

一种同步人脸运动跟踪与表情识别算法

於 俊^{1,2}, 汪增福^{1,2,3}, 李 睿^{2,3}

(1. 中国科学技术大学语音及语言信息处理国家工程实验室, 安徽合肥 230027; 2. 中国科学技术大学自动化系, 安徽合肥 230027; 3. 中国科学院合肥智能机械研究所, 安徽合肥 230031)

摘 要: 针对单视频动态变化背景下的人脸表情识别问题, 提出了一种同步人脸运动跟踪和表情识别算法, 并在此基础上构建了一个实时系统. 该系统达到了如下目标: 首先在粒子滤波框架下结合在线外观模型和柱状几何模型进行人脸三维运动跟踪; 接着基于生理知识来提取人脸表情的静态信息; 然后基于流形学习来提取人脸表情的动态信息; 最后在人脸运动跟踪过程中, 结合人脸表情静态信息和动态信息来进行表情识别. 实验结果表明, 该系统在大姿态和丰富表情下具有较好的综合优势.

关键词: 人脸运动跟踪; 人脸表情识别; 流形学习; 粒子滤波

中图分类号: TB391.9 **文献标识码:** A **文章编号:** 0372-2112 (2015)02-0371-06

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2015.02.025

A Simultaneous Facial Motion Tracking and Expression Recognition Algorithm

YU Jun^{1,2}, WANG Zeng-fu^{1,2,3}, LI Rui^{2,3}

(1. National Laboratory of Speech and Language Information Processing, University of Science and Technology of China, Hefei, Anhui 230027, China; 2. Department of Automation, University of Science and Technology of China, Hefei, Anhui 230027, China; 3. Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei, Anhui 230031, China)

Abstract: In view of facial expression recognition from monocular video with dynamic background, a real-time system was proposed based on the algorithm in which facial motion is tracked and facial expression is recognized simultaneously. Firstly, online appearance model and cylinder head model were combined to track 3D facial motion from video in framework of particle filtering; secondly, the static knowledge of facial expression was extracted through facial expression anatomy; thirdly, the dynamic knowledge of facial expression was extracted through manifold learning; fourthly, facial expression was retrieved by fusing the static knowledge and dynamic knowledge during facial motion tracking process. The experiments results confirmed the advantage on facial expression recognition even in the presence of significant head pose and facial expression variations of this system.

Key words: facial motion tracking; facial expression recognition; manifold learning; particle filter

1 引言

在传统的基于人脸的人机交互中, 人脸的身份识别是其中一个重要的需求. 进一步, 为了构建面向自然人机交互的人机情感交互系统, 人脸表情识别需要进行深入的研究.

1.1 相关工作

人脸表情识别的主要任务是通过输入人脸图像的分析来判断相关人所处的情感状态. 绝大多数的人脸表情识别方法基于对引发表情的情感的认识, 把表情划分为若干类. 例如, 文献[1]把表情划分为愤怒, 高兴, 惊讶, 悲伤, 恐惧, 厌恶等六类. 在这些方法中, 人脸表情识

别的过程就是判断输入人脸图像属于哪一类的过程.

人脸表情具有规则的结构但变化极其丰富, 因此建模人脸表情的变化特性是非常困难的. 为了分析这类高度非线性的复杂变化模式, 目前主要通过相关措施将之变换或降维到易于处理的空间来处理. 例如, 其中的一种有效方法是流形学习. 流形学习是从高维采样数据中恢复低维流形结构, 并求出相应的嵌入映射, 以实现维数约简. 它是从观测到的现象中去寻找事物本质和数据内在规律的有效方法.

人脸表情识别方法根据所采用的几何模型可以分为基于 2D 几何模型的方法和基于 3D 几何模型的方法. 前者由于相机视角的限制, 识别效果有限; 后者由于

不受限于相机的视角,是目前研究的热点^[2].

人脸表情识别方法根据获取表情信息量的方式可以分为静态的方法和动态的方法.前者根据人脸表情在每个视频帧中的空域变化模式来识别每个视频帧中的表情^[3],这里每个视频帧是相互独立的;后者根据人脸表情在每个视频帧之间的时域变化模式来分类表情^[8],这里每个视频帧之间是相互依存的.

