



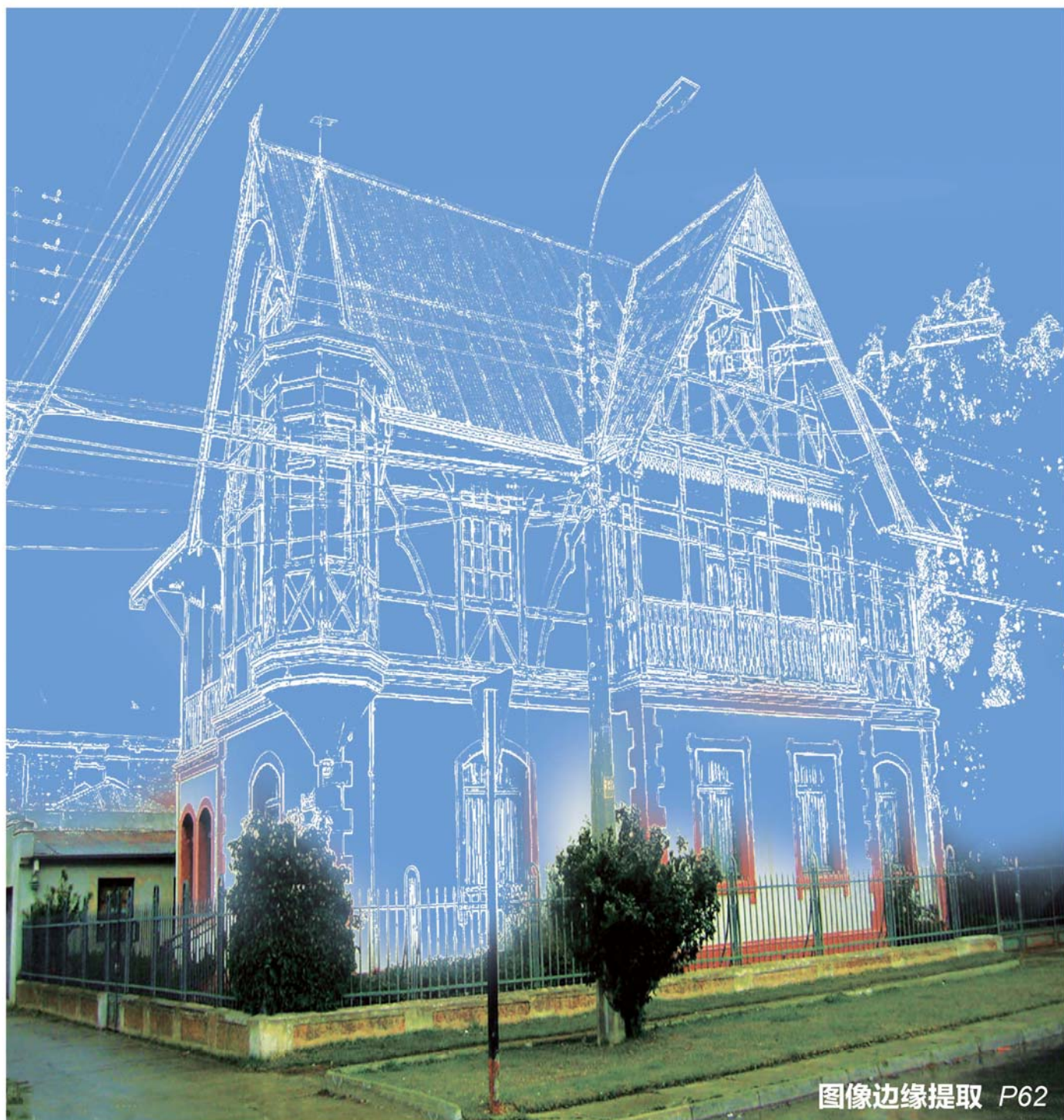
JOURNAL OF IMAGE AND GRAPHICS

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

2014
01
VOL.19

ISSN1006-8961
CN11-3758/TB



图像边缘提取 P62



第19卷第1期（总第213期）

2014年1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

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

2013 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@irsa.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@irsa.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

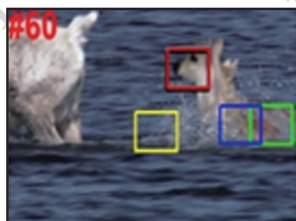
ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

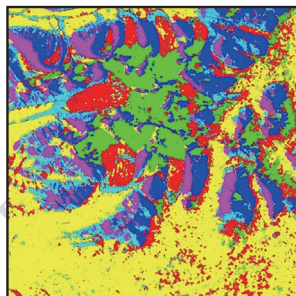
国内定价 50.00元



基于 L_2 范数最小化的实时目标跟踪(第36页)



数字地球渲染中单精度浮点数误差改正(第119页)



全极化合成孔径雷达影像地形纠正及其在雪冰制图中的应用(第150页)

