

# 基于区域特征的快速人脸检测法\*

卢春雨, 张长水, 闻 芳, 阎平凡

清华大学 自动化系, 智能技术与系统国家重点实验室, 北京 100084

**文 摘** 提出一种新的基于区域特征的快速人脸检测方案, 相比传统马赛克方法明显改进: 重新设计了基于器官的人脸分块策略, 使规则制订直观可信; 采用块形状自适应技术以提高对不同脸形的适应性, 改善了检测算法的适用范围; 引入有效的粗检测和层次化局部搜索技术, 大大提高了检测速度。仿真表明, 这种方法对人脸数目、尺寸、脸形、光照、背景复杂性无限制, 适用于不同质量尤其是恶劣质量的图像。

**关键词** 人脸识别; 人脸检测; 区域特征; 马赛克

**分类号** TP 391.41

人脸识别是一种重要的个人身份鉴别方法, 具有直接、友好、方便的特点<sup>[1]</sup>。而一个完全自动的人脸识别任务的第一步就是要找到人脸, 即人脸检测, 本文即研究静态图像中的人脸检测问题

目前各种人脸检测方法可粗略地分为自顶下和自底上两大类, 前一类方法把人脸视为一个整体模式直接进行检测, 如马赛克方法<sup>[2]</sup>、特征脸方法<sup>[3]</sup>、纹理图(SGLD)方法<sup>[4]</sup>、连续对称方法<sup>[5]</sup>等; 而后者则从检测人脸的各局部器官入手, 然后根据人脸器官相对位置关系的先验知识进行综合, 如分层神经网络方法<sup>[6]</sup>等。目前公认普遍性最好的是马赛克方法(又称为镶嵌图方法)<sup>[1]</sup>, 其基本思路就是将人脸分为若干低分辨率马赛克子块, 利用人脸各子块之间的相互关系进行人脸检测。本文在此基础上对于人脸定量模型和规则制订的直观性、各种脸型的适应性、检测速度, 提出一系列改进措施。相比以前的研究, 我们设计的系统在检测速度和检测效果上均有明显的改进。

## 1 自动人脸检测系统描述

自动人脸检测系统工作流程框图由粗检测、候选区域合并及在尺度因子控制下的若干局部环节构成, 如图1所示。尺度空间的遍历控制是为了解决待检测人脸的尺寸未知问题, 也称为尺度空间递进搜索。粗检测环节包括两部分处理, 一是利用检测环境的一些已知信息, 如人脸皮肤的颜色、头部的运动、背景构成情况(如颜色、特定标记)等先验知识, 迅速将人的头部从背景中大致分割出来; 二是结合面部器官先验信息进行显著器官粗检测, 进而找到人脸可能存在的大致范围。旋转估计环节是为解决歪头人脸检测的特殊情况。而灰度均衡环节用来解决不同人脸图像摄取环境下的光照变化问题。经过前述环节的规格化处理后, 得到一些候选图像块, 每一块中可能包含一个正面、直立、光照大致均匀的人脸图像。对此图像块进行层次化局部搜索和块形状自适应调整, 之后完整计算马赛克局部统计量, 并应用三分图人脸模型规则最终判断该输入图像块是否为人脸。对所有可能的尺度空间和子图像块都进行了操作之后, 进入最后一个候选区域合并环节。应用基于区域特征的马赛克方法, 同一个人脸可能在几个邻近尺度或在一个尺度的几个临近位置同时被检测出