因为人脸表情的空域信息和时域信息对于表情识别都是非常重要的,所以将静态方法和动态方法结合起来显然能够改善识别的效果.

传统的基于视频的人脸表情识别策略是先跟踪后识别.然而,由于跟踪和识别的结果是相互影响的,故如果利用这种相关性来同时进行跟踪和识别将能够增加结果的鲁棒性^[12,13].

特别的,对于人脸运动跟踪,它可以通过基于特征的方法或基于外观模型的方法来完成.基于特征的方法由于受限于特征的稳定性会产生运动漂移的问题^[14];基于外观模型的方法相对来说会更加稳定,而构建外观模型可以通过确定性方法^[15]、离线统计方法^[2,16]或在线统计方法^[17]来完成.其中,在线统计方法因其可以根据当前情况进行更新,从而可以取得更好的效果.但它会在姿态变化较大时跟踪失效.

另外,因为简单几何人头模型,比如柱状人头模型^[15,20],椭球^[21]和人头相似形状体^[22]等,具有低维参数性和个体无关性,且采用整个头部的纹理来匹配,所以适合于大姿态下的人脸全局运动跟踪.因此如果能将之与在线外观模型结合起来将能得到更好的结果.

1.2 论文创新点

本文从系统的角度研究动态变化背景下的人脸表情识别问题,提出了一种同步人脸运动跟踪和表情识别的方案(图1).

相对于目前的人脸表情识别系统^[3,4,8,9],该系统具有以下特性:结合人脸表情的静态和动态信息来进行表情识别,比仅仅采用其中一个提高了识别率.

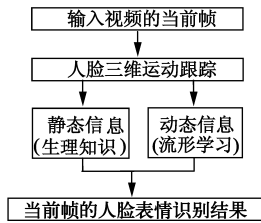


图1 系统框架

2 人脸三维运动跟踪

如引言所述,本文结合简单几何人头模型和在线外观模型各自的优势^[23]进行人脸三维运动跟踪.

2.1 基于在线外观模型的人脸运动跟踪

为了跟踪人脸的三维运动,本文通过 CANDIDE3^[24]三维模型与输入人脸图像匹配来完成.其中,CANDIDE3的人脸运动参数定义为:

$$\mathbf{b} = [\theta_x, \theta_y, \theta_z, s, t_x, t_y, \boldsymbol{\alpha}^T]^T \quad (1)$$

这里, $[\theta_x, \theta_y, \theta_z, s, t_x, t_y]^T$ 表示全局运动参数, $\boldsymbol{\alpha}$ 表示局部运动参数.

匹配的过程分为如下几个部分:

首先是从输入人脸图像中提取观测量.基于人脸三维模型,将输入人脸图像规范化到几何归一化人脸图像(GNI)^[16]后,从中提取两种观测量^[23]:颜色值和当前帧的 GNI 和首帧的 GNI 之间的光照比图像的 Gabor 小波变换系数.

其次是构建在线外观模型,并据此融合多种观测量.分别将提取的两种观测量(y_t 和 G_t)建模为 3 个分量的混合高斯模型^[17],并定义混合高斯模型的更新方程.例如对于第一种观测量:

$$\begin{aligned} m_{i,t} &= (1 - \beta) m_{i,t-1} + \beta m_{i,t-1} N(y_{t-1}; \mu_{i,t-1}, \sigma_{i,t-1}^2) \\ \mu_{i,t} &= (1 - \beta) \mu_{i,t-1} / m_{i,t} + \beta y_{t-1} m_{i,t-1} / m_{i,t} \\ \sigma_{i,t}^2 &= (1 - \beta) \sigma_{i,t-1}^2 / m_{i,t} + \beta y_{t-1}^2 m_{i,t-1} / m_{i,t} - \mu_{i,t-1}^2 \end{aligned} \quad (2)$$

$m_{i,t}$ 为第 i 个分量的权值, $\mu_{i,t}$ 为第 i 个分量的均值, $\sigma_{i,t}$ 为第 i 个分量的均方差, β 是遗忘因子.

