

13 209-214

一种基于肤色和模板的人脸检测方法*

康学雷 邵[✓]凌 张立明

(复旦大学电子工程系工控实验室, 上海, 200433)

TP391.41

摘要 提出了一种基于肤色区域和多种模板结合的人脸检测方法,该方法能在复杂背景下很好的检测出人脸位置并且不受人脸旋转或侧转的限制,本文方法比传统的方法检测速度快、漏检率低,可以在实时系统中应用。

关键词 人脸检测, 肤色, 模板匹配。

图像识别 模式识别

A METHOD USING SKIN-COLOR AND TEMPLATE FOR AUTOMATIC FACE DETECTION*

KANG Xue-Lei SHAO Lin ZHANG Li-Min

(Department of Electronics Engineering, Fudan University, Shanghai 200433, China)

Abstract A combined method for face detection was proposed using both skin-color detection and multi-template matching. Computer simulation shows good results in complex background without any restriction for face position. Compared with traditional ones, this method has high speed and low mis-detection rate, so it can be used in the real-time system.

Key words face detection, skin-color, template-matching.

引言

在静态数字图像中自动检测人脸的算法在90年代初得到较为广泛的研究,出现了很多解决方案,如根据人脸结构构造神经网络,根据学习得到的权的组合输出来判定人脸^[5];设计一个人脸的通用模板,用多尺度的方法在全图像中进行检索^[3];用运动学原理和边缘检测的方法找出人脸的轮廓所在^[4]。所有这些都是基于数字图像的亮度空间的,并且存在以下两个问题:(1)运算量太大,通常要对整个图像进行多尺度的非线性插值和全搜索;(2)很难解决旋转和侧转的人脸检测。因此有人提出了一种基于肤色的检测方法^[1,2],这种方法利用了数字图像的彩色信息,大大提高了检测的速度,基本能做到实时检测并且不受人脸旋转或人侧转的影响,但由于肤色空间和其他颜色空间的交叉性,已有的方法只能限制在简单的背景,在复杂背景图像中误检率非常高。本文提出了一种将肤色检测和多种模板匹配相结合的方法,其结果在检测速度上大大高于神经网络、边缘检测和模板匹配的方法,

而误检率又远低于肤色检测等方法,为实时系统的实际应用提供了手段。

1 肤色检测

人脸检测传统的方法多是在亮度空间进行,信息仅有灰度的变化,没有任何区域或比例的限制,所以必须做多尺度空间的全搜索,计算量非常大,而利用色度信息则可大大降低搜索区域,其中肤色信息是最为直接有效的。常用的肤色检测空间有以下几种:RGB归一化、HIS、CYMK、CbCr方法等。在本文的实验中选用CbCr^[6]空间作为肤色分布统计的映射空间,该空间的优点是受亮度变化的影响较小,且是两维独立分布,能较好的限制肤色分布区域。

1.1 肤色统计分布

经过统计证明^[1]不同人种的肤色区别主要受亮度影响,受色度影响较小。在实验中选择人脸眼睛以下,嘴唇以上的狭长块作为肤色的统计区域,如图1所示。

为了使统计肤色分布更具有代表性,实验选取了不同肤色、不同光照的人脸图像作为样本。图2为20

* 国家自然科学基金(编号:39870194)资助项目
稿件收到日期 1999-10-28,修改稿收到日期 1999-11-24

* The project supported by the National Natural Science Foundation of China (No. 39870194)
Received 1999-10-28, revised 1999-11-24