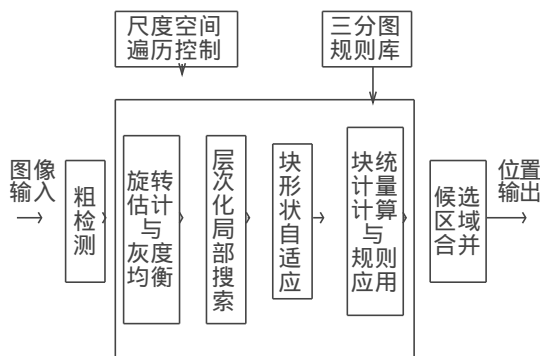


图1 人脸检测流程图

收稿日期: 1998-05-11

第一作者: 男, 1972年生, 博士研究生

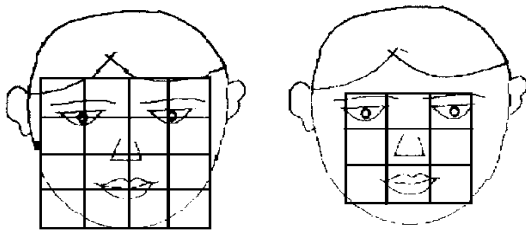
\* 基金项目: 国家“八六三”高技术项目, 863-306-03-

来. 如何从这些候选者中选出最合理的分割, 成为后处理中一个不可忽视的问题. 我们采用的是最小矩形并结合马赛克块自适应的方法. 候选区域合并环节的输出为可能包含人脸的图像块

## 2 策略的选择及技术的优化研究

### 2.1 新的标准脸分块策略

文 [1] 中提出的标准脸假设是基于四分图的 (如图 2(a)所示). 本文提出的标准脸假设如图 2(b)所示, 可视为一种广义的三分图. 相比之下, 这种基于器官的分块方法, 充分利用了人脸器官的自然分布, 规则制订过程中对先验知识的利用更加直观, 并使马赛克子块对脸形的自适应操作成为可能, 从而提高了检测效率和鲁棒性. 文 [1] 对四分图和八分图分别制定规则. 我们的方法中, 只利用了这一广义的三分图, 没有利用更高分辨率 (如六分图) 的信息. 这是考虑到不同人面部器官分布的差异性及不同表情下面部器官的塑性变形. 分辨率越高, 制订规则需要考虑的因素就越复杂, 其有效性和合理性也更难以保证.



(a) 四分块策略 (b) 基于器官的分块  
图 2 分块策略对比

我们设计的基于器官的马赛克人脸模型如图 2(b)和图 3所示: 子块 0 对应于左眼区域 (包括眼眉部分), 子块 2 对应于右眼区域, 子块 4 对应于鼻子区域, 子块 7 对应于嘴巴, 子块 (3, 6, 5, 8) 对应于两侧脸颊. 我们假定子块 (0, 3, 6) 与子块 (2, 5, 8) 具有大致相同的宽度, 子块 (1, 4, 7) 宽度相同, 所有的块高度大致相同. 经过块形状自适应后, 上述分割将有些变化, 但保持同一行三个子块高度相同, 同一列三个子块宽度相同. 如前所述, 由于脸形的不同, 对圆脸的圆脸和方形脸采用方形的子块分割是合理而有效的. 对明显瘦长形脸、扁宽形脸等, 则应采用相应的矩形分块策略, 以免由于无法形成合理的马赛克覆盖而造成检测漏报. 考虑到多数情况下, 人脸可能的纵横比在 0.8~1.3 之间, 而马赛克方法对小范围内纵横比的变化不敏感, 可简单量化为 3~5 个量

级, 但即使如此, 运算量也会成比例增加, 因此我们采用了块形状的自适应技术 (见后文) 来缓和这一矛盾.

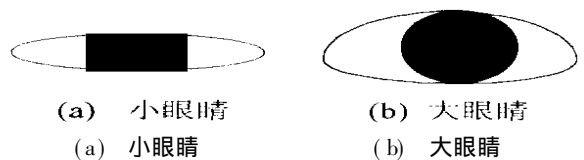
|   |   |   |
|---|---|---|
| 0 | 1 | 2 |
| 3 | 4 | 5 |
| 6 | 7 | 8 |

图 3 马赛克分块

### 2.2 粗检测

由于块统计量的计算过程比较费时, 为尽可能快速的完成检测过程, 进行高效率的粗检测十分必要. 粗检测环节包括利用检测环境信息的粗检测和利用人脸器官信息的粗检测两个子过程, 后一过程更具有普遍意义. 为强调这是一个独立性很强的环节, 在图 1 中我们把它放在了尺度空间的遍历控制之外. 显然, 如果我们能通过某种方法稳定地检测到人脸模型大马赛克块的左上角, 则搜索可局限在这个左上角子集合上进行, 从而显著减小搜索空间, 降低计算量, 称这种能被稳定检测的人脸区域上的代表点 (区域) 为基准点或锚点. 理想的锚点应该是在不同表情下位置比较稳定, 并且易于检测的点. 如眉角、眼角、嘴角等, 眼球、鼻孔在某些简单摄像条件下最易于检测, 我们的粗检测方法就基于眼球检测.