对两种观测量进行融合的策略是基于混合高斯模型的观测似然度来融合: $p(y_t/b_t) \cdot p(G_t/b_t)$.

再次,对人脸运动时会发生的遮挡情况进行处理.计算 CANDIDE3 三维模型中每个 3D 面片重心点与视点的连线与该面片法线的夹角的余弦值,如果为负,则该 3D 面片被遮挡了.对于被遮挡的 3D 面片,因为对应的 GNI 中 2D 面片的像素值是无法获取的,所以由前一时刻 GNI 中 2D 面片的像素值和在线外观模型来估计得到.

$$\begin{cases} P_t^{(i,j)} = P_{t-1}^{(i,j)} + \mu_{t-1}^{(i,j)} + \sigma_{t-1}^{(i,j)}, & P_{t-1}^{(k,j)} \geq \mu_{t-1}^{(k,j)} \\ P_t^{(i,j)} = P_{t-1}^{(i,j)} + \mu_{t-1}^{(i,j)} - \sigma_{t-1}^{(i,j)}, & P_{t-1}^{(k,j)} < \mu_{t-1}^{(k,j)} \end{cases} \quad (3)$$

$P_t^{(i,j)}$ 为在 t 时刻的被遮挡的 3D 面片对应的 GNI 中 2D 面片的第 (i,j) 位置的像素值.

2.2 以在线外观模型匹配的结果初始化简单几何人头模型

因为柱状人头模型匹配的性能受初始值准确性的影响很大,所以当得到在线外观模型匹配得到的全局运动参数后,将它们作为后续柱状人头模型匹配的初始值.其次,将在线外观模型在图像上的投影区域作为后续柱状人头模型匹配的模板图像.

2.3 以简单几何人头模型匹配的结果初始化在线外观模型

本文选取文献[15]中的柱状人头模型作为简单几何人头模型. 匹配可通过最小化柱状人头模型在当前帧和前一帧中的共同部分之间的像素值差来完成. 将匹配得到的全局运动参数作为后续在线外观模型匹配的初始值.

3 人脸表情识别

3.1 基于静态信息来建模人脸表情

基于人脸表情的生理结构, Ekman^[1]定义了相应的量化规律, 即每种表情由哪些肌肉产生, 每个肌肉是如何作用产生特定的表情, 各个肌肉如何协同产生特定的表情. 我们将这些量化规律作为静态信息.

在跟踪得到人脸运动参数后, 基于静态信息计算人脸运动参数属于每个表情的概率, 进而得到各个表情的识别概率.

3.2 基于动态信息来建模人脸表情

3.2.1 流形学习

流形学习分为线性流形学习算法和非线性流形学习算法, 非线性流形学习算法包括等距映射(Isomap), 拉普拉斯特征映射(Laplacian Eigenmaps, LE), 局部线性嵌入(Local Linear Embedding, LLE)等. 而线性方法则是对非线性方法的线性扩展, 如局部保持投影(Locality Preserving Projections, LPP), 邻域保持嵌入(Neighborhood Preserving Embedding, NPE)等.

文献[25]应用一种有监督学习 LPP 算法来得到一个表情子空间, 在子空间内, 将不同个体的表情数据在一个统一流形上对齐. 他们的方法只考虑 6 种基本表情, 通过传统的有监督流形学习, 在同一类别的基本表情数据中定义邻近关系来简单地对齐属于不同个体的表情数据, 没有考虑表情的模糊性. 而表情具有模糊性和多义性, 对不同的类别具有不同的模糊隶属度, 应该被模糊地分为多类, 同一类别中的表情往往也有很大的差异, 比如大笑和微笑. 因此, 不能够按照传统的方式来进行有监督学习.

文献[26]应用一个可分解的生成模型在流形上将内容和形式分开来表示动态物体. 他们的方法假定由不同个体完成的类似表情可以严格地相互匹配, 并按这种匹配关系将不同个体的子流形转换为一个统一的表情流形. 同一个体的一种表情可以由多种多样的方式来完成, 比如笑的表情可以闭住嘴巴或张开嘴巴等. 所以, 同一种表情在高维图像空间会形成不同的流形, 但这些流形并不一定都是同形、同构的, 因此, 在表情流形之间进行严格的匹配是比较困难的.

