

基于 Gabor 特征和混合高斯模型的人脸表情分析

刘伟锋¹,汪增福²,卢纪丽³

LIU Wei-feng¹,WANG Zeng-fu²,LU Ji-li³

1.中国石油大学(华东)信息控制工程学院,山东 东营 257061

2.中国科学技术大学 自动化系,合肥 230027

3.枣庄学院,山东 枣庄 277160

1.College of Information and Control Engineering,China University of Petroleum(East China),Dongying,Shandong 257061,China

2.Department of Automation,University of Science and Technology of China,Hefei 230027,China

3.Zaozhuang University,Zaozhuang,Shandong 277160,China

E-mail:liuwfxy@ustc.edu

LIU Wei-feng,WANG Zeng-fu,LU Ji-li.Facial expression analysis based on Gabor features and Gaussian Mixture Model.Computer Engineering and Applications,2008,44(26):191-195.

Abstract: This paper analyzes the nature of the facial expressions and proposes a new facial expression analysis method using Gaussian Mixture Model(GMM) based on Gabor wavelet and principal component analysis.First,the GMM parameters of each facial expression are formed.The probability of every input facial image belongs to one facial expression class can be computed based on the GMM.Then the expression analysis and recognition of facial images is achieved.The experiments on a facial expression database JAFFE have been conducted.The distribution of different expressions is analyzed,and the qualitative and quantitative description of the facial expression is accomplished.The experimental results show that the facial expression analysis method proposed in this paper can rationally represent the daily facial expressions.

Key words: facial expression analysis;facial expression recognition;Gaussian Mixture Model(GMM);Gabor features

摘 要:首先对人脸表情的特点进行分析,提出了利用 Gabor 小波特征、主分量分析(PCA)结合混合高斯模型的人脸表情分析方法,并在人脸表情数据库 JAFFE 进行了实验。通过对不同表情的分布规律进行实验分析,实现了对表情的定性/定量分析。实验结果表明,提出的人脸表情分析方法能够对人脸表情进行恰当的表达和描述。

关键词:表情分析;表情识别;混合高斯模型;Gabor 特征

DOI:10.3778/j.issn.1002-8331.2008.26.059 **文章编号:**1002-8331(2008)26-0191-05 **文献标识码:**A **中图分类号:**TP391.41

1 前言

表情在人际交往中起着非常重要的作用。研究表明,人的情感会通过其面部表情自然地流露出来。从这个意义上来说,人脸表情识别的主要目的就是通过对输入人脸图像的分析来判断相关人所处的情感状态。近年来,众多学者从不同的角度提出了许多表情识别方法。按照所使用图像的不同,这些方法大体可以分为两类:基于动态图像的方法^[1-5]和基于静态图像的方法^[6-10]。其中,绝大多数方法基于对引发表情的情感的认识,把表情划分为若干类。例如,P.Ekman^[11]把表情划分为愤怒,高兴,惊讶,悲伤,恐惧,厌恶等6类。在这些系统中,人脸表情识别的过程就是判断输入人脸图像属于哪一类表情的过程。然而,这种做法和人的日常经验并不完全符合。实际生活中有些表情往往表现为几种基本表情的混合。例如,惊喜同时具有惊讶和高兴的特征。因此,硬性地将这些表情归属于某一类并不是非常可取的做法。

表情是由人的情感所引发的。目前,有代表性的情感理论主要有两种:离散情感空间理论和连续情感空间理论^[12]。前者把情感划分为类别数在4到8之间的若干个类别^[11],该理论使用上比较方便,也易于理解;然而,将人的情感和人脸表情简单归类的做法并不是很恰当。研究表明,实际中的许多表情往往表现出两种或两种以上的情感特征。与离散情感空间理论不同的是,后者认为情感空间是一个连续空间^[13-14],从感知的角度来看,该理论比较切合实际情况,但已有的相关工作并没有对人脸表情所对应情感状态的表示方式做出更多有实质意义的表述^[15],这给实际应用留下了很多亟待解决的问题。

鉴于此,本文利用 Gabor 小波特征和主分量分析(PCA),结合混合高斯模型对人脸图像中不同基本表情比重的分布规律进行分析,该方法不仅可以对表情进行识别,更能进一步对表情的不同基本表情成分进行分析。