图像处理和编码

图像阈值化的自适应粗糙熵方法

吴涛 0001

利用局部极值点特性的多聚焦图像快速融合

杨达, 王孝通, 徐冠雷, 吴俊波 0011

摄像机位姿的高精度快速求解

李书杰, 刘晓平 0020

整数变换实现大容量图像无损信息隐藏

邱应强 0028

图像分析和识别

基于 L_2 范数最小化的实时目标跟踪

齐美彬, 杨勋, 杨艳芳, 陆磊, 蒋建国 0036

利用七宫格的遮挡车辆凹性检测与分割

刘万军, 陈虹宇, 姚雪, 曲海成 0045

基于体素的人手结构自动估测

姚争为, 潘志庚 0054

结合高斯加权距离图的图像边缘提取

贾迪, 孟祥福, 孟琰, 董娜, 方金凤 0062

基于标定的任意半径柱面上2维条形码畸变校正

王伟, 何卫平, 郭改放, 牛晋波, 曹西征 0069

面向惯性约束聚变实验靶图像的快速椭圆检测

吴倩, 邹伟, 徐德, 张峰 0076

融合MBP和EPMOD的人脸识别

闫海婷, 王玲, 李昆明, 刘机福 0085

遮挡目标的分片跟踪处理

张彦超, 许宏丽 0092

图像理解和计算机视觉

多尺度分析的运动注意力计算

刘龙, 樊波阳 0101

多尺度空间判别性概率潜在语义分析的场景分类

季海峰, 高隽, 郑鹏, 王婧 0109

计算机图形学

数字地球渲染中单精度浮点数误差改正

李尚林, 郑利平, 张静, 杨柳青 0119

医学图像处理

小波与双边滤波的医学超声图像去噪

张聚, 王陈, 程芸 0126

Contourlet变换系数加权的医学图像融合

张鑫, 陈伟斌 0133

遥感图像处理

滑坡高分辨率遥感多维解译方法及其应用

鲁学军, 史振春, 尚伟涛, 周和颐 0141

全极化合成孔径雷达影像地形纠正及其在雪冰制图中的应用

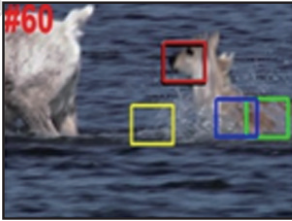
符思涛, 李震, 田帮森 0150

结合邻域相关影像与最大相关性最小冗余性特征选择的面向对象变化检测

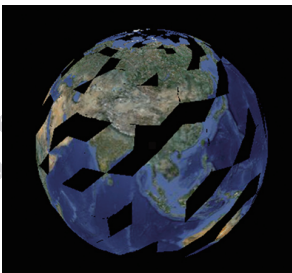
邹利东, 潘耀忠, 朱文泉, 周公器, 李宜展 0158

CONTENTS

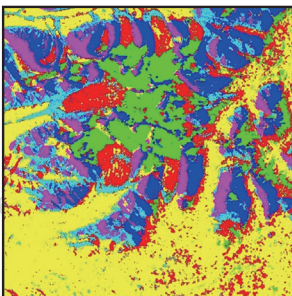
JOURNAL OF IMAGE AND GRAPHICS



Real-time object tracking based on L_2 -norm minimization(P36)



Correction method of single-precision floating-point error for digital Earth rendering(P119)



Topographic correction of polarimetric synthetic aperture radar and its application in snow and ice mapping(P150)

Image Processing and Coding

- Adaptive rough entropy method for image thresholding 0001
Wu Tao
- Multi-focus image fusion based on properties of local extrema 0011
Yang Da, Wang Xiaotong, Xu Guanlei, Wu Junbo
- An accurate and fast algorithm for camera pose estimation 0020
Li Shujie, Liu Xiaoping
- Large capacity lossless information hiding in images using integer transform 0028
Qiu Yingqiang

Image Analysis and Recognition

- Real-time object tracking based on L_2 -norm minimization 0036
Qi Meibin, Yang Xun, Yang Yanfang, Lu Lei, Jiang Jianguo
- Detection and segmentation of occluded vehicles based on seven grids 0045
Liu Wanjun, Chen Hongyu, Yao Xue, Qu Haicheng
- Automatic calibration of hand structures based on voxel data 0054
Yao Zhengwei, Pan Zhigeng
- Image edge extraction combined with a Gaussian weighted distance graph 0062
Jia Di, Meng Xiangfu, Meng Lu, Dong Na, Fang Jinfeng
- Restoration of distorted 2D barcode marked on arbitrary radius cylinder based on calibrated camera 0069
Wang Wei, He Weiping, Guo Gaifang, Niu Jinbo, Cao Xizheng
- Fast ellipse detection for target images in ICF experiment 0076
Wu Qian, Zou Wei, Xu De, Zhang Feng
- Face recognition by fusing MBP and EPMOD 0085
Yan Haiting, Wang Ling, Li Kunming, Liu Jifu
- Fragments tracking under occluded target 0092
Zhang Yanchao, Xu Hongli

Image Understanding and Computer Vision

- Multi-scale analysis based motion attention computation 0101
Liu Long, Fan Boyang
- Scene classification based on multi-scale spatial discriminative probabilistic latent semantic analysis 0109
Ji Haifeng, Gao Jun, Zheng Peng, Wang Jing

Computer Graphics

- Correction method of single-precision floating-point error for digital Earth rendering 0119
Li Shanglin, Zheng Liping, Zhang Jing, Yang Liuqing

Medical Image Processing

- Despeckling for medical ultrasound images based on wavelet and bilateral filter 0126
Zhang Ju, Wang Chen, Cheng Yun
- Medical image fusion based on weighted Contourlet transformation coefficients 0133
Zhang Xin, Chen Weibin

Remote Sensing Image Processing

- The method and application of multi-dimension interpretation for landslides using high resolution remote sensing image 0141
Lu Xuejun, Shi Zhenchun, Shang Weitao, Zhou Heyi
- Topographic correction of polarimetric synthetic aperture radar and its application in snow and ice mapping 0150
Fu Sitao, Li Zhen, Tian Bangsen
- The object-oriented change detection based on neighborhood correlation images and the minimum-redundancy-maximum-relevance feature selection 0158
Zou Lidong, Pan Yaozhong, Zhu Wenquan, Zhou Gongqi, Li Yizhan

中图法分类号: TP391.9 文献标识码: A 文章编号: 1006-8961(2014)01-0054-08

论文引用格式: 姚争为, 潘志庚. 基于体素的人手结构自动估测[J]. 中国图象图形学报, 2014, 19(1): 54-61. [DOI: 10.11834/jig.20140107]

基于体素的人手结构自动估测

姚争为, 潘志庚

杭州师范大学数字媒体与人机交互研究中心, 杭州 310036

摘要: 目的 人手结构估测是建立个性化3维人手模型以及进行人手姿态估测的重要步骤之一。由于人手骨架为表皮所覆盖, 普通摄像机无法直接获得其结构, 因此常常需要人员参与估测过程, 无法真正实现自动化, 这必然会增加估测的时间和估测的复杂度。**方法** 本文从人手纹理与人手关节位置相关性出发, 利用一套图像分割和图像增强技术提取出大致的人手连接结构; 再利用多视点3维建模、3维拟合技术和粒子群优化算法, 优化之前得到的人手连接结构。**结果** 经实验证明该方法能成功地提取出人手结构, 估测的每个阶段均无需人员参与。**结论** 与之前方法相比, 该方法真正实现了估测过程自动化, 减少了估测时间和估测误差。