文献[27]观察到个体可能的人脸形变图像都位于

原始高维图像空间中的一个低维光滑流形上, 采用 LLE 和 Lipschitz 嵌入方法来学习表情流形的结构, 分析表情的混合性和动态演变性. 然而, 他们的方法不能很好地将不同个体的表情流形在子空间内对齐, 统一到一个流形上, 只对有限个体的人脸表情数据作了探讨.

如引言所述, 目前流形学习应对人脸表情这类复杂变化的效果较好, 且由于拉普拉斯特征映射所选取的主特征点能够保持其在原始流形中的局部几何结构, 便于以后的聚类, 故本文采用 LE 进行特征降维以选取关键的主特征.

LE 的基本思想是寻求一个能够保持流形局部特性的映射, 使在高维空间流形中距离很近的点投影到低维空间后距离也很近.

考虑高维空间 D 中的 k 个样本点 X_1, X_2, \dots, X_k 所构成的高维流形, 现要找出其中所嵌入的低维流形 d . 亦即, 找出一个映射 F , 使得 $Y_i = F(X_i)$, $i = 1, 2, \dots, k$. 其中, Y_1, Y_2, \dots, Y_k 是 d 中的 k 个点.

根据谱图理论, 若数据均匀采集于高维空间中的低维流形, 则低维流形上的 Laplacian-Beltrami 算子能够由图上的 Laplacian 算子逼近. 所求解图的 Laplacian 矩阵的特征矢量即为流形上的 Laplacian-Beltrami 算子的离散逼近. 因此, 通过构建邻接图并求解其所对应的 Laplacian 矩阵的特征值即可获得嵌入的低维流形 d .

根据上述分析, 将高维空间 D 中的每个样本点 X_i 看做一个标号为 i 的节点, 则 k 个样本点 (X_1, X_2, \dots, X_k) 依次对应于 k 个节点 $(1, 2, \dots, k)$. 由这 k 个节点 $(1, 2, \dots, k)$ 构造一个加权邻接图 $G = (V, E)$, 其中 V 是顶点的集合, E 是连接相邻顶点的边的集合. 求解邻接图 G 所对应的 Laplacian 矩阵的特征矢量所构成的流形即可获得嵌入的低维流形 d . 具体实现步骤如下:

(1) 构建邻接图 G . 采用 k 近邻法或者 ϵ 法, 将相邻近的节点用一条边连接起来. 如果节点 i 与节点 j 之间相当的“近”, 则 $G_{ij} = 1$;

(2) 建立权值矩阵 W . 具体为: 如果节点 i 与节点 j 相连, 那么

$$W_{ij} = e^{-\|x_i - x_j\|^2/t}, t \in R \quad (4)$$

否则 $W_{ij} = 0$.

(3) 计算特征值 λ 和特征矢量 f . 求解下式中的特征值 λ 及其所对应的特征矢量 f

$$Lf = \lambda Df \quad (5)$$

其中, D 为对角矩阵, 对角线元素是权值矩阵 W 所在列或行的元素之和, $D_{ii} = \sum_j W_{ij}$ 或 $D_{jj} = \sum_i W_{ij}$. $L = D - W$ 是对称半正定矩阵, 称 L 为拉普拉斯矩阵. 求解式(6)的 k 个特征值并按照大小进行排序, $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{k-1}$, 将 $\lambda_0, \lambda_1, \dots, \lambda_{k-1}$ 依次代入式(5), 得到这些

特征值所对应的特征矢量 f_0, f_1, \dots, f_{k-1} . 由于特征值 $\lambda_0 = 0$ 对应的特征矢量 f_0 为常矢量, 它将整个流形都映射为同一点, 故舍之. 这样, 按照某种规则可得到 d 个从小到大依次排列的特征值所对应的特征矢量 f_1, f_2, \dots, f_d . $F = (f_1, f_2, \dots, f_d)$ 即为经过特征降维后所获得的低维空间的一组基矢量. 对于高维空间 D 中的任意一个样本点 X_i , 其映射关系为:

$$X_i \rightarrow (f_1(i), f_2(i), \dots, f_d(i)) \quad (6)$$

图 2 示例了对于某些特征数据的 LE 降维过程.