作者简介:刘伟锋(1979-),男,博士,CCF会员,研究方向为表情识别、图像处理;汪增福(1960-),男,博士,教授;卢纪丽(1979-),女,硕士。

收稿日期:2007-11-06 **修回日期:**2008-02-29

2 基于 Gabor PCA 的混合高斯模型

2.1 Gabor 特征和 PCA 分析

Gabor 小波核函数具有与人类大脑皮层简单细胞的二维反射区相同的特性,能够捕捉对应于空间频率(尺度)、空间位置及方向选择性的局部结构信息,由于上述特点,它在人脸表情识别中得到广泛的应用^[9,16-18]。本文选择人脸图像的 Gabor 小波变换系数作为表情特征。

图 1 给出了人脸表情图像 Gabor 滤波的结果示例。其中,(a)为所采用的 5 尺度 8 方向 Gabor 小波滤波器的余弦部分,(b)为带表情的输入人脸图像,(c)为表情图像(b)的 Gabor 滤波结果。图像经 Gabor 变换后的维数是非常庞大的,而人脸的表情特征主要分布在上半脸的眼部周围和下半脸的嘴部周围。因此,选择如图 2 所示的位于嘴巴、眼睛和眉毛等部位点的小波系数作为表情特征。

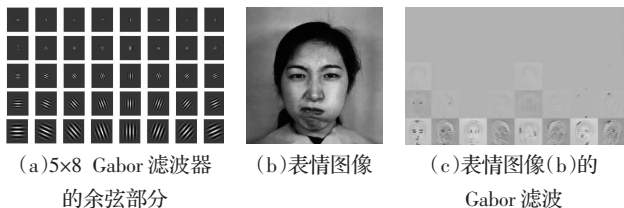


图 1 Gabor 滤波器及表情图像的 Gabor 滤波结果

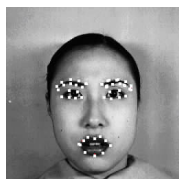


图 2 表情特征点位置

用 f_p 表示 p 点的 5 尺度 8 方向 Gabor 小波系数向量,将表情图像 42 个点的 Gabor 小波系数排成的列向量 $G=[f_{p1}, f_{p2}, \dots, f_{p42}]^T$ 作为训练样本,通过 PCA 方法构造特征子空间。设 G_i 为第 i 个样本向量,共有 m 个样本,构造矩阵 $A=(a_1, a_2, \dots, a_m)$,其中, $a_i=G_i-\bar{G}$, $\bar{G}=\frac{1}{m}\sum_{i=1}^m G_i$,计算协方差矩阵 $C=\frac{1}{m}AA^T$ 的特征值 $\delta_1, \delta_2, \dots$ 和特征向量 g_1, g_2, \dots ,令 $X=(X_1, X_2, \dots, X_r)$ 为最大 r 个特征值 ($\delta_{x1}^2 \geq \delta_{x2}^2 \geq \dots \geq \delta_{xr}^2$) 对应的特征向量组成的基向量矩阵,则可以通过公式(1)将每一个样本向量 G 映射到特征子空间中的 B :

$$B=X^T G \tag{1}$$

2.2 混合高斯模型

混合高斯模型(Gaussian Mixture Model, GMM)是单一高斯概率密度函数的延伸,能够平滑的近似任意形状的概率分布。

为了定量分析表情的不同成分,本文提出用混合高斯模型表示特征子空间的表情分布,通过每类表情在特征子空间中的分布确定混合高斯模型的参数,进而利用混合高斯模型对表情进行识别和分析。

本文采用 P.Ekman^[19]的表情类别表示方式,将表情分为 7 类(包括中性表情)。针对每类表情,对其在特征子空间各基向量的分布进行混合高斯拟合,混合高斯分布概率为 $p_j=\sum_i t_i p(u_i, sig_i)$,

其中 u_i, sig_i, t_i 分别表示混合高斯模型的第 i 个高斯函数的均值、标准差和权重, $p(u_i, sig_i)$ 表示均值为 u_i 、标准差为 sig_i 的高斯分布函数, j 是基向量序号。