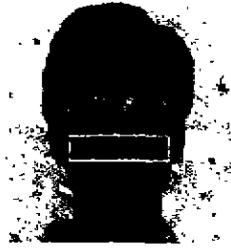


图1 肤色采样区域
Fig. 1 Region for skin-color data collection

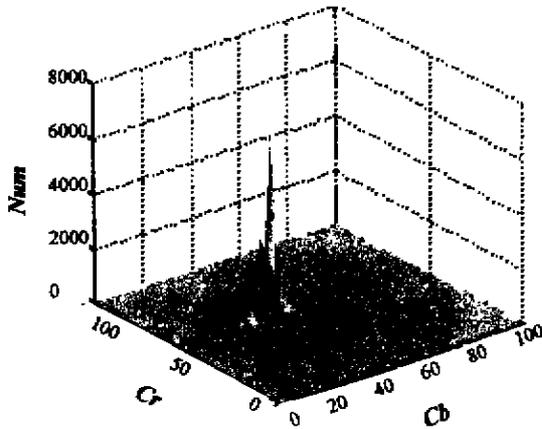


图2 肤色分布
Fig. 2 Skin-color distribution

幅蒙古人种和 20 幅高加索人种的人脸肤色统计分布图(样本分别来自复旦大学电工系工控实验室和美国普度大学图像实验室, Num 为该位置的样本点数)。

由图 2 可知, 在 CbCr 空间内肤色分布相比于整个空间而言仅是很小的一个区域, 并且受人种的影响不大。利用直方图的方法, 将统计所得的数据中去除与中心区域不连续的区域块和分布概率小于 0.2% 的点, 可以得到肤色分布区域。据此得到服从分布的图像掩膜 M_1 , 式中 $Cb(i, j)$ 和 $Cr(i, j)$ 为归一化到 $[0, 1]$ 的图像的色度分量, 并且我们得到的分布与文献[1]中的统计值相符, 即

$$M_1(i, j) = \begin{cases} 1, & Cb(i, j) \in [0.43, 0.55] \text{ and} \\ & Cr(i, j) \in [0.53, 0.62] \\ 0, & \text{else} \end{cases} \quad (1)$$

1.2 肤色检测后处理

在图像掩膜 M_1 内会存在两种误差情况, 一种是由于噪声而引入的被误判为肤色的孤立点或小块孤立区域, 另一种是由于人脸局部区域如眼睛、嘴唇等非肤色区域而导致的整个人脸的非全填充。为了消除这些

误差必须进行腐蚀和膨胀的操作, 操作算子是基于各方向对称的明可夫斯基算子^[6], 具体描述如下。

1.2.1 为降低腐蚀、膨胀对噪声的敏感度, 在 M_1 中统计所有 4×4 的块, 各块属于肤色的象素点数范围为 0~16, 并据此建立缩小 16 倍的模板 M_2 , 即

$$M_2(x, y) = \begin{cases} 1, & \sum_{i=0}^3 \sum_{j=0}^3 M_1(4x+i, 4y+j) \geq 14 \\ 0, & \sum_{i=0}^3 \sum_{j=0}^3 M_1(4x+i, 4y+j) < 14 \end{cases} \quad (2)$$

1.2.2 腐蚀操作用于去除孤立噪声, 设式中 Θ 为腐蚀操作算子, A' 为腐蚀结构元素, A' 结构描述见式(3):

$$M_3(i, j) = M_2(i, j) \Theta A' = \begin{cases} 0, & \sum_{k=-1}^1 \sum_{l=-1}^1 M_2(i+k, j+l) \leq 5 \text{ and} \\ & M_2(i, j) = 1 \\ M_2(i, j), & \text{else} \end{cases} \quad (3)$$

1.2.3 膨胀操作用于填充脸部的非肤色区域, 使整个脸部为全填充的连通区域。设式中 \oplus 为膨胀算子, B' 为膨胀结构元素, B' 结构描述见式(4):

$$M_4(i, j) = M_3(i, j) \oplus B' = \begin{cases} 0, & \sum_{k=-1}^1 \sum_{l=-1}^1 M_3(i+k, j+l) \geq 2 \text{ and} \\ & M_3(i, j) = 0 \\ M_3(i, j), & \text{else} \end{cases} \quad (4)$$

1.2.4 设 W 为在 M_4 中所有全填充的连通区域的集合, n 为区域的个数, 则有

$$W = \{W_i | W_i \cap W_j = \varnothing, i \neq j, i = 1, 2, 3, \dots, m\}. \quad (5)$$

