

基于双目视觉的人脸视差快速获取方法

刘丹华, 赖海滨, 王晓晓

(厦门大学 计算机科学系, 福建 厦门 361005)

摘要: 提出一种在平行双目视觉系统下快速获取致密的人脸视差图的方法。首先对获取的双目图像进行极线校正和人脸的精确定位,降低立体匹配的搜索范围;然后针对人脸这一特殊场景,采用基于盒状滤波的快速区域立体匹配算法对人脸区域进行匹配,得到人脸的初始视差图;最后利用金字塔匹配模型来获取高分辨率的人脸致密视差图。实验结果表明,该方法在保证所获取人脸视差精度的前提下,大大降低了时间复杂度,有较好的实用价值。

关键词: 双目视觉;人脸定位;立体匹配;快速区域匹配;金字塔模型;人脸视差图

中图分类号: TP391 **文献标识码:** A **文章编号:** 1009-3044(2012)07-1586-04

Fast Method of Face Disparity Map Acquisition Based on Binocular Vision

LIU Dan-hua, LAI Hai-bin, WANG Xiao-xiao

(Department of Computer Science, Xiamen University, Fujian Xiamen 361005, China)

Abstract: A fast method of acquiring dense facial disparity map based on binocular vision is presented. Firstly, in order to reduce searching region of stereo matching, the binocular images are processed by polar rectification and face area location. Secondly, according to the special scene of face, a fast region matching algorithm with box filtering is applied to gain an initial face disparity map. Lastly, a dense disparity map with high resolution is acquired by using pyramid matching model. Experimental results show that the proposed method reduces the time complexity greatly on the premise of acquiring dense facial disparity map, and thus to have a high application value

Key words: binocular vision; face location; stereo matching; fast region-matching; pyramid model; face disparity map

1 概述

双目立体视觉技术^[1]通过两幅对应的二维图像信息来重建三维数据,被广泛应用于机器人自动导航、地理信息系统、生物医学成像以及虚拟现实等领域,是计算机视觉的重要研究内容,由于其获取成本低廉,尤其适用于对人脸模型精度要求不是特别高的场合。

立体匹配是双目立体视觉系统中用来获取视差图的一个重要环节,它的本质是给定一幅图像中的一点,寻找另一幅图像中的对应点,使得这两点为空间同一物体点的投影,进而求出二者的视差,以恢复场景的深度信息,并进一步求得该点的空间坐标。匹配结果会受到许多因素的影响,如,光照因素、镜头畸变、物理特性、噪声干扰等^[2]。目前研究的立体匹配算法大部分都是针对复杂的空间场景获取大致的视差图,专门研究人脸立体匹配的相关文献还较少。本文旨在研究匹配精度高、计算速度快的人脸立体匹配方法,从而获得高分辨率的人脸视差图,为后期人脸三维重建提供良好的基础。

与一般的立体场景不同,人脸有着自身的特点,如表面变化相对平缓、场景信息不够丰富、深度信息差异不大、肤色等,这些特点也就决定了要获取可用的人脸视差图需要采用特定的立体匹配方法。本文首先采用文献^[3]中提到的方法精确提取左右图像的人脸区域,以减少视差搜索的范围。其次采用文献^[4]提出的改进区域立体匹配算法对人脸区域进行立体匹配,得到人脸的初始视差图。最后,为了获取高分辨率人脸视差图,在金字塔匹配模型^[5]基础上,引入盒状滤波^[6]方法对时间进行优化,大大降低了时间复杂度。

2 人脸区域精确定位

在进行立体匹配之前,通过极线校正和人脸的定位,一方面缩小了视差的搜索范围,另一方面也指导了匹配的方向,利于提高匹配精度。本文的极线校正方法是通过右图进行像素的上下调整,使左右图片像素处于同一个水平线上。对人脸的定位主要通过两个步骤,第一步是人脸区域粗定位,本文采用了由 Paul Viola^[7]提出的基于 AdaBoost Cascade 算法。第二步是人脸区域精确定位,其依据主要是通过肤色检测滤除头发、首饰等非肤色区域的干扰,并通过形态学的处理方法形成精确的人脸连通区域,该区域即是本文中进行立体匹配的区域。肤色检测的方法采用了基于 YCbCr 颜色空间的肤色聚类算法^{[6][7][8]}。图 1 显示了部分精确定位的人脸区域模板图。

收稿日期:2012-02-04

作者简介:刘丹华(1987-),女,福建永春人,硕士,主要研究方向为计算机视觉和图像识别;赖海滨(1985-),男,福建平和人,硕士,主要研究方向为计算机视觉和图像识别;王晓晓(1986-),女,河南孟州人,硕士,主要研究方向为图像处理和人工智能。