PCA 是一种基于目标统计特性的正交变换,它具有一个重要的性质就是变换后产生的新分量正交或不相关,而符合高斯分布的随机变量独立性和不相关性是等价的。所以每类表情在所有基向量的联合分布即为各基向量概率分布的乘积,用 P_E 表示表情类别 E 所有基向量的联合分布,则 $P_E=\prod_{j=1}^r P_j$, P_j 是在第 j 个基向量的分布。即每类表情在特征子空间的混合高斯分布:

$$P_E=\prod_{j=1}^r P_j=\prod_{j=1}^r (\sum_i t_i p(u_i, sig_i)) \tag{2}$$

于是,给定特征子空间中一个表情点 e ,用 P_E^e 表示表情属于基本表情类别 E 的概率,根据公式(2),将 e 代入各混合高斯分布模型可得 e 属于相应基本表情类的概率 P_E^e ,用 P^e 表示表情 e 属于不同基本表情类别的概率,则 P^e 可以记为 $P^e=[P_{an}^e, P_{di}^e, P_{fe}^e, P_{ha}^e, P_{ne}^e, P_{sa}^e, P_{su}^e]$,其中 an, di, fe, ha, ne, sa, su 分别对应于基本表情类别愤怒、厌恶、恐惧、高兴、中性、悲伤、惊讶。

P^e 能够衡量表情 e 与不同基本表情的相似程度,通过 P^e 可以实现表情 e 的识别和分析。 P^e 分量值越大,表明表情 e 与其对应的表情类别相似度越高,可以据此判定表情属于相似度最高的表情类别来完成识别过程;同时,根据 P^e 各个分量的相对比较,可以分析 e 表情中各表情类别的成分,以得到表情 e 是由哪些基本表情混合而成的。

3 实验分析与讨论

JAFFE^[18]是一个人脸表情图像数据库,该库包括 10 个日本女性 7 种表情(包括愤怒、厌恶、恐惧、高兴、悲伤、惊讶 6 种基本表情和中性表情)共 219 幅人脸图像,其中每人每种表情图像有 3~4 幅。选择 195 幅图像作为训练样本进行 PCA 训练。

计算表明,当 $r=10$ 时,特征子空间能够保留原样本空间的大部分信息,即 $\sum_{i=1}^r \delta_{xi}^2 \approx 0.8 * \sum_j \delta_j^2$,因此这里令 $r=10$ 。

然后用混合高斯模型对每类表情在特征子空间中的分布进行统计。首先拟合每类表情在各基向量的混合高斯分布,实验中发现用 1~3 个高斯函数就能够获得比较好的拟合。图 3 给出的是愤怒表情在不同基向量的混合高斯模型参数和其概率分布,然后根据公式(2)可以得到愤怒表情在特征子空间的混合高斯分布 P_m 。表 1 给出了其他各类表情不同基向量的混合高斯模型参数,同时也能得到各类表情在特征子空间的混合高斯分布 P_E 。

最后,计算各表情图像 e 属于不同表情类别的概率 P^e 。

将各样本的 P^e 最大分量对应的基本表情类别作为表情识别的结果。表 2 给出了本文方法和 Gabor PCA 方法的表情识别率的对比。本文方法的识别率比仅采用 Gabor 小波特征+PCA 的方法有明显的大幅提高。

比较 P^e 的各个分量,对表情 e 中各表情类别的成分进行分析。图 4 是表情图像的表情概率分布 P^e 的示例,从图 4 中可以看出当表情 e 属于某表情类别 E 的概率远大于其他表情类别时

表 2 表情识别率对比

	AN	DI	FE	HA	NE	SA	SU	(%)
Gabor PCA	60	67	80	56	67	56	89	
本文方法	93	90	93	90	85	96	100	

