

# 基于多频域 LBP-TOP 的人脸表情识别

孔 健, 詹永照

(江苏大学计算机科学与通信工程学院, 镇江 212013)

**摘 要:** 针对人脸表情时空域特征信息的有效提取问题, 提出一种多频域 LBP-TOP 与最大间隔球形支持向量机相结合的人脸表情识别算法。采用小波分解增强序列各帧的图像信息, 对同频率的子图像序列提取分块改进的 LBP-TOP 特征, 采用最大间隔球形支持向量机进行样本的训练及识别。实验结果证明, 该方法能有效提取运动的表情特征, 识别率高, 同时符合实时性要求。

**关键词:** 模式识别; 人脸表情识别; 小波分解; 二元局部模式; 最大间隔球形支持向量机

## Facial Expression Recognition Based on Multi-frequency LBP-TOP

KONG Jian, ZHAN Yong-zhao

(School of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang 212013)

**【Abstract】** According to the problem of effective extraction of facial expression information in space-time domain, one kind facial recognition method based on multi-frequency Local Binary Patterns from Three Orthogonal Panels(LBP-TOP) features and Maximal-margin Spherical-structured Support Vector Machine(MSSVM) is proposed. It adapts wavelet decomposition to enhance information of each frame in image sequence. It extracts improved LBP-TOP features of sub-images on the same frequency. MSSVM is applied for sample training and recognition. Experimental result indicates that, this method can extract movement expression feature more effectively, as well as recognition rate is better, and it meets the requirement of real-time.

**【Key words】** pattern recognition; facial expression recognition; wavelet decomposition; local binary pattern; Maximal-margin Spherical-structured Support Vector Machine(MSSVM)

### 1 概述

人脸表情识别是指利用计算机对人脸的表情信息进行特征提取, 按照人的认识和思维方式加以归类和理解, 进而从人脸信息中去分析理解人的情绪, 是计算机视觉研究的重要组成部分, 它的应用领域主要包括: 智能人机接口的设计与完成, 人工智能, 情感机器人的设计与实现等。

目前, 用于人脸表情特征提取的几大流行方法有: Gabor 小波<sup>[1]</sup>, AAM<sup>[2]</sup>, 光流<sup>[3]</sup>等。文献[4]提出的 VLBP+SVM 方法不仅计算简单, 而且能有效提取表情的纹理和形变特征, 同时也在一定程度上提高了表情识别的速度。但传统的 VLBP 特征仅考虑空间域的二元局部模式, 忽略了时域上的表情信息, 对表情描述不够精确。LBP-TOP(Local Binary Patterns from Three Orthogonal Panels)<sup>[4]</sup>特征结合时空域角度考虑, 从 3 个正交平面提取表情图像序列的动态纹理特征, 能够更好地表达人脸表情的实质信息。然而, 传统的均匀 LBP 将至多有 2 次 1 和 0 之间转变的二进制数视为均匀模式, 实验研究表明, 在某些人脸图像中, 均匀模式并非占主导地位, 传统的 LBP 并不一定能反映出人脸图像的主要结构信息。同时, 传统 LBP 仅提取图像高频域的纹理特征, 忽略了低频的信息。

因此, 本文提出一种基于多频域图像的改进 LBP-TOP 表情识别方法, 不仅保留非均匀模式, 而且实现了 LBP 编码的旋转不变性, 并将最大间隔球形支持向量机(Maximal-margin Spherical-structured Support Vector Machine, MSSVM)<sup>[5]</sup>应用到模型的训练和样本分类识别中, 以取得更好的识别效果。

### 2 基于多频域图像序列的LBP-TOP特征提取

#### 2.1 表情图像小波分解

本文所提出的算法是在提取表情图像的多频率特征的基础上实现的, 对各帧图像进行小波分解, 保留一级分解的逼近图像经过小波分解, 原始图像被分解成 4 个子图像。由于表情特征主要体现为形状变化和纹理变化特征, 可舍去子图 HH, 仅考虑在 LL、LH、HL 子图像上提取特征。

#### 2.2 LBP-TOP算子

在标准的 VLBP 中, 参数  $P$  决定了特征向量的大小。当  $P$  较大时, LBP<sup>[6]</sup>算子二进制位将变长; 当  $P$  较小时, 特征向量会减小, 但同时意味着丢失更多的表情信息。与此同时, VLBP 仅提取各帧内部信息, 表情动作的时域信息被忽略。为了解决上述问题, 提出一个简单的算子, 在 XY、XT 和 YT 3 个正交面上提取级联 LBP 特征, 统计这 3 个方向上的表情变化信息。通常, 一个图像序列被认为是以时间轴  $T$  为中心的 XY 面的堆叠, 但却忽视了图像序列也可以是以  $Y$  轴为中心的 XT 面的堆叠, 或是以  $X$  轴为中心的 YT 面的堆叠。如图 1 所示, XT 平面和 YT 平面提供了关于时空过渡的大量信息。

