

基于相关型图像传感器3维人脸成像的3维AAMs人脸识别方法的研究

詹曙¹⁾ 常虹¹⁾ 蒋建国¹⁾ Ando S²⁾

¹⁾(合肥工业大学计算机与信息学院, 合肥 230009) ²⁾(日本东京大学信息科学学院, 东京 113-8656)

摘要 主动外观模型是基于统计分析建立物体2维模型的有效方法,它融合了目标的形状和纹理信息。在基于相关型图像传感器3维人脸成像的基础上,提出了一种建立3维人脸模型的方法,该方法利用由相关型图像传感器得到的深度信息和与之对应的亮度信息将2维AAMs扩展为3维AAMs,融合人脸的形状、纹理和深度信息来构建3维人脸模型。人脸识别实验结果表明,该方法在不同人脸姿态、表情和光照条件下识别效果要优于Eigenface和2维AAMs。

关键词 人脸识别 相关型图像传感器 3维人脸成像 3维主动外观模型

中图法分类号:TP391 文献标识码:A 文章编号:1006-8961(2008)10-2059-04

Improved 3D AAMs for Facial Recognition Based CIS 3D Facial Imaging

ZHAN Shu¹⁾, CHANG Hong¹⁾, JIANG Jian-guo¹⁾, Ando S²⁾

¹⁾(School of Computer & Information, Hefei University of Technology, Hefei 230009)

²⁾(Department of Information Physics and Computing, the University of Tokyo, Tokyo 113-8656 Japan)

Abstract Active Appearance Models (AAMs) is an effective statistical method to build 2D model for an object, which combines shape and texture information. A novel method for building 3D face model is proposed, which makes use of the depth information and corresponding intensity information generated by correlation image sensor (CIS), and extends 2D AAMs to 3D AAMs. The proposed improved AAMs fuse the shape, texture and depth information of face to build 3D face model. In facial recognition experiments which using 3D facial images based on CIS imaging system, the improved 3D AAMs model shows better recognition result than traditional AAMs algorithms and Eigenface.

Keywords facial recognition, CIS (correlation image sensor), 3D facial imaging, 3D AAMs(active appearance models)

1 引言

人脸识别作为一种生物识别技术,在安全控制、身份鉴别等方面都有着巨大的潜力和应用价值。2维人脸识别采用人脸亮度图像进行识别,易受到光照变化、人脸姿态变化和表情变化的影响,很难构成鲁棒性强的识别系统。而3维数据带有人脸的原始几何信息,3维人脸识别能够有效克服上述因素的影响。

近年来,3维人脸识别已经成为计算机视觉,模

式识别和人工智能领域的研究热点。当前3维数据的捕获技术主要有如下3类:基于立体视觉,结构光和激光扫描仪的捕获方式^[1]。基于立体视觉的捕获技术是通过融合不同角度的2维图像来合成3维人脸模型或3维图像,通过各种光照模型,人脸表情模型和姿态估计去消除光照条件,姿态变化和表情变化带来的影响^[2]。结构光技术是采用一组模式光,如光点、光条或光面,投影到人脸,由模式光的形变可以得到人脸表面的深度信息^[3]。激光扫描仪获取的3维数据较精确,但是仪器价格昂贵,采集速度也较慢。

基金项目:教育部博士点基金项目(20060359004);教育部留学归国人员科研启动基金(413117)

收稿日期:2008-07-21;改回日期:2008-07-26

第一作者简介:詹曙(1968~),男,副教授。2000于中国科技大学获信号与信息处理专业博士学位。主要研究方向为图像分析、计算机视觉、3维成像系统和模式识别。E-mail: shu_zhan@hfut.edu.cn

3 维人脸模型的建立是 3 维人脸识别的前提。目前,常用的是网格模型^[4]和统计模型^[5]。网格模型是将人脸表示为许多三角形面片的组合,面片的顶点为网格节点。由于节点和三角形面片数目直接关系到模型的精确度,所以一般网格模型需要处理大量的数据,这使得建模的时间较长。统计模型是利用统计方法从大量样本中得到人脸共性,建立统计意义上的人脸模型。由于数据来源的问题,建立 3 维人脸统计模型常用的不是人脸的直接物理量,这必然会影响到模型的精确度。

