

一种加权矩形提取表情特征的方法*

林庆^{1,2}, 赵翔¹, 胡瑞瑞¹, 詹永照¹, 王士同^{2,3}

(1. 江苏大学 计算机科学与通信工程学院, 江苏 镇江 212013; 2. 南京理工大学 计算机系, 南京 210094; 3. 江南大学 信息学院, 江苏 无锡 214122)

摘要: 表情特征提取是面部表情识别的一个关键步骤。针对目前特征提取效率低的情况, 通过分析 Gabor 特征提取的性质和积分图像计算效率高的特点, 提出一种利用积分图像技术和矩形模板计算面部特征点特征的方法, 模板模拟 Gabor 的多尺度性, 每个模板定义相应的权值, 表情图像按照 Gabor 的各个方向旋转, 使用旋转图像积分图和加权模板而不是在积分图上旋转模板提取特征点的特征值, 最后将此特征值用于表情分类。实验结果表明, 该方法在识别结果相当的情况下极大地提高了特征提取的效率。

关键词: 加权矩形; 积分图像; 矩形模板; 特征提取; 表情识别

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1001-3695(2010)06-2364-03

doi:10.3969/j.issn.1001-3695.2010.06.105

Weighted rectangle method for expression feature extraction

LIN Qing^{1,2}, ZHAO Xiang¹, HU Rui-rui¹, ZHAN Yong-zhao¹, WANG Shi-tong^{2,3}

(1. College of Computer Science & Communication Engineering, Jiangsu University, Zhenjiang Jiangsu 212013, China; 2. Dept. of Computer, Nanjing University of Science & Technology, Nanjing 210094, China; 3. School of Information Technology, Jiangnan University, Wuxi Jiangsu 214122, China)

Abstract: The expression feature extraction is a key step in facial expression recognition. In view of the existing low efficiency in feature extraction, by analyzing the Gabor qualities and the high computational efficiency of integral image, this paper presented a method that combined integral image with rectangle template to obtain the value of facial feature points. These templates with homologous weight simulated the multi-scale of Gabor and expression image rotated via every direction of Gabor. Obtained the values of feature points not by rotating templates on rotation integral image but the joint of weighted templates and rotation integral image. Finally used these feature values for expression classification. The experimental results show that this method improves efficiency while extracting facial feature and does not affect the result heavily.

Key words: weighted rectangle; integral image; rectangle templates; feature extraction; expression recognition

0 引言

随着人机交互领域的发展, 面部表情识别越来越成为模式识别领域的研究热点。人脸表情含有丰富的信息, 它是除了语言之外重要的信息交流方式, 实际上, 表情也是最直接、最显著的信息交流方式。

表情特征提取是表情识别的一个关键步骤, 特征提取的好坏对识别结果产生巨大的影响。目前, 特征抽取方法主要分为基于表观的特征提取和基于几何关系的特征提取, 无论采用哪种提取方法, 快速有效地提取表情特征是特征提取的关键。

积分图像技术^[1,2]是一种快速计算矩形窗口内灰度值的方法。它与矩形窗口的大小无关, 只需要搜索矩形窗口四个顶点的值就可以求出窗口内像素的灰度值总和, 计算全部是加减运算, 速度快。因此, 在需要多次求取矩形窗像素的灰度值总和时, 利用积分图像技术可以快速计算, 提高计算效率。

本文利用积分图像的性质, 将要计算的窗口看做带有权值

的不同大小的矩形覆盖在图像上, 模拟 Gabor 滤波器^[3,4]的多尺度多方向性提取表情特征, 从而提高特征提取的效率。

1 积分图像的概念

首先, 假设原图像中任意一点 (i, j) 的灰度值为 $p(i, j)$, 则积分图像^[1,2]中任意点 (x, y) 的积分值 $ii(x, y)$ 表示该点左上角区域灰度值之和, 如图 1(a) 所示, 定义为

