



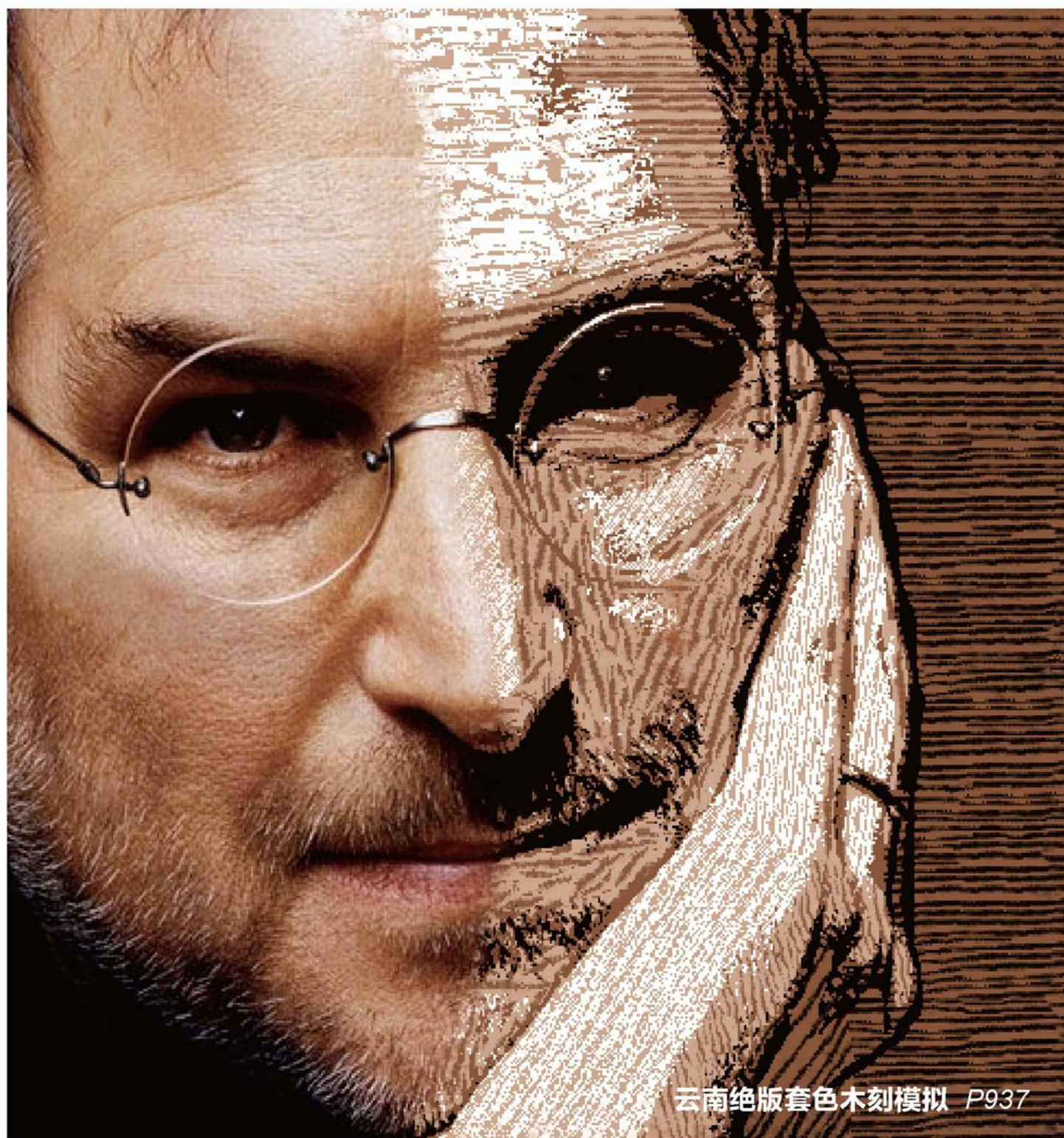
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2015
07
VOL.20

ISSN1006-8961
CN11-3758/TB



云南绝版套色木刻模拟 P937



第20卷第7期（总第230期）

2015年7月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院遥感与数字地球研究所

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 李小文

编辑出版《中国图象图形学报》编辑出版委员会

邮政信箱 北京9718信箱

邮 编 100101

电子信箱 jig@radi.ac.cn

电 话 010-64807995

网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

（邮政信箱：北京399信箱 邮编：100048）

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen

Editor, Publisher Editorial and Publishing Board of Journal of

Image and Graphics

P.O.Box 9718, Beijing, P.R.China

Zip code 100101

E-mail jig@radi.ac.cn

Telephone 010-64807995

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation

(P.O.Box 399, Beijing 100048, P.R.China))

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

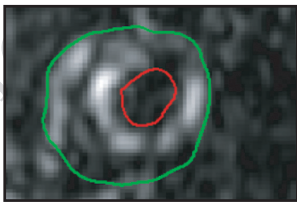
ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

国内定价 50.00元



基于颈动脉粥样硬化斑块的多
序列MR图像去噪与配准
(第0871页)



光滑拟合模型的IIPBDR图像
修复算法(第0895页)