由 Cootes 等人提出的主动外观模型(active appearance models, AAMs)^[6-9]融合了物体的形状和纹理信息,采用统计分析的方法建立先验模型,是一种较好的建立 2 维模型的方法。针对 2 维人脸识别的不足,本文提出了一种新的 3 维人脸模型建立方法,它利用相关型图像传感器得到的人脸深度图像和与之对应的灰度图像将 2 维 AAMs 扩展为 3 维 AAMs。人脸识别实验结果表明,用该方法进行人脸识别效果要优于传统的 AAMs 算法。

2 3 维人脸实时成像系统

本文利用一种基于相关型图像传感器(correlation image sensor, CIS)的 3 维人脸实时成像系统^[10,11]获取 3 维人脸数据。该系统可以获取足够的深度和表面信息,解决 3 维数据获取不足的问题。CIS 是成像设备,它可以得到表示外部参考信号与附带光源之间相关性的相关图像,同时还可以得到亮度图像。CIS 得到的深度图像具有如下优点:(1)对光照变化具有较强的鲁棒性;(2)对姿态的变化具有较强的鲁棒性;(3)剔除替代者可靠性高。系统配置如图 1 所示,可以得到密集的,与表面反射系数无关的像素点法向图,以及由相关型图像传感器和 3 个调制光得到的亮度图像。利用相关积分可以消除背景的影响。从获取的法向图中,假设周期性边界条件且表面是光滑的,通过最小化积分量来重构深度图像:

$$J = \iint \{ (f_x - p)^2 + (f_y - q)^2 \} dx dy$$

其中, p 是测量得到的 x 方向微分, q 是 y 方向微分。在下面内容中,利用上述重构的 3 维信息来构建人脸识别系统。

三相调制光将人脸表面法向编码到反射光的幅度和相位中,相关型图像传感器输出人脸表面每一像素的幅度和相位。实验系统如图 2 所示,输出图像大小为 64×64 ,帧率为 30Hz,调制光的频率为 800Hz。图 3 是相关型图像传感器的图片,系统得到的图像包括亮度图像如图 4(a)和深度图像如图 4(b)所示。

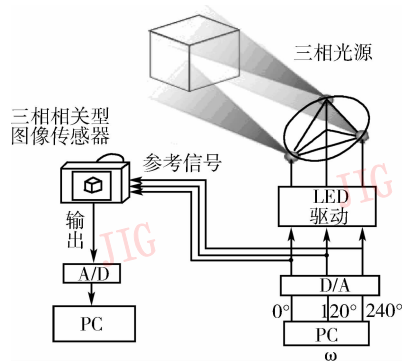


图 1 3 维实时成像系统示意图

Fig. 1 The scheme of 3D real-time imaging system

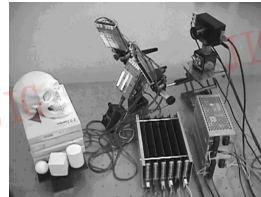


图 2 3 维人脸实时成像实验系统实物图

Fig. 2 The experiment system of 3D real-time facial imaging

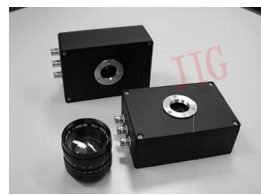


图 3 相关型图像传感器

Fig. 3 The correlation image sensor



(a)亮度图像 (b)深度图像

图 4 3 维人脸图像

Fig. 4 The 3D facial image

亮度图像即为通常的灰度图像,深度图像反映了目标到参考平面的距离信息。用 VTK (visualization toolkit)来显示深度图像如图 5 所示。这样 3 维人脸数据可以在 3 维坐标中测量,如图 6 所示。