2 人脸参数检测和模板的设计实现

经过肤色检测我们得到人脸的待选区域 W , 但由于在复杂背景下, 可能会存在很多和人脸肤色相近或相同的区域被误判为肤色区域(见图 7 和 8), 也可能由于裸露的手臂或腿部被误检测为人脸(见图 7(a)和(c)). 为此, 在本文中我们提出了基于人脸比例大小的参数判定和两个基于人脸局部特征的模板, 其中第一个模板是基于图像的两阶梯度空间的, 后一个模板是基于亮度空间的, 并通过模板的匹配程度来进一步限制人脸区域。

2.1 人脸比例参数检测和区域链表的建立



图3 人脸上界块
Fig. 3 Upper border of face

虽然也有人^[3]提出过对人脸参数的描述但都过于复杂,并且对人脸状态有很多限制.我们提出的参数判别方法则简单有效,基本不需要运算量,且平均能减少待选区域达30%.

2.1.1 我们将坐标原点设在图像的左上角,由上至下由左至右扫描 W 中的每个区域 W_i ,找出上边界块(见图3),并将其值设为区域在该列的长度,区域内其余块标为FF.

2.1.2 如图3所示无论是处于正脸或是侧转的情况下总可以找到一个位于额头附近的相对最高点,建立人脸待选区域链表 S ,链表的数据结构为相对最高点的坐标 (x, y) 、长度 h 、宽度 w ,其中 w 为由上至下约 $h/6$ 到 $h/2$ 处区域的最大宽度.将所有相对最高点的坐标加入链表 S 中.

2.1.3 按照人脸的比例关系,按下式条件删除不符合条件的链表节点 S_i 和待选区域 W_i ,得到 W' 和修正后的链表 S ,删除集为

$$S_{off} = \{S_i | h < w \text{ or } h > 4w \text{ or } h < 4, \\ h \in S_i, w \in S_i\}. \quad (6)$$

式中 f 为删除操作, $h < 4$ 表明小于 16×16 的人脸在本方法内不予考虑.

2.2 基于梯度的模板

在这里我们提出了基于两阶梯度图像的模板,用



图4 梯度模板
Fig. 4 Gradient template

于排除一些和肤色相近的平坦或纹理区域.我们都知道两阶梯度是反映边缘分布情况的,在人的脸部区域中边缘最明显的是眼睛、鼻子以及嘴巴(见图4),考虑到要处理侧脸部旋转的情况,在实验中我们选用了鼻子和嘴巴为局部特征的判别区域,具体步骤如下.

2.2.1 待选区域 W' 和整个图像的亮度分量 L 的交集,用拉普拉斯算了计算两阶梯图像 D .由于仅对人脸待选区域进行卷积,运算量并不很大,即

$$W' = \{W_i | W_i \in W, i = 1, 2, 3, \dots, k, k < n\} \\ D = \nabla^2(W' \cap L) \quad (7)$$

2.2.2 设计梯度模板如图4所示,该模板大小将随区域大小做自适应调整,设中间区域为 mid,左右两边分别为 left 和 right.

$$\overline{mid} = \sum_{m=h/3}^{5h/6} \sum_{n=-w/6}^{w/6} D(x+m, y+n), \\ \overline{left} = \sum_{m=h/2}^{2h/3} \sum_{n=-w/3}^{-w/6} D(x+m, y+n), \\ \overline{right} = \sum_{m=h/2}^{2h/3} \sum_{n=w/6}^{w/3} D(x+m, y+n). \quad (8)$$

式中 $x, y, h, w \in S_i$.

2.2.3 若中间块的梯度值小于左右两边则删除该区域 W_i 和链表中对应的 S_i . 设 a 为比例常数(实验中取为1.1),删除集为

$$S_{off} = \{S_i | \overline{mid} < \overline{left} \cdot a \text{ and } \overline{mid} < \overline{right} \cdot a\}. \quad (9)$$

实验结果发现对于简单背景的图片,仅用梯度模板就能得到很好的效果(见图7(b)).

