

基于双目视觉的三维人脸重建方法

林景亮^a, 陈岳林^b

(桂林电子科技大学 a. 机电工程学院; b. 信息科技学院, 广西 桂林 541004)

摘 要: 提出了一种基于双目视觉的三维重建方法, 无须昂贵设备和通用模型。从校正后的图像中提取有效人脸重建区域以降低整体耗时。改进了 Realtime 局部立体匹配算法和基于种子视差约束的区域生长算法, 融合 Realtime 阈值排序和置信度排序进行区域生长, 提高了种子像素提取的可靠性和降低区域生长误匹配的可能性。最后, 研究了纹理映射技术, 提高了重建模型的逼真度。实验结果表明, 该方法能够产生逼真光滑的三维人脸模型。

关键词: 三维人脸模型; 双目视觉; 摄像机标定; 人脸检测; 立体匹配; 区域生长

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1001-3695(2012)01-0397-04

doi:10.3969/j.issn.1001-3695.2012.01.110

3D face reconstruction using binocular vision

LIN Jing-liang^a, CHEN Yue-lin^b

(a. School of Mechanical & Electrical Engineering, b. School of Information & Technology, Guilin University of Electronic Technology, Guilin Guangxi 541004, China)

Abstract: This paper proposed an algorithm to generate 3D face model using binocular vision. It extracted reconstruction region from the rectified images to reduce the system's consuming. Besides, image matching accomplished based on improved realtime matching algorithm, and a region growing algorithm where realtime-threshold and reliability of seed pixels were sorted to guarantee that the sufficient growing pixels be selected and minimize the erroneous of matching possibility. Finally, a texture mapping method was used to enhance the fidelity of the reconstruction model. Experimental result shows that the proposed method can generate a smooth and vivid 3D face model.

Key words: 3D face model; binocular vision; camera calibration; face detection; stereo matching; region growing

0 引言

将三维人脸模型用于人脸识别是计算机视觉领域的一个研究热点。由于人脸的复杂可变性, 三维人脸重建一直是一个非常具有挑战性的课题。现有的三维人脸重建方法可归为两大类, 即直接三维人脸重建和间接三维人脸重建方法。直接三维重建方法又可以分为两种: a) 基于特殊设备的三维人脸重建方法, 如激光扫描仪、结构光、CT; b) 基于普通摄像机图像的三维人脸重建方法, 它只需要一张或者多张人脸图像即可, 此类方法可统一归类为 Shape from X (X 表示运动 motion、立体 stereo、明暗 shading 等)^[1]。间接的三维人脸重建方法其实是基于特殊设备的三维人脸重建方法的改进, 它只需要利用特殊设备建立起一个通用三维人脸模型, 然后利用不同人脸的主要特征点, 如鼻子、眼睛、嘴巴和人脸轮廓线等, 对通用的三维人脸模型进行调整和纹理映射, 从而重建出特定的三维人脸模型^[2]。

第一类的 a) 方法是目前较为常用的三维人脸重建方法, 技术成熟, 如 Beumier^[3] 和 Blanz 等人^[4] 就分别利用结构光和激光扫描仪建立起三维人脸数据库。此类方法有个共同的缺点, 即设备昂贵或者是对使用场合要求很高, 故一直没有得到

广泛应用。第二类方法是对第一类中 a) 方法的改进, 只需要前期重建好一个通用模型即可, 然后再根据特定人脸的特征对通用模型进行调整, 从而得到需要的三维人脸模型, 如 Guan^[2] 和 Zhang 等人^[5] 分别基于 Candide 模型和形变模型进行重建。这种方法在通用人脸模型和被重构的特定人脸相似时取得非常好的效果, 但当它与特定人脸差别较大, 将可能出现非常大的错误^[6]。第一类 b) 方法虽然也是对每个人脸直接重建, 由于它基于普通摄像机图像, 且对使用场合要求低, 使用方便, 近几年来备受关注。国内沈晔湖等人^[7] 和贾贝贝等人^[8] 成功利用立体图对建立出逼真的三维人脸, 国外 Park 等人^[9] 和 Hos-sain 等人^[10] 也发表了类似的研究成果。

