

# 视频序列中的人脸检测与定位算法研究

俞扬信, 严云洋

(淮阴工学院 计算机工程系, 江苏 淮安 223001)

**摘 要:**在研究和分析各种人脸检测与定位算法的基础上,并以计算简单、速度快、能精确地提取视频序列中的人脸为原则,提出了一种新的基于运动信息与边缘投影函数相结合的视频序列中的人脸检测与定位算法。该算法设计了双阈值 Sobel 算子进行边缘检测,该算子检测到的图像边缘清晰、细致、噪声少;提出了平方投影函数,该投影函数不但可区分均值相同的区域,而且可区分方差相同的区域。将边缘函数与投影函数结合起来设计的人脸检测与定位算法简单实用。

**关键词:**人脸检测;人脸定位;边缘投影;Sobel 算子;投影函数

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2009)02-0109-03

## Algorithm Study of Face Detection and Location in Video Sequence

YU Yang-xin, YAN Yun-yang

(Department of Computer Engineering, Huaiyin Institute of Technology, Huai'an 223001, China)

**Abstract:** On the basis of study and analysis of prior face detection and location algorithm, a new face detection and location algorithm based on the moving information with the edge projection function of combining the video sequence is proposed. It is to calculate a simple, fast, accurate face as the principle. A two threshold sobel operator is presented in the paper which can detect a thinner edge and non-sensitive to noise. A square projection function is designed. The function not only can distinguish regions which have a same mean, but also can distinguish regions which have a same variance. The proposed face detection algorithm combined edge image and projection function has the advantages of simple and practicality.

**Key words:** face detection; face location; edge projection; Sobel operator; projection function

## 0 引言

人脸检测与定位由于其在视频监控、人脸识别等领域的广泛应用,已成为当前计算机视觉检测研究中的一个热点<sup>[1]</sup>。然而,众多因素的影响如人脸大小、位置、角度、姿态等的变化,以及人脸表情、遮挡及运动带来的遮挡、光线改变带来的人脸外观上的变化都会给人脸检测造成干扰。人脸检测的挑战性在于<sup>[2]</sup>:(1)人脸角度。人脸图像与相关的人脸角度有很大关系,如正面、45度、侧面、向上看、向下看时人脸呈现出不同的外观。另外,一些脸部特征如眼睛或鼻子在人脸的某些角度下可能只有部分显露或完全被遮挡。(2)一些特征的有无。一些脸部特征如胡子、络腮胡、眼睛等可能有,也可能没有。而这些特征对人脸的形状、颜色

和大小都有较大的影响。(3)脸部表情。人脸的外观直接受到脸部表情的影响。(4)遮挡。人脸的某些部位可能被其它物体遮挡。当图像中有一群人时,一些人脸可能被另一些人脸遮挡。(5)图像角度。人脸图像也直接受到摄像机光轴角度的影响。(6)成像条件。成像时,一些因素如光线(光谱,光源分布和强度)都会影响到人脸的外观。随着多媒体技术的发展,对视频图像中人脸的处理日渐重要。

## 1 人脸检测与定位方法

近年来在人脸检测与定位领域的常用方法为<sup>[3]</sup>:

(1)基于知识的方法(Knowledge-Based Methods)。基于知识的方法运用人类认知典型人脸的知识,获取脸部特征之间的规则,该方法主要可用于人脸定位。图像中的人脸都有两个对称的眼睛、一个鼻子和一个嘴巴,可以很容易地建立简单的规则来描述人脸的这些特征及它们之间的联系。例如特征之间的关系可以用相对距离和位置来表示。该方法的问题在于将人脸知识转换为充分定义的规则时存在难度。如果

收稿日期:2008-05-19

基金项目:江苏省自然科学基金(08KJB52001);江苏省自然科学基金资助项目(06KJD520024);淮安市科技发展基金资助项目(HAC07063)

作者简介:俞扬信(1970-),男,江苏泰州人,硕士,讲师,研究方向为智能化信息处理技术;严云洋,博士,教授,研究方向为智能化信息处理技术、模式识别。

规则太严格,则因为不能满足所有的规则而导致检测失败;如果规则太一般化又会产生误检测。此外,因为不能列举所有的情况,所以该方法无法处理人脸角度的变化。

(2)特征不变性方法。这种算法的目标是寻找即使当姿态、视角或光照条件改变时也不变的人脸结构特征,然后利用它们定位人脸。使用的特征有脸部特征、纹理和皮肤颜色。使用脸部特征的方法一般是先用边缘算子提取脸部特征如眉毛、眼睛、鼻子、嘴巴等,然后建立统计模型来描述这些特征之间的联系以检验人脸的存在性。该方法存在的问题是图像特征在光线改变、有噪声和遮挡的情况下会受到很大的干扰。人脸的特征边界也不明显,阴影反而会产生较强的边缘,这些都会产生干扰。

