

# 基于局部二元模式的人脸表情识别<sup>①</sup>

朱晓明, 姚明海

(浙江工业大学 信息工程系, 杭州 310023)

**摘要:** 提出了一种基于局部二元模式(LBP)和局部保全投影(LPP)相结合的面部表情识别方法。使用 LBP 算子对图像分块处理, 综合人脸局部和整体的特征; 再使用 LPP 对表情特征降维, 最后采用支持向量机对面部表情分类。在日本女性人脸表情库上实验表明, 本文提出的方法有更好的识别率和更快的识别速度。

**关键词:** 面部表情识别; 局部二元模式; 局部保全投影; 支持向量机

## Facial Expression Recognition Based on Local Binary Pattern

ZHU Xiao-Ming, YAO Ming-Hai

(Department of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

**Abstract:** A method combining Local Binary Pattern(LBP) with Locality Preserving Projections(LPP) was presented for face expression recognition. First, the algorithm processes facial expression images with LBP operator and integrates local and global characteristics of the face. Then Locality Preserving Projections(LPP) method is used for feature dimension reduction and SVM for classification. The algorithm is experimented in Japanese Female Facial Expression database(JAFFE database). The results show that the proposed algorithm is more effective.

**Keywords:** facial expression recognition; local binary pattern(LBP); locality preserving projections(LPP); support vector machine(SVM)

### 1 引言

面部表情是人类进行情感交流的重要途径, 通过表情的交流, 人们可以感知对方的情绪变化和情感的波动, 人们可通过表情准确而微妙地表达出自己的思想感情, 也可通过表情辨认对方的态度和内心世界。

人脸表情识别是基于视觉信息对脸部的运动以及脸部特征的形变进行分类, 主要包括脸部的检测, 脸部特征的提取和表情的分类, 其中特征提取和分类是研究的重点和难点。在特征提取上目前已有许多的应用典范。最早是 Ekman 和 Friesent 提出的面部动作编码系统(Facial Action Coding System, FACS)<sup>[1]</sup>, 系统将人脸划分为若干个运动单元来描述面部动作, 这些运动单元显示了人脸运动与表情的对应关系。6 种基本表情和 FACS 的提出具有里程碑的意义, 这成为后来表情识别研究工作的基础。目前, 提取表情特征的主

要方法有: 基于 Gabor 小波, 基于活动外观模型(Active Appearance Model, AAM), 基于主成分分析法(Principle Component Analysis, PCA), 基于局部二元模式(Local Binary Pattern, LBP)等。其中, 基于 Gabor 小波的方法会产生相当多的冗余信息<sup>[2]</sup>, 增加了特征降维的难度, 不适合建立快速高效的人机交互系统。基于 AAM 的方法虽然能获得较为可靠的表情特征, 却存在着计算复杂、初始参数难以确定等缺点。PCA 降维更多关注的是图像整体, 难以区分局部表情的不同。相比于 Gabor 小波特征, LBP 特征能够对灰度图像中局部邻近区域的纹理信息进行快速的度量和提取, 而且维数小很多, 同时在一定程度上抑制光照的影响<sup>[3]</sup>。本文对局部二元模式特征获取途径做了一些改进, 结合了图像整体和局部的关系, 同时将获得的特征采用局部保全投影, 大大降低了数据维数, 提高

① 基金项目: 国家自然科学基金(61070113)

收稿时间: 2010-09-27; 收到修改稿时间: 2010-10-14

了表情识别速度。最后通过 SVM 在 JAFFE 库上做了实验,结果表明该方法相对于传统的 LBP 特征分类法在正确率和时效上有明显的提高。

## 2 LBP算子

LBP 算子最初由 Ojala 和 Harwood 等人于 1996 年提出<sup>[4]</sup>,并将其成功应用于纹理分类。因为 LBP 算子计算简单,经过 LBP 算子滤波后所获得的特征拥有很强的分辨能力,这使得后来的研究者将其引入其它的应用领域并且取得了很好的效果。基本的 LBP 算子可用图 1 进行描述。将 3×3 邻域周围 8 个像素与中心像素值进行比较,不小于中心点赋值为 1,小于中心点则为 0,然后按顺时针或逆时针方向排列得到一个 8 位二进制数,再将其转化为十进制数,将该值来代替原始图像中的每一个像素。

