

视频人脸检测与特征点定位研究

张晓薇, 张文俊

(上海大学 影视艺术技术学院, 上海 200072)

摘要:基于对视频中人脸的检测和特征点定位算法的学习与研究,文中提出了一种结合连续均值向量和主动形状模型的新方法。该方法首先对视频帧进行局部连续均值量化变换处理,快速地检测出似人脸区。然后用改进ASM方法来定位所检测出人脸的特征点,使用双阈值Sobel算子进行边缘检测,提取边缘信息,当搜索时利用这些边缘信息来改变特征点的权值,使带有更多边缘信息的像素点更有可能被选为最佳候选点。该方法可以快速地对面脸做出检测,并能很好地定位出人脸特征点。

关键词:人脸检测;特征点定位;连续均值量化;主动形状模型

中图分类号:TP309

文献标识码:A

文章编号:1673-629X(2012)05-0202-03

Study of Face Detection and Feature Points Location in Video Sequence

ZHANG Xiao-wei, ZHANG Wen-jun

(School of Film & TV Arts and Technology, Shanghai University, Shanghai 200072, China)

Abstract:In the basic study and analysis of the human face detection and feature point location algorithm in video, it proposes a method. First determine the like-face area in the video frame with local SMQT characteristics; Then position the detected human face feature point with the modified ASM, using double threshold Sobel operator for edge detection to extract edge information, using the edge information to change the weight of the feature points when searching, to make more likely of the possibility that the pixels with stronger edge information is selected as the best candidate point. This method was tested to show good positioning of facial features based on fast detection.

Key words:face detection; feature points location; local SMQT; ASM

0 引言

视频作为一种最有价值信息已经被广泛接受。而视频人脸的检测与定位,作为人脸信息处理过程中的一个关键技术,已经发展成为一个相对独立的研究方向,研究方法和研究手段也在逐步成熟^[1]。

而基于视频流的人脸分析,已成为当前视觉检测和模式识别^[2]研究中的一个热点,其目的是在视频中自动检测、定位及跟踪,在各领域具有广泛的应用价值。如安全监控、身临其境的协作环境、多媒体游戏、人类计算机交互、视频会议、视频编码等。在一个人脸识别系统中,人脸检测是需要的第一步。

近年来,对人脸视频处理的提高已经使得面对面的交流在现实世界应用中成为可能。

1 人脸检测

由于脸部大小、位置、角度和一些其他因素,以及面部表情、光线、遮挡和运动等,会引起在视频中检测人脸的影响,造成干扰^[3]。人脸的多样化,如表情、发型、眼镜等都使人脸检测问题变得复杂。而文中以简单快速为原则,采用局部连续均值量化变换对视频帧进行快速的人脸检测,确定最初的一个似人脸区。

局部SMQT特征:

连续均值量化^[4]使用展现图像的底层结构的方法,自动集中改进图像细节。SMQT特性将会被应用到一幅图像的局部区域来提取光敏感特征。

局部地区可以用几种方式来定义。例如,一种直接的方法是把图像分成预定大小的图像块;另一种方法可以是通过在一个固定点的半径圆上的插值来提取数值。不管怎样,一旦这个局部地区被定义了,那么它就是一个像素值的集合。

设 x 为一个像素, $D(x)$ 为在一幅图像中对于局部地区 $|D(x)|=D(x)$ 时的像素集。那么考虑该局部地区的SMQT变换为:

收稿日期:2011-09-10;修回日期:2011-12-15

基金项目:上海大学创新基金(A16-0114-11-008)