基于暗原色先验的图像快速去
雾(第0914页)



个性化定制的虚拟健身系统设
计与实现(第0953页)

图像处理和编码

融合视觉感知特性的HEVC率失真优化

费马燕, 彭宗举, 李持航, 陈芬, 郁梅, 蒋刚毅 0857

结合空间信息的模糊聚类侧扫声纳图像分割

李阳, 庞永杰, 盛明伟 0865

基于颈动脉粥样硬化斑块的多序列MR图像去噪与配准

李军伟, 汪晓妍, 张剑华, 何俊丽, 管秋, 陈胜勇 0871

小波变换与纹理合成相结合的图像修复

张东, 唐向宏, 张少鹏, 黄俊泽 0882

光滑拟合模型的IIPBDR图像修复算法

鱼江, 翟东海, 聂洪玉, 李同亮 0895

图像分析和识别

噪声方差未知条件下的视频目标跟踪

陈金广, 任冰青, 马丽丽, 温静 0906

基于暗原色先验的图像快速去雾

曾浩, 尚媛园, 丁辉, 周修庄, 付小雁 0914

人体运动分析的实例学习方法

李毅, 刘兴川, 孙亭 0922

图像理解和计算机视觉

多参数加权的无缝纹理映射算法

刘彬, 陈向宁, 薛俊诗 0929

交互式云南绝版套色木刻数字化合成

王雪松, 李杰, 侯剑侠, 徐丹, 普园媛 0937

计算机图形学

基于像素的多边形等距区域子分算法

严志刚, 寿华好 0945

个性化定制的虚拟健身系统设计与实现

赵汉理, 孟庆如, 韩丽贞, 潘志庚 0953

遥感图像处理

LANDSAT-8卫星影像长条带数据处理

李亚林, 李山山, 冯钟葵, 张洪群, 陈勃 0963

MapReduce模式下高光谱图像端元提取算法加速

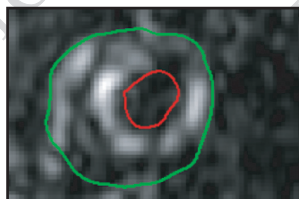
曲海成, 籍瑞庆, 刘万军, 梁雪剑 0973

基于GA-BP神经网络模型的登革热时空扩散模拟

李卫红, 陈业滨, 闻磊 0981

CONTENTS

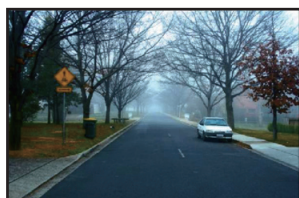
JOURNAL OF IMAGE AND GRAPHICS



Multiple sequence MR image de-noising and registration method based on carotid atherosclerosis plaque (P0871)



Improved IIPBDR algorithm based on smooth-fitting model (P0895)



Fast image haze removal base on dark channel prior (P0914)



Design and implementation of the personalized virtual fitness system(P0953)

Image Processing and Coding

- Human perception features based rate distortion optimization for HEVC
Fei Mayan, Peng Zongju, Li Chihang, Chen Fen, Yu Mei, Jiang Gangyi 0857
- Side-Scan sonar image segmentation via fuzzy clustering with spatial constrains
Li Yang, Pang Yongjie, Sheng Mingwei 0865
- Multiple sequence MR image de-noising and registration method based on carotid atherosclerosis plaque
Li Junwei, Wang Xiaoyan, Zhang Jianhua, He Junli, Guan Qiu, Chen Shengyong 0871
- Image inpainting based on combination of wavelet transform and texture synthesis
Zhang Dong, Tang Xianghong, Zhang Shaopeng, Huang Junze 0882
- Improved IIPBDR algorithm based on smooth-fitting model
Yu Jiang, Zhai Donghai, Nie Hongyu, Li Tongliang 0895

Image Analysis and Recognition

- Video target tracking algorithm with the noise variance unknown
Chen Jinguang, Ren Bingqing, Ma Lili, Wen Jing 0906
- Fast image haze removal base on dark channel prior
Zeng Hao, Shang Yuanyuan, Ding Hui, Zhou Xiuzhuang, Fu Xiaoyan 0914
- Example-based approach for human motion analysis from videos
Li Yi, Liu Xingchuan, Sun Ting 0922

Image Understanding and Computer Vision

- Seamless texture mapping algorithm based on multi parameter weighted
Liu Bin, Chen Xiangning, Xue Junshi 0929
- Interactive digital synthesis of Yunnan out-of-print woodcuts
Wang Xuesong, Li Jie , Hou Jianxia, Xu Dan, Pu Yuanyuan 0937

Computer Graphics

- Pixel based domain subdivision algorithm for polygon offsetting
YAN Zhigang, SHOU Huahao 0945
- Design and implementation of the personalized virtual fitness system
Zhao Hanli, Meng Qingru, Han Lizhen, Pan Zhigeng 0953