通过观察发现, 在适当的分辨率下, 人眼的眼球大致可视为一个白背景上的椭圆形实心黑块 (对眼球的反光可通过中值滤波加以抑制), 图 4(a)(b) 分别表示小眼睛 (眯眼) 和大眼睛 (瞪眼) 的情况. 这种实心小黑块可用中央凹 DoG 算子 (见图 5) 进行相关检测. 粗检测首先形成一个特定尺度下锚点置信度图, 图中每个点的值表示该尺度下, 该点周边地区与人眼球模型的相似程度. 对各尺度下锚点置信度图进行局部阈值二值化, 得到粗检测的锚点集合, 它将指导候选区域的筛选工作.



(a) 小眼睛 (b) 大眼睛  
图 4 眼球的简化模型

选择粗检测算法的主要考虑是要快速和简便, 上述模板匹配形式的人眼粗检测, 由于利用人眼的信息量较少, 筛选效率不是很高, 直接进行自适应和

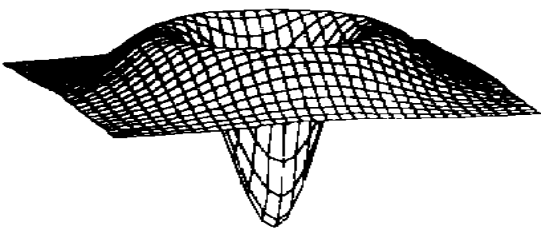


图 5 中央凹 DoG算子

统计量计算仍比较费时,因此我们引入了层次化局部搜索对锚点集合进行一步筛选。

2.3 层次化局部搜索方法与块自适应

层次化局部搜索是为进一步减小搜索空间、提高计算效率而设计的。若粗检测得到的锚点集合中点数仍比较多时,则应先进行层次化局部搜索过滤筛选,使尽可能少而精的锚点子集送到下一级块形状自适应环节。具体而言,是通过不断引入子块级的局部通用规则知识,使锚点集不断精练。比如在子块的层次,计算以当前锚点集合中每一点为锚点时,相应子块 0 与子块 3 的平均梯度差做为该点成为新锚点的置信度水平,从而得到一个新的锚点置信图。对此锚点置信图上进行局部二值化,得到更小的修正锚点子集。对此小子集进行完整的块形状自适应和块统计量计算并运用规则进行人脸检测。

人有多种自然脸型如圆脸、方脸、瓜子脸等。不同脸型下面部器官分布的差异非常大,采用固定的分块策略(如方形或长方形)很难满足各种脸型情况,为此,引入了块尺寸、形状的自适应技术,动态的适应脸形的变化。它分为两个层次,首先对块内马赛克子块划分进行调整,之后进行马赛克大块整体的扩张和收缩。后者主要用于最后的候选区域合并过程,而前者是在每次完整计算子块统计量之前进行。具体操作时,这一环节是在整体分块尺寸大致不变的情况下,按某种经验主义的能量函数对左右两列子块的宽度和上下两行子块的高度进行小范围调整。文 [7] 给出一个调整子块 (2, 5, 8)宽度的简单能量函数如下式所示:

$$f(x)=G_w[2]-G_w[5].$$

它实际上反映的是块 2 与块 5 的梯度对比度。在一定范围内,通过使上述梯度对比度能量函数最大化,实现对马赛克子块 (2, 5, 8)的宽度进行局部的调整。同样道理可实现对马赛克子块 (0, 3, 6)的宽度调整。