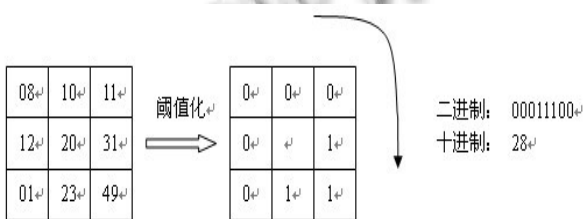


图 1 LBP 算子示例

LBP 可用不同尺寸的邻域来表示。它由半径为 R 的圆周上对称分布的 P 个邻点构成, P 和 R 可取任意值,如图 2 所示, LBP 算子就是在此环形邻域内描述纹理。一般来说,比较细腻的纹路采用小的尺寸算子,粗糙的纹路采用大的尺寸算子。表达式如下:

$$LBP(P, R) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (1)$$

其中,  $g_p$  ( $p=0, 1, \dots, P-1$ ) 指代半径为 R 的圆周上对称分布的 p 个邻点的灰度值,  $g_c$  指代环形邻域中心的灰度值。符号函数  $S(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$ , 通过给每个符号函数

数赋予一个权重因子,得到描述局部图像纹理的 LBP 码。但是由式(1)得到的二值数据对噪声反应比较敏感。在实际应用中,图像中存在的噪声会严重限制 LBP 算子提取特征的效果。因为 LBP 算子考虑的是像素点间的关系,如果  $s(g_a - g_b) = 0$ , 那么  $s(g_b - g_a) = 1$ 。为了降低白噪声对图像的影响,对符号函数稍作修改:

数赋予一个权重因子,得到描述局部图像纹理的 LBP 码。但是由式(1)得到的二值数据对噪声反应比较敏感。在实际应用中,图像中存在的噪声会严重限制 LBP 算子提取特征的效果。因为 LBP 算子考虑的是像素点间的关系,如果  $s(g_a - g_b) = 0$ , 那么  $s(g_b - g_a) = 1$ 。为了降低白噪声对图像的影响,对符号函数稍作修改:

$$s(x) = \begin{cases} 1, |x| \geq C \\ 0, |x| < C \end{cases}, \text{ 其中 } C \text{ 是阈值常数。}$$

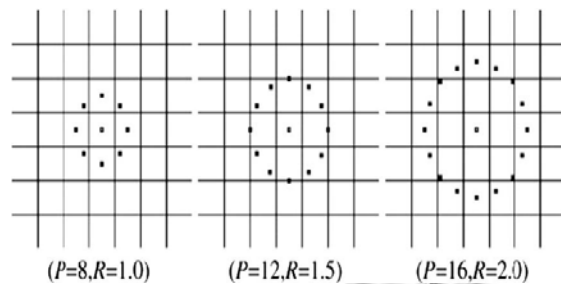


图 2 LBP 不同半径尺寸

对于一个 P 值为 8 的 LBP 算子来说,共有 256 种模式。根据统计,256 种模式中有相当一部分在图像中出现的机率很小。于是研究人员对 LBP 二进制编码 0-1 或 1-0 的变化次数有所限制,规定总变化次数不超过两次,从而提出了一种 LBP 均匀模式,用  $LBP_{P,R}^{u_2}$  表示。例如:编码(10101001)0-1 或 1-0 的变化次数为 6 次则不在均匀模式之列,而编码(00010000)只变化了 2 次,为均匀模式。对于非均匀模式,统一用一个编码来表示。这样能大大减少模式的数量。以 P 值为 8 的 LBP 算子为例,其均匀模式为 58 个,将所有的非均匀模式用一个编码来表示,则所有的模式只有 59 种。也有研究人员提出了具有旋转不变性的高级 LBP 算子<sup>[5]</sup>,用式(2)表示。

$$ALBP(P, R) = \min(Cir(LBP(P, R), i) | i = 0, \dots, P-1) \quad (2)$$

函数 Cir 将 P 位的环形二进制码 LBP(P, R) 逆时针移位 i 次。由式(2)可知,当一种模式是由另一种模式旋转一定角度后得到时,ALBP 把这两种模式视作同一类型,显示出 ALBP 的旋转不变性。考虑到实际处理效果和效率,本文采用 LBP 均匀模式。经过 LBP 均匀模式算子计算得到的图像如图 3 所示。