关键词: 个性化; 连接结构; 体素; 人手模型; 手姿态估测

Automatic calibration of hand structures based on voxel data

Yao Zhengwei, Pan Zhigeng

Digital Media and HCI Research Center, Hangzhou Normal University, Hangzhou 310036

Abstract: **Objective** Calibration of hand structures is one of important steps to create an individualized three-dimensional hand model and estimate hand-pose. Because the hand structure is covered by the hand surface, the common cameras cannot directly get the hand structure, and people were required to participate in the calibration of hand structure, which cannot be completely processed by the computer. That makes the calibration complex and more time is needed. **Method** In order to acquire an automatic calibration of a hand structure, a set of image segmentation and enhancement techniques were used to find the link structure of the hand, based on the relationship of hand prints and joints in the individual's hand. The techniques of 3D modeling using multiple cameras, 3D fitting, and the Particle Swarm Optimization algorithm were used to optimize the link structure of hand. **Result** Proved by the experiments, the calibration method proposed in this paper can successfully extract the structure of the hand, and each stage of the calibration process doesn't need people participating in. **Conclusion** Comparing with the original method, the calibration method proposed in this paper is completed by the computer. As a result the calibration time is saved and errors are reduced.

Key words: individuation; link structure of hand; voxel data; hand model; hand-pose estimation

0 引言

交互自然性是人机交互发展的必然趋势, 而无

连接的人手交互是交互自然性的一个重要体现^[1-3]。对于一些复杂的人手交互, 如抓取、拧动物体等, 操作的真实性和准确性往往由人手模型的个性化程度所决定。系统的人手模型对观测到的输入

收稿日期: 2013-04-16; 修回日期: 2013-06-29

基金项目: 国家自然科学基金面上项目(61173124); 浙江省教育厅年度科研项目(Y201120954)

第一作者简介: 姚争为(1977—), 男, 讲师, 2010年于上海大学获计算机应用技术专业博士学位, 主要研究方向为增强现实, 人机交互。

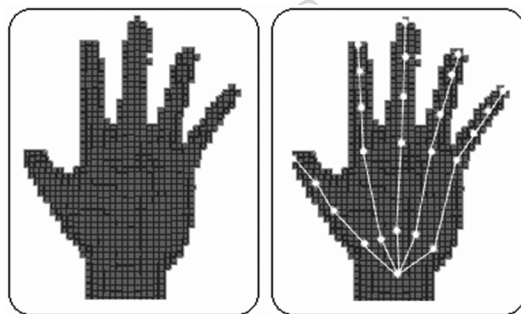
E-mail: yzwiron@hznu.edu.cn

数据反映越准确, 手部姿态的估测就越准确, 所以要为不同用户建立个性化的人手模型^[4-5]。这样有助于减小系统的人手模型和输入数据之间的偏差, 提高交互的真实感和准确性。用于手部姿态估测系统的人手模型主要有: 体素模型^[6]、可变形表面模型^[7]、纸板箱模型^[8]和截断二次型模型^[9]等。到底采用哪种模型由处理速度、精度和技术成本所决定。本文采用的是基于体素的个性化人手模型。

目前有许多国内外学者对个性化人手模型的构建进行研究。有些研究者采用精确的测量仪器, 如 CT (computerized tomography) 或 MRI (magnetic resonance imaging), 来获得高精度的手模型。如, Kurihara 等人^[10]用 CT 扫描所得的医学图像来构建真实的手模型。他们通过对手在不同姿态时的扫描来估测手的结构, 以及利用多边形网格来构建人手的表面模型, 这样构造的手模型无论在静止还是移动时看起来都很真实。Rhee 等人^[11]利用 MRI 扫描得到体数据, 进而调整手模型, 他们的手模型严格遵循运动学原理能很好地反映手骨的结构和皮肤的形变。尽管上述方法符合人手的运动学原理, 且拥有逼真的外观, 但扫描需要花费大量的时间, 而且扫描设备极其昂贵。目前还有一种方式是利用多视点轮廓图像来生成手的体素模型, 并通过 3 维拟合来估测手骨的结构和人手姿态。这种方法虽然在精度上不及前面所提到的方法, 但费用少, 速度快。Hattori 等人^[12]和 Albert 等人^[13]都是利用这类方法来创建个性化的人手模型。前者先建立一个非个性化的通用手模型, 然后在测量过程中根据测得的手的体数据 (用户手的实际形状与姿态信息), 对通用模型进行不断调整, 实现个性化和姿态估测。后者无需建立通用手模型, 但需要人工选取手骨的各连接段的长度, 因此操作不够简便。

本文所描述的基于体素的人手结构自动估测方法, 是在基于 Albert 等人^[13]提出的基于体素的个性化手模型标定方法的基础上建立的。手模型是由手的连接结构和手的表面结构两部份组成, 如图 1 所示。连接结构代表了手的骨架, 模型中的每段连接都有表面结构中的点与之对应, 对连接段加虚拟力就会移动手指, 改变人手模型的形状。所以连接结构直接影响着手的姿态和形状。

文献[13]中, 用户将手完全张开, 操作者通过多视点摄像机捕获完整的手部图像, 然后提取图像轮廓, 并转换成手的体素模型, 由于每一个体素都被



(a) 由体素构成的手模型 (b) 融合了连接结构的手模型

图1 人手模型

Fig. 1 Hand models

看做一个顶点, 所以这些体素数据可代表手的真实形状。然后, 操作者人工地在体素模型表面绘制出手的各连接段长度, 再让用户将各手指逐个缓慢地弯曲, 将取到的图像转化成对应的体素模型进行拟合。为了获得更精确的连接段长度, 系统对所有的连接段及关节角度不断进行调整, 每一个调整都定义为一个新的组合, 并用代价函数对其进行评估, 评估值最高的组合就是最终的连接结构。

