

一种适合于表情分析的人脸二值边缘图像的提取

宋加涛^{1),2)} 刘济林¹⁾ 池哲儒²⁾ 王蔚³⁾

¹⁾(浙江大学信息与通信工程研究所, 杭州 310027)

²⁾(香港理工大学电子及资讯工程系, 香港) ³⁾(宁波工程学院电子系, 宁波 315016)

摘要 边缘图像是人脸图像的一种重要表征方法, 它对人脸图像的分析有重要的作用。提出了一种提取人脸二值边缘图像的新方法。该方法利用小波变换进行图像的高频重构, 并利用其多尺度分析特性进行图像边缘的提取, 包括两次二值化过程和一次去噪过程。提取的边缘图像具有较高的质量, 同一人脸部件的边缘连通性好, 不同部件的边缘粘连现象少, 且具有较好的光照鲁棒性。将所提取的二值边缘图像用于AR库和Yale库图像的人脸4种动作单元的识别, 识别率在93%以上, 初步表明其适合于人脸图像的表情分析。

关键词 人脸图像 边缘提取 二值化 去噪 表情分析 小波变换

中图分类号: TP391.4 **文献标识码**: A **文章编号**: 1006-8961(2004)10-1182-08

Edge Extraction of Face Images for Expression Analysis

SONG Jia-tao^{1),2)}, LIU Ji-lin¹⁾, CHI Zhe-ru²⁾, WANG Wei³⁾

¹⁾(Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027)

²⁾(Department of Electronic and Information Engineering, the Hong Kong Polytechnic University, Hong Kong)

³⁾(Department of Electronics, Ningbo University of Technology, Ningbo 315016)

Abstract Edge map is an important representation of human face image. It is of significance for automatic face image analysis. This paper presents a new method for the extraction of face image edges and a scheme for facial expression analysis based on the resulting binary edge images (BEI). Making use of the multi-resolution property of wavelet transform, the edge detection method includes two binarization steps and a noise removing step. Our method provides a robust solution to face image edge extraction, in particular, if changes of lighting conditions are encountered. Experimental results show that different components in a BEI are well separated and pixels of the same face component are connected well. Preliminary experimental results on facial expression analysis on the subsets of AR and Yale face databases show that recognition rates of above 93% can be achieved when our facial analysis scheme is used for recognizing four action units.

Keywords face image, edge extraction, binarization, noise removing, facial expression analysis, wavelet transform

1 引言

边缘是图像的重要组成部分, 它存在于图像灰度的突变处, 反映了图像中对象的形状、尺寸、组成结构等重要信息, 因此, 边缘图像常被用作目标分割、检测、识别和跟踪等^[1~6]。

检测图像边缘的方法很多^[1], 最经典的是Canny, Sobel and Prewitt算子以及高通滤波方法。

最近几年, 很多其他的方法, 比如高斯滤波器^[7]、神经网络、小波变换等也广泛用于图像边缘的提取。而就人脸图像分析领域而言, 常用的人脸图像边缘提取方法有Soble算子^[3~6]和Gabor滤波方法^[8]。前者属于从空域中进行边缘提取, 而后者则是在频域内提取边缘。比如, 文献[3]~[6]用Sobel算子提取人脸图像的二值边缘, 并将结果用于人脸识别, 取得了较好的效果。文献[5]还给出了一种基于Sobel算子的局部自适应二值化(locally adaptive threshold)

基金项目: 浙江省自然科学基金(602118); 宁波市青年博士基金(02J20102-02)

收稿日期: 2004-01-08; 改回日期: 2004-06-14

人脸边缘提取算法,简称 LAT 算法。Gabor 小波变换广泛应用于人脸图像的识别^[8]和表情分析^[9,10],它有很强的边缘表征能力,但用该方法得到的是灰度边缘,而不是二值边缘。

不同方法所得到的边缘图像的质量不同,而不同的应用对边缘图像质量的要求也不同,到目前为止,没有一种边缘提取方法能满足所有应用的要求^[1]。比如,用 LAT 算法提取的人脸二值边缘图像具有噪声少的特点,因此,当用于人脸识别时,效果较好,但它同样存在边缘的连通性差的缺点,而且在图像分辨率较低时,人脸部件之间的粘连现象也较严重,因此,不适合做人脸表情分析。为了实现人脸表情的自动分析,提出了一种基于小波多尺度分析特性^[11]的人脸二值边缘检测新方法。该方法包括人脸图像的高频重构、单尺度下二值边缘提取、二次边缘提取、边缘图像去噪等过程。用该方法得到的二值边缘图像中人脸部件的细节清晰、噪声少,同一部件的边缘连通性好,而不同部件边缘之间的粘连性少,逼真地反映了原始灰度图像所表达的人脸表情。该方法还具有较好的光照鲁棒性。通过用 AR^[12]和 Yale^[13]人脸库图像进行人脸 4 种动作单元^[14](action units, 或简称 AU)的识别实验,取得了较高的识别率。