表1 各类表情的混合高斯模型参数

厌恶(DI)	基向量1	$u_1=-85.202\ 9, sig_1=12.045\ 3, t_1=0.321\ 5; u_2=-27.501\ 6, sig_2=8.942\ 3, t_2=0.678\ 5$
	基向量2	$u_1=-8.019\ 6, sig_1=28.695\ 4, t_1=0.721\ 0; u_2=50.246\ 6, sig_2=38.031\ 6, t_2=0.279\ 0$
	基向量3	$u_1=-72.903\ 7, sig_1=17.944\ 5, t_1=0.213\ 4; u_2=0.630\ 2, sig_2=22.136\ 5, t_2=0.786\ 6$
	基向量4	$u_1=-47.202\ 7, sig_1=13.630\ 2, t_1=0.171\ 1; u_2=10.290\ 4, sig_2=17.238\ 7, t_2=0.547\ 8; u_3=46.334\ 9, sig_3=12.265\ 0, t_3=0.281\ 1$
	基向量5	$u_1=-10.051\ 6, sig_1=29.138\ 1, t_1=1$
	基向量6	$u_1=20.351\ 1, sig_1=26.466\ 7, t_1=1$
	基向量7	$u_1=-7.947\ 2, sig_1=31.135\ 5, t_1=1$
	基向量8	$u_1=-61.455\ 8, sig_1=31.262\ 0, t_1=0.172\ 8; u_2=-10.416\ 0, sig_2=20.422\ 9, t_2=0.827\ 2$
	基向量9	$u_1=-23.982\ 3, sig_1=20.143\ 6, t_1=0.209\ 9; u_2=14.231\ 2, sig_2=12.446\ 6, t_2=0.790\ 1$
	基向量10	$u_1=-0.209\ 1, sig_1=19.481\ 9, t_1=1$
恐惧(FE)	基向量1	$u_1=-52.847\ 2, sig_1=14.125\ 3, t_1=0.207\ 3; u_2=34.682\ 4, sig_2=20.300\ 2, t_2=0.792\ 7$
	基向量2	$u_1=-30.003\ 9, sig_1=36.759\ 3, t_1=0.635\ 6; u_2=59.926\ 5, sig_2=42.661\ 7, t_2=0.364\ 4$
	基向量3	$u_1=-69.284\ 4, sig_1=25.617\ 0, t_1=0.248\ 3; u_2=25.188\ 6, sig_2=23.044\ 3, t_2=0.751\ 7$
	基向量4	$u_1=18.395\ 5, sig_1=26.439\ 9, t_1=1$
	基向量5	$u_1=-3.037\ 0, sig_1=34.715\ 0, t_1=1$
	基向量6	$u_1=24.421\ 1, sig_1=27.698\ 3, t_1=1$
	基向量7	$u_1=5.889\ 7, sig_1=28.070\ 3, t_1=1$
	基向量8	$u_1=0.872\ 2, sig_1=26.937\ 4, t_1=1$
	基向量9	$u_1=15.007\ 3, sig_1=21.463\ 7, t_1=1$
	基向量10	$u_1=-35.866\ 1, sig_1=23.244\ 4, t_1=0.125\ 2; u_2=7.882\ 9, sig_2=16.307\ 6, t_2=0.874\ 8$
高兴(HA)	基向量1	$u_1=-46.674\ 2, sig_1=14.131\ 2, t_1=0.369\ 4; u_2=28.781\ 6, sig_2=22.730\ 8, t_2=0.630\ 6$
	基向量2	$u_1=-39.911\ 5, sig_1=40.302\ 5, t_1=0.735\ 1; u_2=48.345\ 7, sig_2=5.711\ 9, t_2=0.264\ 9$
	基向量3	$u_1=-24.579\ 2, sig_1=43.102\ 7, t_1=0.570\ 0; u_2=37.130\ 3, sig_2=27.695\ 8, t_2=0.430\ 0$
	基向量4	$u_1=-12.047\ 1, sig_1=40.608\ 3, t_1=1$
	基向量5	$u_1=17.754\ 9, sig_1=41.995\ 1, t_1=1$
	基向量6	$u_1=-29.100\ 4, sig_1=23.667\ 2, t_1=1$
	基向量7	$u_1=0.960\ 4, sig_1=29.754\ 3, t_1=1$
	基向量8	$u_1=-5.541\ 3, sig_1=30.119\ 0, t_1=1$
	基向量9	$u_1=-45.136\ 1, sig_1=28.035\ 2, t_1=0.623\ 8; u_2=16.352\ 4, sig_2=12.337\ 1, t_2=0.376\ 2$