(3)模板匹配方法。该方法首先存储几个标准模式的人脸描述整个的人脸或人脸的独立特征,然后计算标准人脸与输入人脸之间的相关性进行人脸检测。这种方法既可用于人脸定位,也用于人脸检测,简单可行。但是,它不能适应人脸在尺度、角度和形状上的改变。

(4)基于外观的方法(Appearance - Based Method)。与模板匹配不同的是,该方法从一组训练图像中学习模型或模板以捕捉人脸外观上的可变性,这种方法主要是用于人脸检测。基于外观的方法依赖统计分析和匹配学习来寻找相关的人脸和非人脸特征。

(5)变形方法。这种方法将脸部的可变性分解为形状模型和形状归一化纹理模型,如活动形状模型、活动外观模型、动态连接结构。这种方法非常适合精确的定位。但是,人们没有公布广泛的评估结果。

## 2 视频序列中的人脸检测与定位算法

人脸检测和定位是人脸自动识别系统的第一步,算法的优劣直接影响着后续的特征提取及识别的效果。视频序列中的人脸检测与定位更是一个活跃的研究区域。根据以上对现有方法的分析可知,使用多种特征相结合的方法比单独使用某个特征进行人脸检测和定位更加有效。

### 2.1 算法原理及过程

文中以计算简单、容易实现为原则,提出了一种将运动信息与边缘投影函数相结合的视频序列人脸检测与定位算法。

文中设计的算法将人脸检测与定位分为 4 步:

(1)利用特征对象法进行运动目标的检测与分割,得到只包含运动目标的特征对象图像,消除了背景的混乱干扰。

(2)提取运动目标的边缘图像。边缘图像携带了原始图像的丰富信息,且具有对光照不敏感的特点。针对经典边缘算子对噪声敏感,提取的边缘较粗等缺点,设计了双阈值 Sobel 算子,该算子提取的边缘清晰、噪声少。

(3)计算边缘图像的投影函数。投影函数用于区域分割时具有不受区域大小控制的优点,适于进行多尺度人脸的检测。针对现有投影函数对均值或方差敏感的缺点,设计了对均值和方差都不敏感的平方投影函数计算投影。

(4)进行区域分割,确定似人脸区的位置。人脸上由于五官的存在使得人脸具有较大的灰度变化率。尤其是人脸的五官分布具有一定的比例,因此具有相邻的较大灰度变化区间。利用这个特点可进行人脸检测。由于实际计算到的投影函数具有多个起伏,从中能够分离出多个相连区域,所以利用主元素分析法确定检测到的各个区域是否是人脸。

### 2.2 双阈值 Sobel 算子检测图像边缘

边缘图像携带了原始图像的丰富信息,且具有对光照不敏感的特点。因此利用边缘图像可处理光照改变条件下的人脸检测<sup>[4]</sup>。

经典的边缘提取算子有 Sobel 算子, Canny 算子, 拉普拉斯算子等。其中, Sobel 算子属于空间域的边缘检测方法,它检测出的边缘较光滑,且在噪声存在时也能检测出边缘,但不能较好地消除噪声;使用传统的 Sobel 算子检测边缘时的另一个缺点是检测出的边缘线较粗,即像素点过多。阈值较高时保留的信息较多,但将非边缘也保留下来;阈值较低时又去掉了有用的信息,使边缘断续。为此,设计了双阈值,保留两个阈值之间的像素点;然后设计了细化算法,以消除边缘中的过多像素及噪声。具体步骤为:

(1)设高阈值为  $ThreH$ , 低阈值为  $ThreL$ , 梯度图像为  $G(x, y)$ , 设置处于两个阈值之间的像素值为 1, 其它的设置为零。阈值可根据经验来确定,或使用统计方法等确定。

$$G(x, y) = \begin{cases} 1 & ThreL < G(x, y) < ThreH \\ 0 & \text{other} \end{cases}$$

(2)将处于两个阈值范围内的点作为候选点  $G(x, y)$ , 选择从候选点出发的大小不同的两个区域 Small 和 Round, 求每个区域的像素值之和  $SumSmall$  和  $SumRound$ , 即

$$G(x, y) = \begin{cases} 1 & SumSmall = n, SumRound > (n, a) \\ 0 & SumSmall = n, SumRound < (n, a) \end{cases}$$

其中,  $n$  是小区域中的像素数,  $a$  是小于大区域像素数的一个常数, 根据经验可定为 2。

经过上述处理后得到的边缘图像边缘较细,轮廓清晰。

图1所示为传统Sobel算子检测出的边缘图像与双阈值Sobel算子的比较。依次为原始图像、传统Sobel提取的边缘图像、双阈值Sobel提取的边缘图像效果。由图中看出,传统Sobel算子检测出的边缘较粗,而双阈值Sobel算子检测出的人脸边缘则较细。