2 人脸图像边缘的提取

2.1 单尺度下边缘提取

小波变换是一种信号分析的“数学显微镜”,它可以同时从时域和频域对信号进行分析。从频域上看,小波变换相当于用品质因素恒定的滤波器组对信号进行滤波处理,滤波器组由一个低通滤波器和一个高通滤波器组成。利用小波变换的多分辨率分析,可以将原始信号分解为一个低频分量和若干各高频分量(如图 1 所示),其中每一层的频谱分量是

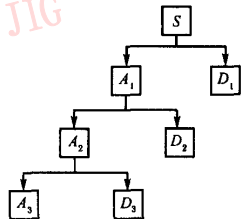


图 1 小波变换多尺度分解

(S 为原始信号, A_i 为低频分量, D_i 为高频分量)

通过对上一层的低频分量做进一步的分解而得到。

图 2(b)、(c)、(d)3 列分别显示了在尺度 1、3、5 下得到的高频重构图像及其对应的灰度直方图,其中,尺度 V 下的高频重构图像是指通过舍弃小波分解中第 V 层的低频分量 A_V , 而只用第 $V, V-1, \dots, D_1$ 层的高频分量, 即 $D_V, D_{V-1}, \dots, D_1$ 重构得到的图像。与图 2(a)列的原始灰度图像相比较,可以看出,在尺度较小(如尺度 1~3)时,由于舍弃的低频分量较多,直方图中像素灰度的动态范围大大减小,形成一个强峰,表现在图像中就是整个图像的灰度较原始图像更均匀,甚至基本消除了人脸区域和周边之间的灰度差异,图像中更多体现的是由高频分量决定的原始图像中的一些边缘,包括脸部轮廓以及人脸中眼睛、嘴巴、鼻子等部件的轮廓。随着尺度的增加,低频分量也逐渐增加,高频重构图像及其直方图又逐渐趋向于原始灰度图像。

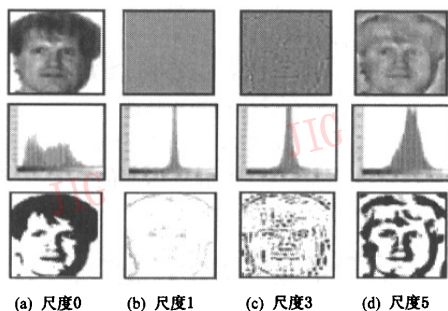


图 2 高频重构图像、相应的直方图以及单尺度下的二值边缘图像(从上至下)

仔细分析高频重构图像及其直方图,可以发现,直方图具有很好的对称性,并且比强峰位置较低的灰度值,对应的正是高频重构图像中边缘处的一些较黑的区域。因此,取高频重构图像的灰度均值偏下一点的值作为阈值对其进行二值化,人脸的轮廓以及眼睛、嘴巴等部件的边缘便可从中抽取出来。

图 2 中第 3 行是采用上述方法分别从原始灰度图像和在 1、3、5 3 个尺度下的高频重构图像中抽取的二值边缘图像(binary edge image, 简称 BEI),其中,对原始图像实际上相当于用全局单阈值对其进行二值化。从中不难发现,图 2(a)列的二值图像中右半部分的前景(即黑色)像素明显比左半部分多,这其实反映的是光照的影响,而图 2(b)、(c)、(d)3 列的二值边缘图像中光照的影响大大减少了。另外,

还可以看到,在小尺度时,人脸部件的边缘并不清晰,也不完整,它们主要由一些并不粘连但靠得很近的小连通分量组成,而随着尺度的增加,这些小连通分量逐渐连接起来,部件变得逐渐清晰和完整。但当尺度进一步增加,比如在尺度5时,二值边缘图像中部件的细节却变模糊了,部件之间的粘连也多了。

2.2 单尺度边缘图像的融合及其二值化

为了获得高质量的二值边缘图像,将各单尺度下获得的二值边缘图像进行融合,然后实行二次二值化,具体步骤如下:

(1) 计算各二值边缘图像之和,即

$$X = \sum_{i=1}^K X_i \quad (1)$$

其中, X_i 为在尺度 i 下得到的二值边缘图像, K 是小波分析中所用的最大尺度值。显然,这样得到的图像 X 是一个多值图像,其像素的取值介于0和 K 之间。

(2) 取 $0 \sim K$ 之间的某个值作为阈值 T , 对 X 进行二值化, 所得结果即为多尺度融合的二值边缘图像, 用 I_{BEI} 表示, 即

$$I_{BEI}(x, y) = \begin{cases} 1 & X(x, y) \geq T \\ 0 & \text{其他} \end{cases} \quad (2)$$

其中, $1 \leq x \leq M, 1 \leq y \leq N$, M 和 N 分别是图像的行数和列数。在实验中, K 的值为6, 阈值 T 取3, 所得的二值边缘图像如图3(a)所示。



图3 二值边缘去噪

2.3 二值边缘图像的去噪

比较图3(a)和图2中第3行中的各个图像, 可以看到, 经过两次边缘提取, 所得的BEI中人脸各部件比较完整, 眼睛的边缘、眼球和眼白等细节非常清晰, 而且基本消除了光照的影响。然而, 在边缘提取过程中也引入了一些噪声。为此, 构造了用人脸图像的另一种二值边缘图像, 即用LAT算法得到的二值边缘图像(为便于区别, 称为 binary edge map, 简称BEM)作为模板, 对用上述方法得到的BEI进行噪声“滤除”的算法。

BEM的获取包括以下几个步骤: 首先用4个方向的Sobel算子和一个低通滤波模板^[5]分别和原始灰度图像做卷积, 其结果记作 $Z_k(x, y)$ ($k=1, 2, 3, 4$) 和

$G(x, y)$, 然后计算每个像素的LAT值, 记为 I_{LAT}

$$I_{LAT}(x, y) = \frac{\max_{k=1 \sim 4} \{ |Z_k(x, y)| \}}{G(x, y)} \quad (3)$$

最后, 用下列规则对 I_{LAT} 进行二值化, 所得结果即为原始图像的BEM, 记作 I_{BEM} :

$$I_{BEM}(x, y) = \begin{cases} 1 & I_{LAT}(x, y) > 1 \\ 0 & \text{其他} \end{cases} \quad (4)$$

用上述方法得到的BEM如图3(b)所示。可以看到, 虽然BEM中人脸及其部件的边缘中常存在断点(即不连续), 但是, 相对于图3(a)所示的BEI, 该边缘图像在眼睛、嘴巴等部件周围区域很少存有噪声, 而且, 如果对原始灰度图像预先进行灰度滤波操作, 则BEM中的噪声可进一步减少。以此少噪声的BEM作为模板, 对BEI进行去噪操作。算法如下:

(1) 将BEI和BEM的对应元素相乘, 得到图像 P , 即

$$P = I_{BEI} \otimes I_{BEM} \quad (5)$$

其中, 符号“ $\otimes$ ”表示两个矩阵对应位置上的元素相乘。

(2) 找出BEI中的所有8连通分量, 并将每个连通分量中的元素分别用整数 $1, 2, \dots, N_c$ 标识, 这里 N_c 是连通分量的数目, 标识的结果保存在矩阵 L 中, 即

$$L(x, y) = \begin{cases} k & (x, y) \in \text{连通分量 } k \\ 0 & (x, y) \in \text{背景} \end{cases} \quad (6)$$

($1 \leq k \leq N_c$)

(3) 将 P 和 L 中的对应元素相乘, 结果保存在矩阵 Q 中, 即

$$Q = L \otimes P \quad (7)$$

(4) 按下列规则逐个处理BEI中的各个连通分量: 如果非零元素 $L(x, y)$ 在 Q 中不存在, 则说明BEI中对应位置上的元素是噪声, 应该除去, 即

$$I_{BEI}(x, y) = \begin{cases} 0 & L(x, y) \neq 0 \\ & \text{且 } L(x, y) \notin Q \\ I_{BEI}(x, y) & \text{其他} \end{cases} \quad (8)$$