图1 人脸区域模板图

3 人脸区域匹配算法

3.1 初始视差图获取

传统的基于区域的立体匹配算法,是对匹配图中的每一个特征基元,对目标图中所有视差搜索范围内的特征基元计算相似性,并从中选择最优的匹配基元,其衡量标准为相似性度量算子的值达到最大或最小。常用的相似性度量算子有像素差绝对值和(SAD)、像素差平方和(SSD)、零均值归一化互相关(ZNCC)等。实验表明^[1],ZNCC算子的抗噪能力优于SAD和SSD算子。本文在立体匹配时采用了ZNCC算子作为相似性度量算子,它定义如公式(1)。

$$ZNCC(x,y,d) = \frac{\sum_{i,j \in W} [(I_l(x+i,y+j) - \bar{I}_l)(I_r(x+i+d,y+j) - \bar{I}_r)]}{\sqrt{\sum_{i,j \in W} (I_l(x+i,y+j) - \bar{I}_l)^2 \sum_{i,j \in W} (I_r(x+i+d,y+j) - \bar{I}_r)^2}} \quad (1)$$

其中ZNCC(x,y,d)表示分别以左图像素点(x,y)和右图像素点(x+d,y)(d为当前视差值)为中心点时的区域相关值,W为匹配窗口, \bar{I}_l 和 \bar{I}_r 是待匹配的两个区域中的所有像素灰度平均值。

直接运用传统的区域匹配方法进行人脸视差图获取不仅速度低,而且误匹配点多。文献^[3]考虑人脸的特性,只针对精确定位后的人脸区域进行立体匹配,并设定一个视差基准值,缩小了视差搜索范围。另外,对该人脸区域做了高斯滤波、灰度均衡化和图像锐化,在降低噪声和亮度影响的同时保持图像边缘的细节,并引入了左右一致性验证,减少了误匹配点数。

图2为利用文献^[3]方法获取的初始视差图。同传统的区域匹配方法相比,文献^[3]方法能有效提高匹配速度,并且可以获得较精确的初始视差图,从而能有效指导后期高分辨率的视差图的获取。

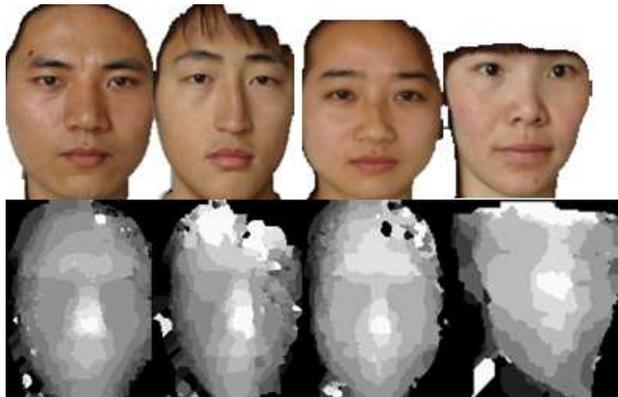


图2 文献[3]获取的初始视差图

3.2 金字塔匹配模型

为了获取高分辨率的人脸视差图,本文采用了金字塔匹配模型^[10]。金字塔结构(Pyramid)是立体匹配中的一种加速技巧。在进行立体匹配时,先在低分辨率图像上进行粗匹配,之后将匹配结果作为下一层图像的视差预测初值,以此来指导该层的匹配,依次逐级处理,直到获取底层的视差图。金字塔匹配模型减少了高分辨率图像匹配时的视差搜索范围,同时也降低了误匹配发生的概率。

本文采用了平均金字塔匹配模型,即高一层图像某个位置的像素值是由低一层图像对应的2x2区域像素求平均值得到的。如图3所示。具体的匹配步骤如下,以左图匹配右图为例:

- (1) 建立N层图像金字塔,本文取N=4。依次为第3层(528x396)、第2层(1056x792)、第1层(2112x1584)、第0层(4224x3168,原始图像大小)。
- (2) 利用3.1法获取顶层视差图。
- (3) 对于第M层视差图(2 ≤ M ≤ 0,非顶层),其视差计算由第M-1层指导。具体如下:假设对于第M-1层某一位置(x,y),视差为d,它在第M层映射2x2的区域,这个区域中每一位置的初始视差均为2xd,其在右图的视差搜索范围为(2xd-t, 2xd+t),t为某个预设值。依此类推,直至生成第M层视差图。
- (4) 对于M=0层视差图,即为最终的视差图。该视差图具有最高分辨率。本文获取的底层视差图分辨率约为100万个像素左右。