2.3 基于亮度的模板

传统的人脸模板,都是要求人脸处于正面位置,缺乏通用性.1994年Sinha在模式识别应用中提出了比例模板的概念,分析人脸的局部特征,可以发现比例模板应于人脸检测也非常合适.为了处理人脸侧转和旋转的情况,本文提出了用于检测眼睛的比例模板,该模板包含有左右对称的两部分,图5给出了左眼模板.

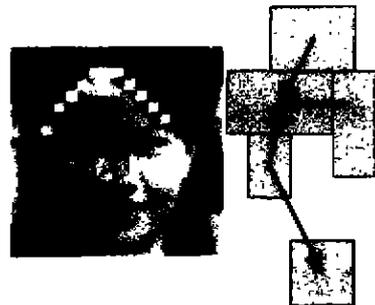


图5 亮度模板
Fig. 5 Luminance template

在亮度空间内,人眼睛的平均灰度要低于额头、鼻梁和脸颊,图5箭头方向为灰度下降方向,为了处理人脸侧转的情况,须对人脸的左右眼分别检测,只需检测到一只眼睛即可判定为人脸。由于人脸的大小是不固定的,因此要对图像进行非线性插值,做多尺度空间的判别,实际算法如下:(1)选取人脸待测区域,根据人脸待选区域链表 S ,以 $S_i(x,y)$ 为额部中心点,以 $h \times h$ 为整个检测区域。(2)区域缩放,为减少运算量,模板要保持不变,将图像变换到适合模板的大小,如图5所示左右模板大小为 12×14 ,考虑到容错性,最小的搜索区域定为 20×20 ,区域的缩放其实是一个重采样和插值的问题。在这里我们对灰度图像采用三次插值法^[5]。(3)在插值后的图像区域内,根据模板的比例,求出每个块的平均灰度,如果按箭头所示各块的灰度比率大于常数 a (在实验中取为1.1),则认为已经找到了眼睛,对整个人脸区域由上到下由左到右逐一匹配,直到找到眼睛为止。

我们都知道多尺度运算和全搜索非常消耗运算量,为此这里提出了几点对于提高速度的算法改进:(1)人脸待选区域链表 S 中本身就包含了人脸长宽比例的信息,可将区域先缩为 20×20 ,进行模板匹配,若找不到眼睛再将区域放大 b 倍(实际实验中 b 取1.3)进行更广尺度的匹配,实验表明一次尺度扩大已足够;(2)改进搜索优先级,如左眼模板由左至右搜索,右眼模板由右至左搜索;(3)采用丢弃算法,在灰度比率中若有一个不符,则直接开始下一个位置的匹配。

若图像较暗或过饱和时,亮度空间的模板匹配可能无法取得很好的效果,这时可通过光线校正和直方图均衡来予以纠正。

3 系统介绍和检测步骤

在这里我们将给出整个系统的框图如图6,整个系统共有7个模块,流程描述如下:空间映射将输入静态图像的RGB分量映射为LCbCr分量;肤色检测利用色度分量CbCr检测肤色分布得到掩膜 M_1 (见公式(1));后处理利用公式(2)~(5)得到可能包含人脸的区域集合 W ;比例参数测定根据我们提供的参数提取法则提取区域比例参数,进一步缩小待选区域 W , W' 约为 W 的 $2/3$;对由 D 步得到的 W' 和图像亮度分量 L 的交集做拉普拉斯运算,这样平均运算量将降低为原来的0.5%;梯度模板匹配在梯度图像上用我们设计的模板进行匹配运算(见式(8),(9)),这样将基本却除误检测中的平坦和纹理区域;亮度模板匹配根据修正的区域链表 S 在亮度分量上用提出的基于眼睛