研究表明, 人脸表情图像的主要模式并非总是均匀模式, 因此, 丢弃非均匀模式可能会造成信息的丢失。与此同时,

**基金项目:** 国家自然科学基金资助项目(60673190)

**作者简介:** 孔 健(1986—), 男, 硕士研究生, 主研方向: 图像处理, 模式识别; 詹永照, 教授、博士、博士生导师

**收稿日期:** 2010-02-15 **E-mail:** justus\_kuu@126.com

传统的 LBP 算子并不是旋转不变的, 因为中心点周围的近邻点的强度值将随着圆形结构的旋转而改变。对应的模式也会不同。

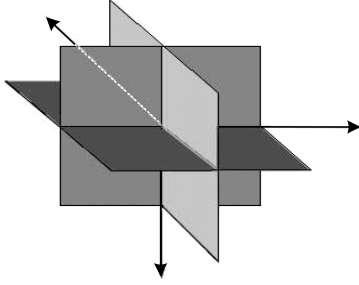


图1 LBP-TOP 结构

$T/\text{帧}$

根据上述研究提出一种改进的 LBP, 定义为  
 $LBP(P, R) = \min(Rot(LBP(P, R), n))$

其中,  $P$  是近邻点个数;  $R$  是圆形半径;  $n=1, 2, \dots, P-1$ 。

函数  $Rot$  将  $P$  位的环形二进制码  $LBP(P, R)$  逆时针旋转  $n$  次。由上式可见, 当一种模式是由另一种模式旋转一定角度后得到时, LBP 把这 2 种模式视作同一类型, 并将 LBP 扩展到立体空间中, 从 3 个不同平面提取 LBP 特征。不但抓住了时空域的信息, 同时实现了表情图像的旋转不变性。

### 2.3 多频率图像分块LBP-TOP特征提取方法

由于纹理方向不可预知, 对于帧之间的二元局部模式, 采用椭圆形的结构来定义中心点的所有邻域点, 而不是直接的线性结构。定义在各正交面提取的 LBP 编码, 这样就从 3 个纹理方向同时考虑了纹理和动作特征, 融合了空域和时域统计信息。不同正交平面计算 LBP 值可以选择不同数量的邻域点。

在  $X-Y-T$  坐标系中, 计算各点的坐标, 从 3 个平面得到像素的二进制编码并进行逆时针旋转, 求取最小特征值。采用分块法对一定区域中提取 LBP 特征, 对分块中的每个点的 LBP 特征值进行相加得到该分块的特征。

## 3 最大间隔球形支持向量机

最大间隔球形支持向量机(MSSVM)<sup>[5]</sup>采用类似支持向量机的建模结构, 力图用一个超球来分离正负类样本, 寻找类间最大分类间隔以提高分类器的泛化能力, 能够实现正负类类间间隔的增大和正类类内体积的减小, 从而提高分类器的分类能力。

MSSVM 如图 2 所示, 正类样本用+号表示, 负类样本用-号表示。

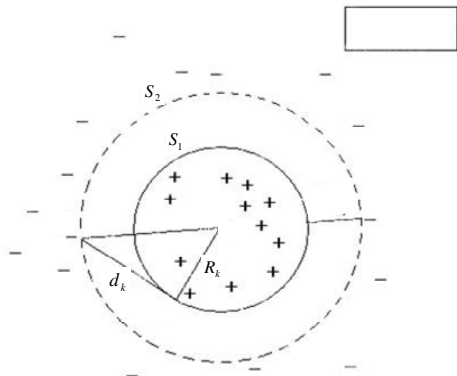


图2 最大间隔球形支持向量机(MSSVM)

设 2 类样本为超球可分, 存在 2 个同心超球  $S_1$ 、 $S_2$ , 球

心为  $O$ ,  $S_1$  为小超球,  $S_2$  为大超球,  $S_1$  将正类样本包裹其中,  $S_2$  将负类样本排除其外。由图 2 可知, 正负类类间间隔为  $\sqrt{R_k^2 + d_k^2} - R_k$ , MSSVM 的目标为最大化类间间隔, 等价于最大化  $d_k^2$  和最小化  $R_k^2$ , 并建立相应的数学模型和定义类判决规则。

模型首先通过非线性变换  $\phi$  将训练样本点映射到一个高维特征空间, 然后在此特征空间中找到  $K$  个最小半径超球, 使所有的  $k$  类正类样本都被包裹在第  $k$  个超球中, 而负类样本被排除在外。