图4为在经过金字塔匹配后得到的各层视差图,第一列为人脸模板图,从左到右依次为顶层到底层的人脸视差图,由实验结果得,经过金字塔匹配后得到的底层视差图精度比较高,也比较平滑,跟真实视差比较接近。

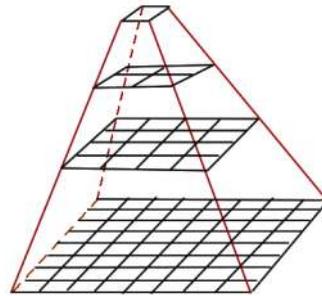


图3 平均金字塔模型

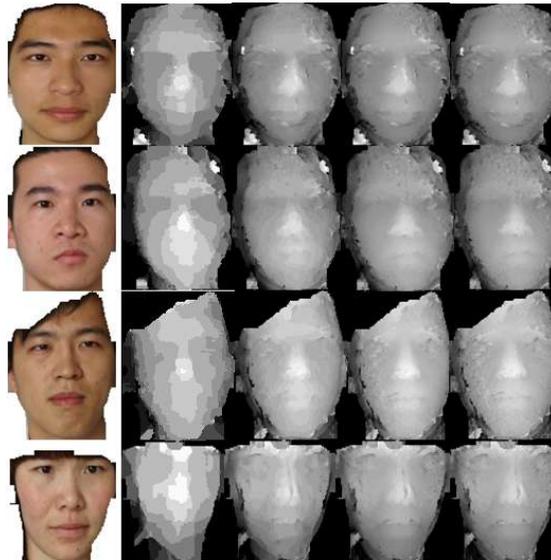


图4 金字塔匹配后的各层视差图

3.3 快速区域立体匹配

虽然3.1节对立体匹配的速度和精度方面做了一些改进,又引入了金字塔匹配模型,底层的人脸视差图分辨率达到了100万像素左右,在精度上已经满足后续特征提取的需要。但随着金字塔层数的增多,时间复杂度呈指数级增长,难以满足实时系统的要求。因此本文引入了盒状滤波方法来降低相似性度量算子的计算复杂度。

传统的相似性度量算子如SAD、SSD、ZNCC存在着大量的冗余运算,且随着匹配窗口的增大,这种冗余运算所带来的时间复杂度体现得越明显。结合文献^[41],本文推导了ZNCC算子的快速算法。

假定匹配窗口 W 大小为 N ($n \times n$), $S_1(x, y)$ 代表窗口 N 内像素灰度值的和, $S_2(x, y)$ 代表窗口 N 内像素灰度值平方的和, $S_{12}(x, y, d)$ 代表左图以点 (x, y) 为中心的窗口与右图以点 $(x + d, y)$ 为中心的窗口对应像素灰度乘积的和。具体的表达式如下所示:

$$S_1(x, y) = \sum_{i, j \in W} I(x + i, y + j) \tag{2}$$

$$S_2(x, y) = \sum_{i, j \in W} (I(x + i, y + j) * I(x + i, y + j)) \tag{3}$$

$$S_{12}(x, y, d) = \sum_{i, j \in W} (I_l(x + i, y + j) * I_r(x + i + d, y + j)) \tag{4}$$

将公式(1)的分子、分母展开,并代入 $S_1(x, y)$ 、 $S_2(x, y)$ 、 $S_{12}(x, y, d)$ 后,ZNCC (x, y, d) 可化为下式:

$$ZNCC(x, y, d) = \frac{S_{12}(x, y, d) - \frac{S_{1l} * S_{1r}}{N}}{\sqrt{(S_{2l} - \frac{S_{1l}^2}{N})(S_{2r} - \frac{S_{1r}^2}{N})}} \tag{5}$$

由上式可知,只要求出 $S_1(x, y)$ 、 $S_2(x, y)$ 、 $S_{12}(x, y, d)$, 即可求出 ZNCC (x, y, d) 的值。下面以 $S_1(x, y)$ 的求解为例,窗口大小设为 5×5 , 推导其快速求解方法。

$$S_1(x + 1, y) = S_1(x, y) + \sum_{j=-n}^n [I(x + n + 1, y + j) - I(x - n, y + j)] = S_1(x, y) + U_{S1}(x + 1, y) \tag{6}$$

其中,

$$U_{S1}(x + 1, y) = U_{S1}(x + 1, y - 1) + I(x + n + 1, y + n) - I(x - n, y + n) - I(x + n + 1, y - n - 1) + I(x - n, y - n - 1) \tag{7}$$