然而, 该方法也有不少缺陷: 首先, 需要人为干预, 即在初始化时手动选取各连接段的长度, 这会给操作者制造不便。其次, 由于该算法在拟合时采用穷举法, 每次调整幅度设为 5° 和 4 mm, 因此算法的执行效率和精度都有待提高。本文对该方法进行了改进, 提出了基于体素与手部纹理特征的人手结构估测方法, 实现了估测过程自动化, 提高了估测精度和效率。

1 初步提取手的连接结构

1.1 手部纹理与连接模型的关系

经过仔细观察分析, 不难发现人的手部纹理与手指各关节的位置有一定的相关性, 如图 2 所示, 除拇指外的其他手指都有 3 条比较清晰的纹线 (大拇指为两条), 第一、二条纹线在手指的远侧指关节 (DIP) 和近侧指关节 (PIP) 处, 因此从指尖到第一条纹线、从第一条纹线到第 2 条纹线可代表不同指骨的长度。掌关节 (MP) 节点的确定有所不同, 但同一手指的 MP 节点和 DIP、PIP 节点基本处在同一直线上 (拇指除外), 在本文中这 4 条直线分别被称为 Index、Middle、Ring、Pinkie。通过手骨结构和大量手部纹理样本分析可知, MP 节点可按以下方法初步

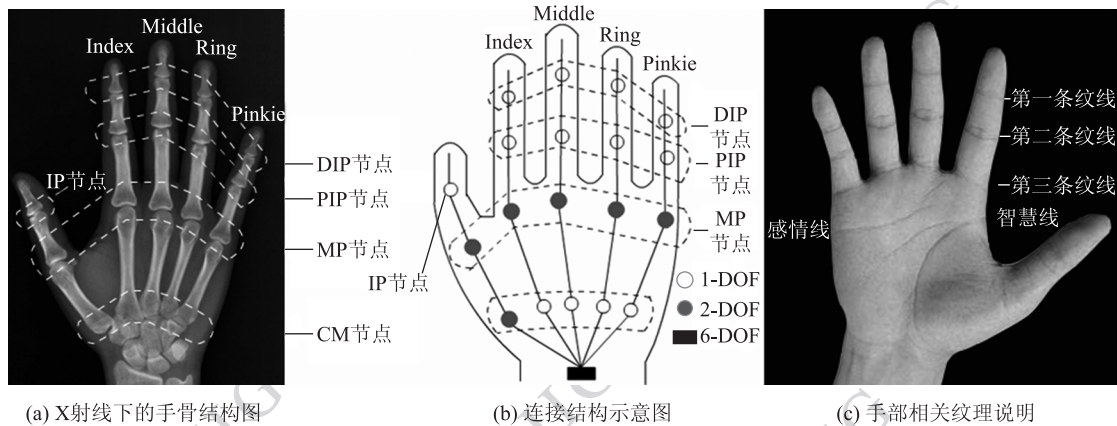


图2 手部结构与纹理

Fig. 2 Hand structure and hand prints

确定。

查找食指 MP 节点:在智慧线与手轮廓的交点处作食指第三条纹线的平行线,该平行线与 Index 的交点定为食指 MP 节点。

查找中指 MP 节点:在食指 MP 节点处作中指第三条纹线的平行线,该平行线与 Middle 的交点定为中指 MP 节点。

查找小拇指 MP 节点:在感情线与手轮廓的交点处作小指第三条纹线的平行线,该平行线与 Pinkie 的交点定为小指 MP 节点。

查找无名指 MP 节点:在小指 MP 节点处作无名指第三条纹线的平行线,该平行线与 Ring 的交点定为无名指 MP 节点。

查找大拇指 MP 节点:大拇指第 2 条纹线中点。

文献[6,13]都将 CM 节点列入了手的连接结构中,该结构中的所有节点都视为具有一定的自由度并能参与手部运动,在个性化手模型的标定中也参与运算。但根据人手的运动特征,除大拇指的 CM 节点有两个自由度的运动能力外,其余 4 根手指的 CM 节点基本无法独立运动,因此可忽略除大拇指之外的其余 4 根手指的 CM 节点,这样连接模型从 31 个自由度简化为 27 个,连接段从 24 段简化为 20 段,这利于缩减标定时间。

要提取连接结构必须先提取手部轮廓,计算指尖、指根,找出手部纹理流出点。

1.2 手部轮廓的提取

采用基于区域的水平集方法(C-V 模型)提取手部轮廓^[14]。水平集方法将图像上的 2 维曲线看成是 3 维隐函数 Φ 的零水平集。曲线在随时间演

化时遵循特定的关于 Φ 的能量函数,该能量函数在曲线位于图像中目标边缘时取到最小值,从而得到图像中物体的轮廓。该方法不依赖图像梯度信息,能使曲线在演化过程中任意断开或合并,具有很好的抗噪性。由于手部轮廓提取不只在获取连接模型中使用,在后面的多视点 3 维手模型构建中也要用到,所以为了减小背景阴影的影响,对原有算法做了改进。

首先可通过大量肤色样本建立 HSV 颜色空间下的肤色 3 维高斯概率分布 $N(\mu, \Sigma)$, μ 为均值向量, Σ 为协方差矩阵。然后计算每个像素 x 在该肤色概率分布中的概率

$$P_{\text{skin}}(x) = \frac{1}{2\pi \sqrt{|\Sigma|}} \times \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) \quad (1)$$

对结果归一后,肤色像素接近 1,而非肤色像素接近 0,将其替代能量函数中的原有的图像灰度,可从图像中将人手分割出来。修正后的能量函数为

$$F(c_1, c_2, \phi) = \mu \int_{\Omega} \delta(\phi(x, y)) |\nabla \phi(x, y)| dx dy + \lambda_1 \int_{\Omega} |P_{\text{skin}}(x, y) - c_1|^2 \cdot |H(\phi(x, y))| dx dy + \lambda_2 \int_{\Omega} |P_{\text{skin}}(x, y) - c_2|^2 \cdot |1 - H(\phi(x, y))| dx dy \quad (2)$$