图1 传统Sobel算子与双阈值Sobel算子的比较

### 2.3 平方投影函数

边缘图像虽然属于二值图像,但包含了丰富的信息,且具有对光照不敏感的特点;投影函数检测出的区域不受区域大小的限制。因此,将两者结合起来可设计出对光照、表情、大小都具有鲁棒性的人脸定位方法<sup>[5]</sup>。

利用积分投影函数可以分割灰度不同的区域,但当各区域的灰度均值相同时,积分投影函数无能为力;方差投影函数能够检测出均值相同的不同区域,但检测不出方差相同的区域。针对上述两种投影函数的缺点,文中提出了一种既能检测均值相同区域,又能检测方差相同区域的平方投影函数:

$$p_v^2(x) = \frac{1}{y_2 - y_1} \sum_{y_1 \leq y_i \leq y_2} [F(x, y_i) - S_v(x)]^2$$

$$p_h^2(x) = \frac{1}{x_2 - x_1} \sum_{x_1 \leq x_i \leq x_2} [F(x_i, y) - S_h(y)]^2$$

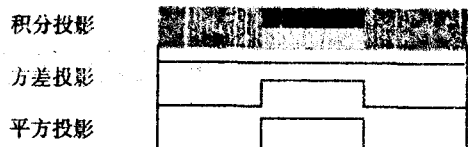
其中,  $F(x, y)$  是待检测的图像,  $p_v^2(x)$  是垂直平方投影,  $p_h^2(x)$  是水平平方投影,  $S_v(x)$  和  $S_h(y)$  分别是垂直积分投影和水平积分投影:

$$S_v(x) = \int_{y_1}^{y_2} F(x, y) dy$$

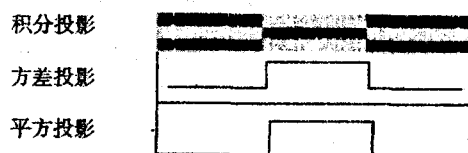
$$S_h(y) = \int_{x_1}^{x_2} F(x, y) dx$$

平方投影函数综合了灰度均值和方差的变化,具有更强的适应性。

如图2所示,图2(a)中第一个图是均值相同的三个区域,用积分投影检测不出区域的变化,但用方差投影和平方投影可以检测出来;图2(b)中的第一个图是方差相同的三个区域,用方差投影检测不出区域的变化,但用积分投影和平方投影可以检测出来。即平方投影函数由于在公式中考虑到了均值和方差,所以能检测出令积分投影或方差投影失效的不同区域。



(a) 均值相同的区域投影结果



(b) 方差相同的区域投影结果

图2 积分投影函数、方差投影函数、平方投影函数的比较

### 2.4 似人脸区域检测

如图3所示,在边缘图像中,人脸上由于五官边缘的存在,使得人脸比头发、服饰等具有更高的灰度变化率,因此检测平方投影函数中灰度变化率最高的区域便可以得到人脸的精确位置。有时衣服的纹理较丰富,灰度变化率会超过人脸的灰度变化率,这时定位的结果不但包含人脸,还包含了一部分肩部,即检测出的人脸面积过大。为了解决这个问题,可以设置一个判断准则:当相邻两个具有较大变化率的区域之间的距离超过一定值时,取区域大的一个为人脸。

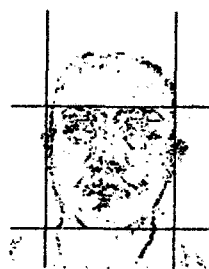
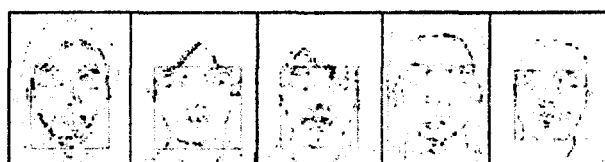