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} p(x', y') \quad (1)$$

从左到右、从上到下计算积分图像, 随着像素矩阵的增大, 计算量呈线性增长。也可根据下面递推式

$$ii(x, y) = p(x, y) + ii(x-1, y) + ii(x, y-1) - ii(x-1, y-1) \quad (2)$$

从左到右、从上到下计算。

利用积分值可以计算积分图像上任意大小矩形区域的像素和, 无论矩形大小, 只需取出矩形四个顶点的积分值即可。设矩形窗内灰度值总和为 W , 矩形四个顶点分别为 (i_1, j_1) , (i_2, j_2) , (i_3, j_3) , (i_4, j_4) , 如图 1(b) 所示, 则有

收稿日期: 2009-09-20; 修回日期: 2009-11-17 基金项目: 国家自然科学基金资助项目(60673190)

作者简介: 林庆(1962-), 男, 福建厦门人, 副教授, 硕导, 主要研究方向为数据挖掘、模式识别与人工智能(milam0479@126.com); 赵翔(1980-), 男, 硕士研究生, 主要研究方向为模式识别; 胡瑞瑞(1982-), 女, 硕士研究生, 主要研究方向为图像处理、模式识别与数据挖掘; 詹永照(1962-), 男, 教授, 博导, 主要研究方向为分布式系统、多媒体技术、模式识别; 王士同(1964-), 男, 教授, 博导, 主要研究方向为人工智能/模式识别、神经模糊系统、生物信息学及其应用。

$$W = \ddot{u}(i_4, j_4) + \ddot{u}(i_1 - 1, j_1 - 1) - \ddot{u}(i_2, j_2 - 1) - \ddot{u}(i_3 - 1, j_3) \quad (3)$$

因此,无论需要计算多大矩形区域内的像素和,均可以利用矩形四个点求出,只需要常数时间,大大提高了计算效率。

2 表情特征提取

表情变化主要集中在五官上,根据表情变化的特征,主要在最能表征表情变化的位置标定特征点,各个特征点代表的意义所起的作用不同。人脸表情特征点的选取应该基本代表人脸表情变化的信息,可以在脸部增加一些点来提高识别率,但是由于这些点不容易标定,标定错误不但降低效率,而且影响识别率^[5]。本文只在眉毛等处标定 26 个特征点,利用主动形状模型(active shape model, ASM)确定特征点的位置^[6]。

Gabor 滤波器有效应用在特征提取和模式识别领域。Gabor 滤波器利用不同尺度、不同方向的小波函数与图像卷积来提取特征,在时间域和空间域有很好的局部性,而且能够提取表情图像的细节信息,但是计算量大,这在实时性和准确率之间产生了瓶颈,在保证准确率的情况下,如何提高特征提取效率显得尤为重要。本文模拟 Gabor 小波多尺度、多方向的性质,在多方向积分图像基础上,利用多尺度的加权矩形模板提取表情特征。

矩形模板的尺度性主要体现在高 h 和宽 w ,它们分别由 Gabor 核函数的 k_v, λ 确定。其中: $k_v = 2^{-(v+2)/2} \pi, \lambda = \pi / (2k_v)$ ($v=0,1,2$)。经过实验, $h = k_v / 2, w = \lambda / 2$,随着 v 值的不同,得到三个不同大小的矩形模板,将 k_v 与 λ 的比值作为矩形模板的权值 ω 。 k_v, λ 和 ω 的关系如表 1 所示。

表 1 k_v, λ 和权值 ω 的关系

$k_v = 2^{-(v+2)/2} \pi$	$\pi/2$	$\pi/2^{3/2}$	$\pi/4$
$\lambda = \pi / (2k_v)$	1	$2^{1/2}$	2
$\omega = k_v / \lambda$	$\pi/2$	$\pi/4$	$\pi/8$

计算矩形模板内像素的灰度值时,模板模拟 Gabor 核函数的多方向性,每个模板依次旋转六个方向(由 $\pi/6$ 到 π ,步长为 $\pi/6$),如图 2 所示。