去噪以后的BEI如图3(c)所示。可以看到, 与图3(a)相比较, BEI中人脸部件周围的噪声基本上消除了。

3 对BEI的质量评价

3.1 BEI的特点

利用AR和Yale两个人脸数据库中的图像来进行二值边缘的提取实验, 它们代表了两种不同分

分辨率的图像,其中,AR 图像人脸区域的大小约 400×300 ,而 Yale 图像的人脸区域约 210×170 。选用同一个人但不同光照以及不同表情的图像进行实验,所得二值边缘图如图 4 所示。为了比较,图中同时给出了用 LAT 算法得到的二值边缘图像。可以看到,用本文方法所得的 BEI 图像具有以下特点:

(1)部件的形状和轮廓清晰而完整,部件之间的粘连较 BEM 大为减少,这有助于人脸部件的分割以及关键特征点,比如瞳孔、眼角点以及嘴角点的定位。

(2)部件的细节非常清楚,特别在分辨率较高情况下。比如,眼睛和嘴巴的张开或闭合状态表现得很清楚,眼球能够非常容易地从眼白和眼睛的上下边框中分辨出来;

(3)具有很好的光照鲁棒性。比如,在光照较强的情况下,BEM 图像中有些人脸部件,如嘴巴和鼻子已无法完整地分割出来,而在 BEI 中则清晰可见。

(4)充分反映了原始图像中包含的类别差异和表情变化。

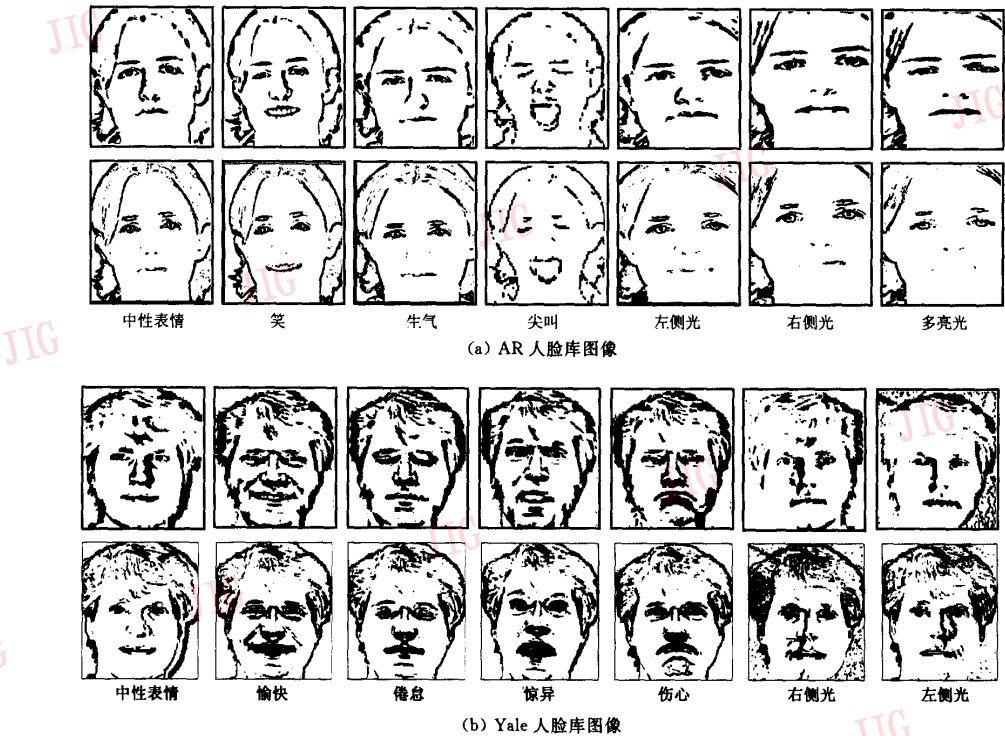


图 4 BEI 及其与 BEM 的对照
(两个库的上面一行:用本文方法得到的 BEI;下面一行:用 LAT 方法得到 BEM)

3.2 图像质量的定量评价

为了更为客观地评价二值边缘图像的质量,采用一种可归属为基于优秀(goodness-based)^[15]一类的方法,选用二值边缘图像中的“连通分量数”、“孔数”(所谓“孔”,就是指二值图像中被前景像素包围的少量背景像素)以及“所有连通块的边界像素数和其总像素数之比”(下文简称“边总比”)共 3 个参数来对图像质量进行评价。由于在人脸分析中常采用的是人脸部件,故设计了基于 BEI 的图像分割算法,将人脸部件所在的区域从整个人脸图像中分割出来。所有的参数统计工作均在该子图像中进行。