Remote Sensing Image Processing

- Study on processing method of LANDSAT-8 long strip satellite imagery
Li Yalin, Li Shanshan, Feng Zhongkui, Zhang Hongqun, Chen Bo 0963
- Acceleration of hyperspectral image endmember extraction based on MapReduce pattern
Qu Haicheng, Ji Ruiqing, Liu Wanjun, Liang Xuejian 0973
- Simulation of spatio-temporal diffusion of dengue fever based on the GA-BP neural network model
Li Weihong, Chen Yebin, Wen Lei 0981

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2015)07-0922-07

论文引用格式: Li Y, Liu X C, Sun T. Example-based approach for human motion analysis from videos[J]. Journal of Image and Graphics, 2015, 20(7): 0922-0928. [李毅, 刘兴川, 孙亭. 人体运动分析的实例学习方法[J]. 中国图象图形学报, 2015, 20(7): 0922-0928.] [DOI: 10. 11834/jig. 20150708]

人体运动分析的实例学习方法

李毅, 刘兴川, 孙亭

中国电子科技集团公司第二十八研究所, 南京 210007

摘要: 目的 面向实时、准确、鲁棒的人体运动分析应用需求,从运动分析的特征提取和运动建模问题出发,本文人体运动分析的实例学习方法。**方法** 在构建人体姿态实例库基础上,首先,采用运动检测方法得到视频每帧的人体轮廓;其次,基于形状上下文轮廓匹配方法,从实例库中检索得到每帧视频的候选姿态集;最后,通过统计建模和转移概率建模实现人体运动分析。**结果** 对步行、跑步、跳跃等测试视频进行实验,基于轮廓的形状上下文特征表示和匹配方法具有良好的表达能力;本文方法运动分析结果,关节夹角平均误差在 5° 左右,与其他算法相比,有效提高了运动分析的精度。**结论** 本文人体运动分析的实例学习方法,能有效分析单目视频中的人体运动,并克服了映射的深度歧义,对运动的视角变化鲁棒,具有良好的计算效率和精度。

关键词: 运动分析;实例学习;形状上下文;统计建模

Example-based approach for human motion analysis from videos

Li Yi, Liu Xingchuan, Sun Ting

The 28th Research Institute of China Electronics Technology Group Corporation, Nanjing 210007, China

Abstract: **Objective** Human motion tracking from monocular image sequences is a challenging work in computer vision. It also has many penitential applications, such as human computer interface, computer animation, and intelligent video surveillance. Methodologies of example-based human motion tracking from monocular camera are explored in this study to meet timeliness, accuracy, and reliability requirements of human motion tracking for real applications. We focus on two main aspects: visual feature extraction and human motion modeling. **Method** Based on an example database that is constructed with motion capture data, an example-based approach for human pose estimation from monocular image sequences is proposed. First, we use a motion detection method to extract human region from images. Then, an edge-tracking method is used to detect human silhouette from human region. Second, shape context is used to describe the human silhouette detected from video frames, and candidate poses are obtained from the example database by silhouette matching. Third, we build probability and statistical model of human motion and conducted pose estimation from these candidates. **Result** Experimental results on walking, running, and jumping videos demonstrate that shape context-based silhouette representation and matching method can effectively extract human visual feature from image. Compared with other methods, the proposed method can tackle orientation ambiguity problem effectively. Moreover, it is invariant to viewpoints. **Conclusion** In this paper, we proposed an example-based method for human motion analysis. Shape context is used for visual feature extraction and matching. Probability and statistical model of motion are used for pose estimation. Experimental results on different

收稿日期:2014-12-18;修回日期:2015-03-20

基金项目:“十二五”装备预先研究项目(51306030101);江苏省自然科学基金青年项目(20140078, 20140079)

第一作者简介:李毅(1984—),男,工程师,2013年于南京大学获计算机应用技术专业博士学位,主要研究方向为计算机视觉、计算机图形学。E-mail: njulanty@gmail.com

Supported by: Project supported by the Natural Science Foundation of Jiangsu Province, China (20140078, 20140079)

types of motion demonstrate that the proposed method can analyze 3D pose from videos effectively, thereby increasing the efficiency and accuracy of human motion analysis. Moreover, the proposed method can solve the orientation ambiguity problem, and is invariant to viewpoints.

Key words: motion analysis; example-based approach; shape context; statistical modeling

0 引言

基于视频的人体运动分析在视频监控、人机交互、计算机动画等领域有着重要的应用,近年来对它的研究得到了广泛的关注^[1-2]。但复杂条件下图像特征提取、2维图像特征到3维人体姿态映射的多义性、人体运动模式的复杂性等因素,使得该问题研究面临极大挑战。

当前,基于视频的人体运动分析方法主要包括生成式(Generative)方法和判别式(Discriminative)方法。生成式方法^[3]利用预定义的人体模型,通过优化人体模型投影和视觉输入的相似度函数实现人体运动分析。尽管分析结果精确,但该方法需要人工初始化且计算复杂度高,难以满足实际应用的实时性要求。相反,判别式方法^[4]不显式定义人体模型,而通过训练数据的学习建立2维视觉输入到3维人体姿态之间的直接映射。通常,判别式方法的计算开销小于生成式方法,且无需姿态初始化,具有更广泛的应用价值。实例学习方法^[5]是判别式方法的一种,其通过在姿态实例库中检索与输入图像中的视觉特征最相似的姿态,并采用插值等方法得到运动分析结果。由于实例学习方法实现简单且实时性好,因而得到研究者的重视。然而,鲁棒图像特征提取和映射多义性引起的姿态误判是实例学习方法研究面临的主要难题。