式中, Ω 表示整个图像域; δ 是 Dirac 函数; x, y 为像素坐标; H 为 Heaviside 函数; $P_{\text{skin}}(x, y)$ 为像素在该肤色 3 维高斯概率分布中的概率; ϕ 为水平集函数;

λ_1, λ_2 是自由参数, 实验取 $\lambda_1 = \lambda_2 = 1$; c_1, c_2 为

$$c_1 = \frac{\int_{\Omega} P_{\text{skin}}(x, y) H(\phi(x, y)) dx dy}{\int_{\Omega} H(\phi(x, y)) dx dy} \quad (3)$$

$$c_2 = \frac{\int_{\Omega} P_{\text{skin}}(x, y) (1 - H(\phi(x, y))) dx dy}{\int_{\Omega} (1 - H(\phi(x, y))) dx dy} \quad (4)$$

另外, C-V 模型要求水平集函数在迭代过程中需保持为符号距离函数, 因此 C-V 模型经多次迭代后需要对水平集函数重新初始化以满足条件, 这项工作非常耗时, 为了加快计算速度, 参照文献[15]引入惩罚项, 最终的能量函数修改为

$$F'(c_1, c_2, \phi) = F(c_1, c_2, \phi) + v \int_{\Omega} \frac{1}{2} (|\nabla \phi - 1|)^2 dx dy \quad (5)$$

式中, v 为自由参数, 取 $v = 1.5$ 。对式(5)求一阶变分, 便可得到 ϕ 随时间的演化方程。

1.3 指尖、指根的获取

利用水平集方法获得手轮廓的点集合后, 参照文献[16]采用第1类边界条件的3次样条插值法对轮廓点进行拟合, 找到序列点变化趋势, 对插值后的曲线计算每一个边界点的曲率, 再采用 db4,

8 阶离散小波变换 (db 为 daubechies 小波的简称, db4, 8 中的 4 和 8 表示 db 小波的阶次) 检测极大值点, 这些极大点除去起点和终点, 剩下的即为指尖和指根点, 将这些点按顺序存储, 奇数为指尖偶数为指根。

1.4 流出点的提取

手指的第一、二、三条纹线都贯穿手指并从两侧流出, 智慧线和感情线也都汇聚于手掌一侧流出, 因此确定了流出点的位置后, 手部纹线也基本得以确定。

由于纹线在局部范围内方向不变, 从该方向出发的累积灰度值比较高, 所以在手轮廓附近不断选取 25×25 的小窗口, 采用方向投影算法, 从 $0^\circ, 45^\circ, 135^\circ, 180^\circ$ 这 4 个方向投影, 找出流出点。为了使纹线更加清晰, 流出点的提取更加准确, 在方向投影前, 必须对图像进行预处理: 反色、priwitt 锐化、二值化处理; 锐化后会产生较多噪点, 使用自适应中值滤波去噪, 效果如图 3 所示。

经上述处理过程, 连接结构已能大致确定, 但并不精确, 因此在接下来的过程中需根据用户的动作做进一步估测, 以获得各连接段长度的最佳值。

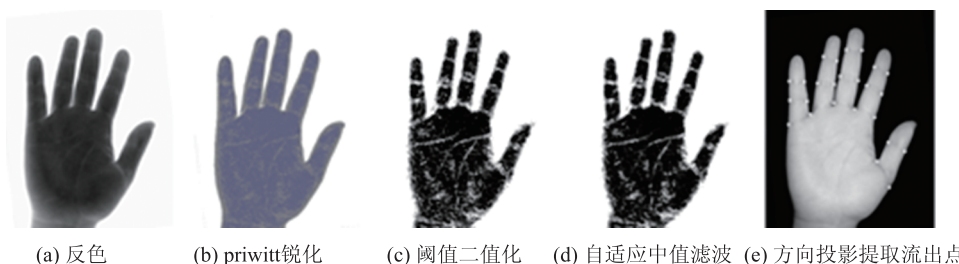


图3 流出点提取处理过程

Fig. 3 Procedure of getting the outflow point

2 优化手的连接结构

在求指骨最佳长度与手部姿态时, 需要用到两套体数据。第一套体数据是手先前姿态的体数据。第二套是手的移动体数据。最开始手掌打开, 五指完全伸展; 移动时手指逐个向手掌缓慢弯曲。先前姿态的体数据就是融合了连接结构的手的表面模型。然后通过调整手的连接结构来改变模型姿态来拟合移动体数据模型以找出最优值。

参照文献[13], 手的连接结构的优化步骤如下:

- 1) 从多视点摄像机获取完全张开的手的图像, 并将其转换成轮廓图。
- 2) 将轮廓图转换成体数据模型, 即手的表面结构模型。
- 3) 将前面得到的初始连接结构融入表面结构模型。
- 4) 获取手指缓慢移动图像, 并将这些图像转换成轮廓图。

5) 将这些轮廓图像转换成时间序列体数据模型。

6) 调整手的连接结构, 使得先前的手模型逐渐拟合移动体数据模型, 步骤如下:

(1) 逐渐地调整关节角度 ($\pm 5^\circ$) 和各个连接段的长度 (± 4 mm);

(2) 计算代价函数;

(3) 重复步骤(1), 直到尝试了所有角度和连接段长度的组合。

7) 找出最优的手的连接结构, 该连接结构能使代价函数取到最大值。

步骤(2)中代价函数为 $f(\mathbf{M}(\mathbf{\Gamma}, \mathbf{\Theta}), \mathbf{V})$, 其中, f 是移动体数据 $\mathbf{V}$ 和计算模型 $\mathbf{M}$ (融合了手的连接结构和表面结构的体素模型) 之间的比值, 而计算模型 $\mathbf{M}$ 是由 $\mathbf{\Gamma}$ 和 $\mathbf{\Theta}$ 来定义的, $\mathbf{\Gamma}$ 是连接长度 ($l_1, l_2, \dots, l_n$) 的集合, $\mathbf{\Theta}$ 是关节角度 ($\theta_1, \theta_2, \dots, \theta_m$) 的集合, n 和 m 分别是连接段和自由度的个数。体素数据和计算模型之间的比值最高的组合就是所要求的手指连接长度标定值。