完成对脸型的块形状自适应后,就可以完整地计算各种马赛克块统计量,并应用马赛克规则进行人脸图像块的筛选。

2.4 块统计量的选择与规则制订

人脸灰度图像上包含的信息很多,寻找并采用计算简便、稳定反映各器官马赛克子块区别的统计量对规则制订的难度影响非常大。人的脸认知和检测过程是在丰富知识模型指导下进行的,完全由计算机实现要遇到很多困难。经过反复实验比较,我们采用基于水平方向梯度强度和梯度方向的区域统计量,它们的稳定性和区别性都非常强。

马赛克方法除了前述若干技术(效率与鲁棒性)实现的考虑之外,最终决定检测性能(虚警、漏报联合损失)的是马赛克规则,它是标准人脸模型的量化表现。一方面该模型要有较大的适用范围,即较好的模型弹性以降低漏报率,另一方面要尽可能的严格,增强约束力以降低虚警。马赛克规则的制订首先要考虑对如下多样性的支持和容许:

- 1) 脸形的影响,对圆脸和方形脸,子块 6, 8 基本在脸的范围之内,对瓜子脸而言,这两个子块将不确定地包括以外的区域,影响规则制订。
- 2) 器官相对面积的影响: 大眼睛小嘴巴与小眼睛大嘴巴、瞪眼发怒与开口大笑的块统计量分布差别很大,头部俯仰对鼻子形状很有影响,进而造成块统计量的大范围波动。
- 3) 面部轮廓的不稳定性: 脸部的上边缘受发型影响很大,两侧的边缘随扭脸角度和背景高亮度的不同而变化,而下颌的边缘由于与颈部相连,通常难于检测。
- 4) 光照的影响,侧光影响灰度分布对称性,并产生阴影边缘,影响马赛克上的梯度分布。暗光导致对比度过低,影响梯度和灰度计算的有效程度。
- 5) 各种眼镜的影响。

基于水平方向梯度强度的普遍性较好的一些规则如表 1 所示,其中  $D(i)$ 表示块  $i$  的平均梯度。

| 表 1 一些普遍性比较好的规则          |                                    |  |
|--------------------------|------------------------------------|--|
| 内容                       | 说明                                 |  |
| 1 $D(0)>D(1,3)$          | 子块 0 的梯度大于子块 1, 3 的梯度              |  |
| $D(2)>D(1,5)$            | 子块 2 的梯度大于子块 1, 5 的梯度              |  |
| $D(7)>D(3,5)$            | 子块 7 的梯度大于子块 3, 5 的梯度              |  |
| 2 $D(0+2+7)>0.5D(0+5+7)$ | 眼、嘴巴梯度之和占(除块 6, 8)全部子块梯度之和的 50% 以上 |  |
| 3 $D(0)\approx D(2)$     | 梯度对称性                              |  |
| $D(3)\approx D(5)$       |                                    |  |
| $D(6)\approx D(8)$       |                                    |  |

事实上,由于规则库不参与前面图像处理的过程

程,有很强的独立性,因此可以有针对性地对具体应用环境和图像质量建立专门的规则库;此外,可把更加细致的器官验证规则引入这一环节,从而提高检测的准确度。

### 3 仿真结果

仿真结果如图 6 所示。图 6(a)是 MIT 人脸库

中的一张照片,代表长形脸和戴眼镜的情况;图 6(b)是一张照片经扫描得到的图像,上面有多个人脸;图 6(c)是 CMU 人脸库中的一张照片,代表复杂背景的情况;图 6(d)是直接在实验室环境下用摄像头直接摄取的图像,包括两个不同大小的人脸。上面 4 种情况下,均准确的找到了人脸。