作者简介:张晓薇(1986-),女,浙江嘉兴人,硕士研究生,研究方向为视频图像处理、变形动画;张文俊,博士,教授,研究方向为数字媒体技术与应用、数字新媒体、网络通信技术及计算电磁学等。

$$\text{SMQT}_L: D(x) \rightarrow M(x) \quad (1)$$

这个结果对增益和偏差是不敏感的,而这些特性在考虑图像整体强度 $I(x)$ 的形成时是可取的,其中 $I(x)$ 是反射率 $R(x)$ 和照度 $E(x)$ 的乘积。另外,相机的影响可以建模为一个增益因子 g 和一个偏差度 b 。因此,该图像的模型可描述为:

$$I(x) = gE(x)R(x) + b \quad (2)$$

一般情况,反射性和照明的分离是一个病态呈现的问题。普遍解决这个问题的方法包括假设 $E(x)$ 是空间平滑的。进一步,在选择局部地区,如果照度被认为是常数,那么 $E(x)$ 由下式给出:

$$E(x) = E, \forall x \in D \quad (3)$$

对于给出的方程(3)的有效性,局部地区的SMQT会产生照明和摄像头的不敏感性特性^[5]。这表示对一个特定的 L 层,所有的具有相同结构的局部形式会产生相同的SMQT特征,如图1所示。

使用局部SMQT特征可能产生的模式数量是 $(2^L)^D$ 个。例如,图1中, 4×4 模式在 $L=1$ 有 $(2^1)^{4 \times 4} = 65536$ 个可能模式。

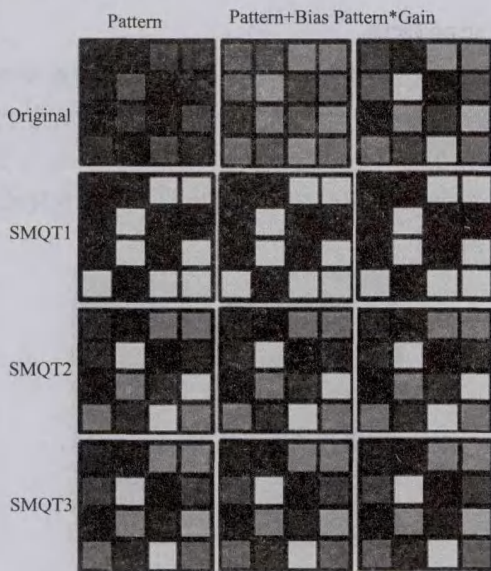


图1 4×4 的局部SMQT模式

对于输入视频,首先进行预处理,然后对于视频帧进行局部SMQT,对视频序列中的人脸进行检测实验,该方法能较快地确定出似人脸区,如图2所示。

不能说这样的检测方法是最好的,像其他还有许多好的技术,比如Ada-boost^[6]算法。但原则是方便快捷,所以在这个实验中,局部SMQT是一个更好的选择。

2 人脸特征点定位

自动并且准确的人脸特征点定位是很困难的,如果定位的偏差相对较大时,就会引起图像的错误定位。

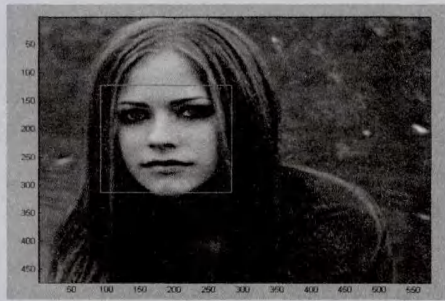


图2 基于局部SMQT的人脸检测

经典ASM^[7]被描述为一维剖面图上在每个标记点和一个线性点分布模型上计算一般距离。ASM是一种属于基于特征匹配方法的模型,在形状近似复合图像特征时,它和图像交互,把形状弯曲到特征。ASM首先在手动标记图像集上训练,训练后就能用ASM来搜寻人脸上的特征。

2.1 ASM模型

ASM首先需要在手工标记好的图像集中训练,完成训练后就可以用它对人脸进行特征点定位。

ASM方法主要有以下3个步骤^[8]:

1) 图像训练集标定^[9]。

随机选择 N 幅人脸图像,在边缘手动标记 n 个点作为训练数据。这些 n 个边缘点一般标记在那些可以代表目标区域的地方,像眼睛和嘴巴的最高点和最低点。标定的点为:

$$S_i = (x_{i1}, y_{i1}, x_{i2}, y_{i2}, \dots, x_{in}, y_{in})^T \quad (4)$$

$i = 1, 2, \dots, N$

其中 (x_{ij}, y_{ij}) 代表了第 i 幅图像的第 j 个轮廓点的坐标; n 代表了每幅图像标记的点数; N 代表训练图像数目; 每一幅图像的 S_i 代表一个形状向量。

2) 建立模型。

边缘标定是在各自不同的样本图像上进行操作,在得到图像的形状向量后要对其进行归一化。采用主成份分析(PCA)的思想对标定点的坐标值进行正交变换,得到主要变化参数。

为了从训练样本中获得平均轮廓向量:

$$\bar{S} = \frac{1}{N} \sum_{i=1}^N S_i \quad (5)$$

N 个形状向量的协方差阵:

$$C = \frac{1}{N} \sum_{i=1}^N (S_i - \bar{S})(S_i - \bar{S})^T \quad (6)$$

计算协方差 C 的特征值以及特征值对应的特征向量,并且把特征值按照降序排列,选取 k 个最大特征值 $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$, 得到对应的特征向量 $P = \{P_1, P_2, \dots, P_k\}$, P 代表主成份。

主成份分析后,任意的人脸形状向量 S 可以被表示成一个平均形状向量和 k 个形状向量的一个线性组

合 $s = \bar{s} + pb$, 因为 P 是正交的, 所以 b 可以被表示成 $b = p^T(s - \bar{s})$, 其中向量 b 是形状模型的参数, p 是特征子空间基, s 是目标形状。不同的 b 对应不同的形状, 并且它们的变换是在一定的范围内。

3) 定位模型。

ASM 使用了基于轮廓的灰度匹配^[10]方法, 沿着每个轮廓点的轮廓法线方向进行采样作为一个灰度特征, g_{ij} 代表了第 i 幅图像的第 j 个样本点的灰度特征。所有图像的第 j 点的灰度特征平均值为:

$$\bar{g}_j = \sum_{k=1}^M g_{kj} \quad (7)$$

而协方差矩阵为:

$$G_j = \frac{1}{M} \sum_{i=1}^M (g_{ij} - \bar{g}_j)(g_{ij} - \bar{g}_j)^T \quad (8)$$

灰度平均和协方差作为该点的匹配特征。对每个特征点作相同的采样分析, 整个轮廓的统计特征就可以获得。ASM 使用轮廓的灰度特征执行一个迭代过程, 在每步的迭代搜索中, 调整参数 b 来改变当前模型的位置, 使得模型和物体的实际轮廓逐渐匹配, 并且最终获得特征点的位置信息。

2.2 ASM 算法改进

人脸上的定位标记点和定位特征点是相同的, 因为标记点标出了人脸特征。寻求 ASM 的改进, 因为它常常不精确而且有时会完全失败。边缘图像带有大量的原始图像信息, 并且有对光不敏感的特征。因此, 视频序列中人脸检测带有的光条件改变, 可以用边缘图像来处理^[11]。

像 Sobel 算子、Canny 算子、拉普拉斯算子等都是经典的边缘提取算子。其中 Sobel 算子属于空间域的边缘检测方法, 它检测出的边缘相对比其他算子检测的边缘较光滑, 而且在图像存在噪声时也能检测出边缘, 但是并不能较好地消除噪声, 使用传统 Sobel 算子的另外一个缺点是检测出的像素点过多, 使边缘线过粗。阈值设置的高时保留的信息多, 同时将不需要的非边缘信息也保留了下来; 阈值设置的低时又去掉了有用的信息, 使边缘有间断, 不连续。