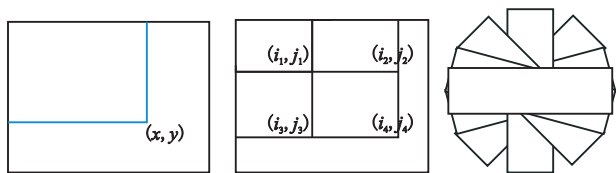


图 1 (x,y)的积分值区域与矩形W的灰度值区域
图 2 六个方向的矩形模板

对于旋转角度为 ϕ 的矩形模板,利用式(2)对原始图像计算积分图;然后将旋转 ϕ 的模板中心在积分图上移动,对准每个特征点,通过每个模板的宽、高计算模板的四个顶点位置坐标,判断所围区域内像素的灰度值总和。当 $\phi = \pi/2, \pi$ 时,矩形模板与图像的 x, y 轴平行(或垂直),此时模板覆盖的像素点容易求得,矩形区域的灰度值总和由式(3)求得,计算简单快速;但是当 $\phi = \pi/6, \pi/3, 2\pi/3, 5\pi/6$ 时,矩形模板与图像分别成 ϕ 角,四个顶点的位置确定以后,矩形的边落在部分像素方格上,这些方格的取舍根据插值方法确定,通过判断重新确定有缺口的矩形或是锯齿形矩形,这在一定程度上增加了运算时间,降低了运算效率。此外,为了便于利用式(2),对覆盖区域尽可能最大化地划分为与 x, y 轴平行(或垂直)的矩形块,计算每个矩形块的灰度值之和,某些像素点可能不属于任

意矩形块,这在判断模板覆盖区域内灰度值总和时容易遗漏,显然对结果引起一定的误差,甚至产生错误的结果。

从上面分析可知,当旋转模板时,为了确定某个旋转模板覆盖的像素,首先要判断矩形边缘的像素方格,然后计算模板覆盖范围内最大矩形块的灰度值和不属于任意矩形块的像素点,这不仅增加了运算量,还将引起较大的误差。

为了减小误差,这里提出变换旋转对象的思想,将需要旋转六个方向的模板始终保持垂直状态,将图像旋转 ϕ ,然后计算旋转 ϕ 角度的图像积分图,由于旋转后图像范围变大,原图像范围之外的像素灰度值设为 0,这样不影响旋转后图像积分图的值,原图像范围之外的像素灰度值为 0,由图像范围增大而增加的计算量可以忽略;将方向不变的垂直模板在旋转后的积分图上移动,依次对准要提取的特征点位置,此时,只需要判定矩形四个顶点所在的行和列,边缘与顶点所在的行列相同,这与旋转模板的方法相比节省了边缘像素点的判定时间;最后利用式(3)计算矩形区域内灰度值总和,不需要将模板覆盖的区域分块,相比旋转模板节省了分块时间。

利用积分图像技术,结合矩形模板提取图像表情特征值时,矩形模板保持垂直不变,将要提取特征的图像根据需要旋转 ϕ ,然后在新的图像范围内计算积分图。旋转后范围增大的部分用 0 填充,这样不仅不影响积分图的生成,而且尽可能降低了增加范围的计算时间。最后将垂直模板的中心匹配新的积分图的特征点,此时计算特征点在新坐标位置处的像素灰度值和。模板与原图像的旋转对照图如图 3 所示。圆点表示特征点,矩形框表示矩形模板,左图表示旋转矩形模板 ϕ ,右图表示矩形模板不变,旋转原图 ϕ 。