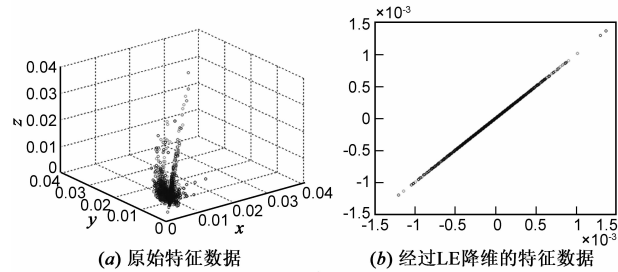


图 2

3.2.2 获取人脸表情动态信息

本文在 CMU Cohn-Kanade DataBase^[28] 上获取表情的动态信息. 该表情视频库中每个序列的首帧为中性表情, 后续帧的表情幅度依次增大, 到末帧该种表情的幅度达到最大值. 首先利用前述的人脸运动跟踪算法获取每个序列的 α_i 序列 (图 3): $\alpha_{(1)}, \dots, \alpha_{(T)}$, T_i 是第 i 个序列的帧数; 接着利用动态时间规整 (DTW)^[29] 将 T_i 校正到同一个维度 T 进而得到标准训练序列 $\alpha_{(1)}, \dots, \alpha_{(T)}$.

在得到标准训练序列后, 首先利用拉普拉斯特征映射得到其所嵌入的低维流形; 接着考虑到让该流形更适合于识别, 对其做线性判别分析 (Linear Discriminant Analysis, LDA); 最后将 LDA 后的结果作为动态信息.

基于上面获得的动态信息, 采用最近邻分类器 (Nearest Neighbor, NN) 得到各个表情的识别概率.

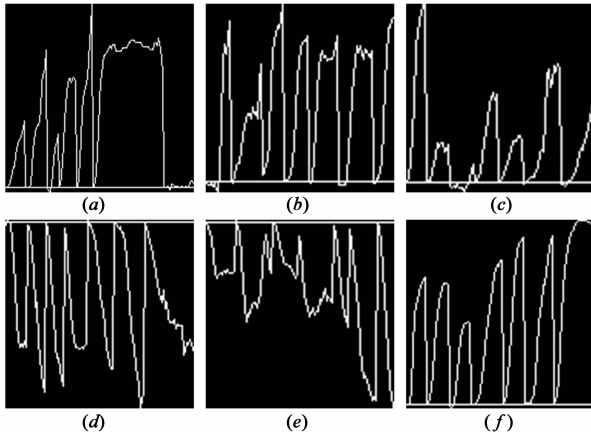


图 3 从视频中跟踪得到的 α_i 中分量的变化过程

3.3 结合静态信息和动态信息来识别人脸表情

根据 3.1 节由静态信息得到各个表情的识别概率和根据 3.2 节由动态信息得到各个表情的识别概率, 将上述概率的乘积作为各个表情的最终识别概率, 最大值对应的表情为表情识别结果.

4 实验结果与分析

4.1 人脸运动跟踪

从图 4 可见, 在姿态和表情变化较大的情况下, 本文的跟踪算法也能取得较好的结果.



图 4 人脸运动跟踪结果

4.2 人脸表情识别

4.2.1 人脸表情识别结果

实验数据为 CMU Cohn-Kanade DataBase^[28] 中视频和捕捉的视频. 在上述实验数据中, 采用 90% 的部分来训练, 10% 的部分来检验算法性能. 图 5 为表情识别结果图像. 从中可见, 本文的人脸表情识别算法取得了较好的结果.

4.2.2 评测流形学习的性能

为了评测 LE 的性能, 我们也用主成份分析 (Principal Component Analysis, PCA) 进行降维, 进而比较了 LE + LDA 和 PCA + LDA 的表情识别率.

由表 1 可见, 当降维后子空间的维数较少 (低于 20) 的时候, LE + LDA 的识别率要大于 PCA + LDA 的识别率. 这也证明了 LE 在低维空间上的有效性, 该有效性在数据的维数非常高的时候尤其显得重要.