针对 ASM 方法的缺陷和不足, 文中对其进行了改进: 利用边缘信息来改变特征点搜索过程的权值, 使边缘信息较强的像素点被选为最佳候选点的可能性更大。

为此, 为了保留两个阈值之间的像素点, 文中采用了双阈值^[12]; 然后通过细化来消除边缘中的过多像素及噪声。具体步骤为:

1) 设置一个高阈值为 TH , 一个低阈值为 TL , 梯度图像为 $G(x, y)$, 处于两个阈值之间的像素值为 1, 其它为零。

$$G(x, y) = \begin{cases} 1 & TL < G(x, y) < TH \\ 0 & \text{other} \end{cases} \quad (9)$$

2) 将处于高低两个阈值范围内的点作为候选点 $G(x, y)$, 从候选点出发, 选择两个大小不同的区域 S 和 B , 求这两个区域的像素值之和 SumS 和 SumB , 即

$$G(x, y) = \begin{cases} 1 & \text{SumS} = n, \text{SumB} > (n, a) \\ 0 & \text{SumS} = n, \text{SumB} < (n, a) \end{cases} \quad (10)$$

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。经过上述处理后得到的边缘图像边缘相对较细, 轮廓也较为清晰, 如图 3 所示。

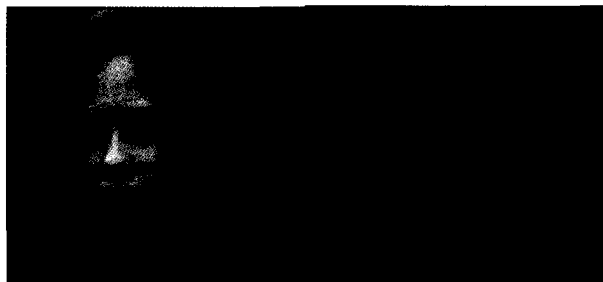


图 3 原始图像, Sobel 边缘图

3 实验结果

通过实验, 采用局部 SMQT 能较为简单快速地对视频序列中的人脸进行检测。针对主动形状模型 ASM 存在的不足, 对其进行了改进处理, 加入了边缘信息约束, 并可以较准确定位出较多的人脸特征点, 如图 4 所示。

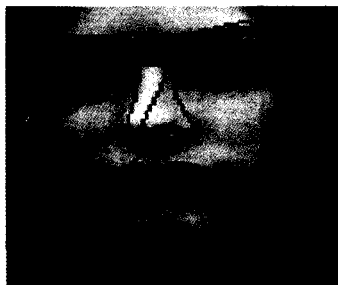


图 4 改进 ASM 的人脸特征点定位

4 结束语

文中展示了连续均值量化变换与主动形状模型相结合的方法, 对视频序列中的人脸进行检测定位。而对于主动形状模型算法, 可以增加手工标记点使定位更准确, 或者减少某些变量来加快速度, 又或改进初始模型, 比如改进纹理模型等。

人脸检测是一个富有挑战性的课题, 从目前的发展以及工作中发现的一些问题来看, 如何缩短检测定位时间、消除光照影响, 同时能保证定位的准确性等方面仍然值得继续探索。

(下转第 208 页)

文档的片段,而不破坏其层次结构。当其中一部分数据发生变化时,不需要重发整个结构化的数据^[12]。这不仅避免了重建 WEB 页,还减轻了服务器的负担,这种数据更新方式称之为“粒状更新”。基于 DB2 9.5 兼具有处理 SQL 和 XQuery 的两种不同解析方式的特点,便于对关系数据和 XML 数据的灵活处理,故选其为 EMR 系统的数据库管理平台,DB2 的数据存储图如图 3 所示。

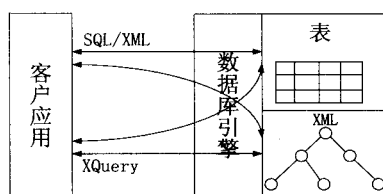


图 3 SQL 或 XQuery 透明查询

4 结束语

论文将 XML 技术应用于医疗临床系统,定义了以病人为中心的 EMR 数据结构,重点论述了 EMR 的数据存储策略。由于 XML 关系数据库的模式是二维的表结构,因此相应的存储策略就是将 XML 文档的树形结构转换为表结构。