上述连接结构的优化算法采用穷举法, 需罗列所有组合。如果组合数较少, 这也不失为一个简便方法, 但手部结构比较复杂, 经过简化后自由度仍有 27 个, 连接段 20 段, 因此组合数过多。此外, 在迭代中每个连接段和节点的调整幅度定为 4 mm 和 5° , 这样的调整过于粗略, 会影响连接结构估测的准确性。因此采用粒子群算法优化搜索过程, 步骤如下:

1) 初始化 初始化搜索点的位置 x_i^0 及其速度 v_i^0 。 x_i^k 是粒子 i 在第 k 次迭代中的值, 它的维数为手指连接长度与关节角度的数量总和。 v_i^k 是粒子 i 在第 k 次迭代中的速度, 即每次的调整幅度; 为防止粒子远离搜索空间, 将粒子的速度限制在一定的范围内 (± 2 mm、 $\pm 3^\circ$), 初始值就在该范围内随机产生。将全局极值和每个粒子的个体极值设为 0, 将粒子数设为 30。

2) 评价每一个粒子 计算粒子的适应度值, 如果好于该粒子当前的个体极值, 则将 p_{best} 设置为该粒子的值 (p_{best} 记录该粒子获得的各维度的最好极值), 且更新个体极值。如果所有粒子中最好的个体极值好于当前的全局极值, 则将 g_{best} 设置为该粒子的值 (g_{best} 记录的是所有粒子中最好的个体极值), 记录该粒子的序号, 且更新全局极值。适应度函数定义为代价函数, 即 $f(\mathbf{M}(\mathbf{\Gamma},$

$\mathbf{\Theta}), \mathbf{V})$ 。

3) 粒子的更新 对每一个粒子的速度和值进行更新, 即

$$v_{id}^{k+1} = v_{id}^k + c_1 \text{rand}_1^k (p_{\text{best } id}^k - x_{id}^k) + c_2 \text{rand}_2^k (g_{\text{best } id}^k - x_{id}^k) \quad (6)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (7)$$

式中, v_{id}^k 是粒子 i 在第 k 次迭代中第 d 维的速度; c_1 、 c_2 是学习因子, 分别用来调节向全局最好粒子和个体最好粒子方向搜索的最大步长, 令 $c_1 = 2, c_2 = 2$; $\text{rand}_1, \text{rand}_2$ 是 $[0, 1]$ 之间的随机数; x_{id}^k 是粒子 i 在第 k 次迭代中第 d 维的值, $p_{\text{best } id}^k$ 是粒子 i 在第 d 维的个体极值点的值; $g_{\text{best } id}$ 是整个群在第 d 维的全局极值点的值。

4) 判断搜索是否结束 当迭代次数大于 100, 则停止迭代, 输出最优解, 否则转到步骤 2)。

3 实验分析

实验所用计算机采用 Intel 酷睿 i5 四核处理器 (主频 3.3 G), 内存为 4 G。采集手纹所用的摄像机的分辨率为 800×600 , 捕获手部轮廓图的多视点摄像机的分辨率为 320×240 。构建的立方体检测空间大小为 $60 \text{ cm} \times 60 \text{ cm} \times 60 \text{ cm}$, 在框架上分别安置了 4 台摄像机来捕获手部图像。

图 4 为通过计算拟合后的轮廓点的曲率以及小波变换来获得指尖、指根的手轮廓点集处理图。图 5 为求得初始连接结构的整个过程。图 6 为通过 4 台安装在不同方位的摄像机获取手部图像来建立基于体素的人手模型。图 7 为连接结构的优化过程 (体数据模型拟合过程)。图 8 为采用粒子群方法优化前后的效果对比。

为了更好地检验实验效果, 在实验前先对人手进行测量 (在图 9 中称为标准测量), 然后再用文献 [13] 的方法 (在图 9 中称为人工辅助测量) 进行测量, 最后再用本文方法进行测量 (在图 9 中称为自动测量)。三者的比较结果如图 9 和表 1 所示, 通过比较可以看出本文方法的测量结果与真实人手还是比较接近的。与文献 [13] 相比, 虽然在误差小于 5 mm 区间内的百分比相同, 但在 5 mm 至 10 mm 区间, 百分比有所提高, 在大于 10 mm 区间的百分比则大大降低。本文方法的误差平均值为 3.75 mm, 文献 [13] 的方法为 5.083 3 mm。