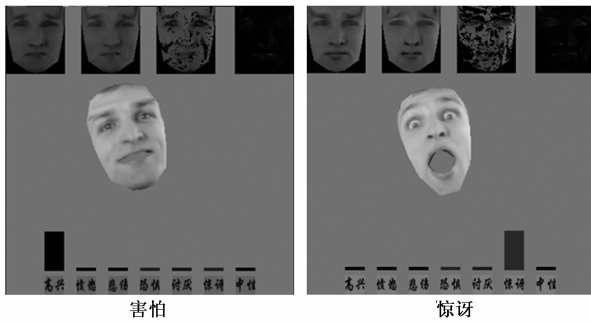


图5 人脸表情识别结果

表 1 分别采用 LE 和 PCA 降维至不同维空间后的表情识别率

LE/PCA 降维后空间的维数	表情识别率
10/10	96.97/92.46
15/15	90.64/84.86
20/20	87.65/64.78

4.2.3 客观评测与比较

分别利用本文算法及文献[3]、文献[4]、文献[8]、文献[9]提出的算法进行实验. 表情识别率与平均每帧耗时如表 2 所示.

表 2 分别采用本文算法的平均每帧耗时和表情识别率

算法	平均每帧耗时	表情识别率
本文算法	4.4s	82.3%
文献[3]的算法	3.3s	70%
文献[4]的算法	3.4s	72.7%
文献[8]的算法	4.4s	78.4%
文献[9]的算法	4.9s	79.2%

由表 2 可见,与所比较的其它算法相比,本文算法的表情识别率最高.

5 结论与展望

本文针对单视频动态变化背景下的人脸表情识别问题,提出了一种同步人脸运动跟踪和表情识别算法,并在此基础上构建了一个实时系统. 实验结果表明,该系统具有较好的综合优势.

针对本文的实验结果提出进一步研究方向:

寻找措施来减少人脸表情识别中光照和个体相关性的影响. 建立增量流形学习的方法使得当来新训练数据时不需要重复整个训练过程.

参考文献

[1] P Ekman, W V Friesen. Manual for the Facial Action Coding System[M]. Palo Alto, CA, USA: Psychologists Press, 1978.
 [2] J Sung, T Kanade, D Kim. Pose robust face tracking by combining active appearance models and cylinder head models[J].

International Journal of Computer Vision, 2008, 80(2): 260 – 274.
 [3] M Bartlett, G Littlewort, C Lainscsek. Machine learning methods for fully automatic recognition of facial expressions and facial actions[A]. ICSMC[C]. Vancouver, BC: IEEE Xplore, 2004. 1574 – 1579.
 [4] Wang Y, Ai H, Wu B, Huang C. Real time facial expression recognition with Adaboost[A]. ICPR[C]. Vancouver, BC: IEEE Xplore, 2004. 1683 – 1688.
 [5] Tang F Q, Deng B. Facial expression recognition using AAM and local facial feature[A]. ICNC[C]. Vancouver, BC: IEEE Xplore, 2007. 632 – 635.
 [6] Hammal Z, Couvreur L, Caplier A. Facial expression classification: An approach based on the fusion of facial deformations using the transferable belief model[J]. IJAR, 2007, 46(3): 542 – 567.
 [7] Kotsia I, Zafeiriou S, Pitas L. Texture and shape information fusion for facial expression and facial action unit recognition[J]. PR, 2008, 41(3): 833 – 851.
 [8] Cohen I, Sebe N, Garg A, Chen L, Huang T S. Facial expression recognition from video sequences: Temporal and static modeling[J]. CVIU, 2003, 91(1 – 2): 160 – 187.
 [9] Zhang Y, Ji Q. Active and dynamic information fusion for facial expression understanding from image sequences[J]. TPA-MI, 2005, 27(5): 699 – 714.
 [10] Zhu Z, Ji Q. Robust pose invariant facial feature detection and tracking in real time[A]. ICPR[C]. Vancouver, BC: IEEE Xplore, 2006. 1092 – 1095.
 [11] Tong Y, Wang Y, Zhu Z W. Facial feature tracking using a multi-state hierarchical shape model under varying face pose and facial expression[A]. ICPR[C]. Vancouver, BC: IEEE Xplore, 2006. 283 – 286.
 [12] North B, Blake A, Isard M, Rittscher J. Learning and classification of complex dynamics[J]. PAMI, 2000, 22(9): 1016 – 1034.
 [13] Zhou S, Krueger V, Chellappa R. Probabilistic recognition of human faces from video[J]. CVIU, 2003, 91(1 – 2): 214 – 245.
 [14] Wei Zhang, et al. Real time feature based 3-D deformable face tracking[A]. ECCV[C]. Berlin, Heidelberg: Springer-Verlag, 2008. 720 – 732.
 [15] M L Cascia, et al. Fast, reliable head tracking under varying illumination: An approach based on registration of texture mapped 3D models[J]. TPAMI, 2000, 22(4): 322 – 336.
 [16] J Ahlberg. Model-based coding: Extraction, coding, and evaluation of face model parameters[D]. Sweden: Linköping University, 2002.
 [17] S Zhou, et al. Visual tracking and recognition using appearance – adaptive models in particle filters[J]. TIP, 2004, 13