本文的方法类似于第一类 b) 方法, 重建的人脸模型主要是针对人脸识别应用。首先搭建立体视觉图像采集平台并利用人脸检测算法从摄像机图像中提取有效的人脸重建区域, 然后结合立体匹配算法和区域生长算法计算人脸图像的视差图, 再基于双目视觉原理从视差图恢复人脸的三维信息, 最后利用 OpenGL 平台进行三维人脸点云重建和纹理映射, 从而得到逼真的三维人脸模型。

1 摄像机标定与极线校正

摄像机标定是指根据空间场景中某一点的实际位置和它

在摄像机中图像位置的几何关系来求取摄像机的内外参数。它是整体系统的基础,不仅可以利用它来反求图像的三维信息,而且利用摄像机参数能够校正摄像机镜头畸变和图像极线,给下阶段的立体匹配算法提供极线约束。

本文基于自建的双目摄像机图像采集与处理平台,即两台 SONY SSC-DC493BP 彩色摄像机、一块为了保持摄像机在同一平面而设计的铁板、一个三脚架(角度可调)、一块天敏 VC404P 图像采集卡和一台 PC 机,利用张正友^[11]的摄像机标定方法进行标定。首先对每个摄像机进行标定求得各自的内参,然后利用两摄像机间的相互关系求取旋转矩阵和平移向量。计算结果如下:

左摄像机参数

焦距: $f_{\text{left}} = [1071.75541 \quad 1066.07323] \pm [9.75465 \quad 9.93934]$
主点: $c_{\text{left}} = [319.75886 \quad 225.50577] \pm [11.27794 \quad 10.04209]$
畸变: $k_{\text{left}} = [-0.43304 \quad 0.05668 \quad 0.00182 \quad -0.00306 \quad 0.00000] \pm [0.06569 \quad 1.04706 \quad 0.00179 \quad 0.00239 \quad 0.00000]$

右摄像机参数

焦距: $f_{\text{right}} = [1034.81116 \quad 1029.68733] \pm [9.56592 \quad 9.73642]$
主点: $c_{\text{right}} = [344.70964 \quad 267.16798] \pm [11.01645 \quad 11.46205]$
畸变: $k_{\text{right}} = [-0.52872 \quad 1.08664 \quad -0.00438 \quad -0.00582 \quad 0.00000] \pm [0.05827 \quad 0.62085 \quad 0.00214 \quad 0.00318 \quad 0.00000]$

两摄像机之间的关系

旋转矩阵 $R = [0.02056 \quad 0.03503 \quad 0.02391] \pm [0.01208 \quad 0.01414 \quad 0.00094]$
平移向量 $T = [-148.67654 \quad -0.52136 \quad -2.30818] \pm [0.95583 \quad 0.63729 \quad 6.48110]$

求得参数后,根据极线几何原理对摄像机图像进行极线校正,实验结果如图 1 所示。其中,图 1(a)是原始摄像机图像,(b)是极线校正后图像,图像中的水平线表示极线。

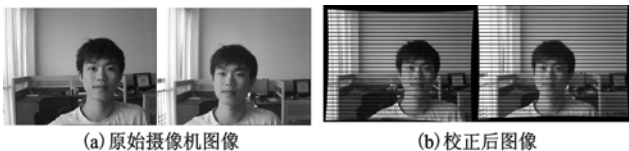


图1 图像校正

2 人脸重建区域提取

在人脸识别中,三维人脸模型的有效区域主要是额头以下下颌以上、左右耳之间的区域,故从摄像机图像中提取人脸重建区域不但可以减少系统的整体耗时,而且可以在一定程度上提高立体匹配阶段的匹配精度。

基于 Haar 特征是目前人脸检测最有效的方法之一,但通过实验发现其存在几个明显的缺点:a)对于背景中与人脸结构类似的物体常被误检为人脸;b)对于有遮挡的不完整人脸常被漏检;c)对于图像像素较低的图像漏检情况非常严重,甚至出现无法检测的现象。这种情况可以用基于肤色模型的方法进行补充,而且图像像素较低时,此类方法依然具有很好的鲁棒性,且对人脸姿态、表情等具有很好的稳定性^[12]。于是文本结合基于肤色模型和 Haar 特征的两种方法的优点,针对三维人脸重建中的特殊应用,提出了一种融合的人脸重建区域提取方法。

算法步骤如下:

- 对输入图像预处理。
- 利用基于 Adaboost 算法的人脸检测方法进行快速检测。

c)对每个检测得到的人脸区域,利用基于肤色模型的方法进行验证,当人脸区域中没有肤色特征或者肤色区域过小,则认为是非人脸区域。假如检测到的某两个或以上的人脸区域有超过 1/4 的重叠区域,则认为是同一个人脸。

d)为了保证检测后的左右人脸区域极线对齐,本文提出了一种基于左图像坐标约束的方法,即检测到左摄像机图像中人脸区域后,将其左上角坐标提取保存并作为右摄像机图像中人脸检测的坐标约束,让它们的区域在 y 轴方向上保持一致。效果如图 2 所示。

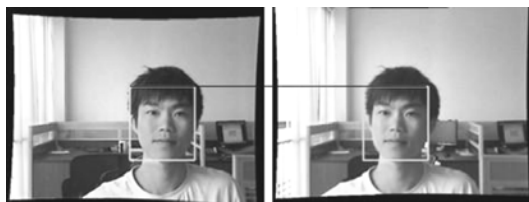


图2 校正后人脸检测实验结果

e)对检测到的人脸肤色区域进行形态学处理,提取更加精简的区域,提取过程如图 3 所示。其中图 3(a)是通过检测算法提取的人脸区域;(b)是基于肤色模型进行相似度计算效果图;(c)是对肤色似然图进行二值化处理效果图;(d)是对二值化图像进行形态学处理效果图;(e)是处理后得到的人脸区域;(f)是提取的最终人脸重建区域。

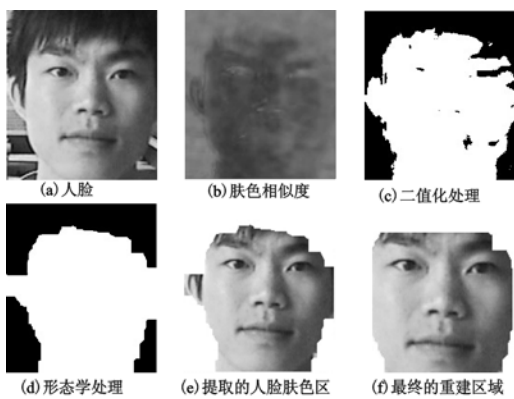


图3 重建区域提取

3 立体匹配与区域生长

3.1 匹配代价计算

匹配代价是整个立体匹配算法的基础,实际是对图像像素灰度值的相似相度量,根据代价来源可以分为基于像素点的匹配代价和基于图像区域的匹配代价。基于像素点匹配代价的有灰度差的平方 SD(squared intensity differences)、灰度差的绝对值 AD(absolute intensity differences)等。基于图像区域的匹配代价有差值绝对值和(sum of absolute differences SAD)、差值平方和(sum of squared differences, SSD)和正交化互相关(normalized cross correlation)NNC 算法,分别见式(1)~(3)。

$$SAD(x, y, d) = \sum_{i=-m}^m \sum_{j=-n}^n |I_L(x+i, y+j) - I_R(x+i+d, y+j)| \quad (1)$$

$$SSD(x, y, d) = \sum_{i=-m}^m \sum_{j=-n}^n [I_L(x+i, y+j) - I_R(x+i+d, y+j)]^2 \quad (2)$$

$$NCC(x, y, d) =$$

$$\frac{\sum_{i=-m}^m \sum_{j=-n}^n [I_L(x+i, y+j) - \overline{I_L(x, y)}][I_R(x+i+d, y+j) - \overline{I_R(x+d, y)}]}{\sqrt{\sum_{i=-m}^m \sum_{j=-n}^n [I_L(x+i, y+j) - \overline{I_L(x, y)}]^2 [I_R(x+i+d, y+j) - \overline{I_R(x+d, y)}]^2}} \quad (3)$$

其中:

$$\overline{I_L(x,y)} = \frac{\sum_{i=-m}^m \sum_{j=-n}^n I_L(x+i,y+j)}{(2m+1)(2n+1)}$$

$$\overline{I_R(x,y)} = \frac{\sum_{i=-m}^m \sum_{j=-n}^n I_R(x+i+d,y+j)}{(2m+1)(2n+1)}$$

