**学校编码：**10384

**学号：**23020091152757

厦门大学

# 硕士学位论文

# 面向虚拟试衣的人体骨架及关节点的提取

# Human Skeleton and Joints Extraction for Virtual Clothing Fitting

罗伟

指导教师：杨晨晖

专业名称：计算机应用技术

答辩日期：2012年6月

# 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（　　）课题（组）的研究成果，获得（　　）课题（组）经费或实验室的资助，在（　　）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年　月　日

# 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文(包括纸质版和电子版)，允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于:

(　　)1.经厦门大学保密委员会审查核定的保密学位论文，于　年　月　日解密，解密后适用上述授权。

(　　)2.不保密，适用上述授权。

(请在以上相应括号内打“✓”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。)

声明人(签名):

年　　月　　日

# 摘　要

人体运动捕捉是计算机视觉领域倍受关注的一个研究热点，在智能视频监控、视频分析、动画、游戏、医学诊断和人机交互等领域均有广阔的应用前景。它包括人体的标定与跟踪和人体动作的识别与理解两个主要内容。其中，人体的标定与跟踪是运动识别和理解的基础，在人体运动捕捉中起着关键性的作用。因此，本文以虚拟试衣系统为背景，研究人体的骨架化技术与人体关节点的定位，具有重要的理论价值和实际意义。

骨架是描述物体形状和拓扑结构的一种有效手段，广泛应用于人体的描述。基于距离变换的骨架化算法对人体的骨架有较好的效果，但无法保证骨架的连通性。本文利用图像梯度的性质，通过梯度化距离变换图，着重突出了潜在的骨架点；利用距离值和梯度值，寻找关键点；引入基于梯度的最短路径算法，连接所有关键点，形成物体的曲线骨架，解决骨架连通性问题。在此基础上，本文还提出基于梯度的优化方法，综合 、 、 和 四个方向的梯度图，形成新的包含所有潜在骨架点的梯度图。通过预设的阈值消除大部分非骨架点；利用轮廓法提取由骨架点组成的“粗”中脊线；通过拓扑细化得到“细”中脊线。最后，利用“细”中脊线构建最短路径算法所需的邻接矩阵，大大地减少算法的运行时间。

为了满足无标记的要求，本文利用人体的特性，通过寻找“边缘点”与主干点之间的路径，把人体曲线骨架分割成四肢和脊柱5个骨架分枝；利用人体模型的先验知识，提取每个骨架分枝的主要关节点；连接每个骨架分枝的关节点，形成人体刚体骨架。实验表明，关节点的位置与人体的真实关节点位置基本吻合，能够满足虚拟试衣中人体自动建模跟踪的需要。

**关键词：**运动捕捉；骨架提取；梯度；关节点

# Abstract

Human motion capture is a hot research in the field of Computer Vision. This research has various application prospects in intelligent video surveillance, video analysis, animation, computer games, medical diagnostics, human-computer interaction, and so on. It includes two main components: human calibration and tracking, and human action recognition and understanding. Among them, the human calibration and tracking is the basis of recognition of human movement, and also plays a key role in the human motion capture. Therefore, based on virtual clothing fitting system, this paper has the important theoretical value and practical significance, through studying human skeletonization and main joints location.

Skeleton is an effective means to describe the shape and topology of an object. It is used to describe the human body widely. There is a good effect on human skeleton for skeleton algorithm based on distance transform, but it cannot guarantee the connectivity. This paper finds the gradient of distance transform image to highlight the potential skeleton points, using the nature of image gradient; looks for the critical points taking advantage of distance value and gradient; connects all the critical points though introducing shortest path based gradient; forms the curve skeleton of object; solves the skeleton connectivity problem. In addition, this paper also proposes a optimized method based on gradient. It combines the gradient images of and four directions and forms a new gradient image including all the potential skeleton points. This method presents a threshold to eliminate most of non-skeleton points; extracts “rough” skeleton line in the method of contour; uses topology thinning to get the “small” skeleton line; and builds the adjacency matrix of the shortest path algorithm using “small” skeleton line finally. It reduces the running time greatly.