- (11): 1491 – 1506.
- [18] Lui Y M, et al. Adaptive appearance model and condensation algorithm for robust face tracking[J]. TSMC Part A, 2010, 40(3): 437 – 448.
- [19] F Dornaika, F Davoine. Simultaneous facial action tracking and expression recognition in the presence of head motion [J]. IJCV, 2008, 76(3): 257 – 281.
- [20] Xiao J, Moriyama T, Kanade T. Robust full-motion recovery of head by dynamic templates and registration techniques[J]. IIJST, 2003, 13: 85 – 94.
- [21] Basu S, Essa I, Pentland A. Motion regularization for model-based head tracking [A]. ICPR [C]. Vancouver, BC: IEEE Xplore, 1996. 611 – 616.
- [22] Malciu M, Preteux F. A robust model-based approach for 3D head tracking in video sequences[A]. ICAFGFR [C]. Vancouver, BC: IEEE Xplore, 2000. 169 – 174.
- [23] 於俊, 汪增福. 一种鲁棒高精度的人脸三维运动跟踪算法[J]. 计算机研究与发展, 2014, 51(4): 802 – 812.
Yu Jun, Wang Zeng-fu. A robust and high accurate 3D facial motion tracking algorithm[J]. Journal of Computer Research and Development, 2014, 51(4): 802 – 812. (in Chinese)
- [24] J Ahlberg. CANDIDE3-an updated parameterized face[R]. Technical Report LiTH-ISY-R-2326, Sweden; Department of Electrical Engineering, Linköping University, 2001. 1 – 16.
- [25] C Shan, S Gong, et al, Dynamic facial expression recognition using a bayesian temporal manifold model[A]. British Machine Vision Conference [C]. Vancouver, BC: IEEE Xplore, 2006. 2364 – 2369.
- [26] K Lee, J Ho, M Yang, D Kriegman, Video-based face recognition using probabilistic appearance manifolds [A]. CVPR [C]. Vancouver, BC: IEEE Xplore, 2003. 1784 – 1789.
- [27] Y Chang, C Hu, M Turk, Probabilistic expression analysis on manifolds [A]. CVPR [C]. Vancouver, BC: IEEE Xplore, 2004. 520 – 527.
- [28] T Kanade, J Cohn, Y Tian. Comprehensive database for facial expression analysis [A]. ICAFGFR [C]. Vancouver, BC: IEEE Xplore, 2000. 46 – 53.
- [29] M Muller. Information Retrieval for Music and Motion [M]. Springer Berlin Heidelberg, 2007.

作者简介



於俊男, 1983 年生于安徽滁州. 中国科学技术大学语音及语言信息处理国家工程实验室副研究员. 研究方向为人机情感接口, 智能机器人.

E-mail: harryjun@ustc.edu.cn



汪增福男, 1960 年生于安徽合肥. 教授, 研究方向为计算机视听觉, 智能机器人.

E-mail: zfwang@ustc.edu.cn