, happiness, neutral, sadness and surprise. The local binary pattern(LBP) operator is used to extract face appearance features. A two-stage classification method is proposed. At the first (coarse classification) stage, two expression candidates from initial seven are selected. At the second (fine classification) stage, one of the two candidate classes is verified as final expression class by the Chi square statistic. Our algorithm is tested on the Jaffe database and the recognition rate for a new person's expression is 77.9%, which is higher than all the other reported methods.

**【Key words】** Facial expression; Local binary pattern(LBP); Expression classification; Chi square statistic

### 1 概述

人脸面部表情识别是计算机视觉领域一个富有挑战性的研究课题, 在人机接口、心理学研究等方面具有重要意义。由于静态面部图像包含表情信息较少, 因此从静态图像中识别表情比从图像序列中进行难度更大。在有些情况下, 单幅图像中包含的信息足以用来识别表情, 而且在很多应用中也都需要识别单幅图像中的表情。近年来出现了很多识别静态图像中表情的方法<sup>[1]</sup>。

Gao等人用线性边缘图(LEM)描述表情<sup>[2]</sup>, 把面部表情识别转换成了通过LEM方法描述出的脸部之后, 再与画线的表情模型进行匹配。Shinohara和Otsu<sup>[3]</sup>把高阶局部自动相关(HLAC)特征和Fisher权值图结合起来形成一个特征向量, Fisher判别式分析法常用于表情识别。Lyons等人<sup>[4]</sup>提出了一种基于模板的表情识别方法, 用二维Gabor小波表示脸部, 利用主分量分析法(PCA)压缩数据, 用线性判别分析法(LDA)进行表情分类。Buciu等人<sup>[5]</sup>将Gabor小波(或ICA)和SVMS结合起来识别面部表情。Fasel等人用的是神经网络方法来进行表情的学习和识别。然而, 这些方法或者需要人工定位很多面部关键点, 不能实现自动化的识别, 不便于实际应用, 或者识别率很低, 均不能很好地实现对表情的识别。

为了克服现存的特征信息提取不全面和对表情识别率不高等问题, 本文采用能够很好地进行纹理分析的局域二值模式(LBP)技术进行面部特征提取, 并在基于模板的分类方法基础上进行改进, 采用由粗略到精确的表情分类方案。在粗分类阶段, 比较测试样本与7种表情模板间的差异, 选择2个与测试样本最相似的表情作为候选表情; 在细分类阶段, 进

一步确定候选表情中的一种作为最终分类结果。

### 2 面部特征提取

近年来, 用于纹理分析中的局域二值模式(LBP)技术已经成功地应用到人脸的识别, 本文采用LBP提取面部特征。图1描述了基本的LBP算法。图1(a)是一个3×3矩形区域, 以其中中心像素灰度值为阈值, 比较其邻域8个像素的灰度值的大小, 产生该区域的二进制模式(如图1(b)), 该模式可用二进制码表示。将二进制码转换成十进制数, 即得中心像素的LBP码。图像中各像素的LBP码对应了其邻域的灰度分布情况。图像中某区域LBP码的直方图可用来描述该区域的纹理结构。在图1中, 二进制码为01101101, LBP值为109。



(a)初始邻域 (b)二进制模式

图1 基本的LBP变换

本文采用与文献[6]类似的方法进行面部特征描述:

(1)将脸部图像划分成若干(根据多次实验结果, 选择将图像划分成10×8=80块)不重复的区域(见图2(a))。

(2)计算每一区域的LBP码直方图。

**基金项目:** 芬兰 CIMO 基金资助项目; 国家留学基金资助项目; 西北工业大学“英才计划”基金资助项目

**作者简介:** 崔洁(1981-), 女, 硕士, 主研方向: 面部表情识别; 冯晓毅, 博士、副教授

**收稿日期:** 2006-03-22 **E-mail:** lcjcici@126.com

(3)各个区域的 LBP 直方图构成特征向量,描述脸部的纹理分布(见图 2(b))。在图 2 中,脸部图像被分为 80 个小区域。

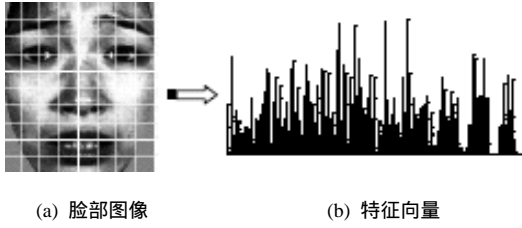


图 2 人脸特征提取

本特征提取方法基于如下考虑:表情总是伴随着面部纹理及特征区域(如眉毛、眼睛、嘴巴等)几何形状的变化而变化,基于 LBP 技术的特征描述可以反映这种变化。

### 3 表情的 2 级分类

在基于模板的表情分类中,模板可能是一幅图像,或者是一个特征向量。一般而言,模板形成过程中会使一些重要的面部表情特征由于脸部未对准、个人表情差异较大等原因被削弱<sup>[3]</sup>,因此基于模板的方法识别率不高。但是基于模板的分类方法简单、易于实现。

为提高基于模板的表情分类方法的识别率,本文采用由粗略到精细的 2 步分类方案。在粗略分类阶段,7 类表情的分类问题被转换为 2 类表情的分类问题。在精确分类阶段,进行两类表情的判断,得到最终结果。