如何有效提取和表示视频中的人体特征是实例学习方法面临的首要难题。当前,使用较多的图像特征包括颜色^[6]、边缘、剪影^[7]等。其中剪影特征因提取简单且不受服饰颜色影响而得到广泛使用。Poppe^[8]对剪影特征及其描述方法进行了对比研究,比较分析了形状上下文方法、Fourier描述子、Hu不变矩在人体上半身姿态库上的运动分析结果,并探讨了动作类型、人体模型、视角变化对运动分析结果的影响。另外,多特征融合方法研究也得到重视^[4]。尽管研究者开展了大量工作,但人体特征受到光线变化、背景噪音等多种因素影响,使得视频中

人体特征的提取和表示仍然是研究面临的重要难题。

如何使用运动时序信息提高运动分析的精度是实例学习方法研究的另一个难题。在已有工作中,Andreas等人^[9]基于动态森林模型提出一种改进的非参数化非线性的Markov模型,并从理论分析和实验角度证明了其建模能力优于Markov模型;Andreas等人^[10]将非参数化的稀疏表示的贝叶斯网络模型用于描述复杂人体运动,该模型对非训练数据集中的数据具有泛化表示能力。以上方法对如何利用运动时序信息进行了有效探索,但由于人体运动的复杂性使得人体运动分析仍极具挑战。如何有效对运动时序信息进行表示和建模,还有待更多的研究。

针对实例学习方法的以上难题,提出一种基于形状上下文描述^[11]和运动建模^[12]的人体运动分析的实例学习方法。该方法的特点表现在:首先,基于相关系数方法^[13]提取人体轮廓,并用改进的形状上下文描述子进行轮廓特征描述和姿态检索,具有良好计算效率;其次,采用加权的Markov方法建立姿态间的转移概率模型和姿态统计模型,有效利用了运动的时序信息,提高了运动分析精度。本文方法能有效克服映射的深度歧义,对视角变化鲁棒,能准确分析单目视频中的人体运动。

1 人体运动分析的实例学习方法

基于实例学习的视频人体运动分析方法流程如图1所示。该方法的核心思想是从人体姿态库中检索与输入视频中人体最相似的姿态,并利用时空约束实现运动分析,包括以下步骤:首先,通过运动检测提取输入视频中的人体轮廓;其次,利用人体运动捕获数据构造人体姿态库和人体轮廓库,使用形状上下文描述子抽取轮廓特征,并为视频每帧建立姿态候选集;最后,建立转移概率模型和姿态统计建模,实现运动分析。

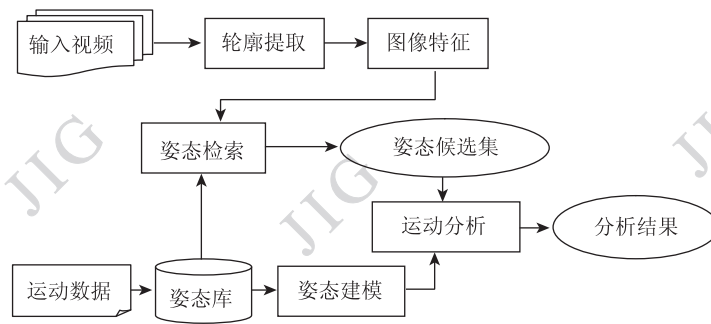


图1 人体运动分析的实例学习方法框架

Fig. 1 Flowchart of the method

1.1 视频轮廓提取和采样

视频中的人体特征提取是运动分析的基础,但光照、阴影、背景变化等因素使得人体特征提取困难。本文采用了一种基于相关系数的方法提取人体运动区域。基于相关系数的人体运动区域提取公式为

$$R(I_t(x, y), I_{t-1}(x, y)) = \frac{[m[W(I_t(x, y))W(I_{t-1}(x, y))] - m[W(I_t(x, y))]m[W(I_{t-1}(x, y))]]}{\sqrt{v[W(I_t(x, y))]v[W(I_{t-1}(x, y))])}} \quad (1)$$

$$I_t^M(x, y) = \begin{cases} 1 & R(I_t(x, y), I_{t-1}(x, y)) < \sigma_s \\ 0 & \text{其他} \end{cases} \quad (2)$$

式中, $R(I_t(x, y), I_{t-1}(x, y))$ 表示连续两帧图像对应像素点的相似性,其通过计算对应像素点在 W 窗口中的点的相关系数得到, $m[W]$ 和 $v[W]$ 分别表示点 (x, y) 在窗口 W 邻域中的点的均值和方差。通过设置一个阈值 σ_s , 当相似度小于阈值 σ_s 时, 判为前景, 否则为背景。 $I_t^M(x, y)$ 为运动检测结果。该方法能有效克服背景中的小的变动, 抑制阴影。