图 5 深度图像显示结果

Fig. 5 The depth image visualized result

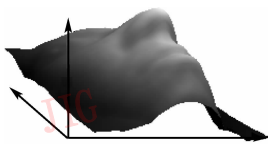


图 6 3 维人脸参考系

Fig. 6 The reference frame of 3D face

3 3 维 AAMs

3.1 AAMs

AAMs 融合了物体的形状和纹理信息来建立外观模型。AAMs 的建立主要有 3 部分,即 shape 模型、texture 模型和 Appearance 模型的建立^[6-9]。在本文的方法中,增加深度模型将 2 维 AAMs 扩展为 3 维 AAMs。深度信息由与亮度图对应的深度图像得到。

3.2 3 维 AAMs

(1) 形状模型

shape 描述了目标的几何信息,可以用一系列点的坐标来表示。在人脸上调定反映面部关键特征点,shape 模型描述各个标定点的分布。如图 7 所示,选取了 68 个特征点。对于一幅标定了 n 个特征点的人脸图像,其 shape 向量 S 可以用一个向量表示: $S = (s_1, s_2, \dots, s_{2n})$ 。训练集中的 N 幅图像就可以用 N 个向量表示: $X = (S_1, S_2, \dots, S_N)$ 。通过形状对齐来把训练集中的所有形状对齐到同一坐标系下,之后对对齐后的形状做 PCA (principal components analysis) 分析,得到统计形状模型:

$$S = \bar{S} + P_s b_s \quad (1)$$

式中, \bar{S} 为平均形状, b_s 为形状参数, P_s 为 X 的协方差矩阵的特征向量构成的矩阵。

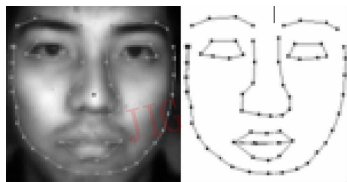


图 7 人脸标定及点分布模型

Fig. 7 Labeled face and the PDM (points distribution model)

(2) 深度模型

深度信息是通过深度图像采样得到的,深度图像存储了人脸表面到参考面的距离信息。将所有人脸变形到平均形状并对深度图像采样,得到不同图像中采样出来的像素一一对应的深度样本。然后对深度样本进行对齐,即归一化深度样本,以减小不同参考平面对深度数据的影响。对归一化后的深度进行 PCA 分析,得到深度模型:

$$d = \bar{d} + P_d b_d \quad (2)$$

式中, \bar{d} 是归一化后的平均深度向量, P_d 是特征矩阵, b_d 为深度模型的参数。

(3) 纹理模型

纹理指人脸的颜色信息,通过对人脸亮度图像的采样获得。与获取深度样本类似,将所有人脸变形到平均形状并采样亮度图像,得到不同图像中采样出来的像素一一对应的纹理样本。然后进行纹理对齐,即归一化纹理样本,以消除成像因素的影响。对归一化后的纹理进行 PCA 分析,得到纹理模型:

$$g = \bar{g} + P_g b_g \quad (3)$$

式中, \bar{g} 是归一化后的平均纹理向量, P_g 是特征矩阵, b_g 是纹理模型参数。

(4) 外观模型

3 维外观模型是把形状模型,深度模型和纹理模型融合在一起得到的。将形状模型,深度模型和纹理模型的参数链接起来,构成矢量:

$$b = \begin{pmatrix} W \begin{pmatrix} W_{sd} b_s \\ b_d \end{pmatrix} \\ b_g \end{pmatrix} = \begin{pmatrix} W \begin{pmatrix} W_{sd} P_s^T (S - \bar{S}) \\ P_d^T (d - \bar{d}) \end{pmatrix} \\ P_g^T (g - \bar{g}) \end{pmatrix} \quad (4)$$

式中, W 和 W_{sd} 为对角阵,用来调整 b_s , b_d 和 b_g 之间量纲的不同。对 b 做 PCA 分析进一步消除形状,深度和纹理之间的相关性,得到混合的外观模型:

$$b = \Phi m \quad (5)$$

式中, Φ 是特征矩阵, m 是外观模型的参数,这样就可以通过改变向量 m 中元素的值来改变形状、深度和纹理。

3.3 AAMs 的匹配过程

AAMs 对目标图像进行匹配的思想是通过调整模型参数 m ,姿态参数 t 使合成的图像与搜索到的图像之间的灰度差别最小,即使下式的误差最小化:

$$\Delta = \|\delta_g\|^2 = \|\mathbf{g}_{\text{sample}} - \mathbf{g}_{\text{model}}\|^2 \quad (6)$$