3.2 立体匹配与种子像素提取

常用的立体匹配算法可以分为两大类:局部匹配与全局匹配方法^[13]。Realtime 算法是 Hirschmuller^[14]于 2001 年提出的比较有代表性的局部立体匹配算法,是相对于 WTA 算法的一种非常有效的改进,它使用多窗口叠加匹配代价进行匹配,以便在遇到视差不连续的区域可以由邻域的相似性来保证匹配的精度。算法思想如下:

a) 假设 C_0 为中央窗口的匹配代价, C_{li} 为其 8 个周围窗口的匹配代价。对这 8 个值进行排序,取其中最优的 4 个值代入式(4)进行计算,最终得到的 C 值就是中心点的最终叠加匹配代价。

$$C = C_0 + C_{1i1} + C_{1i2} + C_{1i3} + C_{1i4} \quad (4)$$

b) 使用交叉检验法(crosscheck method)进行视差验证,如图 4 所示。两条实线分别表示两图像中的对应极线(经极线校正后),对于参考图像中的一个像素点 A ,其可能的匹配点将落在匹配图的 C 和 D 之间,即 CD 的距离为搜索范围。通过匹配,假设最优匹配代价为 B 点,则 B 为 A 的匹配点。然后以 B 为待匹配点,同样地在参考图中(EF 范围内)搜索最优匹配代价的点;假如是 A 点,则认为 B 点是 A 点的可靠匹配点;假如为其他点,则认为 A 没有可靠的匹配点,如此循环匹配直到参考图中所有的像素点都参加过匹配。

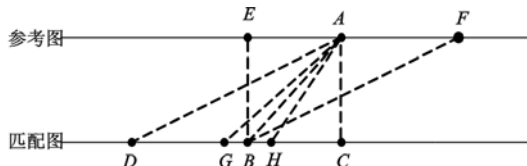


图4 交叉匹配

c) 相似点滤除。如图 5 所示,经过交叉检验后, A 点与 B 点之间的匹配代价为 C_1 , 匹配图中位于 C 、 D 之间的所有点中除去 B 、 G 、 H (G 、 H 为 B 的相邻像素点) 三点外与 A 点的匹配代价最小值为 C_2 。

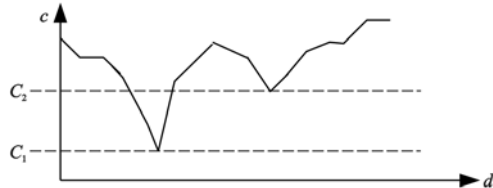


图5 相似点滤除

令匹配代价的相对偏差为 C_d , 则

$$C_d = \frac{C_2 - C_1}{C_1} \quad (5)$$

设定阈值 T_e , 如果 $C_d < T_e$, 则点 B 仍被定为不可靠匹配点, 通过相似点滤除将大大减少这方面的误差, 提高算法的整体精度。

根据算法的思想, 本文从下面两点进行改进:

a) 对式(4), 把相邻像素点和邻域窗口融合作为叠加代

价, 即 C_0 为中央窗口的匹配代价, C_{li} 为 8 个周围窗口的匹配代价, B_{li} 为 8 个邻域像素的匹配代价, 分别对 C_{li} 和 B_{li} 进行排序, 各取其最优的值代入式(4)得

$$C = C_0 + C_{1i1} + C_{1i2} + C_{1i3} + C_{1i4} + B_{1i1} + B_{1i2} + B_{1i3} + B_{1i4} \quad (6)$$

这样, 就可以把区域和特征点的匹配代价很好地融合在一起, 从而提高匹配过程的鲁棒性, 降低误匹配情况。

b) 实验发现, 单使用 SSD 作为匹配代价时, 利用 WTA 算法得到的误匹配区域在利用 Realtime 算法匹配时同样存在, 本文将其称为单匹配代价的“匹配惯性”。也就是无论是 WTA 还是 Realtime 算法, 当它们使用相同的匹配代价时 (SSD), 某些误匹配点会在同一块人脸图像区域存在。于是, 对于 Realtime 算法, 本文在第一次匹配和交叉验证的两次匹配过程使用不同匹配代价的策略, 即初次匹配利用 SSD, 交叉验证过程使用 NCC。实验发现这样的匹配策略能降低误匹配情况, 提高匹配的精度, 算法匹配效果如图 6 所示。