运动检测得到的人体区域可能存在空洞和背景噪音点等情况, 本文通过噪音点去除, 形态学运算等方法对运动检测图像进行后处理。对检测得到的人体区域, 通过 8 邻域边界跟踪方法得到轮廓点, 采用等间距采样方法对轮廓点进行采样, 结果如图 2 所示。

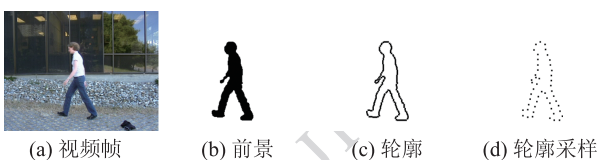


图2 视频中人体轮廓提取和采样

Fig. 2 Silhouette and sampling points of human

1.2 形状上下文计算与匹配

形状上下文是基于轮廓的形状特征提取方法, 其基本思想是: 对于给定的一个形状, 对其轮廓进行采样得到一组离散的点集 $P = \{p_i | i = 1, 2, \dots, n\}$ 。以其中任意一点 p_i 为参考点, 计算 p_i 相对于其他 $n-1$ 个点的极对数直方图, 即

$$h_{p_i}(l) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(l)\} \quad (3)$$

式中, $h_{p_i}(l)$ 为点 p_i 的极对数直方图, $\#\{\}$ 表示计数, $(q - p_i) \in \text{bin}(l)$ 表示相对于 p_i , 点 q 属于形状直方图的第 l 个分量。 $l = \phi r$ 为方向参数 ϕ 和距离参数 r 的乘积。对整个点集 P , 分别计算其他所有点的极对数直方图, 构成轮廓的形状上下文描述。

模板半径 R 对形状上下文的描述能力具有较大影响, 采用模板半径为

$$R = \alpha \bar{R} = \alpha \frac{1}{n(n-1)} \sum_{\Omega} \|\vec{p_i p_j}\| \quad (4)$$

$$\Omega: \{p_i, p_j \in P, i \neq j\}$$

式中, $\bar{R}$ 为归一化半径, α 为比例因子。 α 越小, 局部性描述越强, α 越大, 全局性描述越强。

对于两个图像轮廓点集 $P = \{p_i | i = 1, \dots, n\}$ 和 $Q = \{q_i | i = 1, \dots, n\}$, 两个点 p_i 和 q_i 的距离定义为

$$c(p_i, q_j) = \chi^2(h_{p_i}, h_{q_j}) = \frac{1}{2} \sum_{l=1}^L \frac{[h_{p_i}(l) - h_{q_j}(l)]^2}{h_{p_i}(l) + h_{q_j}(l)} \quad (5)$$

采用二分图匹配算法计算轮廓间相似度的算法复杂度为 $O(n^3)$, 本文在不降低匹配精度的情况下采用平均 Hausdorff 距离 (MHD) 方法计算形状上下文相似度, 使算法复杂度降为 $O(n^2)$, 提高了匹配效率, 具体如下:

定义点集 P 上的点 p 到点集 Q 的相似度为 C_{pQ} , 最小匹配相似度为 $C(p, Q)$ 、 $C(q, P)$, 即

$$\begin{aligned} C(p, \mathbf{Q}) &= \min_{q \in \mathbf{Q}} C_{pq} = \min_{q \in \mathbf{Q}} c(p, q) \\ C(q, \mathbf{P}) &= \min_{p \in \mathbf{P}} C_{qp} = \min_{p \in \mathbf{P}} c(q, p) \end{aligned} \quad (6)$$

则轮廓 $\mathbf{P}$ 和 $\mathbf{Q}$ 的相似度为

$$D(\mathbf{P}, \mathbf{Q}) = \max(C_{PQ}, C_{QP}) \quad (7)$$

式中, C_{PQ} 和 C_{QP} 为轮廓点集的有向匹配相似度, 即

$$\begin{aligned} C_{PQ} &= \frac{1}{n} \sum_{p \in \mathbf{P}} C(p, \mathbf{Q}) \\ C_{QP} &= \frac{1}{n} \sum_{q \in \mathbf{Q}} C(q, \mathbf{P}) \end{aligned} \quad (8)$$

通过形状上下文描述图像中人体轮廓特征, 采用简化的轮廓匹配方法进行姿态检索, 有效提高了姿态检索的效率。

1.3 概率转移和姿态统计建模

采用的人体模型包括 22 个关节, 人体姿态用人体根关节的位置和朝向, 以及其他 21 个关节的朝向表示。人体运动由一系列人体姿态组成, 表示为 $\mathbf{P}_i = (p_i^{(1)}, r_i^{(1)}, r_i^{(2)}, \dots, r_i^{(n)}, \dots, r_i^{(22)})$ 。其中 $p_i^{(1)} \in \mathbf{R}^3$ 和 $r_i^{(1)} \in \mathbf{R}^3$ 表示根关节的位置和方向, $r_i^{(n)} \in \mathbf{R}^3$, $n=2, 3, \dots, 22$ 表示第 n 个关节的朝向。