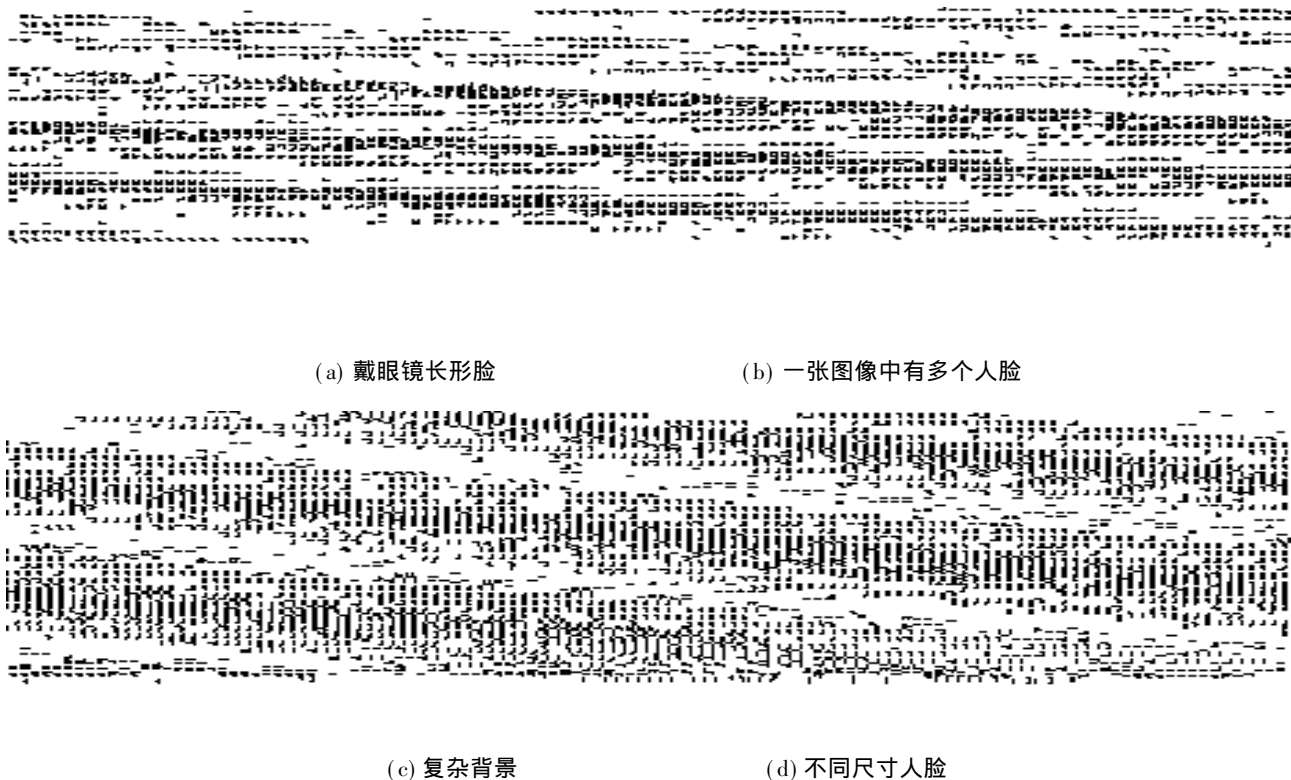


图 6 人脸检测仿真结果

在我们自行设计的一个包括不同质量、不同来源、不同背景复杂程度共计 90 张图像 110 张正面人脸的小型复杂图库中,检测到真实人脸 95 个,检测到虚假人脸 19 个。在 Pentium 200, 64 MB 内存 PC 机上对简单背景 256 $\times$ 256 分辨率中等人脸尺寸检测时间为 15 s 左右。漏报主要发生在深色眼镜、侧光、找不到眼睛等类型的图像中。

### 4 结束语

在传统马赛克方法的基础上,我们重新设计了人脸模型分块方案使规则的制订更加直观可信;通过采用块形状自适应技术,提高了检测算法对不同脸形的适应性,拓宽了检测适用范围。实验表明,改进后的方法适用于多人、不同尺寸、表情和光照、复杂背景等各种情况,使漏报率和虚报率同时达到较低的水平,取得了良好的检测效果。