从而有,

$$S_1(x + 1, y) = S_1(x, y) + U_{S1}(x + 1, y - 1) + I(x + n + 1, y + n) - I(x - n, y + n) - I(x + n + 1, y - n - 1) + I(x - n, y - n - 1) \tag{8}$$

当已知 $S_1(x, y)$ 的值时,求解 $S_1(x + 1, y)$ 的时间复杂度由 $O(n^2)$ 变成常数时间复杂度,所需的额外操作只需要对图像的第一行和第

一列按常规方法计算求得 $S_1(x, y)$ 的值,从而获得迭代变量 $U_{s_1}(x, y)$ 的初值 $U_{s_1}(x, 0)$,如图5所示,深蓝色框为 $S_1(x+1, y)$ 对应的 5×5 窗口,四个椭圆标记的框代表由 $S_1(x, y)$ 计算 $S_1(x+1, y)$ 需要用到的4个值。同理,我们可以推出 $S_2(x, y), S_{12}(x, y, d)$ 的快速计算公式。

假设图像大小为 $M \times M$,利用传统ZNCC公式计算整幅图像视差的时间复杂度为 $O(M^2n^2)$,而利用改进的ZNCC公式计算整幅图像视差的时间复杂度为 $O(Mn^2+4M^2)$,相比原来的方法降低了一个等级。

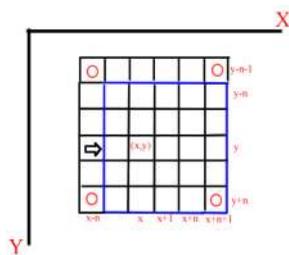


图5 盒状滤波图

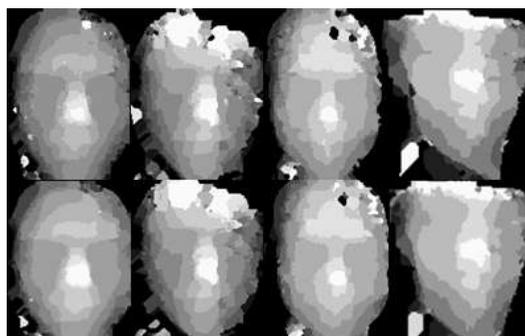


图6 文献[3]与本文算法的初始视差图

图6给出快速区域立体匹配与文献^[3]立体匹配的效果对比,第一排为文献^[3]获取的人脸顶层视差图,第二排为快速区域立体匹配算法获取的人脸视差图。实验结果表明快速区域立体匹配在精度上不仅没有降低,还有所提高。

在速度方面,为了定量验证采用该改进公式之后计算性能的改变,下面以计算左图顶层金字塔图像(528×396)人脸区域中每个肤色点在右图中的前5个最相似匹配点总耗时为例,并变换匹配窗口大小,对比改进前后两者计算速度上的差异。实验数据结果见表1:

表1 ZNCC算子改进前后计算时间复杂度 单位:s

匹配窗口大小	传统 ZNCC 算法	ZNCC 加速算法
9×9	1.36	0.156
11×11	1.89	0.165
13×13	2.67	0.172
15×15	3.48	0.188

(1) 与传统的ZNCC公式相比,利用ZNCC加速公式的时间复杂度下降了很多,且随匹配窗口大小的增大而增大的幅度也比较低。

(2) 随着计算区域的增大,即对于越底层的金字塔(相应计算区域也越大)ZNCC加速公式的优势将越来越明显;

(3) 当采用左右一致性验证准则时,如果应用ZNCC加速公式,由于在正向匹配(左图匹配右图)时已经保存了部分逆匹配(右图匹配左图)时所需要计算的值,这部分不需要冗余计算,在这种情况下,将进一步降低时间复杂度。

为了体现本文提出的快速区域立体匹配算法时间上的优势,表2给出了文献^[3]和本文算法在同台主机下获取各层视差图后耗费的总时间情况,本文算法大大降低了时间复杂度,较能满足实时系统的性能要求。

表2 文献[3]和本文算法获取各层视差图耗费的总时间 单位:s

分辨率大小	文献[3]算法	本文算法
912×1256	87.54	13.172
1016×1416	99.92	16.359
904×1208	75.30	11.953
768×778	44.91	6.516

4 结论

本文以人脸的三维重构为背景,研究如何快速获取高分辨率的人脸致密视差图,利用极线校正和人脸精确定位减少视差搜索

(下转第1638页)

用户对扩展资源学习的时间越长,表明当前资源的学习能给前一资源带来进一步的补充学习,因此两资源的关系就越近。

4)得到 θ 后便可以对权值 k 进行维护,权值的计算方法如公式二:

$$\text{公式二: } k = \begin{cases} k + 1; (\theta \geq 1) \\ k + 0.5; (1 \geq \theta \geq 0.5) \\ k; (\theta \leq 0.5) \end{cases}$$