式中, $\mathbf{g}_{\text{sample}}$ 为实际图像采样纹理归一化后的值, $\mathbf{g}_{\text{model}}$ 是当前模型得到的纹理。在本文的方法中,将目标函数改进为深度误差和纹理误差的和:

$$\Delta = \|\delta_{dg}\|^2 = \|\mathbf{d}_{\text{sample}} - \mathbf{d}_{\text{model}}\|^2 + \lambda^2 \|\mathbf{g}_{\text{sample}} - \mathbf{g}_{\text{model}}\|^2 \quad (7)$$

式中, λ 为调整深度与灰度比重的系数,实验中取值为 10。与 2 维 AAMs 类似,采用线性模型来简化模型参数 m 和姿态参数 t 的优化。假设 δ_{dg} 与 δ_m 和 δ_t 之间存在一定的线性关系,即

$$\delta_m = R_m \delta_{dg}, \delta_t = R_t \delta_{dg} \quad (8)$$

R_m 与 R_t 可以利用多变量线性回归算法计算出来。3 维 AAMs 的搜索流程如下:

(1) 给出模型参数和姿态参数的初始值 m_0 和 t_0 , 计算出当前的 d_{model} , g_{model} 和 δ_{dg} ;

(2) 分别在深度图像和亮度图像中对当前模型形状覆盖的区域采样, 并归一化, 得到 d_{sample} 和 g_{sample} ;

(3) 计算误差 $E_0 = \|\delta_{\text{dg}}\|^2 = \|\delta_{\text{d}}\|^2 + \lambda^2 \|\delta_{\text{g}}\|^2$;

(4) 利用 $\delta_m = R_m \delta_{\text{dg}}$ 和 $\delta_t = R_t \delta_{\text{dg}}$ 计算模型参数 m 和姿态参数 t 的调整;

(5) 设置 $k = 1$, 并更新模型参数和姿态参数:

$m = m_0 - k\delta_m$, $t = t_0 - k\delta_t$;

(6) 重复步骤(2)、(3), 计算新的误差 E_i ;

(7) 如果 $E_i > E_0$, 尝试利用 $k = 0.5, 0.25$, 重复步骤(5)、(6); 如果 E_i 与 E_0 的差异足够小, 则匹配结束;

4 实验

实验包括3部分, 分别利用 Eigenface 和 2 维 AAMs、3 维 AAMs 建立的模型进行人脸识别, 并比较了实验结果。用 CIS 采集的图像包括 7 个人, 每人 16 幅, 其中 8 幅为亮度图像, 另 8 幅为与之对应的深度图像。这些人脸图像包括不同姿态: 抬头/低头; 微侧; 不同表情: 睁眼/闭眼, 微笑; 不同人脸细节: 是否戴眼镜, 以及不同光照。图像尺寸为 64×64 。深度图像仅用在 3 维 AAMs 的实验中。对于每个人的 8 幅亮度图像, 其中 4 幅作为训练样本, 另外 4 幅作为测试集中的人脸图像。这样在实验中有 28 个训练样本, 28 个待识别人脸图像; 在 3 维 AAMs 的实验中还包括对应的深度图像。训练样本中的亮度图像采用 68 个特征点进行标定(如图 7 所示)。

表 1 为分别采用 Eigenface, 2 维 AAMs 和 3 维 AAMs 进行人脸识别的实验结果。对于测试集中的 28 个人脸图像, 3 维 AAMs 算法可以全部识别正确, 此方法优于另外两种方法。

表 1 识别结果比较

Tab. 1 Recognition result comparison

| 模型类别 | 识别结果 |
|-----------|-------|
| Eigenface | 20/28 |
| 2 维 AAMs | 25/28 |
| 3 维 AAMs | 28/28 |