对于连通分量数,在不考虑眼镜、胡子等情况下,包括眉毛、眼睛、鼻子和嘴巴以及部分人脸表面纹理在内,其理想值应该是介于 6 至 12 之间。该值越大,说明图像中的噪声越多,或者是人脸部件边缘中的断点越多,有些部件被分隔成了几个部分;该值太小,则可能是属于多个部件的边缘被粘连在一起,或者可能是有些部件被分割成了背景。对于孔数,理想情况下只有眼睛上下边框和眼珠之间、瞳孔中央的高反光点以及张开的嘴巴才可能存在,其理想值应该是一个较小的数(<10),该值太大,则说明很多本应该属于前景的像素被分割成了背景像素,这不

仅在主观效果上大大降低了二值边缘图像与原始灰度图像的相似程度,而且很可能导致在表情分析中眼睛和嘴巴的张闭状态变得难以确定。至于边总比,则其值越小,说明连通分量的内部像素越多,部件的形状越饱满,否则,说明它们是一些只有一个或者两个像素宽的线状图形,而对于这种线状图形,显然,他们被断点分割成几部分的概率也更大。

虽然单纯从技术上讲,图像中的孔、孤立噪声、断点或者粘连现象等并不难消除,但究竟哪一个应该真正被删除,则很难找到一种自适应判断机制。比如,为了消除孤立噪声,则很可能将一些人脸部件消除掉,而为了将孔消除,则又可能将眼珠和眼睛边框之间的眼白消除掉。因此,如果二值边缘图像中存在

的孔、断点及粘连现象越少,则实际应用时所需要的后处理操作也越少、越简单。

从 AR 库和 Yale 库中分别任意抽取 80 幅图像,对 BEI 和 BEM 就上述几个参数分别进行统计,所得结果如图 5 所示。其中,每个子图中的横坐标表示统计时所用的连通分量的最小面积值,比如最小面积值 5,是指只有面积大于或等于 5 的连通分量才作为有效的连通分量被统计进去,小于该值的则作为噪声不被统计。图 5(a)~(c)分别给出了 BEI 中连通分量数的均值(ANCC)、孔数均值(ANH)和边总比均值(ARBT)随最小面积的变化情况。作为比较,BEM 的有关数据也同时给出。

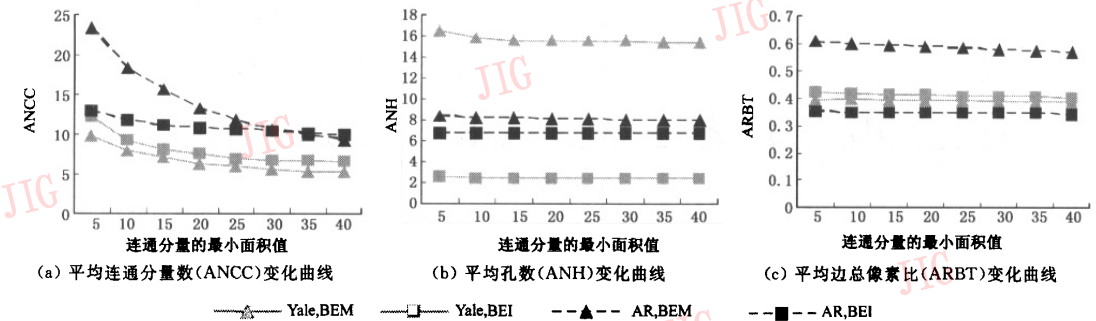


图5 两种二值边缘图像质量的定量评价

对于 AR 图像,从图 5(a)可以看出,虽然 BEI 和 BEM 中的连通分量数都随面积的增加而减小,但前者减小的速度明显比后者慢;在用小面积进行统计时,前者的连通分量数明显比后者少,而用大面积统计时,则趋于相反,并且在面积大于 15 以后,BEI 中连通分量数趋于一个常量(约等于 10)。由此说明,在 BEI 中主要的连通分量是 10 个左右面积较大的分量,它们分别对应人脸中各个部件的边缘和一些纹理,小面积的连通分量数很少(1~2 个),而在 BEM 中,连通分量之间面积相差很大,并且小面积的连通分量数目较大。由于在 BEM 中孤立噪声较少,因此,这些小面积的连通分量是由同一部件的边缘被分割成了几部分形成的。也就是说,对于分辨率较高的 AR 图像,用本文方法得到的二值边缘图像(BEI)中,同一部件的边缘连通性很好,不同部件边缘之间粘连很少(因为其均值介于理想范围之内);而用 LAT 算法得到的二值边缘图(BEM)中,同一部件的边缘连通性较差。