图3 人脸边缘图像

由于边缘图像对光照、表情不敏感,投影函数能够适应不同的图像大小,因此检测出的人脸位置精确,且不受光照、表情、图像大小的影响。图4所示为使用平



(a) 原始图像



(b) 边缘图像定位结果

图4 对边缘图像的人脸定位结果

原始特征空间。

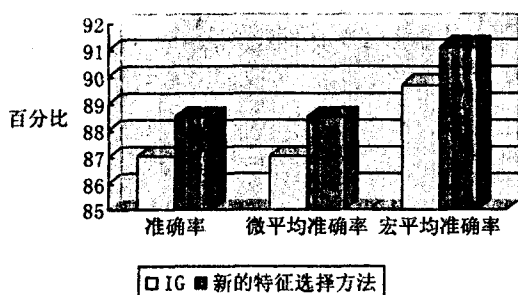


图3 特征维数为800时两种方法在准确率等方面的比较

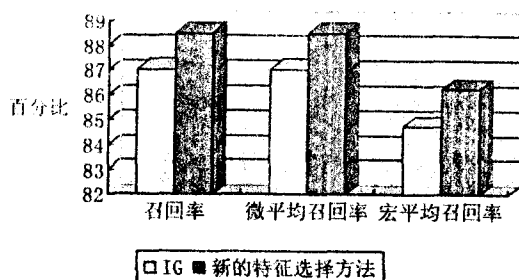


图4 特征维数为800时两种方法在召回率等方面的比较

#### 4 结束语

在文本分类中,如何进行特征选择是一个重要的研究问题。传统的特征选择方法主要集中在寻找与类别相关性大的特征上,而对特征冗余的问题没有给予足够的重视。如果没有消除特征冗余,那么选择出的特征子集并不能最大程度地代表其原始特征空间,显然会对分类性能产生影响。所以,提出了一种基于mRMR模型的特征选择方法,在考虑选择和类别相关

性大的特征词的同时还考虑了特征冗余的消除。实验验证,其分类性能有所提高,虽然计算方法比传统的特征选择方法稍麻烦些。那么,研究该特征选择方法在不同语料集的分类情况以及改变公式(8)中 $I$ 值的计算方法是否出现更好的分类效果,将是未来的研究工作。

#### 参考文献:

- [1] Salton G, Wong A, Yang C S. On the specification of term values in automatic Indexing[J]. Journal of Documentation, 1973,29(4):351-372.
- [2] Fox C. Lexical Analysis and Stoplists[C]//Frakes W B, Baeze - Yates R. In Information Retrieval: Data Structure & Algorithms[s.l.]: P T R Prentice Hall, 1992:102-130.
- [3] Frakes W B. Stemming Algorithms[C]//Frakes W B, Baeze - Yates B. In Information Retrieval: Data Structure & Algorithms. [s.l.]: T P R Prentice Hall, 1992:131-160.
- [4] Yang Yiming, Pederson J O. A Comparative Study on Feature Selection in Text Categorization[C]//Proceedings of the 14th International Conference on Machine learning. Nashville: Morgan Kaufmann, 1997:412-420.
- [5] Ding C, Peng Hanchuan. Minimum redundancy feature selection from microarray gene expression data[C]// Proceeding of Second IEEE Computational Systems Bioinformatics Conference. LosA Lamitos, USA: IEEE Computer Society Press, 2003:523-528.
- [6] Peng Hanchuan, Long Fuhui, Ding C. Feature Selection Based on Mutual Information Criteria of Max - Dependency Max - Relevance and Min - Redundancy[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005,27(8):1226-1238.

(上接第111页)

方投影函数从边缘图像中定位出的人脸。

#### 3 结束语

回顾了近年来在人脸检测与定位领域的一些常用方法,并在研究和分析各种人脸检测与定位算法的基础上,以计算简单、定位准确为原则,提出了一种将运动信息与边缘投影函数相结合的视频序列人脸检测与定位算法。边缘函数和投影函数检测人脸的缺点是易受背景中其它物体的干扰,但从视频序列中提取的特征对象图像具有背景简单、统一的特点,所以可利用边缘函数与投影函数进行后续的人脸检测。设计了双阈值 Sobel 算子进行边缘检测,该算子检测到的图像边缘清晰、细致、噪声少;提出了平方投影函数,该投影函数不但可区分均值相同的区域,而且可区分方差相同

的区域。实验中使用的图像序列包括了室内环境、室外环境下的视频图像序列,运动目标包括单运动目标、具有可分离性的多运动目标。实验结果表明,该方法简单有效,能准确地进行人脸检测及定位。

#### 参考文献:

- [1] 李玉山. 数字视觉视频技术[M]. 西安:西安电子科技大学出版社, 2005.
- [2] 刘党辉, 沈兰荪. 视频运动对象分割技术的研究[J]. 电路与系统学报, 2002,7(3):77-85.
- [3] 严云洋, 郭志波, 杨静宇. 人脸识别特征抽取方法的研究进展[J]. 淮阴工学院学报, 2007,16(3):24-30.
- [4] 杨莉, 张弘, 李玉山. 视频运动对象的自动分割[J]. 计算机辅助设计与图形学学报, 2004,16(3):301-306.
- [5] 严云洋, 高尚兵, 郭志波, 等. 基于视频图像的火灾自动检测[J]. 计算机应用研究, 2008,25(4):1075-1078.