In order to meet the requirements of marker-less motion capture, this paper

divides human curve skeleton into the limbs and spine skeleton five branches, through finding the path between the edge points and torso using human body characteristics; extracts the main key joints from each skeleton branches using a priori knowledge of the human body model; connects the joints of each skeleton branch to form the rigid body skeleton. Experiments show that the joints location fits the real body's joints location basically. And it also can meet the needs of modeling and tracking human body for virtual clothing fitting.

**Keywords:** Motion Capture; Skeleton Extraction; Gradient; Joints

# 参考资料


[1]李豪杰,林守勋,张勇东. 基于视频的人体运动捕捉综述[J]. 计算机辅助设计与图形学学报,2006,18(11): 1645-1651.

[2]Thomas B. Moeslund. Computer Vision-Based Human Motion Capture – A Survey[J]. Computer Vision and Image Understanding, 2001, 81(3): 231-268.

[3]Haritaoglu I, Harwood D, Davis L S. W4:real-time surveillance of people and their activities[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 809-830.

[4]Wang J, Cohen M F. Very low frame-rate video streaming for face-to-face teleconference[C]. Proceedings of IEEE Data Compression Conference, USA: IEEE Computer Soc press, 2005: 309-318.

[5]孙鹏. 单目视频中基于退火粒子滤波的三维人体运动跟踪[D]. 华中科技大学,2009.

[6]F.Multon, L.Hoyet, T.Komura and R.Kulpa. Interactive Control of Physically-valid Aerial Motion: Application to VR Training System for Gymnasts[C]. Proceedings of ACM Virtual Reality and Technology VRST, Newport Beach, CA, USA, 2007: 77-80.

[7]T.B. Moeslund, A. Hilton, V. Kruger. A survey of advances in vision-based human motion capture and analysis[C]. CVIU, 2006, 104: 90-126.

[8]陈坚. 单目视频人体运动跟踪和获取技术研究[D]. 中科院计算所,2005.

[9]Wren Christopher Richard, Azerbaijani Ali, Darrell Trevor, et al. Pfinder: Real-time tracking of the human body[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 780-785.

[10]Nait H, Charif, S.J.McKenna. Head tracking and action recognition in a smart meeting room[C]. Proceedings of 4th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Graz Austri, 2003, 24-31.

[11]Wu Y, Huang T S. Robust visual tracking by integrating multiple cues based on co-inference learning[C]. International Journal of Computer Vision, 2004, 58(1): 55-71.

[12]Shan C F, Wei Y C, Tan T N, et al. Real time hand tracking by combining particle filtering and mean shift. Proceedings of 6th International Conference on Automatic Face and Gesture Recognition. Los Alamitos USA: IEEE Computer Soc press, 2004, 669-674.

[13]谭煌. 网络虚拟服装试衣系统分析和设计[D]. 山东大学,2009.

[14]G.Johansson. Visual motion perception[M]. Scientific American, 1976, 76-88.

[15]E.Sifakis, I.Neverov, R.Fedkiw. Automatic Determination of Facial Muscle Activations from Sparse Motion Capture Marker Data. ACM Transactions on Graphics[J], 2005, 24(3): 417.

[16]晏剑云,郑江滨,李秀秀. 视频运动捕获系统中的多标记点的三维跟踪算法[J]. 计算机工程,2008,34(24): 206-207.

[17]C.Ott, D.Lee, Y.Nakamura. Motion Capture based Human Motion Recognition and Imitation by Direct Marker Control[C]. 2008 8th IEEE-RAS International Conference on Humanoid Robots, 2009: 399-405.

[18]C.W.Chu, O.Jenkins, M.Mataric. Markerless Kinematic Model and Motion Capture from Volume Sequences[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003, 3: 16-22.

[19]C.W.Chu, O.Jenkins, M.Mataric. Towards Model-free Markerless Motion Capture[C]. IEEE International Conference on Robotics and Automation, 2003.

[20]L.Sigal, M.J.Black. Humaneva: Synchronized Video and Motion Capture Dataset for Evaluation of Articulated Human Motion[J]. International journal of computer vision, 2010, 87(1): 4-27.