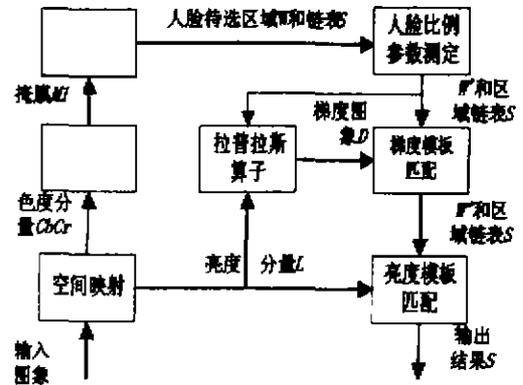


图6 系统框图
Fig. 6 Diagram of system

的模板进行匹配运算,将得到最后的输出结果,即人脸的位置和比例参数。

4 实验结果分析

我们在非统计样本中随机选取了10幅图像,图像分别是在不同背景、不同光照下拍摄的,结果表明所有的人脸都检测出来了,漏检率为0。在这里我们选取3幅有代表性的图像对第一步的结果进行分析,如图7,图7(a)是在复杂背景下拍摄的正脸,背景中有许多和人脸形状相近的动物模型的脸,并且与肤色相近的区域也比较多;(b)中的人脸约旋转 30° ,并且整个人脸处于较暗的光线下,但背景比较简单;(c)是戴眼镜的人脸,并且人穿的衣服背景都和肤色比较接近。

通过肤色检测和后处理,在图7(a)、(b)、(c)中被检测出的人脸待选区域数分别为21、10、14个,且在(a)和(c)中背景物体的边缘也很明显,很难通过传统的方法去掉。

经过人脸参数和梯度模板检测后,在图8(a)、(b)、(c)中被检测出的人脸待选区域数分别为5、1、5个,我们还可以发现对于简单背景的人脸图像如(b),仅用梯度模板就可得到很好的效果。

图9所示(a)和(c)两幅图都成功的检测出了人脸的位置,其中(c)图中有一个误检测点,该点位置背景图案上,实际运用中可通过目标背景分离的方法去除。

在实验的最后,我们用这10幅图像(共含人脸18个)进行了速度的测试并和其他算法作了些比较。在表中,漏检率定义为漏检人脸数比上实际总人脸数,误检率定义为误检的非人脸数比上实际总人脸数。

在算法实现中漏检率和误检率是需要平衡的一对矛盾量,本文的算法是在保证漏检率尽量小的前提下牺牲了部分误检率。

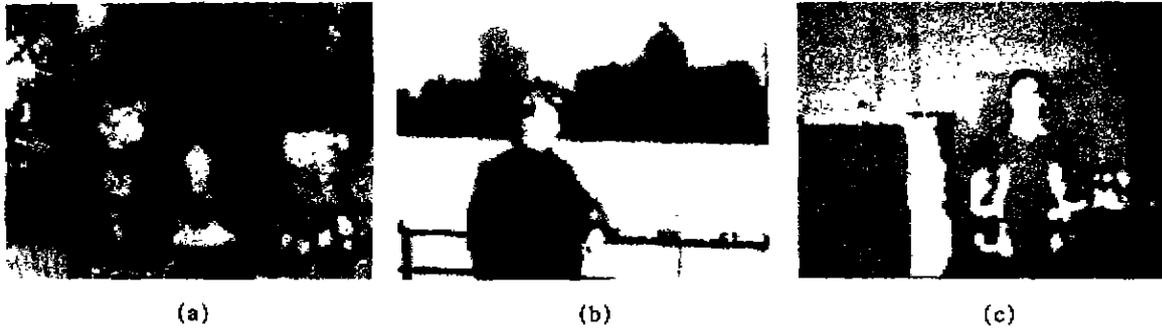


图 7 肤色检测和腐蚀膨胀后的结果
Fig. 7 The results after skin-color detection, erosion and dilation