	基向量10	$u_1=-5.473\ 8, sig_1=22.322\ 1, t_1=0.756\ 2; u_2=33.045\ 3, sig_2=7.927\ 8, t_2=0.243\ 8$
中性(NE)	基向量1	$u_1=29.589\ 8, sig_1=22.884\ 8, t_1=1$
	基向量2	$u_1=-66.148\ 6, sig_1=12.849\ 7, t_1=0.192\ 3; u_2=17.231\ 0, sig_2=33.600\ 7, t_2=0.807\ 7$
	基向量3	$u_1=1.576\ 6, sig_1=44.899\ 1, t_1=1$
	基向量4	$u_1=-49.829\ 6, sig_1=17.100\ 3, t_1=0.379\ 2; u_2=-7.670\ 7, sig_2=8.947\ 6, t_2=0.320\ 8; u_3=51.385\ 6, sig_3=22.648\ 4, t_3=0.299\ 9$
	基向量5	$u_1=-8.142\ 6, sig_1=10.640\ 4, t_1=0.664\ 2; u_2=50.157\ 2, sig_2=26.993\ 2, t_2=0.335\ 8$
	基向量6	$u_1=-22.800\ 3, sig_1=27.513\ 6, t_1=1$
	基向量7	$u_1=-4.013\ 0, sig_1=24.400\ 3, t_1=1$
	基向量8	$u_1=13.351\ 3, sig_1=22.816\ 7, t_1=1$
	基向量9	$u_1=1.601\ 8, sig_1=25.000\ 1, t_1=1$
	基向量10	$u_1=-19.863\ 0, sig_1=31.212\ 9, t_1=0.810\ 7; u_2=30.513\ 9, sig_2=12.527\ 2, t_2=0.186\ 9$
悲伤(SA)	基向量1	$u_1=-60.753\ 8, sig_1=17.689\ 4, t_1=0.596\ 7; u_2=14.854\ 0, sig_2=22.662\ 0, t_2=0.403\ 3$
	基向量2	$u_1=-49.534\ 0, sig_1=29.427\ 7, t_1=0.496\ 4; u_2=8.241\ 4, sig_2=31.706\ 6, t_2=0.503\ 6$
	基向量3	$u_1=-46.273\ 9, sig_1=10.565\ 2, t_1=0.233\ 8; u_2=22.143\ 4, sig_2=23.291\ 7, t_2=0.766\ 2$
	基向量4	$u_1=-2.144\ 7, sig_1=34.706\ 0, t_1=1$
	基向量5	$u_1=-0.678\ 5, sig_1=27.241\ 6, t_1=1$
	基向量6	$u_1=-1.455\ 8, sig_1=27.317\ 2, t_1=1$
	基向量7	$u_1=0.801\ 7, sig_1=28.434\ 6, t_1=1$
	基向量8	$u_1=-0.852\ 7, sig_1=32.583\ 3, t_1=1$
	基向量9	$u_1=-0.164\ 0, sig_1=26.041\ 5, t_1=1$
	基向量10	$u_1=-9.176\ 1, sig_1=34.709\ 8, t_1=1$
惊讶(SU)	基向量1	$u_1=70.662\ 2, sig_1=32.229\ 9, t_1=1$
	基向量2	$u_1=1.298\ 5, sig_1=44.732\ 4, t_1=1$
	基向量3	$u_1=-54.254\ 4, sig_1=18.122\ 7, t_1=0.251\ 4; u_2=41.911\ 4, sig_2=24.266\ 1, t_2=0.748\ 6$
	基向量4	$u_1=3.424\ 3, sig_1=30.477\ 3, t_1=1$
	基向量5	$u_1=-16.137\ 1, sig_1=37.063\ 2, t_1=1$
	基向量6	$u_1=-24.826\ 8, sig_1=24.143\ 2, t_1=0.601\ 2; u_2=45.697\ 3, sig_2=21.989\ 5, t_2=0.398\ 8$
	基向量7	$u_1=-13.091\ 7, sig_1=26.097\ 5, t_1=0.536\ 9; u_2=55.944\ 7, sig_2=12.155\ 7, t_2=0.463\ 1$
	基向量8	$u_1=-6.418\ 2, sig_1=22.944\ 7, t_1=1$
	基向量9	$u_1=0.415\ 5, sig_1=25.118\ 0, t_1=1$
	基向量10	$u_1=5.981\ 9, sig_1=28.360\ 6, t_1=1$