5 结论

AAMs 是一种较好的建立物体 2 维模型的方法, 它利用 2 维图像中目标物体的形状和纹理信息, 经过统计分析建立先验模型。但此模型无法描述 3 维物体。通过增加深度模型, 将 AAMs 扩展到 3 维, 并用于人脸识别实验。实验结果表明, 该方法要优于

Eigenface 和传统的 AAMs 方法。由于增加了 depth 模型, 3 维 AAMs 的建模时间要大于 2 维 AAMs。在 2 维 AAMs 建模过程中, 共有 3 次 PCA 和一次计算对角阵 W 的运算; 而 3 维 AAMs 的建立有 5 次 PCA 和两次计算对角阵 W 和 W_s 的运算; 这样 3 维 AAMs 的建模时间大约是 2 维 AAMs 的两倍。实验在 PC 机上进行, 配置为 CPU: Pentium4 2.93GHz; 内存: 256MB。实验中 3 维 AAMs 和 2 维 AAMs 的建立时间分别为 2 389.7s 和 1 142.7s, 证实了上述推测。实验中 PCA 运算的精确度设定为 95%, 在此条件下, 3 维 AAMs 和 2 维 AAMs 模型参数 m 的维数分别为 10 和 14, 这也说明了 3 维 AAMs 对物体的描述要比 2 维 AAMs 精确。在 2 维图像中, 纹理信息可以清楚地描述物体, 但无法描述现实 3 维世界中的物体; 深度信息可以提供 3 维距离信息。通过融合形状, 深度和纹理信息, 3 维 AAMs 能够更加精确地描述真实的物体。

参考文献 (References)

- Akarun L, Gökberk B, Salah A A. 3D face recognition for biometric applications [A]. In: Proceedings of the 13th European Signal Processing Conference [C]. Antalya, Turkey, 2005: 358 ~ 362.
- Xiong He-jin, Zhu Jia-lu. Study on stereo face recognition [J]. Computer Engineering, 2005, 31(8): 14 ~ 16. [熊和金, 朱家禄. 人脸立体模式识别 [J]. 计算机工程, 2005, 31(8): 14 ~ 16.]
- Mei Yi-jun, Wang Yuan-qing. Study of human faces capture and detection system based on structured-lights [J]. Computer Measurement & Control, 2007, 15(6): 720 ~ 735. [梅益君, 王元庆. 基于结构光的人脸图像获取与检测 [J]. 计算机测量与控制. 2007, 15(6): 720 ~ 735.]
- Duan Hong, Long Fei, Jiang Qing-shan. 3D face modeling method for a certain person [J]. Computer Engineering, 2007, 33(14): 1 ~ 7. [段鸿, 龙飞, 姜青山. 计算机工程, 一种个性化人脸三维建模方法 [J]. 2007, 33(14): 1 ~ 7.]
- Iwasa T, Shima T, Sai M, et al. 3D Eigenfaces for face modeling [A]. In: Proceedings of 5th Asian Conference on Computer Vision [C], Melbourne, Australia, 2002: 888 ~ 893.
- Edwards G J, Lanitis A, Taylor C J, et al. Statistical models of face images-improving specificity [J]. Image and Vision Computing, 1998, 6(16): 203 ~ 211.
- Cootes T F, Edwards G J, Taylor C J. Active appearance models [J]. IEEE Transactions on PAMI, 2001, 23(6): 681 ~ 685.
- Cootes T F, Taylor C J. Statistical models of appearance for medical image analysis and computer vision [A]. In: Proceeding of SPIE Medical Imaging [C], Washington DC, USA, 2001: 236 ~ 248.
- Cootes T F, Edwards G, Taylor C J. Comparing active shape models with active appearance models [A]. In: Proceedings of British Machine Vision Conference [C], Nottingham, UK, 1999: 173 ~ 182.
- Kurihara T, Ando S. Surface orientation imager using three-phase amplitude-modulated illumination and correlation image sensor [A]. In: Proceedings of SPIE [C], San Jose, CA, USA, 2003: 95 ~ 102.
- Shigeru Ando. Consistent Gradient Operators [J]. IEEE Transactions on PAMI, 2000, 22(3): 252 ~ 265.