[21]Yebin Liu, Carsten Stoll, Juergen Gall, et al. Markerless Motion Capture of Interacting Characters Using Multi-view Image Segmentation[C]. IEEE International Conference on Computer Vision(ICCV), 2011, 1249-1256.

[22]Luca Ballan, Guido Maria Cortelazzo. Marker-less motion capture of skinned models in a four camera set-up using optical flow and silhouettes[C]. Proceedings of 3DPVT ’08-the Fourth International Symposium on

3D Data Processing, Visualization and Transmission, GA, USA, 2008.
[23]庄越挺,刘小明. 运动图像序列的人体三维运动骨架重建[J]. 计算机辅助设计与图形学学报,2000,12(4): 245-250.
[24]Carsten Stoll, Nils Hasler, Juergen Gall, et al. Fast articulated motion tracking using a sums of Gaussians body model[C]. IEEE International Conference on Computer Vision(ICCV), 2011, 951-958.
[25]Juergen Gall, Carsten Stoll, Edilson de Aguiar, et al. Motion Capture Using Joint Skeleton Tracking and Surface Estimation[C]. Computer Vision and Pattern Recognition, 2009: 1746-1753.
[26]Min Sun, Murali Telaprolu, Honglak Lee, Silvio Savarese. An Efficient Branch-and-Bound Algorithm for Optimal Human Pose Estimation[C]. In CVPR 2012.
[27]Jamie Shotton, Andrew Fitzgibbon, Mat Cook, et al. Real-Time Human Pose Recognition in Parts from Single Depth Images[C]. In CVPR, 2011.
[28]Leonid Pishchulin, Arjun Jain, et al. Articulated People Detection and Pose Estimation: Reshaping the Future[C]. In CVPR 2012.
[29]A.Rosenfeld, J.L.Pfaltz. Sequential Operations in Digital Picture Processing[J]. Journal of the ACM, 1996, 13(4): 471-494.
[30]刘相滨,邹北骥,孙家广. 基于边界跟踪的快速欧氏距离变换算法[J]. 计算机学报,2006,29(2): 317-323.
[31]刘俊涛,刘文予,吴彩华,原亮. 一种提取物体线形骨架的新方法[J]. 计算机学报,2008,34(6): 617-622.
[32]H.Blum. A transformation for Extracting New Descriptors of Shape. MIT Press, 1967, 362-380.
[33]H.Blum. Biological shape and visual science(part I)[J]. Journal of Theoretical Biology, 1973, 38(2): 205-287.
[34]A.Lieutier. Any open bounded subset of Rn has the same homotopy type than its medial axis[J]. Proceedings of 8th ACM symposium on Solid modeling and applications, New York: ACM, 2003: 65-75.
[35]J.Serra. Image Analysis and Mathematical Morphology. Orlando, FL, USA: Academic Press, 1983: 56-57.
[36]C.Menier, E.Boyer, B.Raffin. 3D Skeleton-based Body Pose Recovery[C]. 3rd International Symposium on Data Processing, Visualization and Transmission(DPVT ’06), 2006: 389-396.
[37]R.Ogniewicz, M.lg, E.T.H.Zurich. Voronoi skeletons: Theory and Applications[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1992: 63-69.
[38]W.Gong, G.Bertrand. A Simple Parallel 3D Thinning Algorithm[C]. 10th International Conference on Pattern Recognition, 1990, 1: 188-190.
[39]G.Borgefors, I.Nystrom, Baja.Gs.Di. Computing Skeletons in Three Dimensions[J]. Pattern Recognition. 1999, 32(7): 1225-1236.
[40]C.Theobalt, M.A.Magnor, H.Theisel, H.P.Seidel. Marker-free Kinematic Skeleton Estimation from Sequences of Volume Data[C]. Proceedings of the ACM Symposium on Virtual Reality Software and Technology, 2004: 57-64.
[41]N.D.Cornea, D.silver, X.Yuan, R.balasubramanian. Computing Hierarchical Curve-skeleton of 3D Objects [J]. The Visual Computer, 2005, 21(11): 945-955.
[42]S.M.Yoon, C.Malerczyk, H.Graf. 3D Skeleton Extraction from Volume Data Based on Normalized Gradient Vector Flow[C]. International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision(WSCG), 2009: 177-182.
[43]苏翠萍. 海洋浮游植物显微图像形状骨架化及相似性分析[D]. 厦门大学,2011.
[44]廖志武. 2-D骨架提取算法研究进展[J]. 四川师范大学学报（自然科学版）,2009,32(5): 676-688.
[45]F.P.Preparata, M.I.Shamos. Computational Geometry[J]. Springer-Verlag, New York, 1990.
[46]T.k.Dey, J.Sun. Defining and computing curve-skeletons with medial geodesic function[C]. Proceedings of the 4th Eurograhpics symposium on Geometry processing, Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2006: 143-152.
[47]Tamal K.Dey, Wulue Zhao. Approximating the Medial Axis from the Voronoi Diagram with a Convergence Guarantee[J], Algorithmica, 2003, 38(1): 179-200.
[48]曾荣军. 基于聚类分析的三维网格骨架提取. 中南大学,2011.
[49]F.Leymarie, Martin D.Levine. Simulating the grassfire transform using an active contour model[J]. IEEE