用人体姿态数据驱动人体模型得到对应姿态的姿态模型, 基于轮廓特征描述和表示方法通过数据库查找, 为输入视频的每帧建立了规模为 K 的姿态候选集。对于查询轮廓 $\mathbf{S}_i$ 检索得到的 K 个最近邻姿态构成的姿态候选集, 可表示为 $\{\mathbf{S}_{i1}, \mathbf{S}_{i2}, \dots, \mathbf{S}_{iK}\}$, $\mathbf{S}_{ik}$, $k=1, \dots, K$ 为 $\mathbf{S}_i$ 对应的一个候选姿态, k 为候选姿态序号。

在姿态候选集构建基础之上, 通过概率转移建模和姿态统计建模实现人体运动分析。

1.3.1 概率转移建模

利用运动的时序信息, 同时考虑轮廓相似度对各候选姿态的影响, 采用带权值的一阶 Markov 过程来约束运动分析, 建立人体运动的概率转移模型。

首先, 定义两个姿态 $\mathbf{P}_i$ 和 $\mathbf{P}_j$ 之间的转移概率为 $P_{ij} = \exp(-\text{Dis}(\mathbf{P}_i, \mathbf{P}_j)/K_p)$, $\text{Dis}(\mathbf{P}_i, \mathbf{P}_j)$ 为两个姿态 $\mathbf{P}_i$ 和 $\mathbf{P}_j$ 之间的距离, 定义为

$$\text{Dis}(\mathbf{P}_i, \mathbf{P}_j) = \sum_{n=1}^{22} w_n \|\log((r_i^n)^{-1}(r_j^n))\|^2 \quad (9)$$

式中, w_n 为各个关节的权重。对关键关节如根关节、膝关节、肘关节、跨关节、头肩关节设权重为 1, 其他关节权重为 0。

其次, 为每个候选姿态 $\mathbf{S}_{ik}$ 设置一个权重 $w_{ik} = \exp(-D(\mathbf{S}_i, \mathbf{S}_{ik})/K_s)$ 。

最后, 综合考虑转移概率和轮廓相似度, 从候选姿态中选择姿态, 姿态选择依据为

$$k = \arg \min(w_{ik} P_{(i-1)(ik)}, k) \quad (10)$$

式中, $P_{(i-1)(ik)}$ 表示的是姿态 $\mathbf{P}_{i-1}$ 到轮廓 $\mathbf{S}_{ik}$ 对应的姿态 $\mathbf{P}_{ik}$ 的转移概率。

概率转移建模综合利用了运动的时序信息和轮廓相似度, 可有效约束运动分析过程。

1.3.2 姿态统计建模

随着运动数据库规模的增大, 概率转移变得复杂, 时间开销增大。为此采用聚类分析的方法, 建立运动数据库中运动数据的姿态统计模型。

首先, 采用 C 均值方法对人体姿态进行聚类, 得到人体姿态的 C 个类, 分别记为 $C_1, C_2, \dots, C_c$ 。

其次, 对于类 C_i 中的某个姿态 $\mathbf{P}_i$, 查询其在原始运动序列中的后继帧 $\mathbf{P}_{i+1}$, 如果 $\mathbf{P}_{i+1}$ 不属于类 C_i 而属于类 C_j , 则将类 C_j 作为类 C_i 的子节点。对类 C_i 中的每个姿态进行同样的操作可以得到以类 C_i 为根节点的聚类树。

最后, 对聚类分析得到的所有类 $C_1, C_2, \dots, C_c$ 构建以其为根节点的聚类树, 这些聚类树构成了聚类森林。

1.3.3 运动分析过程

综上所述, 本文基于概率转移模型和姿态统计模型的运动分析过程归结为:

1) 对姿态库中的姿态建立概率转移模型和姿态统计模型。

2) 已知当前帧姿态, 对其后继帧重构分 3 步: 首先, 通过轮廓匹配得到后继帧的姿态候选集; 其次, 检索后继帧的姿态候选集, 根据姿态统计建模结果, 通过查询当前帧段的聚类树的子节点信息, 剔除候选集中的不属于子节点类的候选数据; 最后, 按照概率转移模型以加权概率最大的候选姿态作为运动分析结果。

3) 视频首帧通过最近邻方法得到。

2 实验结果与分析

采用卡耐基梅隆大学 (CMU) 的运动捕获数据, 使用 Poser 运动合成工具合成人体的 3 维运动, 构建运动姿态库和人体轮廓库。库中包含了人体的步行、跑步和跳跃运动, 每种运动包括了不同的视角和不同的运动模式, 训练和测试数据具体如表 1 所示。测试视频采用了步行运动数据^[14], 以及室外采集的