#### 3.1 表情的粗略分类

该阶段利用卡方值来测量模板和测试样本之间的差异。

设集合  $X$  包含 7 个子集合,  $n$  个训练样本,每个子集合包含一种基本表情的训练样本。设第  $c$  个子集合为

$$X_c = \{x_i^c | 0 \leq i < n_c\}$$

其含  $n_c$  个样本,  $x_i^c$  是其中第  $i$  个样本。有

$$X = \bigcup_{c=1}^7 X_c \quad (1)$$

$$n = \sum_{c=1}^7 n_c$$

设  $m^c$  是第  $c(c=1,2,\dots,7)$  类表情的模板,则

$$m^c = \frac{1}{n_c} \sum_{i=0}^{n_c-1} x_i^c \quad (2)$$

$S$  是测试样本,卡方值  $\chi^2$ <sup>[6]</sup> 为

$$\chi^2(s, m^c) = \sum_{j=0}^{d-1} \frac{(s_j - m_j^c)^2}{s_j + m_j^c} \quad (3)$$

与测试样本差异最小的模板对应的表情可被确定为该测试样本的表情。

在用上述方法进行表情分类的实验中,可以发现,超过 50% 的错误分类结果具有一个共性:与测试样本最匹配的表情不是实际表情,而次匹配的表情为实际表情,而且测试样本与最匹配、次匹配表情间的差异均较小。

这说明上述方法可有效辨别特征差别比较明显的表情,如高兴和惊讶,而对一些特征区别小的表情效果不佳,如不能很好地辨别厌恶和悲伤。为了进一步提高分类效果,在在粗略分类阶段,最匹配和次匹配表情组成候选表情集  $cc = \{c1, c2\}$ , 使得

$$\chi^2(s, m^{c1}) = \min_{1 \leq c \leq 7} \chi^2(s, m^c) \quad (4)$$

$$\chi^2(s, m^{c2}) = \min_{1 \leq c \leq 7, c \neq c1} \chi^2(s, m^c)$$

### 3.2 表情的精确分类

为了正确辨别最匹配和次匹配表情,选用加权卡方值计算其对应模板与测试样本间的差异,即

$$\chi_w^2(s, m^c) = \sum_{j=0}^{d-1} w_j \frac{(s_j - m_j^c)^2}{s_j + m_j^c} \quad (5)$$

加权系数的计算如下:

设  $c1$ 、 $c2$  2 类训练样本的特征向量为  $d$  维,训练样本的个数分别为  $n_1$  和  $n_2$ 。由  $c1$ 、 $c2$  2 类训练样本组成训练集,设其中心为  $\mu$ , 训练样本各自的中心分别为  $\mu^{c1}$ 、 $\mu^{c2}$ , 则

$$\mu_j = \frac{1}{n_1 + n_2} \left( \sum_{i=0}^{n_1} x_j^{c1} + \sum_{i=0}^{n_2} x_j^{c2} \right) \quad (6)$$

$$\mu_j^{c1} = \frac{1}{n_1} \sum_{i=0}^{n_1} x_j^{c1} \quad \mu_j^{c2} = \frac{1}{n_2} \sum_{i=0}^{n_2} x_j^{c2} \quad (7)$$

设 2 类训练样本特征向量第  $j$  维分量的类间方差和类内方差分别为  $\sigma_{Bj}^2$ 、 $\sigma_{Wj}^2$ , 则

$$\sigma_{Bj}^2 = \frac{1}{n_1 + n_2} \left[ \sum_{i=0}^{n_1} (x_{ij}^{c1} - \mu_j)^2 + \sum_{i=0}^{n_2} (x_{ij}^{c2} - \mu_j)^2 \right] \quad (8)$$

$$\sigma_{Wj}^2 = \frac{1}{n_1 + n_2} \left[ \sum_{i=0}^{n_1} (x_{ij}^{c1} - \mu_j^{c1})^2 + \sum_{i=0}^{n_2} (x_{ij}^{c2} - \mu_j^{c2})^2 \right] \quad (9)$$

可得加权系数如下:

$$k_j = \frac{\sigma_{Bj}^2}{\sigma_{Wj}^2}, j = 1, 2, \dots, d \quad (10)$$

$$w_j = \frac{k_j}{\sum_{i=0}^d k_i}, j = 1, 2, \dots, d \quad (11)$$

$w_j$  这个系数描述了最匹配和次匹配表情特征向量各分量的差异程度。式(11)即为加权系数的计算式,代入式(5)可得计算对应模板与测试样本间差异的加权卡方值。

为了进一步加强模板的有效性,在精确分类阶段,采用多个模板来描述表情特征。

精确分类方法如下:(1)对每种基本表情构造多个模板,每个模板对应训练集合中一个人的表情。(2)对每对候选表情对,采用上述方法,确定一组权值。(3)利用加权卡方值测量测试样本与 2 个表情模板的差异。(4)选择差异最小的表情作为最终表情分类结果  $c$ , 满足

$$\chi^2(s, m^c) = \min(\chi^2(s, m^{c1}), \chi^2(s, m^{c2})) \quad (12)$$

## 4 实验分析

本实验采用的数据库是日本女性的面部表情(JAFFE)数据库。该数据库包含了 10 个人的 213 幅图像。每个人做出 7 种基本表情(高兴、悲伤、惊讶、愤怒、厌恶、恐惧、无表情),而且每种表情都做出 3 次或 4 次。

图 3(a)所示为数据库中的一些样本。在实验中,利用 CSU 面部识别评估系统<sup>[7]</sup>中的预处理子系统来对人脸大小进行归一化处理,结果如图 3(b)所示。

在图 3 中,初始图像的分辨率为 256\*256,预处理后的为 150\*128。



图 3 日本女性面部表情数据库图和预处理结果

为了将本方法与之前的方法进行比较,本文采用与其相同的样本训练与测试方法,即选择由 9 个人的 193 幅表情图 (下转第 172 页)