对于 AR 库,从图 5(b)可以看出,BEI 和 BEM

中的孔数都较接近,分别为 6 个和 8 个左右,这和理想值较为接近,并且它们基本不随面积的增大而增加,说明这些孔基本都在大面积的连通分量(即人脸部件)中。就边总比参数而言,图 5(c)显示 BEM 中约为 0.60,而 BEI 中约为 0.35,说明用 LAT 算法所得二值边缘图中边缘主要是一些线状连通分量,而用本文方法所得的二值边缘图中边缘的外围像素仅占其总像素数的 1/3 左右,部件边缘的形状较为饱满。

对于 Yale 库,从图 5(a)可以看到,BEI 的连通分量数比 BEM 中略多,在面积大于 20 以后,前者的均值基本固定在 6.8 个,介于理想范围之内,而后者则继续减低,至 6 个以下,超过了理想范围值的下限。这说明对于分辨率较低的 Yale 图像而言,BEI 中部件的边缘之间虽然可能有一些粘连,但其粘连程度不严重,而且比 BEM 中更轻微。对于该图像库,另一个突出的现象是图 5(b)中所示的两种边缘图像中孔数的差异,BEI 中孔数的均值约 2.5,而 BEM 中约为 15.5,前者介于理想值范围之内,而后

者比理想值大得多,说明在 BEM 中部件边缘的完整性很差。

综上所述,相对于 LAT 算法,在高分辨率时,用本文方法得到的二值边缘图像(BEI)中其属于同一部件的边缘连通性更好,而在低分辨率时,不同部件的边缘之间分离性更好。BEI 中部件边缘的完整性和饱和度在各种分辨率下都好于 BEM 图像。

图 6 给出了在二次边缘提取过程中使用不同的阈值 T 进行二值化所得的图像。可以看出,相对于 T 取 3 的情况,在 T 取 2 时,边缘像素的连通性大大降低,导致一个部件被分割成几个部分,部件的完整性大为降低,这样不利于人脸部件的分割和提取;而当 T 取 4 时,由于过多的像素被分割成前景像素,因此,不同部件之间的粘连性增加,部件的一些细节特征也变模糊了,由此导致部件分割难度增加,特别是不利于部件关键特征点的提取,故 T 的最佳取值为 3。

另外,如果把 AR 图像和 Yale 图像看作是同一

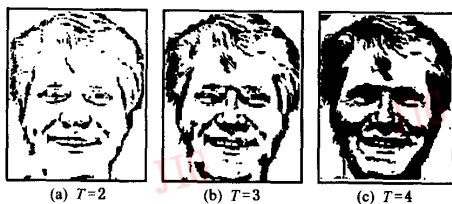


图 6 二次二值化过程中使用不同阈值所得的 BEI

图像金字塔中不同层(即不同分辨率)上的图像,则从图 5(a)可以看到,随着分辨率的提高,BEM 中连通分量的增加远远大于 BEI 中的增加。这主要是因为,在 BEM 中,属于同一部件的边缘在分辨率提高时,被分割成了几部分,而在 BEI 中,连通分量增加值较小,而且造成这种增加的主要原因是由于不同部件之间的粘连减少了,而同一部件的边缘仍然基本保持连通,图 7 给出了 Yale 库图像在不同分辨率下得到的 BEI。由此特性,可以构造基于图像金字塔的部件分割算法,从 BEI 中实现对面脸部件及其特征点的提取。



图 7 不同分辨率下用本文方法获得的二值边缘图像(Yale 库)

4 将 BEI 应用于人脸动作单元的识别

人脸表情的自动识别是人脸图像分析的重要分支之一,它在人机交互方面有着重要的应用,近年来,已经引起人们越来越多的关注。

人脸表情通常用一些人脸上面可以用视觉觉察的运动来解释。在人脸动作编码系统,即 FACS^[14]中,定义了 44 种用于解释人脸表情变化的动作单元(active units, 或者 AU)。由这些动作单元的组,便可以产生各种各样的表情。

为了考察用本文方法所得二值边缘图像(BEI)用于人脸表情分析的可行性,做了一些动作单元识别实验,主要是识别图 8 所示的人脸 4 种动作单元,它们是:AU12(嘴角上翘),AU15(嘴角下拉),AU27(嘴巴大开)和 AU43(眼睛闭合)。