跑步和跳跃运动。实验中,运动检测的算法参数 σ_s 取值为0.95,窗口 W 大小为 5×5 。

表1 训练数据及测试数据集
Table 1 Training and test dataset

运动姿态	帧数	
	训练数据	测试数据
步行	2 808	304
跑步	4 013	139
跳跃	2 969	350

2.1 特征提取与描述

实验研究分析了不同半径和采样点数对形状上下文方法的影响,并比较了形状上下文(SC),傅里叶描述子(FD)和Hu矩的运动分析精度。

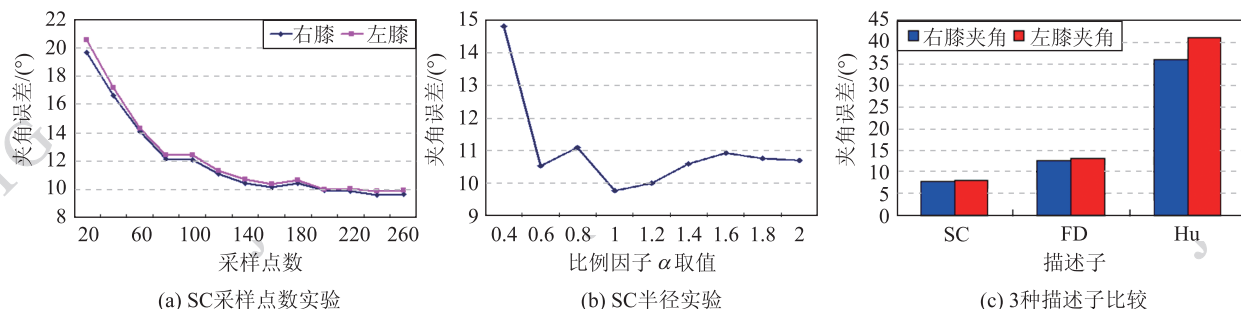


图3 特征表示对比实验

Fig. 3 Experimental results of feature

2.2 运动分析结果

实验比较了LWR方法^[15]、最近邻方法(NN)、Markov方法^[5]以及本文方法对步行、跑步、跳跃运动的分析结果,并给出了不同方法的分析误差。

图4以步行和跳跃动作为例,对比了几种方法对右膝、左膝夹角分析的结果。图4(a)是对人体步行右膝夹角的分析结果,图中黑色虚线和绿色虚线分别表示了检测视频中人体真实的右膝和左膝夹角参数。黄色曲线表示的最近邻方法在第38、90、154帧将右膝夹角分析成了左膝夹角,这主要是由深度歧义引起的;紫色表示的LWR方法由于误差累积,使得在个别帧的运动分析误差提高;Markov方法克服了映射的深度歧义,但误差较大;红色曲线表示的是采用本文方法进行运动分析的结果,其不仅与人体真实右膝夹角参数最接近,而且更平滑,表明了本文方法的有效性。图4(b)是对人体跳跃左膝夹角的分析结果,同样表明了本文方法更准确。图4(b)

实验在跑步视频上进行,采用10近邻方法,即最匹配的10个轮廓对应的姿态参数的均值作为分析结果。

1)图3(a)是在 $\alpha=1$ 时不同采样点的实验。由图3(a)可见,运动分析误差随着采样点数的增而降低,并在采样点数180处逐渐收敛。但随着采样点个数增加,运动分析的计算开销增大,因此本文取采样点数为200。

2)图3(b)是采样点数为200时,不同半径参数的检索结果。由图3(b)可见,当 $\alpha=1$ 时检索误差最低,因此本文取 $\alpha=1$,即平均距离 $\bar{R}$ 作为SC的半径。

3)图3(c)是3种描述子的分析结果,其中SC和FD的采样点数都为200。由图3(c)可见,SC方法的误差最低,FD次之,Hu方法误差最高。

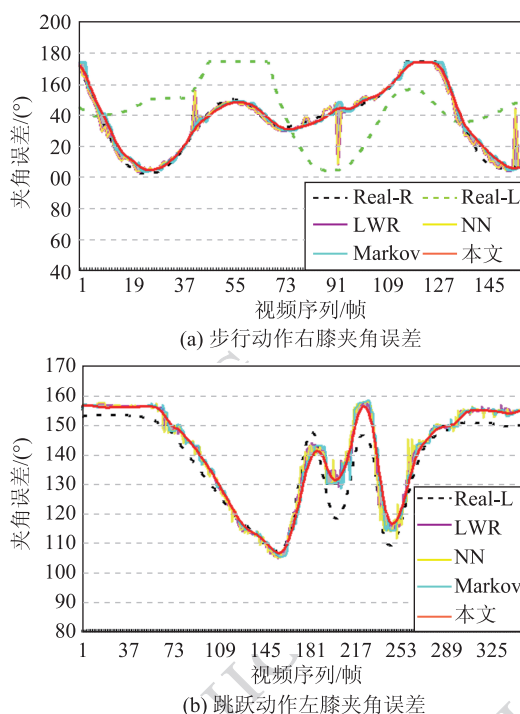


图4 运动分析结果

Fig. 4 Results of motion analysis

中第200帧左右姿态参数与真实数据的偏离主要是由于姿态库不完备性造成的。另外,方法对步行、跑步、跳跃运动的有效性表明了本文方法可以处理视频中不同类型的人体运动。

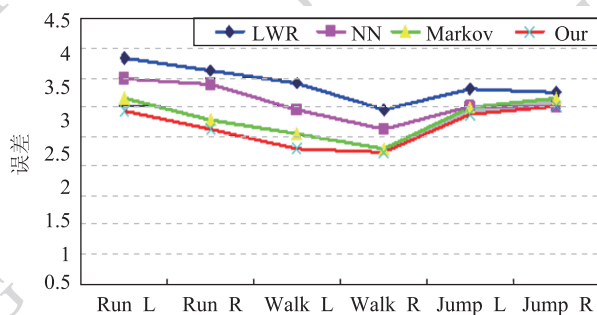


图5 运动分析方法比较

Fig. 5 Comparison of different methods

图5给出了最近邻(NN)方法、LWR方法、Markov方法与本文方法对步行、跑步、跳跃运动分析的误差。由图5可见,本文基于概率转移和姿态统计建模方法的平均误差最低;跑步、步行中左膝比右膝误差高,这主要是因为本文实验采用的是人体侧面由左向右步行和跑步运动所致,这些运动中左膝被遮挡。