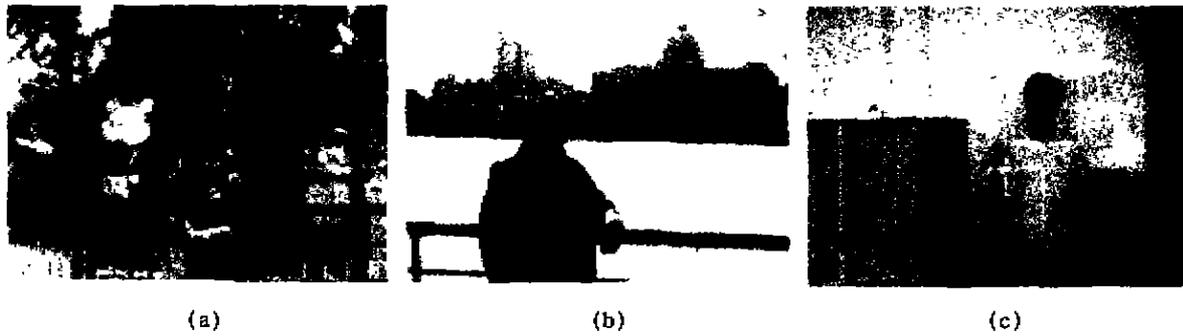


图 8 人脸参数判别和梯度模板匹配后的结果
Fig. 8 The results after face parameter judging and gradient template matching



图 9 眼睛模板匹配后的最后结果
Fig. 9 The final result after eye template matching

表 1 实验结果比较
Table 1 Comparison of the result

系 统	耗时	漏检率	误检率	图像大小	运行机器	限制
肤色+模板	136ms	0%	11%	352×288	Pentium 100	无
肤色+方差	65ms	0%	331%	352×288	Pentium 100	无
神经网络(张林)	1-3min	2.6%	10%	640×480	Pentium 133	正脸
神经网络(Henry)	1-3min	7.1%	41%	640×480		正脸
模板匹配(Brain)	3-5s	29%		128×128	TMS320C40	正脸
素类(Sung)		18.1%	8.4%	640×480		正脸

5 总结

本文介绍了一种在静态数字图像中实现实时检测人脸的方法,这种方法利用色度、梯度和亮度空间,首次将肤色检测和模板匹配结合起来,并提出了梯度模板的概念.目前在基于目标、基于分割的编码(如MPEG-4、MPEG-7、H. 263)中越来越多的需要一些能将感兴趣的目标背景有效分离的方法,而人脸正是通常最令人感兴趣的目标.本文的方法中由于已经通过肤色确定了连通的各区域,所以检测出位置后可以很方便的用区域生长的方法或跟踪链码的方法完成人脸的分割.在实际运用中如果是含有运动信息的图像序列,那么结合运动补偿的方法将大大降低运算量和误检率.另外,本文的方法对数码相机应用中也提供了一种基于目标的自动对焦,以后焦距将不在局限于照片中心,而是可以对准位于任何位置的人脸.总之,本文的方法具有较大的应用背景,我们将在今后针对各种实际应用进行进一步的研究.

REFERENCES

- [1] Douglas Chai, King N Ngan. Face segmentation using skin-color map in videophone applications, *IEEE Trans. CSVT.*, 1999, 9(4): 551—564
- [2] Wang Hualu, Chang Shin-Fu. A highly efficient system for automatic face region detection in MPEG video, *IEEE Trans CSVT*, 1997, 7(4): 615—628
- [3] Roberto Brunelli, Tomaso Poggio. Face recognition: features versus templates. *IEEE Trans PAMI.*, 1993, 15(10): 1042—1052
- [4] Leung T K, Burl M C, Perona P. Finding faces in cluttered scenes using random labeled graph matching, Fifth Intl. Conf. on Comp. Vision, Cambridge, MA, June, 1995
- [5] ZHANG Lin. Face detection, recognition and facial feature extraction in the complex background, master dissertation of Fudan University (张林复杂背景中人脸图像的检测、识别以及特征的提取,复旦大学硕士论文), 1999: 10—14
- [6] CUI Yi. *The Technology and Application for Digital Image Processing*. Beijing: Publishing House of Electronics Industry (崔屹. 数字图像处理技术与应用. 北京: 电子工业出版社), 1997: 201—225, 317—337