### 参 考 文 献

- 1 Chellappa R, Wilson C L, Sirovny S. Human and machine recognition of faces: a survey. *Proceedings of the IEEE*, 1995, 83(5): 705~740
- 2 Yang G Z, Huang T S. Human face detection in a complex background. *Pattern Recognition*, 1994, 27(1): 53~63
- 3 Turk M, Pentland A. Eigenfaces for recognition, *J Cognitive Neuroscience*, 1991, 3(1): 71~86
- 4 Dai Y, Nakano Y. Face-texture model based on SGLD and its application in face detection in a color scene. *Pattern Recognition*, 1996, 29(6): 1007~1017
- 5 Zabrodsky H, Peleg S, Avnir D. Symmetry as a continuous feature. *IEEE Trans. PAMI*, 1995, 17(12): 1154~1166
- 6 Juell P, Marsh R. A hierarchical neural network for human face detection. *Pattern Recognition*, 1996, 29(5): 781~787

7 贾 玮. 图像中人脸检测定位的方法研究: [硕士学位论文]. 北京: 清华大学自动化系. 1996

Regional feature based fast  
human face detection

LU Chunyu, ZHANG Changshui ,  
WEN Fang , YAN Pingfan  
State Key Laboraory of Intelligent  
Technology and Systems,  
Department of Automation, Tsinghua University,  
Beijing 100084, China

**Abstract** Automatic face recognition (AFR) is one of the most attractive and challenging tasks in fields of computer vision and pattern recognition, the first critical step of AFR is face detection. A fast face detection algorithm based on regional feature is proposed. In contrast with the award-winning mosaic method, three main improvements were introduced here 1). organ-based blocking scheme and more intuitive mosaic rule design, 2). sub-block type adaptive technique according to different face shape, 3). fast organ-based rough detection and hierarchical local searching method. Also directional gradient statistics were used as sub-block feature instead of absolute gray value ones. Experimental results show that the algorithm can find the face rapidly with relative high accuracy, with little limit to the complexity of background, lighting condition, face size, person number, resolution of the image etc.

**Key words** face recognition; face detection; regional feature; mosaic;

书 讯

《 Windows 95通信编程》

Windows程序员正面临着建造通信软件的挑战,这些软件要在电话、传真、话音邮件、视频和其它消息服务接口中使用。提供这些服务的应用软件是通往信息超高速公路的入境坡道。《Windows 95通信编程》从基础开始讲解,一直到电话应用程序编程接口(TAPI)和消息应用程序编程接口(MAPI)这两个最重要的Windows通信组件。本书还包括在现有软件应用中增加通信能力的有用忠告。

在《Windows 95通信编程》一书中,你可以学到:

- 各种 Windows 95 API怎样协同工作;
  - 怎样在应用程序中增加拨号和传送消息的能力;
  - 怎样实现用 Win32通信 API进行低级串行数据传输;
  - 怎样通过串口方式连接两台计算机;
  - 怎样用 TAPI设计和控制电话呼叫;
  - 怎样用 MAPI进行传真、电子邮件和文件传送
- 现在,用《Windows 95通信编程》一书来增强你的通信能力吧!
- 本书由清华大学出版社出版。

书 讯

《怎样用计算机写应用文》

本书区别于其他同类书籍的鲜明特点是通俗、实用。作者是中学语文特级教师,成功地探索出快速掌握的学习规律。该书反映了作者由不会到熟练驾驭这一技术的过程中所悟得的化难为易、缩短学习时限的各种“窍门”,及学习过程中容易陷入的各种“误区”。本书能成为一切欲使用计算机人事写作和使用计算机进行文字处理的读者竭诚服务。

本书以计算机中文录入与打印输出、基本编辑与高级编辑、制表及相关问题为主经线,配以应用文的特点、种类、某些应用文体写作的语言忌讳及多种应用文体为练习题或举例对象,示之以范,从而能有效地帮助读者,使初学者尽快入门,使初知者尽快熟练掌握计算机写作技术。

书后附有一张软盘,汇集了各种应用文体的模板,读者只需在文中的×××××处换成你需要的内容,即可写成一篇规范的应用文。

本书由清华大学出版社出版。