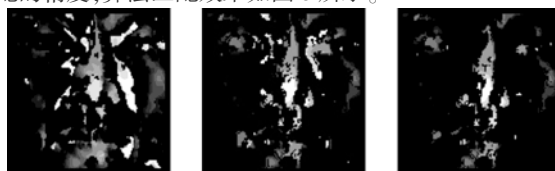


图6 种子视差图

根据视差图原理, 比较可知利用 WTA 算法的种子点视差绘制的视差图 (图 6(a)) 中, 最亮的点并非是鼻子点, 反而是比较平坦的脸颊区域, 从而认定图中比鼻子亮的区域都是误匹配; 而在图 6(b) 中, 利用 Realtime 算法已经很好地显示了人脸的深度差异, 鼻子区域的视差正确计算出来, 但在眼睛边缘区域还存在小部分的误匹配区域; 对于本文改进的算法 (图 6(c)), 鼻子区域和其他区域深度梯度正确, 匹配精度得到了提高, 消除了误匹配区域, 这对于后期的区域生长有着非常重要的意义, 因为误匹配点很可能在区域生长中扩大误匹配区域。

3.3 区域生长

区域生长算法的思想是有利于已经匹配好的一些可靠的种子像素的视差, 在平滑约束或者说种子像素视差约束下, 将其向四周扩张生长至图像的其余部分^[15]。与传统的直接匹配相比, 基于种子像素的视差约束区域生长可以令匹配过程更加快速, 获得的视差图效果更加平滑。目前主要的区域生长算法可以分为基于种子像素约束的区域生长方法^[8]和基于算法的区域生长方法^[16]。本文主要从区域生长的精度和耗时出发, 提出了一种基于阈值和置信度融合约束的立体匹配算法。

a) 首先利用改进的 Realtime 算法进行种子像素提取, 以保证种子像素的可靠性, 与此同时, 保存种子像素的 Realtime 相似点滤除阈值, 并计算其 SSD 匹配代价作为生长置信度。

b) 根据相似点滤除阈值对种子像素进行排序得到一个生长队列, 基于 WTA 算法和种子像素视差约束 ($\Delta d \leq 1$) 进行初次生长 (在生长过程中新的点不用插入到队列中且不用排序, 但要判断其是否与种子点重复, 没有重复则保存下来, 直到没有新增点为止), 同时计算新生长点的匹配代价 (SSD), 并将其保存为置信度。最后, 对其进行置信度排序, 得到一个新的生长队列。

c) 假如新的队列不为空, 选择最优的置信度种子像素作

为最优的生长像素,利用视差约束进行生长,假如新生长的邻域点已经存在,则保存原来的点,重新排序。如果新生长的点为0,则结束生长,输出视差图。

d)删除队列头部种子像素点,重复c),直到没有新的点出现。生长结束后,一些没有匹配点的区域可以通过视差平滑来解决,即对没有匹配的点的像素点,取其相邻点的视差均值作为其视差。

e)进行视差图光滑处理,对视差图的每个像素点进行灰度平滑,从而得到一张光滑完整的人脸视差图像。

4 三维人脸重建实验与分析

求得人脸的完整视差图后,可利用式(7)恢复人脸的三维信息。

$$\begin{cases} X = \frac{Z}{f} \cdot u_1 = \frac{u_1 \cdot b}{d} \\ Y = \frac{Z}{f} \cdot v_1 = \frac{v_1 \cdot b}{d} \\ Z = \frac{b \cdot f}{d} \end{cases} \quad (7)$$

其中: b 代表像素点的视差; f 代表摄像机焦距; d 代表两摄像机间的基线距离。

根据恢复得到的三维点,本文利用 OpenGL(Open Graphics Library)进行人脸的三维点云重建,基于其中的纹理映射技术,利用左摄像机的人脸图像作为重建纹理,成功将二维纹理图像的坐标和三维模型的坐标对应起来,重建出具有真实纹理的逼真模型。重建的三维人脸模型如图7所示。为了便于讨论,文中截取了6张不同角度的三维人脸图像,包括正面(图7(a))、向左旋转(图7(b))、向右旋转(图7(c))、抬头(图7(d))、低头(图7(e))和其他角度图像(图7(f))等。