如图 2 所示, 取离超球表面最近的负类样本点, 它到该超球面的距离即为类间间隔, 最大化此间隔可以让负类样本远离超球面。通过引进一个间隔参数  $d_k$ , 并使其最大化使正负类样本相分离。求解下面带约束条件的二次规划问题, 可以找出半径为  $R_k$  和球心为  $a_k$  的类超球  $S_k$ 。构造 MSSVM 的模型为

$$\begin{aligned} & \text{minimize } R_k^2 - M d_k^2 + \frac{C}{N_k} \sum_{i: y_i = k} \xi_i + \frac{C}{N_{\bar{k}}} \sum_{l: y_l \neq k} \xi_l \\ & \|\phi(x_i - a_k)\|^2 \leq R_k^2 + \xi_i \quad \forall i, y_i = k \\ & \|\phi(x_l - a_k)\|^2 \geq R_k^2 + d_k^2 - \xi_l \quad \forall l, y_l \neq k \\ & \xi_i \geq 0, \xi_l \geq 0, \forall i, l \end{aligned}$$

其中,  $N_k$  和  $N_{\bar{k}}$  分别是第  $k$  类以及非  $k$  类的样本数; 正类样本和负类样本分别用  $i, j$  和  $l, m$  计数;  $\xi_i$  和  $\xi_l$  为松弛因子; 参数  $C \geq 0$  为惩罚因子; 参数  $M \geq 0$  用于两者的折衷。引入 Lagrange 函数来求解对偶变量优化问题。求解对偶二次规划问题, 得到 Lagrange 乘子  $\alpha_i$  和  $\alpha_l$ , 使得球心  $a_k$  作为  $\phi(x_i)$  和  $\phi(x_l)$  的一个线性组合, 可以求得数据点  $\phi(x)$  到第  $k$  类超球球心  $a_k$  的距离, 得到  $a_k$  的值, 随后可以利用 KKT 条件求出第  $k$  类超球半径  $R_k$  和间隔参数  $d_k$ 。

判决规则: 给定一个训练样本的集合构造  $k$  类超球  $S_k$ 。定义相似度函数  $\text{sim}(x, S_k) = -\|\phi(x) - a_k\|^2$ ,  $\phi(x)$  越靠近  $a_k$ ,  $\text{sim}(x, S_k)$  的值越大,  $x$  属于  $k$  类的概率越大。

## 4 实验结果与分析

本文在 Cohn-Kanade 人脸表情库上进行实验。每种人脸表情选择 20 个图像序列作为训练样本进行模型训练, 选择 40 个图像序列作为测试样本, 共 240 个图像序列, 每个序列都有 10 帧图像。

### 4.1 表情图像的预处理

表情图像的预处理包括表情图像子区域的分割以及表情图像的归一化处理。对一个长度为  $N$  (如  $N=10$ ) 的表情图像序列来说, 抽取帧间间隔为  $L$  的 3 帧图像进行特征提取和识别。

在预处理前先采用 Adaboost 算法来检测表情图像中的人脸, 然后采用双线性插值算法实现图像的尺寸归一化, 经过处理后的人脸表情图像大小为  $180 \times 180$  个像素。通过对表情图像灰度直方图的修正从而调整图像的均值和方差来完成灰度归一化处理。

### 4.2 基于小波分解的子图像序列LBP-TOP特征提取

首先, 对预处理之后的人脸表情图像序列进行小波分解, 图 3 为各帧图像小波分解后生成的子图, 其中, 子图 LL 保持了原图像的低频信息, 子图 LH 保持了原图像的垂直边缘

细节,子图 HL 保持了原图像的水平边缘细节,于是形成了不同频率的图像子序列,各序列大小均为 3 帧。由于小波分解后的不同子图包含原图像不同频率的信息,这些子图刻画原图像信息的角度和精细程度是不同的,因此可按照不同的分块大小进行分块特征提取。选取  $12 \times 12$ 、 $9 \times 9$ 、 $6 \times 6$  分块大小,对各频率表情图像序列进行分块实验,经过结果比较,对不同小波分解子图分别采用  $9 \times 9$ 、 $6 \times 6$ 、 $9 \times 9$  分块时,识别率最好,然后提取各频率子图像序列的分块 LBP-TOP 特征,对于二元局部模式的参数选择,图 4 的实验分析表明,当  $P$  均为 8、 $R$  均为 3 时,表情识别率为最高。