文中以 DTD 来定义需要交换的 EMR 文档,通过验证文档的有效性,保证文档内容遵守共同的定义和规范,一方面,实现了病历数据的存储和展现,另一方面,DTD 的集成能将病历中新出现的结构化内容集成到 EMR 系统中。实践应用表明,XML 技术实现了病历的自由输入和结构化输入相结合的可行性,这一数据模型,可作为 EMR 与医院内部其他业务数据库数据

交互的标准,为实现整个医疗信息平台的数据集成奠定了基础。

参考文献:

- [1] 赵洋,李万龙,白杰英.基于本体的电子病历检索系统研究[J].计算机技术与发展,2010,20(3):211-217.
- [2] 覃永胜,林崇健.基于 DB2 pure XML 的电子病历研究[J].中国数字医学,2009(2):14-17.
- [3] Health Level Seven (HL7) Version 3.0[S]. Michigan: Health Level Seven Inc,2003.
- [4] 周健,孙丽艳.面向对象 XML 的存储模式的研究[J].计算机技术与发展,2009,19(3):114-119.
- [5] Graves M. Designing XML Database[M]. 北京:机械出版社,2002:46-47.
- [6] Erdman M, Studer R. How to Structure and Access XML Documents with Ontologies[J]. Data and Knowledge Engineering, 2000,36(3):317-335.
- [7] 庄子明.基于 XML 的数据库技术及应用[J].计算机工程,2002,28(6):119-121.
- [8] 李德有.基于 XML 电子病历数据存贮与转换的研究[D].哈尔滨:哈尔滨理工大学,2007.
- [9] 方翔.XML 文档到关系数据库的直接转换[J].计算机工程,2001,27(1):65-66.
- [10] 曹亮,王茜,卢菩.XML 数据在关系数据库中存储和检索的研究和实现[J].东南大学学报(自然科学版),2002,32(1):124-127.
- [11] 张继光.电子病历特征及其设计方案探讨[J].中国医院管理,2003,23(5):44-46.
- [12] Widom J. Data Management for XML: Research Directions[J]. IEEE Data Engineering Bulletin,1999,22(3):44-52.

(上接第 204 页)

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] 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.
- [3] 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.
- [4] Nilsson N, Nordberg J, Claesson I. Face Detection Using Local SMQT Features and Split Up Snow Classifier[C]// ICASSP. [s.l.]:[s.n.],2007:589-592.
- [5] Nilsson M, Dahl M, Claesson I. The successive mean quantization transform[C]//IEEE International Conference in Acoustics, Speech and Signal Processing (ICASSP). [s.l.]:[s.n.],2005:429-432.
- [6] 王刚,周激流,何坤.基于 AdaBoost 和 LLE 的视频人脸特征提取研究[J].四川大学学报,2008,45(3):512-516.
- [7] 朱杰,唐振民. ASM 与彩色 Gabor 特征相结合的人脸关键特征点提取[J].计算机科学,2001,37(4):265-268.
- [8] 李晓卫,杨邦荣,杨炼.改进的用于人脸面部特征定位的 ASM 方法[J].电脑与信息技术,2008,16(2):14-17.
- [9] 刘洵,张大力.基于 ASM 的 CT 图像序列标记点定位方法研究[J].计算机工程与应用,2005,41(13):180-182.
- [10] Cootes T F, Taylor C J, Cooper D H, et al. Active shape models: Their training and application[J]. CVGIP: Image Understanding, 1995, 61:38-59.
- [11] 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.
- [12] 俞扬信,严云洋.视频序列中的人脸检测与定位算法研究[J].计算机技术与发展,2009,19(2):109-111.