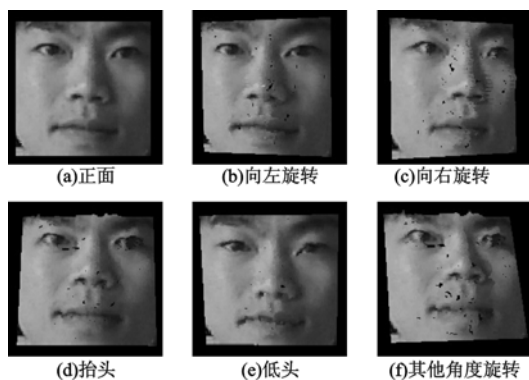


图7 三维人脸重建结果

从三维重建的实验结果可见,正面效果非常好,用于人脸识别时可以直接代替二维人脸图像。此外,通过模型旋转,三维人脸模型在其他角度下识别的优点将是二维图像无法比拟的,因为二维图像没有深度信息。由于在立体匹配阶段为了减少误匹配点,有些置信度过低的点没有找到对应点,而在视差平滑阶段,某些相邻点之间视差相差较大,取其邻域均值后没有达到像素级精度(也就是相邻点视差的差大于1个像素),故当人脸旋转一定角度后,点与点之间的间隙就容易出现,如图7(f)所示的“黑洞”部分。

5 结束语

本文基于普通摄像机图对,成功利用改进的 Realtime 局部

立体匹配算法和区域生长算法获取人脸视差图,然后根据双目视觉原理重建出逼真的三维人脸模型。针对三维人脸模型在人脸识别领域的应用,利用融合的人脸检测算法从摄像机图像中提取有效的人脸重建区域,降低了重建系统的整体耗时,并在一定程度上提高了重建模型的精度。未来将对立体匹配算法和区域生长算法作进一步改善,以消除重建后出现的“黑洞”现象。

参考文献:

- [1] 潘纲. 三维人脸识别若干技术研究[D]. 杭州:浙江大学,2003.
- [2] GUAN Ye-peng. Automatic 3D face reconstruction based on single 2D image[C]//Proc of International Conference on Multimedia and Ubiquitous Engineering. 2007: 1216-1219.
- [3] BEUMIER C, AEHEROY M. Automatic 3D face authentication[J]. Image and Vision Computing, 2000, 18(4): 315-321.
- [4] BLANZ V, VETTER T. A morphable model for the synthesis of 3D faces[C]//Proc of the 26th Annual Conference on Computer Graphics and Interactive Techniques. 1999: 187-194.
- [5] ZHANG Yu, PRAKASH E C, SUNG E. Hierarchical modeling of a personalized face for realistic expression animation[C]//Proc of the IEEE International Conference on Multimedia and Expo. 2002: 457-460.
- [6] CHOWDHURY A R, CHELLAPPA R, VO T, et al. 3D face reconstruction from video using a generic model[C]//Proc of IEEE International Conference on Multimedia and Expo. 2002: 449-452.
- [7] 沈晔湖,刘济林. 利用立体图对的三维人脸模型重建算法[J]. 计算机辅助设计与图形学学报, 2006, 18(12): 1904-1910.
- [8] 贾贝贝,阮秋奇. 双目立体视觉的三维人脸重建方法[J]. 智能系统学报, 2009, 4(6): 513-520.
- [9] PARK U, JAIN A K. 3D face reconstruction from stereo video[C]//Proc of the 3rd Canadian Conference on Computer and Robot Vision. 2006: 41-42.
- [10] HOSSAIN M S, AKBAR M, STARKEY J D. Inexpensive construction of a 3D face model from stereo images[C]//Proc of the 10th International Conference on Computer and Information Technology. 2007: 1-6.
- [11] ZHANG Zheng-you. A flexible new technique for camera calibration[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [12] 雷温奇,宋晓冰,袁美玲,等. 双目视觉中一种人脸立体匹配及视差计算方法[J]. 厦门大学学报:自然科学版, 2009, 48(1): 36-41.
- [13] SCHARSTEIN D, SZELISKI R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. International Journal of Computer Vision, 2002, 47(1/2/3): 7-42.
- [14] HIRSCHMULLER H. Improvements in real-time correlation-based stereo vision[C]//Proc of IEEE Workshop on Stereo and Multi-Base-line Vision. Washington DC: IEEE Computer Society, 2001: 141-148.
- [15] LHUILLIER M, QUAN L. Quasi-dense reconstruction from image sequence[C]//Proc of the 7th Europe Conference on Computer Vision. 2002: 125-139.
- [16] 傅有成,达飞鹏. 一种基于立体匹配的三维人脸重建方法[C]//2008年全国模式识别学术会议 CCPR. 2008.