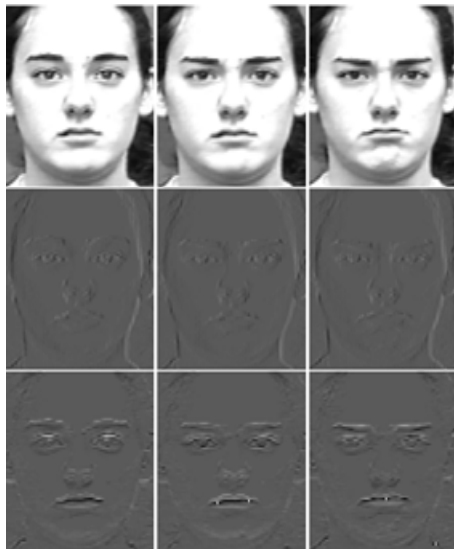


图 3 各帧图像小波分解后子图像

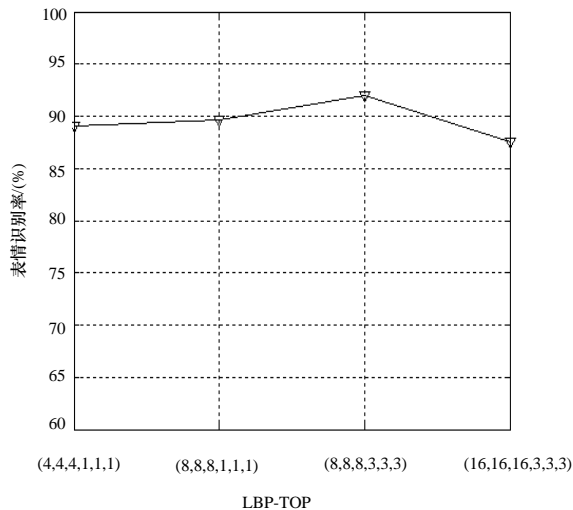


图 4 不同的参数识别率比较

#### 4.3 图像序列表情识别实验与结果分析

为了验证本文提出的表情特征和分类器的有效性,基于统一的训练和测试样本,设计了不同的 5 种方法作对比试验,如表 1 所示。分类器所选核函数均为高斯径向基核函数, $r=10$ ,MSSVM 参数  $C_1=C_2=10$ , $M=1.5$ ,采用一对多方

式构造多分类器,这里固定参数的目的在于只考虑算法的相对性能,而不考虑其绝对性能。

表 1 列出了几组方法各种表情的识别率。由表 1 可看出,提取多频域改进的 LBP-TOP 特征和使用 MSSVM 分类器的识别结果优于其他的识别结果,提取图像序列小波分解后的 LBP-TOP 特征更能有效地反映表情运动的本质,同时利用最大间隔超球支持向量机明显提高了各种表情的分类能力。

表 1 各方法表情识别结果 (%)							
方法	高兴	惊讶	悲伤	害怕	生气	厌恶	平均
VLBP+SVM	90.0	87.5	83.3	83.3	91.7	83.3	86.5
VLBP+MSSVM	91.7	91.7	83.3	88.3	93.3	85.0	88.9
LBP-TOP+SVM	91.7	93.3	85.0	83.3	90.0	88.3	88.6
LBP-TOP+MSSVM	96.7	95.0	88.3	90.0	91.7	86.7	92.0
本文方法	97.5	97.5	91.6	88.3	97.5	87.5	93.3

同时还进行了算法运行耗时的实验,在配置为 Pentium(R) D 3.00 GHz/1 GB 的计算机上进行测试。从预处理、提取特征到分类识别来统计耗时,本文方法平均耗时为 0.073 s。这表明,该方法处理速度快,满足实时图像序列表情识别的要求。

#### 5 结束语

本文提出一种基于 LBP-TOP 和最大间隔球形支持向量机相结合的方法,对人脸表情图像序列进行小波分解,提取不同频率子图像序列 LBP-TOP 特征,使用最大间隔球形支持向量机(MSSVM)进行样本序列的训练和识别。实验结果表明,本文方法能得出较好的平均识别率,并且特征维数明显减少,识别速度快,通过实验的对比表明其有效性。

#### 参考文献

- [1] Zhu Jianxiang, Su Guanda, Li Yingchun. Facial Expression Recognition Based on Gabor Feature and Adaboost[J]. Journal of Optoelectronics Laser, 2006, 17(8): 993-998.
- [2] Zuo Kunlong, Liu Wenyao. Facial Expression Recognition Using Active Appearance Models[J]. Journal of Optoelectronics Laser, 2004, 15(7): 853-857.
- [3] Hsieh C K, Lai Shanghong, Chen Yunchang. Expressional Face Image Analysis with Constrained Optical Flow[C]//Proc. of IEEE International Conference on Multimedia and Expo. Hannover, Germany: [s. n.], 2008: 1553-1556.
- [4] Zhao Guoying, Pietikainen M. Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6): 915-928.
- [5] Hao Peiyi, Chiang J H, Lin Y H. A New Maximal-margin Spherical Structured Multi-class Support Vector Machine[J]. Applied Intelligence, 2009, 30(2): 98-111.
- [6] 崔 洁, 冯晓毅. 基于粗略到精细分类的面部表情识别方法[J]. 计算机工程, 2007, 33(5): 168-172.

编辑 任吉慧