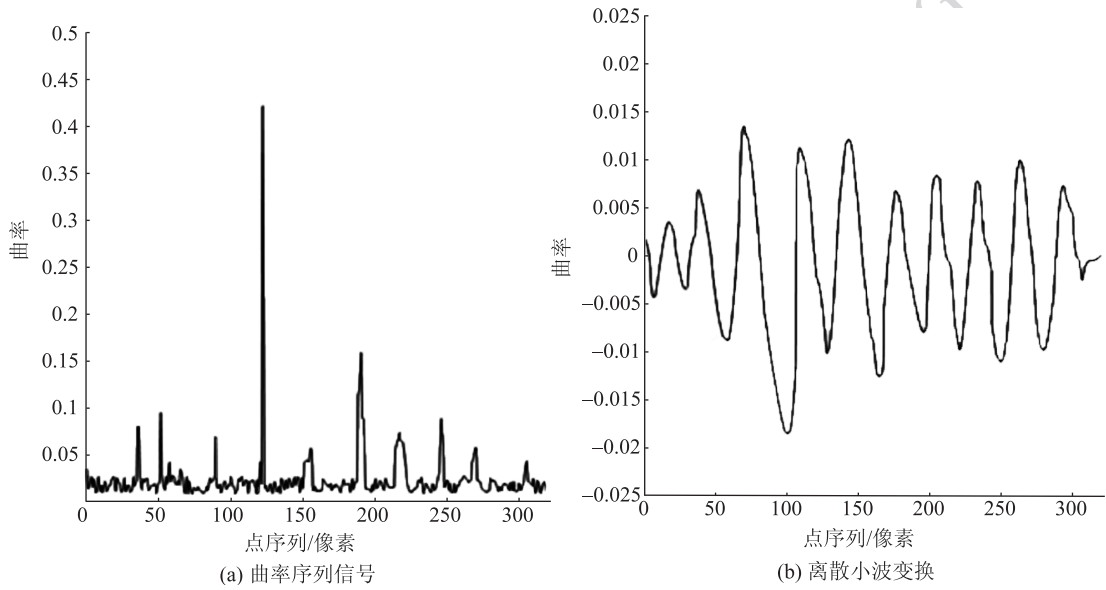


图 4 手轮廓点集处理

Fig. 4 Process the point set of hand contour

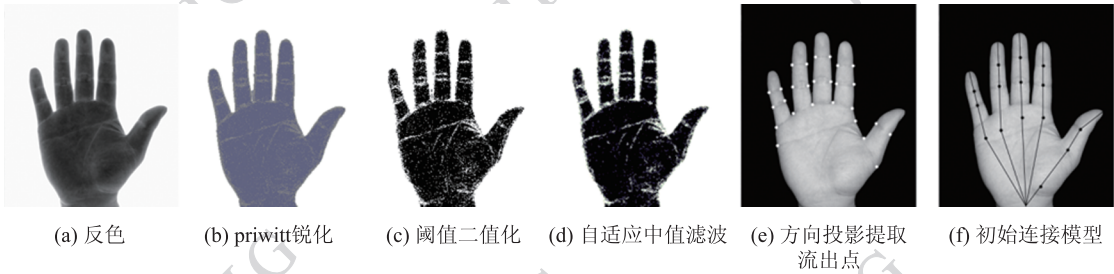


图 5 提取初始人手连接结构过程

Fig. 5 Procedure of getting the initial hand structure

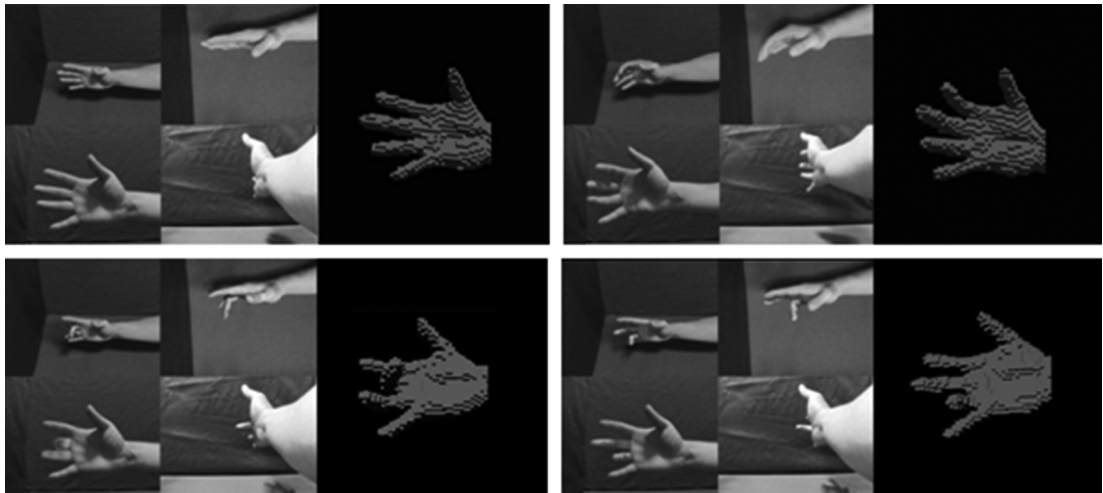


图 6 通过多视点摄像机构建基于体素的人手模型

Fig. 6 Voxel-based hand model created by using multiple cameras

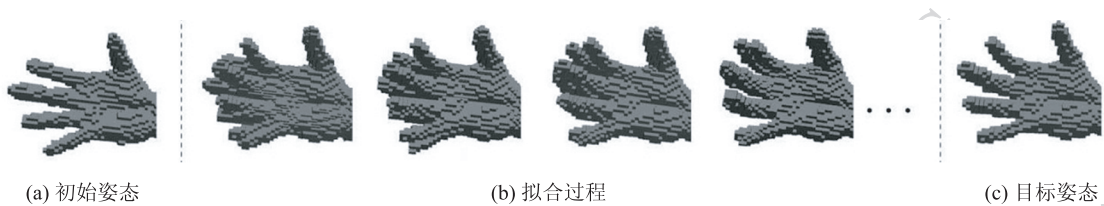


图 7 手的连接结构优化过程
Fig. 7 Procedure of optimizing link structure of hand

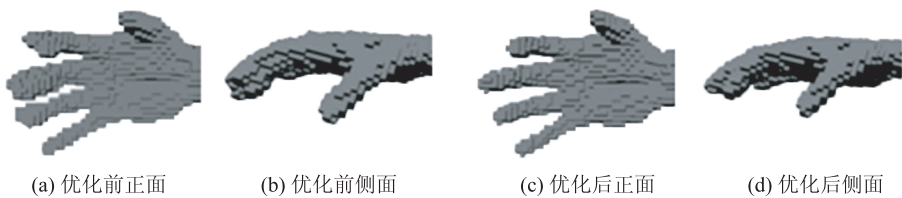


图 8 采用粒子群算法优化前后的效果比较
Fig. 8 Compare the results before and after using the particle swarm algorithm