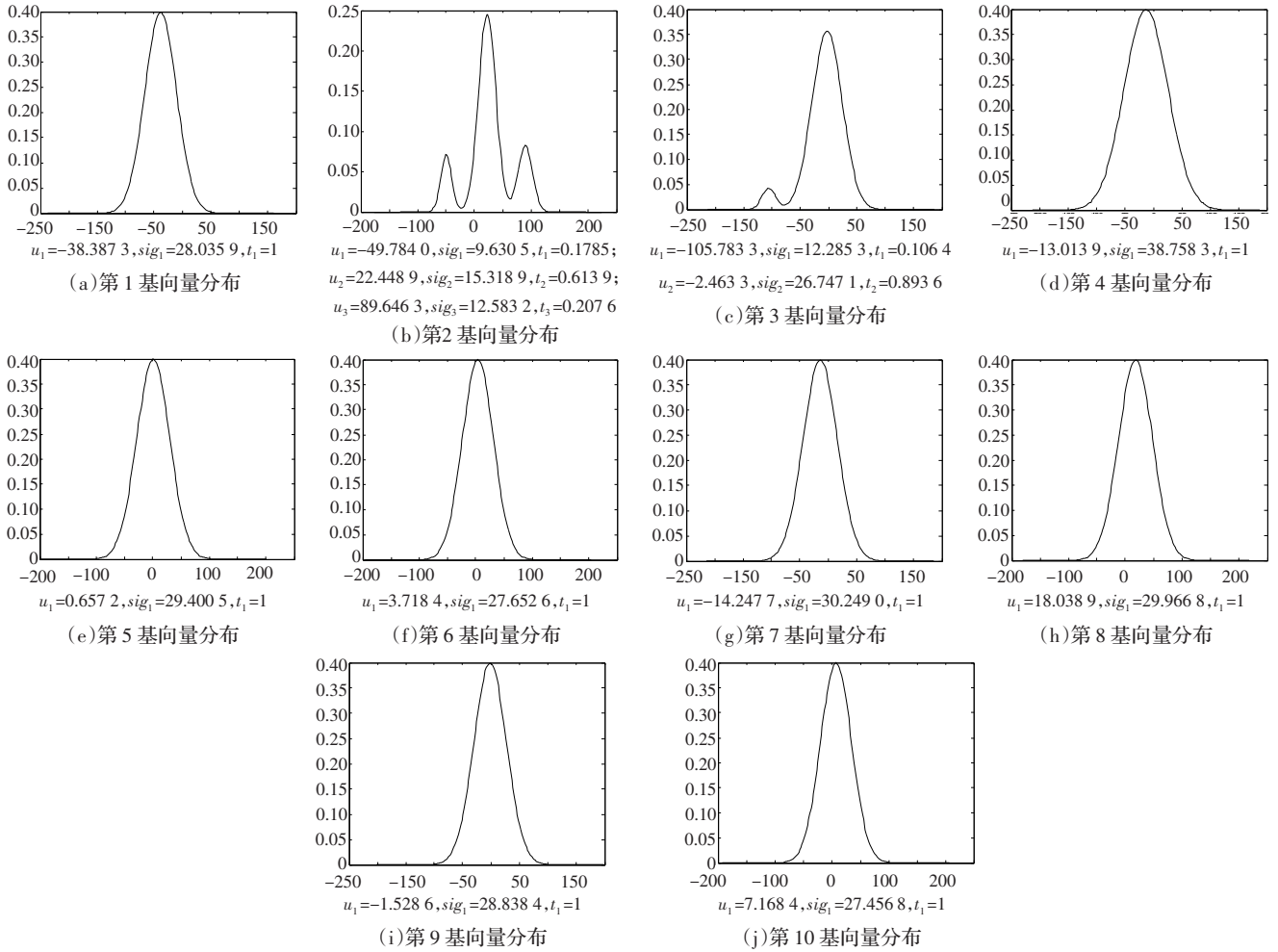


图3 愤怒表情在不同基向量的混合高斯分布

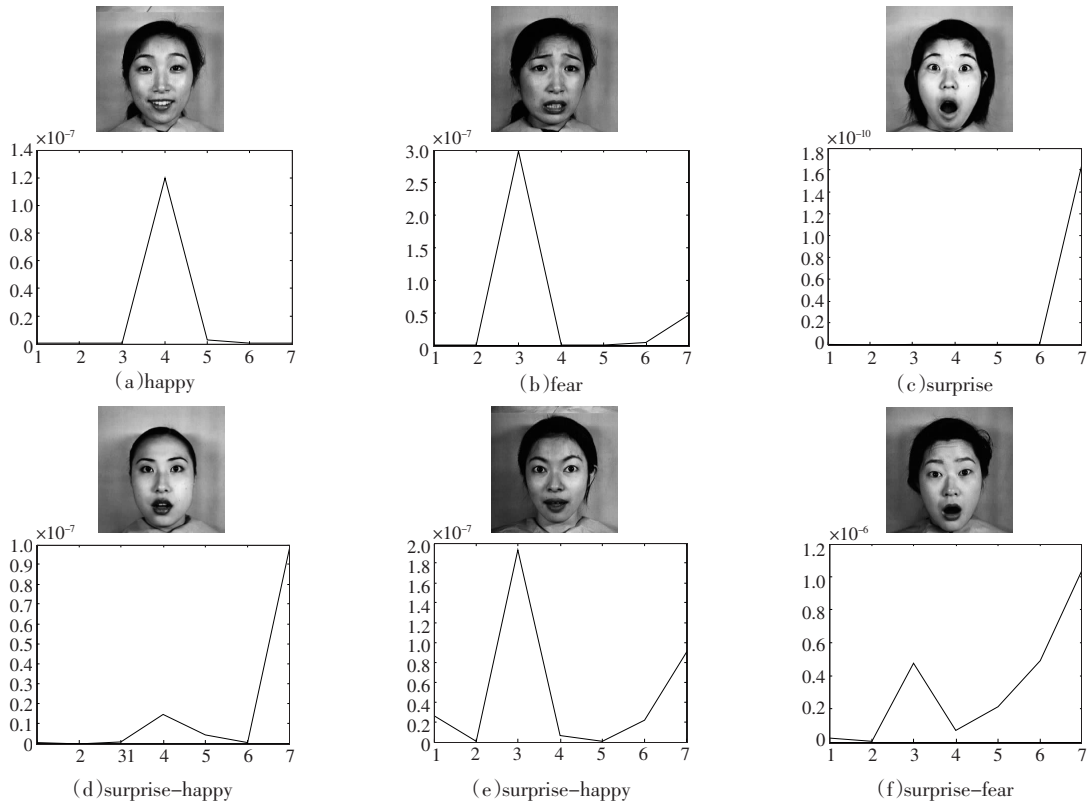


图4 表情图像的表情概率分布示例

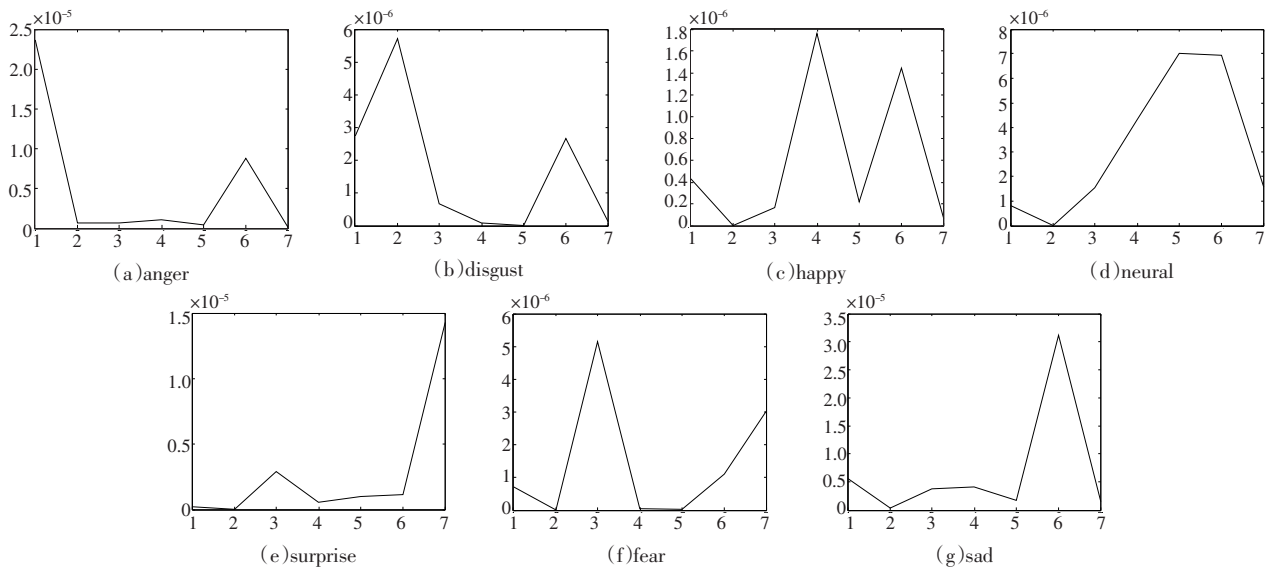


图5 各类表情样本概率的直方图分布

(如图4(a)(b)(c)),表情 e 可以归为表情类别 E ;但当表情 e 与多种表情类别相似时,将表情 e 简单归为哪一类就不太符合实际情况, P^f 各分量大小正体现了表情 e 与不同基本表情的相似程度,如图4(d)(e)(f)中表情 e 与多种基本表情相似,因此其属于各表情类别的概率分布中就有2项或几项都比较大。