在个性化VOD点播系统中,资源之间因相似程度形成了一张网状的图。

比起系统管理员手工对相似资源关联归类 and 通过检索系统找出资源描述相似的资源,建立在认可率上的资源相似度—权值表示法能提供更高的精度,而且不需要管理员的参与。推荐时,按关系“亲密程度”降序排序检索推荐。

3.2 个性化同伴推荐

学习者在学习的过程中可能会遇到疑难,需要他人协助来进行问题解决。谁最能学习者带来帮助,这是个性化同伴推荐要解决的。个性化同伴推荐的思路是:系统保存之前对该资源进行学习的学习者的测验结果,系统把测验结果按高低排序,把测验结果好的其他学习者推荐给当前学习者。这需要系统保存学习者的用户信息和联系方式。“向学习好的同学问问题”这也是我们平时学习中能最快解决学习问题的一种行之有效的方法。实现方法是对自主学习者和该资源测验结果建立它们各自的“实体—属性”关系:

自主学习者= \langle 学习者唯一标识符,联系方式,其它信息 \rangle

资源测验结果= \langle 资源唯一标识符,学习者唯一标识符,测验结果, \dots \rangle

这也就可以通过测验结果找到对应学习者的联系方式,联系方式可以是邮箱、电话、QQ等工具。

4 结束语

本文尝试在基于VOD点播技术的学习系统中,通过向学习者推荐合适自己的学习资源以及能起到帮助的学习同伴,使学习者能快速有效的学习。对于数字化学习平台的探索和研究有助于更好的支持学习。

参考文献:

- [1] 庄捷.流媒体原理与应用[M].北京:中国广播电视出版社,2007:5-41.
- [2] 鄂大伟.多媒体技术基础与应用[M].北京:高等教育出版社,2002:20-22.
- [3] 曾春,邢春晓,周立柱.个性化服务技术综述[J].软件学报,2002,13(10):1-2.
- [4] 舒蓓,申瑞民,王加俊.个性化的远程学习模型[J].计算机工程与应用,2001,37(9):2-6.
- [5] 李书明,田俊.网络学习中个性化学习服务策略研究[J].中国电化教育,2011(6).
- [6] 李广,姜英杰.个性化学习的理论建构与特征分析[J].东北师大学报:哲学社会科学版,2005(3).
- [7] Armstrong T. Multiple intelligence in the classroom[M]. Virginia:Alexandra,1994.
- [8] 王永固,邱飞岳,赵建龙,刘晖.杂志基于协同过滤技术的学习资源个性化推荐研究[J].远程教育,2011,18(13):5-8.

(上接第1589页)

范围,引入盒状滤波方法对区域立体匹配算法进行时间上的优化,最后利用金字塔模型获取致密的人脸视差图,并用实验证明了该算法在保证获得高分辨率人脸致密视差图的同时,大大降低了时间复杂度,能满足实时系统的性能要求,为后期进行三维人脸重建打下良好的基础。

参考文献:

- [1] 党乐.基于双目立体视觉的三维重建方法研究[D].西安:长安大学,2009.
- [2] 雷蕴奇,宋晓冰,等.双目视觉中的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.
- [3] 李庆敏.多特征融合的三维人脸识别技术研究[D].厦门:厦门大学,2010.
- [4] Changming S. A fast stereo matching method[C]. In Digital Image Computing: Techniques and Applications, 1997:95-100.
- [5] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]. IEEE Conf on Computer Vision and Pattern Recognition, Kauai, Hawaii, USA: IEEE Computer Society, 2001, 1(1):511-518.
- [6] 徐战武,朱森良.基于颜色的皮肤检测综述[J].中国图像图形学报,2007,12(3):377-388.
- [7] 陈锻生,刘政凯.肤色检测技术综述[J].计算机学报,2006,29(2):194-207.
- [8] 吴绿芳.双目视觉中的人脸三维数据获取研究[D].厦门:厦门大学,2008.
- [9] 斯蒂格,威德曼,尤里奇.机器视觉算法与应用[M].北京:清华大学出版社,2008.
- [10] 谢耀华,汤晓安,孙茂印.不产生精度截断及数据膨胀的图像金字塔[J].中国图像图形学报,2009,14(6):1070-1074.
- [11] Stefano L. D, Marchionni M, Mattocchia S. A fast area-based stereo matching algorithm[J]. Image Vis. Compt., 2004, 22:983.