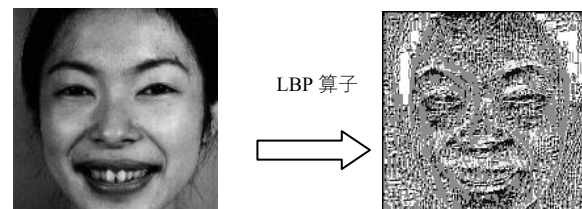


图 3 LBP 算子图像处理实例

## 3 特征提取与LPP降维

当一幅表情图像经过 LBP 算子计算之后,图像的

每个像素的值, 已经不再是原来人脸图像所表示的明暗亮度, 而是 LBP 均匀模式中人为定义的一种模式。得到 LBP 编码图后, 接下去就是特征提取。表情特征应该既是局部的又是整体的, 不同特征部位的运动可能表示不同的表情, 例如眼睛张大, 既可以表示惊讶, 也能指恐惧。所以仅仅看眼睛一部分不能判断出全部表情, 必须结合整体才能准确的识别表情。在一般的 LBP 算法中, 直方图通过计算一定大小窗口内的 LBP 值获得, 通常将原图像分成几大块, 计算各块的 LBP 直方图, 然后将每块的直方图串联起来得到最终的图像特征, 如图 4 所示<sup>[6]</sup>。这种方法过多的注意了图像局部细节, 而忽略了图像整体特征, 使识别率收到影响。本文将图像分成 64 块, 先逐个的计算每个小块的直方图, 再合并小块计算大块的直方图。如此, 最后计算整幅图像的直方图, 将得到的直方图特征全部串联起来得到图像的待识别特征。采用这样的方法提取到的特征, 既包含了局部纹理, 也包含了全局的纹理特征, 较好的解决了表情识别中局部相似而整体不同的问题。经过 LBP 算子计算得到的特征维数十分庞大, 对于后期的训练、识别来说, 大量的冗余数据在吞噬存储空间的同时, 还耗费了大量的计算能力, 而且效果不明显, 很容易产生维数灾难。所以需要特征降维, 从大量的特征中寻找合适的特征, 用尽量少的数目来表示图像数据的本质。局部保权投影 (LPP) 作为一种有效的降维方法, 将脸部图像映射到一个脸部的子空间, 然后进行分析计算。它不同于 PCA 和 LDA 算法, 保留的是图像局部空间结构。研究表明, 人脸图像是在一个非线性的子流形之上, PCA 和 LDA 仅仅关注几何结构, 并不能发现那些非线性的底层结构<sup>[2]</sup>, 所以本文选择 LPP 对特征数据进行降维。

局部保全投影<sup>[7]</sup>主要算法首先是求取训练样本集  $\{x_1, x_2, \dots, x_c\}$  的 LBP 直方图  $X$ , 即人脸图像分割成的  $C$  个子区域的图像直方图。为了避免  $XD X^T$  无法求取逆矩阵, 需要先把直方图映射到 PCA 空间, 使得  $XD X^T$  非奇异, 可以求逆矩阵, 同时也减少了噪声的影响。映射转换矩阵表示为  $W_{PCA}$ , 空间的映射值为  $X^p$ , 即  $X^p = W_{PCA}^T X$ 。然后是构建相似性度量矩阵  $S$ ,  $S_{ij} = \begin{cases} \exp(-\|x_i - x_j\|/a), & \text{情况1} \\ 0, & \text{其他情况} \end{cases}$

情况 1 指:  $x_i$  是  $x_j$  的  $k$  近邻或者  $x_j$  是  $x_i$  的  $k$  近邻,  $a$  是一常数,  $i, j=1, 2, \dots, c$  最后是特征映射。

$$X^p L (X^p)^T W = \lambda X^p D (X^p)^T W \quad (3)$$

其中  $D$  是对角阵,  $D_{ii} = \sum_j S_{ji}$ ,  $L=D-S$ ,  $w_1, w_2, \dots, w_k$  是式 (3) 的结果, 按照它们的特征值  $0 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k$  进行排序, 即

$$W=[w_1, w_2, \dots, w_k]$$

最终的映射转换矩阵  $W_{LBPHP}$  可定义如下:

$W_{LBPHP} = W_{PCA} W$ ,  $W_{PCA}$  为前面的 PCA 映射转换矩阵

$$y_i = W_{LBPHP}^T x_i; \quad i=1, \dots, c.$$

$y_i$  就是样本  $x_i$  的降维 LBP 特征。更多的 LPP 算法细节可详见参考文献[7]。

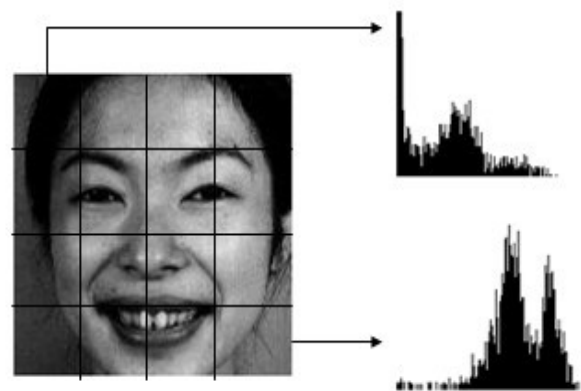


图 4 表情图像特征形成

#### 4 SVM特征分类

支持向量机(Support Vector Machine, SVM)是以统计学习理论为基础, 按照结构风险最小化原则设计的一种模式分类方法。它对两类问题具有很好的分类能力。主要思想是把特征空间通过核函数映射到高维空间, 并在其中寻找最优分类的超平面。选取不同的核函数可以进行不同的特征空间映射。典型的核函数有: 线性核函数、多项式核函数和径向基函数(RBF)核函数。扩展到多类问题时, 一般使用多个二值分类器来解决。常用的一对一的分解策略就是将整个训练样本分解成两两不同的类组合形成的新的训练样本。本文将面部表情分为中性、高兴、生气、厌恶、害怕、伤心、惊讶七种, 根据排列组合的知识可以知道, 共产生了 21 个子训练样本集。通过对这 21 个子样本集分别训练, 产生对应于每个子样本集的训练模型。对于一个给定的表情, 分别用训练好的样本表情模型进行分类, 最后采用投票测率, 选取分类结果中次数出现较多的类别作为分类的表情结果。

### 5 实验结果

本次实验采用 JAFFE 表情库，一共包含 10 位日本女性的 7 种不同表情，每种表情由 2 到 4 张不等的图片组成，选取 210 幅图作为实验数据。将图像 10 等分为 10 组，第一组作为测试集，其余 9 组训练，得到第一个识别结果；然后将第二组作为测试集，得到第二个结果。如此循环 10 次，最后通过计算这 10 次结果的平均识别率来评估算法的识别性能。





每个样本经过剪切和调整到尺寸 64x64，然后进行预处理，消除光线干扰，用均匀 LBP 算子在每个子区域上提取 LBP 直方图并串接起来描述整张人脸的信息。因为将图像分为 3 层，则最后的维数为(64+16+4)\*59=4596 维，但映射到 LPP 空间后就降到 398 维，大大提高了识别速度。实验采用 VC++ 6.0 在 AMD turion64 2.0GHz, 2G 内存的 PC 机上对不同算法做出比较。

表 1 列出了几种不同算法的比较，可以看出，LBP 方法明显优于 Gabor 小波，而结合了 LPP 降维的算法六种表情的识别效果好于传统的基于 LBP 直方图交的算法。其中，高兴和惊讶识别率最高，害怕的识别率最低，容易和高兴的表情混淆。在识别速度方面，降维后大大减少了需要处理的数据，平均识别每幅图像的时间不到 1 秒，而基于传统 LBP 直方图交的方法则需要几十秒。同时考虑到现实环境获得的图像一般像素都不高，表 2 则比较了几种算法在不同尺寸图像中的识别效果。由表 2 可以看出，在各种情况下，本文的算法识别效果都要好于基于 Gabor 小波和传统的 LBP 算法。即使分辨率低至 32\*32，仍然有 84%的识别率，稳定性较好。

表 1 不同方法的表情识别率

	基于Gabor小波(%)	基于直方图交LBP(%)	本文方法(%)
高兴	83	86	91
生气	71	80	84
厌恶	68	67	79
害怕	60	63	77
伤心	61	68	83

表 2 不同分辨率下的识别效果

	 32*32(%)	 48*48(%)	 64*64(%)	 96*96(%)
Gabor	75	78	80	83
传统 LBP	78	80	82	86
LBP+LPP	84	87	89	91

### 6 结论

面部表情识别有广阔的应用前景，吸引了国内外许多研究机构的关注。本文提出的结合局部与整体特征 LBP 的算法，能够有效的获得特征向量，同时通过强大的 LPP 降维筛选，使得识别效率和速度有了进一步的提升。但是要进一步提高识别效果，还需要继续努力。

#### 参考文献

- Ekman P, Ffiesen WV. Facial action coding system(FACs). Manula, Consulting Psychologist Press, 1978:16-23.
- 应自炉,方谢燕.基于局部二元模式的面部表情识别研究.计算机工程与应用,2009,45(29):180-183.
- Feng X, Pietikainen M, Hadid T. Facial expression recognition with local binary patterns and linear programming. Pattern Recognition and Image Analysis, 2005,15(2):546-548.
- Ojala T, Pietikainen M, Harwood D. A comparative study of texture measures with classification based on feature distributions. Pattern Recognition, 1996,29(1):51-59.
- 付晓峰,韦巍.基于高级局部二元模式直方图映射的表情识别.模式识别与人工智能,2009,22(1):123-128.
- Shan CF, Gong SGB, Peter W. McOwanFacial expression recognition based on Local Binary Patterns:a comprehensive study. Image and Vision Computing 2009,27:803-816.
- He XF, Yen SC, Hu YX, et al. Face recognition using laplacianfaces. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005,27(3):328-340.