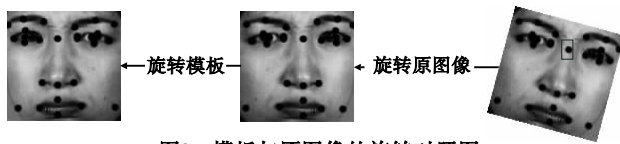


图 3 模板与原图像的旋转对照图

设表情特征点为 (x_c, y_c) ,将每一个矩形模板的中心对准各个特征点,利用积分图像计算矩形区域内的灰度值之和 $RSum(x_c, y_c)$,将权值与灰度值总和作乘法提取特征^[7],定义如下:

$$Y(x_c, y_c) = \omega \times RSum(x_c, y_c) \quad (4)$$

其中: ω 是矩形模板的权值,共有三个大小的矩形模板, ω 共有三个值; $RSum(x_c, y_c)$ 表示矩形区域灰度值总和,通过积分图像计算得到,由式(4)得到各个特征点的特征值。

3 实验分析

本文采用日本女性表情库(JAFFE)进行实验,表情库共有 10 个日本女性提供的 213 幅图像,分为生气、厌恶、恐惧、高兴、中性、悲伤和惊讶七种表情,每人每种表情包含 2~4 幅图像,均为 256×256 的灰度图像。由于本文只研究人脸表情,为了消除头发等的影响,提高实验效率,剪切仅包含全部特征点的人脸图像,经过图像预处理后得到各组图像的纯人脸图像。

实验在 MATLAB 环境下,使用同一类图像的相同特征点,对比 Gabor、模板旋转和图像旋转三个特征提取各自在每个方向上所用的平均时间。下面仅列出 Gabor 和图像旋转在每个方向上的平均时间,如表 2 所示。

表 2 Gabor 与图像旋转在各个方向上的平均时间

方法/ms	$\pi/6$	$\pi/3$	$\pi/2$	$\pi \times 2/3$	$\pi \times 5/6$	π
Gabor	0	30	30	29	30	29
图像旋转	30	30	0	30	29	0

表 2 是两种方法在某一类图像上特征提取所需要的平均时间,分别在不同方向上测试它们所需要的时间。从表中可以看出:a) Gabor 所用的总时间大于图像旋转所用的总时间;b) 角度不同,但每个方向所用时间基本相同,说明角度对提取特征花费的时间影响不大,两者相同;c) 在 $\pi/2$ 和 π 两个方向上,图像旋转与其他几个方向相比,所用时间极少,实验基本测量不到在这两个方向上所用的时间;d) 理论上 Gabor 每个方向所用时间基本相等,但是在 Gabor 的 $\pi/6$ 方向上,与图像旋转在 $\pi/2$ 和 π 方向相同,测量不到时间数据,在 π 上时间不太稳定,原因在进一步研究中。

总体上,旋转图像时,旋转后图像范围变大,需要计算每个像素点旋转后的坐标,耗时较长,但与 Gabor 提取相比,本文的图像旋转方法所用的总时间明显比它少。因此,对数据库图像的所有特征点抽取特征值,本文的方法节省了大量的时间,提高了整个系统的识别效率。

为了验证本文提出的加权矩形旋转图像方法提取的表情特征是否有效,本文将七种表情训练为七个模板,使用最近邻方法^[8]对待测图像进行分类,将待测图像提取的特征分别与七个模板比较,待测图像与哪个模板的距离最近,就把待测图像划分到此类中。实验表明它们的识别性能相当,它们的识别率如表 3 所示。

表 3 两种识别方法的识别率

识别方法	识别率/%
Gabor + 近邻	85.57
旋转图像 + 近邻	83.86

从表 3 中可以看出,旋转图像方法的识别率略低于 Gabor 方法,相差 1.71 个百分点,相差不大。尽管表 2 中每幅图像旋转图像所用的平均时间略少于 Gabor 特征提取所用的时间,但

对于整个数据库而言,提高了表情识别系统的效率,对于识别率要求不太高的情况下,此方法有较好的应用。

4 结束语

本文模拟 Gabor 特征提取的尺度和方向性,提出旋转图像的加权矩形特征提取方法,理论上论证了图像旋转比模板旋转所用的时间少,而且通过在同一环境下,就一类图像的特征提取所需的平均时间进行对比,实验表明,本文方法所需的时间比 Gabor 提取和模板旋转所用的时间少,旋转图像提取的特征与 Gabor 和旋转模板提取的特征,它们用于表情识别的识别准确率相当,从总体上提高了系统的效率,缓解了实时性与准确性之间的瓶颈问题,验证了此种方法的有效性。

参考文献:

[1] VEKSLER O. Fast variable window for stereo correspondence using integral images [C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2003;1556-1561.