2.3 视频测试结果

图6给出了本文方法对单目视频步行、跑步和跳跃运动的分析结果。视频在自然场景中采集,跑步运动包含了人体的转圈运动。由图6可见,本文方法能有效分析单目视频中的人体运动,转向跑步运动的实验结果表明了本文方法对视角变化的鲁棒性。



图6 步行、跑步、跳跃运动分析结果

Fig. 6 Motion analysis results on walking, running, jumping videos

3 结论

针对单目视频人体运动分析,提出一种人体运动分析的实例学习方法。在利用运动捕获数据构建人体姿态实例库的基础上,采用运动检测方法得到视频每帧的人体轮廓,基于形状上下文轮廓匹配方法,从实例库中检索得到每帧视频的候选姿态集,通

过统计建模和转移概率建模实现人体运动分析。该方法能有效克服映射的深度歧义,对视角变化鲁棒,能准确分析单目视频中人体运动,具有良好的效率和精度。

姿态实例库是实例学习方法的基础,实例库中包含的运动种类和数量直接影响了运动分析的结果。构建一个包含各类运动的实例库相当困难,因而针对特定运动构建实例库,研究不同实例库间的

选择和切换机制是可行的研究方向。

参考文献 (References)

- [1] Moeslund T B, Hilton A, Kruger V. A survey of advances in vision-based human motion capture and analysis [J]. Computer Vision and Image Understanding, 2006, 104(2): 90-126.
- [2] Poppe R. Vision-based human motion analysis: an overview [J]. Computer Vision and Image Understanding, 2007, 108(1): 4-18.
- [3] Li Y, Sun Z X. Generative tracking of 3D human motion in latent space by sequential clonal selection algorithm [J]. Multimedia Tools and Applications, 2014, 69(1): 79-109.
- [4] Zhao X, Fu Y, Ning H Z, et al. Human pose regression through multi-view visual fusion [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(7): 957-966.
- [5] Howe N R. Silhouette lookup for monocular 3D pose tracking [J]. Image and Vision Computing, 2007, 25(3): 331-341.
- [6] Sapp B, Toshev A, Taskar B. Cascaded models for articulated pose estimation [C]//Proceedings of the 11th European Conference on Computer Vision. Heraklion, Greece: IEEE Press, 2010: 406-420.
- [7] Ukita N. Articulated pose estimation with parts connectivity using discriminative local oriented contours [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE Press, 2012: 3154-3161.
- [8] Poppe R, Poel M. Comparison of silhouette shape descriptors for example-based human pose recovery [C]//Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition. Southampton, United Kingdom: IEEE Press, 2006: 541-546.
- [9] Andreas M L, Peter V G, Sebastian N. Efficient nonlinear markov models for human motion [C]//Proceedings of the IEEE International Conference on Computer vision and Pattern Recognition. Columbus, Ohio: IEEE Press: 2014, 220-228.
- [10] Andreas M L, Peter V G, Sebastian N. A non-parametric bayesian network prior of human pose [C]//Proceedings of the IEEE International Conference on Computer Vision. Sydney, Australia: IEEE Press, 2013: 1281-1288.
- [11] Mori G, Malik J. Recovering 3D human body configurations using shape contexts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(7): 1052-1062.
- [12] Wang R R, Qiu X J, Wang W Z, et al. A video-driven approach to 3D human animation synthesis [J]. Journal of System Simulation, 2007, 19(8): 1700-1704. [王璐璐, 邱显杰, 王文中, 等. 一种视频驱动的三维人体动画合成方法[J]. 系统仿真学报, 2007, 19(8): 1700-1704.]
- [13] Li Y, Sun Z X, Yuan B, et al. An improved method for motion detection by frame difference and background subtraction [J]. Journal of Image and Graphics, 2009, 14(6): 1162-1168. [李毅, 孙正兴, 远博, 等. 一种改进的帧差和背景减相结合的运动检测方法[J]. 中国图象图形学报, 2009, 14(6), 1162-1168.] [DOI:10.11834/jig.20090624]
- [14] Ormonieit D, Sidenbladh H, Michael J, et al. Learning and tracking cyclic human motion [C]//Proceedings of Advances in Neural Information Processing Systems. Vancouver, Canada: MIT Press, 2001: 894-900.
- [15] Shakhnarovich G, Viola P, Darrell T. Fast pose estimation with parameter-sensitive hashing [C]//Proceedings of the IEEE International Conference on Computer Vision. Nice, France: IEEE Press, 2003: 750-759.