在进行动作单元识别以前,首先进行一些预处理工作,主要包括:(1)对 BEI 图像进一步去噪;(2)

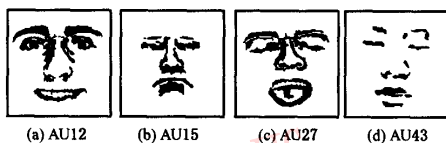


图 8 4 种待识别的人脸动作单元

从 BEI 中定位和分割人脸部件,主要是眼睛和嘴巴;(3)使属于同一部件的边缘更加连通。后两项工作是结合图像金字塔算法完成的。

对于嘴巴,先抽取其两个角点,即图 9(a)中的 A 点和 B 点,然后,求出线段 AB 的中点 O,以及 AB 的垂直平分线与上嘴唇的上边缘和下嘴唇的下

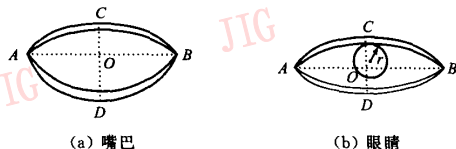


图 9 嘴巴和眼睛模型

边缘的交点 C 和 D , 近似作为嘴唇的最高点和最低点。对于眼睛, 按同样的方法确定其相应的 5 个点。

用于判断动作单元 AU12 存在的条件是: 点 C 位于点 O 的下方; 或者下列公式成立:

$$\frac{|OC|}{|OD|} < \delta_1 \quad \text{且} \quad \frac{|CD|}{|AB|} \leq \delta_2 \quad (9)$$

由于嘴角下拉时, 上嘴唇和下嘴唇并在一起, 此时, 点 D 位于点 O 上方, 依此可以判断出 AU15 的存在。

AU27 存在的检测条件是

$$\frac{|CD|}{|AB|} > \delta_2 \quad (10)$$

当眼睛闭合时, 上眼皮和下眼皮合并在一起, 眼珠被盖住, 因此, 将眼睛部分做纵向投影, 令各投影值为 $h(i)$, $i=1, 2, \dots, N_P$, N_P 为投影值的个数, 当

$$\max_{i=1, 2, \dots, N_P} (h(i)) < \delta_3 \quad \text{或} \quad (\max_{i=1, 2, \dots, N_P} (h(i)) < \delta_4 \quad \text{且} \quad \text{var}(h) < \delta_5) \quad (11)$$

成立时, 说明 AU43 存在。这里 $\max()$ 和 $\text{var}()$ 表示求最大值和方差。式(8)~(10)中的 δ_i ($i=1, 2, \dots, 5$) 是预先设定的阈值, 且 $\delta_3 < \delta_4$ 。下面的实验中, δ_1 取 0.3, δ_2 取 0.6。

从 AR 和 Yale 数据库中挑选 100 幅图像, 组成测试图像集, 用上述方法逐个检测每一幅图像中是否存在某个动作单元, 最终结果如表 1 所示。其中, “正识率”定义为“正确检测到带有某个 AU 的图像数”与“实际带有该 AU 的图像数”之比, “误识率”定义为“错误检测到带有某个 AU 的图像数”与“实际上不带有该 AU 的图像数”之比。该表说明, 对于这 4 个动作单元来说, 基于 BEI 的检测结果是比较令人满意的。

表 1 动作单元(AU)识别结果

AU	实际带有 AU 的图像数	检测到带有 AU 的图像总数	正确检测的图像数	正识率 (%)	误识率 (%)
12	32	32	30	93.75	2.94
15	8	8	8	100	0
27	32	33	31	96.88	2.94
43	13	15	13	100	1.15

注: 总的测试图像数为 100

研究还表明, 上述基于几何特征的动作单元识别方法对于与嘴巴有关的动作单元的识别能取得较好的效果, 但对于与眼睛有关的除 AU43 以外其他动作单元的识别效果较差。这主要是因为眼睛的动作变化较为微妙, 而且眼睛的边缘较多、且较集中,

有些变化很难用简单的几何关系来描述。然而, 从 BEI 本身来说, 它已经很好地反映了原始灰度图像中所隐含的表情变化, 因此, 开发出更为有效的基于 BEI, 或者基于灰度图像+BEI 的表情识别方法, 是一个很值得研究的课题。