[2] 邵平,杨路明. 基于模板分解和积分图像快速 Kirsch 边缘检测 [J]. 自动化学报,2007,33(8):795-800.

[3] 陆慧聪. 面部表情识别系统中表情特征提取与识别算法的研究 [D]. 南京:东南大学,2006.

[4] 张一鸣. 人脸表情识别 [D]. 大连:大连理工大学,2006.

[5] ZHANG Zheng-you. Feature-based facial expression recognition: experiments with a multi-layer perception [R]. [S.l.]:Institut National De Recherche en Informatique ET EN Automatique,1998.

[6] 陈培俊. 基于静态图像的人脸表情识别研究 [D]. 成都:西南交通大学,2007.

[7] CHOI W P, TSE S H, WONG K W, et al. Simplified Gabor wavelets for human face recognition [J]. Pattern Recognition, 2008, 41(3):1186-1199.

[8] DUDA R O, HART P E, STORK D G. 模式分类 [M]. 李宏东,姚天翔,等译. 2 版. 北京:机械工业出版社,2003.

(上接第 2363 页)

[19] FERNÁNDEZ-ESCRIBANO G, CUENCA P, ORIACO-BARBOSA L, et al. A fast intra-frame prediction algorithm for MPEG-2/H.264 video transcoders [C]//Proc of IEEE ICIP. 2005;684-687.

[20] TANG Qiang, NASIOPOULOS P, WARD R K. Compensation of requantization and interpolation errors in MPEG-2 to H.264 transcoding [J]. IEEE Trans on Circuits and System for Video Technology, 2008, 18(3):314-325.

[21] FERNÁNDEZ-ESCRIBANO G, KALVA H, CUENCA P, et al. A fast MB mode decision algorithm for MPEG-2 to H.264 P-Frame transcoding [J]. IEEE Trans on Circuits and System for Video Technology, 2008, 18(3):314-325.

[22] FERNÁNDEZ-ESCRIBANO G, KALVA H, CUENCA P, et al. A first approach to speeding-up the inter mode selection in MPEG-2/H.264 transcoders using machine learning [J]. Springer Multimedia Tools and Application, 2007, 35(2):225-240.

[23] KIM S E, HAN J K, KIM J G. Efficient motion estimation algorithm for MPEG-4 to H.264 transcoder [C]//Proc of IEEE ICIP. 2005:656-659.

[24] AHMAD I, WEI Xiao-hui, SUN Yu, et al. Video transcoding: an overview of various techniques and research issues [J]. IEEE Trans on Multimedia, 2005, 7(5):793-804.

[25] KUO T Y, CHAN Chen-hung. Fast variable block size motion estimation for H.264 using likelihood and correlation of motion field [J]. IEEE Trans on Circuits and System for Video Technology, 2006, 16(10):1185-1195.

[26] KIM J H, KIM B G. Fast block mode decision algorithm in H.264/AVC video coding [J]. Journal of Visual Communication and Image Representation, 2008, 19(3):175-183.

[27] TSAI A C, PAUL A, WANG Jia-ching, et al. Intensity gradient technique for efficient intra-prediction in H.264/AVC [J]. IEEE Trans on Circuits and System for Video Technology, 2008, 18(5):694-698.

[28] SHEN Hui-feng, SUN Xiao-yan, WU Feng. Fast H.264/MPEG-4 AVC transcoding using power-spectrum based rate-distortion optimization [J]. IEEE Trans on Circuits and System for Video Technology, 2008, 18(6):746-755.

[29] SHEN Bo. Perfect requantization for video transcoding [J]. Springer Multimedia Tools and Application, 2007, 35(2):163-173.