图5是各类表情样本概率的直方图分布,从图5中可以看到,表情往往与多种表情类别都有一定的相似度,这是跟生活中的现实表情一致的。所以仅将表情简单划为某一类别是不够的,还需分析其与各表情类别的相似程度,而 P^f 的引入则是达到这个目的的一种方式。

通过实验结果,看到,提出的方法能够很好的对日常生活中的表情进行表示,并能够运用于表情识别和分析。表情与某基本表情较相似,其概率分布 P^f 相应的分量比较大;反之, P^f 相应的分量比较小。当某个分量明显大于其他分量时,可以很容易地判定表情的类别。日常表情往往是几种基本表情的混合, P^f 能够表示表情和各基本表情的相似程度,各基本表情成分的大小能够由表情点 P^f 各分量大小体现。

4 结束语

本文根据日常生活表情往往由多种情感引发的实际特点,提出了一种基于 Gabor 特征和混合高斯模型的表情分析方法,并通过大量的数据给出实验分析。实验结果表明,该方法更符合现实情况,不仅能够给出表情的类别归属,还能够进一步分析组成表情的基本表情成分,比现有的表情识别方法更好地反映表情的实际表现。

参考文献:

- [1] Yacoob Y, Davis L S. Recognizing human facial expressions from long image sequences using optical flow[J]. IEEE Trans PAMI, 1994, 16(6): 636-642.
- [2] Zhu Y, de Silva L C, Ko C C. Using moment invariants and HMM in facial expression recognition[J]. Pattern Recognition Letters, 2002, 23: 83-91.
- [3] Hu T, de Silva L C, Sengupta K A. A hybrid approach of NN and HMM for facial emotion classification[J]. Pattern Recognition Letters, 2002, 23: 1303-1310.

- [4] 金辉, 高文. 人脸面部混合表情识别系统[J]. 计算机学报, 2000, 23(6): 602-608.
- [5] Pardas M, Bonafonte A. Facial animation parameters extraction and expression recognition using Hidden Markov Models[J]. Signal Processing: Image Communication, 2002, 17: 675-688.
- [6] Calder A J, Burton A M, Miller P, et al. A principal component analysis of facial expressions[J]. Vision Research, 2001, 41: 1179-1208.
- [7] Dubussion S, Devoine F, Masson M. A solution for facial expression representation and recognition[J]. Signal Processing: Image Communication, 2002, 17: 657-673.
- [8] Matsugu M, Mori K, Mitari Y, et al. Subject independent facial expression recognition with robust face detection using a convolutional neural network[J]. Neural Networks, 2003, 16: 555-559.
- [9] Zhan Yong-zhao, Ye Jing-fu, Niu De-jiao, et al. Facial expression recognition based on Gabor wavelet transformation and elastic templates matching[C]//Proceedings of the Third International Conference on Image and Graphics, 2004.
- [10] Donato G, Bartlett M, Hager J C, et al. Classifying facial actions[J]. IEEE Trans PAMI, 1999, 21(10): 974-989.
- [11] Ekman P, Friesen W V. Unmasking the face: a guide to recognizing emotions from facial clues[M]. [S.l.]: Consulting Psychologists Press, 1975.
- [12] Picard R W. Affective computing[M]. [S.l.]: The MIT Press, 1998.
- [13] Jin Xue-cheng, Wang Zeng-fu. An emotion space model for recognition of Emotions in Spoken Chinese[C]//ACII2005, Beijing, 2005.
- [14] Smith C, Scott H A. A componential approach to the meaning of facial expressions[M]//The Psychology of Facial Expression. Cambridge: Cambridge University Press, 1997.
- [15] Kappas A. The human face: measurement and meaning[M]. Dordrecht: Kluwer Academic Publishers, 2003.
- [16] Lades M, Vorbruggen J, Buhmann J, et al. Distortion invariant object recognition in the dynamic link architecture[J]. IEEE Trans Computers, 1993, 42(3): 300-311.
- [17] Wiskott L, Fellous J, Kruger N, et al. Face recognition by elastic bunch graph matching[J]. IEEE Trans PAMI, 1997, 19(7): 775-779.
- [18] Lyons M. Coding facial expressions with Gabor wavelets[C]//Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara Japan, 1998: 200-205.