5 结 论

提出了一种基于小波多尺度特性的人脸二值边缘提取新方法, 它包含两次二值化过程, 即首先对单尺度下重构的高频图像进行自适应二值化, 然后将获得的多个边缘图像进行融合, 进行第二次二值化。为了提高边缘图像的质量, 本方法还包含了一个后续的去噪过程, 它是用另外一种含噪少的二值边缘图像作为过滤模板来实现的。定量的评价表明, 用该方法获得的二值边缘图像中在部件边缘的连通性、完整性以及光照鲁棒性等方面都比用基于 Sobel 算子得到的二值边缘图像要好。将所得的二值边缘图像用于人脸表情分析, 通过对 4 种特定的人脸动作单元的识别实验, 取得了较好的效果。

实验中还将用本方法获得的人脸二值边缘图像应用于人脸部件的分割及其特征点的提取, 收到了较好的效果。下一步的工作是构建更为有效的基于 BEI 的人脸表情识别算法, 以及基于 BEI 的人脸识别方法研究等。

参 考 文 献

- 1 Ziou D, Tabbone S. Edge detection techniques——An overview [J]. International Journal of Pattern Recognition and Image Analysis, 1998, 8(4): 537~559.
- 2 Kawaguchi T, Rizon M. Iris detection using intensity and edge information[J]. Pattern Recognition, 2003, 36(2): 549~562.
- 3 Gao Y, Leung M K H. Face recognition using line edge map[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(6): 764~779.
- 4 Gao Y, Leung M K H. Line segment Hausdorff distance on face matching[J]. Pattern Recognition, 2002, 35(2): 361~371.
- 5 Lee S Y, Ham Y K, Park R H. Recognition of human faces using knowledge-based feature extraction and neuro-fuzzy algorithm[J]. Pattern Recognition, 1996, 29(11): 1863~1876.
- 6 Takacs B. Comparing face images using the modified Hausdorff distance[J]. Pattern Recognition, 1998, 31(12): 1873~1881.
- 7 Basu M. Gaussian-based edge-detection methods——A survey [J]. IEEE Transactions on Systems, Man, and Cybernetics——Part C: Applications and Reviews, 2002, 32(3): 252~260.
- 8 Wiskott L, Fellous J M, Kruger N, et al. Face recognition by

- elastic bunch graph matching[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 775~779.
- 9 Fasel B, Luetttin J. Automatic facial expression analysis: a survey[J]. Pattern Recognition, 2003, 36(1): 259~275.
 - 10 Fellenz W, Taylor J, Tsapatsoulis N, *et al.* Comparing template-based, feature-based and supervised classification of facial expressions from static images[A]. In: Proceedings of International Conference on Circuits, Systems, Communications and Computers [C], Nugata, Japan, 1999: 5331~5336.
 - 11 Mallet S. A wavelet tour of signal processing (2 edition)[M]. San Diego, CA, USA: Academic Press, 1998.
 - 12 Martinez A M, Benavente R. The AR Face Database[R]. CVC Technical Report # 24, Purdue University, West Lafayette, Indiana, USA, 1998.
 - 13 Yale University Face Database[DB/OL]. <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.
 - 14 Ekman P, Friesen W V. Facial Action Coding System: the Measurement of Facial Movement [M]. Palo Alto, San Francisco, CA, USA: Consulting Psychologists Press, 1978.
 - 15 Zhang Y J. A survey on evaluation methods for image segmentation[J]. Pattern Recognition, 1996, 29(8): 1335~1346.



宋加涛 1966 年生。副教授。1991 年毕业于杭州大学半导体专业,获硕士学位,现为浙江大学通信与信息系统专业博士研究生。主要研究方向为信号分析、图像处理 and 模式识别。

E-mail: sjt6612@163.com



刘济林 1947 年生。教授、博士生导师。1970 年毕业于清华大学精密仪器系,1981 年毕业于中国科学院研究生院,获硕士学位。曾先后在联邦德国西柏林工业大学、德国国家信息工程中心和美国 UC Berkeley 做访问学者。目前主要从事机器视觉、图像通信、智能交通等方面的研究。在国内外学术刊物上已发表学术论文 70 余篇。



池哲儒 1962 年生。副教授。1982 年、1985 年先后毕业于浙江大学电子工程专业,分别获学士、硕士学位,1994 获澳大利亚悉尼大学电子工程专业博士学位。主要研究方向为图像处理、模式识别、智能计算。发表论文 100 余篇。



王 蔚 1971 年生。实验师。1994 年毕业于浙江工业大学计算机专业,获学士学位。现主要从事电子技术和多媒体课件的开发及应用研究。