Transaction on Pattern Analysis and Machine Intelligence, 1992, 14(1): 56-75.
[50]Stanley Osher, James A Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations[J]. Journal of Computational Physics, 1988, 79(1): 12-49.
[51]B.B.Kimia, A.R.Tannenbaum, S.W.Zucker. Shapes, shocks, and deformations I: The components of two-dimensional shape and the reaction-diffusion space [J]. International Journal of Computer Vision, 1995, 15(3): 189-224.
[52]车武军,杨勋年,汪国昭. 动态骨架算法[J]. 计算机学报,2003,4(14): 818-823.
[53]李玉海,高敬惠,姜子敬,刘国丽. 基于形态学的视频序列人体骨架提取[J]. 中文核心期刊《微计算机信息》（嵌入式与SOC）,2007,27(8-2): 246-247.
[54]袁俊. 数学形态学理论及其在图像处理中的应用[J]. 武汉理工大学,2007
[55]Gonzalez,R.C. 等著,阮秋琦(译). 数字图像处理（第二版）[M]. 北京: 电子工业出版社,2007.
[56]徐奕信. 基于视频序列的三维人体刚体骨架提取算法研究[D]. 哈尔滨工业大学,2010.
[57]Jen-Hi Chuang, Chi-Hao Tsai, Min-Chi Ko. Skeletonization of Three-Dimensional Object Using Generalized Potential Field [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1241-1251.
[58]A.Torsello, E.R.Hancock. A skeletal measure of 2D shape similary[J]. Computer Vision and Image Understanding, 2004, 95(1): 1-29.
[59]李劝男. 灰度图像骨架的提取及在物体检测中的应用[D]. 华中科技大学,2005.
[60]X.Bai, L.J.Latecki, W.Liu. Skeleton Pruning by Contour Partitioning with Discrete Curve Evolution [J]. IEEE transactions on Pattern Analysis and Machine Intelligence, 2007, 29(3): 449-462.
[61]P.Golland, W.Eric, L.Grimson. Fixed topology skeletons[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA: IEEE Computer Society Press, 2000: 10-17.
[62]Wai-Pak Choi, Kin-Man Lam, Wan-Chi Siu. Extraction of the Euclidean Skeleton Based on a Connectivity Criterion [J]. Pattern Recognition, 2011, 36(3): 721-729.
[63]丁颐,刘文予,郑宇化. 基于距离变换的多尺度连通骨架算法[J]. 红外与毫米波学报,2004,24(4): 281-285.
[64]车武军. 距离变换与中轴变换在变形问题中的应用研究[D]. 浙江大学,2003.
[65]C.Wayne Niblack, Phillip B Gibbons, David W Capson. Generating Skeletons and Centerlines from the Distance Transform [J]. CVGIP: Graphical Models and Image Processing, 1992, 54(5): 420-437.
[66]MPEG. MPEG7 CE Shape-1 Part B, available: http://www.imageprocessingplace.com/root_files_V3/image_databases.htm, 2004.