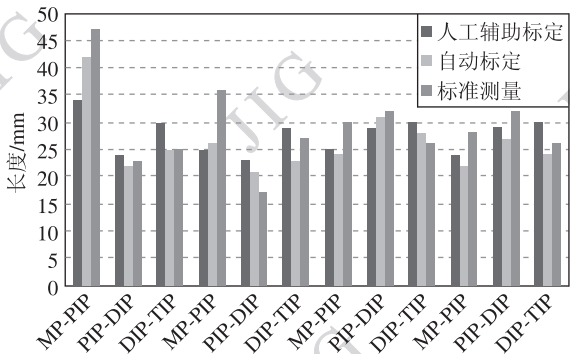


图 9 测量结果比较
Fig. 9 Compare the results of using different methods

表 1 测量错误比较

Table 1 Compare the errors generated in different methods

误差	连接段数量	
	人工辅助测量	自动测量
小于 5 mm	58. 3%	58. 33%
5 ~ 10 mm	25%	33. 33%
大于 10 mm	16. 67%	8. 33%
误差平均值/mm	5. 083 3	3. 75

从实验结果可知,本文改进算法在平均误差上较之原算法有明显改善,这是因为在粒子群优化算法中将粒子的速度限制在了 $\pm 2\text{ mm}$ 、 $\pm 3^\circ$ (原算法采用穷举法,调整幅度为 $\pm 4\text{ mm}$ 、 $\pm 5^\circ$)。但在误差小于 5 mm 区间和 5 mm 至 10 mm 区间,新旧算法的效果差别并不明显,这是由于为了保证响应的实时

性,系统将粒子群算法的终止条件设置为迭代次数大于 100 (适当地增加迭代次数,会提高精度,但会延长响应时间)。

4 结 论

将手的连接结构的自动估测分为两个阶段:连接结构的初步提取和连接结构的优化。第一阶段根据人手纹理与人手各关节位置之间的相关性,用优化的水平集方法提取手的轮廓边界,再用 3 次样条插值法和离散小波变换求出指尖与指根的位置,用一系列图像处理技术突出手部纹理,再采用方向投影提取流出点,最后根据以上信息提取出手的连接结构。第二阶段通过体素模型的构建和 3 维拟合来优化连接结构。在 3 维拟合中引入了粒子群算法优化了搜索过程。有了测量过程的合理划分,以及有效算法的支撑,手的连接结构的估测实现了自动化,估测精度也得到了提升。

然研究仍有可提升空间:没有真正解决大拇指的自动估测问题(因大拇指 CM 节点没有对应的纹理);估测过程仍然是一个独立的程序,没有直接融入到交互活动之中等。

参考文献 (References)

[1] Pan Z G, Li Y, Zhang M M, et al. A real-time multi-cue hand tracking algorithm based on computer vision[C]// Proceeding of

- IEEE Virtual Reality 2010. Waltham-Boston;IEEE Press, 2010: 219-222.
- [2] Stiefelhagen R, Fugen C, Gieselmann R, et al. Natural human-robot interaction using speech, head pose and gestures[C]//Proceedings of IEEE/RSJ Conference on Intelligent Robots and Systems. Sendai;IEEE Press, 2004:2422-2427.
- [3] Feng Z Q, Zhang M M, Pan Z G, et al. 3D-freehand-pose initialization based on operator's cognitive behavioral models[J]. The Visual Computer, 2010,26(6-8): 607-617.
- [4] Zhao W P, Chai J X, Xu Y Q. Combining marker-based mocap and RGB-D camera for acquiring high-fidelity hand motion data [C]//The ACM SIGGRAPH/Eurographics Symposium on Computer Animation Symposium on Computer Animation. Lausanne; The Eurographics Association, 2012:33-42.
- [5] Romero J, Kjellstrom h, Kragic d. Hands in action;real-time 3D reconstruction hands in interaction with objects[C]//Proceedings of IEEE International Conference on Robotics and Automation. Anchorage;IEEE Press, 2010:458-463.
- [6] Etsuko U, Yoshio M, Masakazu I, et al. A hand-pose estimation for vision-based human interfaces[J]. IEEE Transactions on Industrial Electronics, 2003, 50(4):676-684.
- [7] Kofman J, Wu X H, Luu T J, et al. Teleoperation of a robot manipulator using a vision-based human-robot interface[J]. IEEE Transactions on Industrial Electronics, 2005, 52(5): 1206-1219.
- [8] Wu Y, Lin J Y, Huang T S. Capturing natural hand articulation [C]//Proceedings of International Conference on Computer Vision. Vancouver;IEEE Computer Society, 2001:426-432.
- [9] Stenger B, Mendonca P R S, Cipolla R. Model based 3d tracking of an articulated hand [C]//Proceedings of Conference in Computer Vision and Pattern Recognition. Kauai;IEEE Computer Society, 2001:310-315.
- [10] Kurihara T, Miyata M. Modeling deformable human hands from medical images[C]//The ACM SIGGRAPH/Eurographics Symposium on Computer Animation. Grenoble; The Eurographics Association, 2004:355-363.
- [11] Rhee T, Lewis J P, Neumann U, et al. Soft-tissue deformation for in vivo volume animation[C]//Proceedings of the 15th Pacific Conference on Computer Graphics and Applications. Maui; IEEE Computer Society, 2007:435-438.
- [12] Hattori Y, Nakazawa A, Takemura H. Refinement of the shape reconstructed by visual cone intersection using fitting the standard human model[C]//Proceedings of the Information Processing Society of Japan SIG Notes CVIM. Japan; CVMI Society, 2007: 147-154.
- [13] Albert C, Mai M, Etsuko U, et al. Hand pose estimation using voxel-based individualized hand model[C]//Proceeding of 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. Singapore;IEEE Press, 2009:451-456.
- [14] Chan T, Vese L. Active contours without edges [J]. IEEE Transactions on Image Processing, 2001, 10(2):266-277.
- [15] Li C, Xu C, Gui C F. Level set evolution without reinitialization;a new variational formulation[C] //Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. San Diego;IEEE Press, 2005:430-436.
- [16] Shen H, Cheng X, Qu J H, et al. Principle line extraction and restoration based on wavelet theory[J]. Journal of Image and Graphics, 2006, 11(8):1139-1149. [黄申,徐成,屈景辉,等. 基于小波理论对掌纹主线提取和修复[J]. 中国图象图形学报,2006,11(8): 1139-1149.]