

文章编号:1007-2780(2010)04-0593-05

基于特征点匹配的多视图图像校正

程明明, 王 贺, 安 平*, 张 洋, 张兆杨

(上海大学 通信与信息工程学院, 上海 200072, E-mail: cmm@shu.edu.cn)

摘 要: 多视图图像的校正是提高 3DTV 观看质量的关键步骤之一。现有的图像校正算法大多是双目或三目相机校正, 对于平行或弧形相机阵列的多视图图像校正的研究较少, 而且受相机标定参数和相机间距的影响较大。文章提出了一种对未标定的多目相机图像进行校正的方法, 通过寻找匹配特征点, 依据匹配点垂直视差的关系计算校正变换矩阵, 用以对原图像进行校正变换。实验结果显示, 图像间的垂直失配大为减少, 校正后的图像能够提供良好的立体视觉主观测试效果。

关 键 词: 3DTV; 图像校正; 未标定图像; 特征点匹配

中图分类号: TN27 **文献标识码:** A

Multi-View Images Rectification Based on Feature Points Matching

CHENG Ming-ming, WANG He, AN Ping*, ZHANG Yang, ZHANG Zhao-yang

(School of Communication and Information Engineering, Shanghai University,
Shanghai 200072, China, E-mail: cmm@shu.edu.cn)

Abstract: The rectification of multi-view images is one of key techniques to improve the quality of 3DTV viewing. Most of the existing image rectification algorithms are for stereo or triple images, but fewer for images captured by parallel or arc multi-camera arrays, and subject to camera calibration parameters and camera spacing greatly. This paper presents an uncalibrated rectification algorithm for multi-view images captured by a multi-camera array. Without camera calibration, the rectification transformation matrix was calculated by searching for matched feature points and using the relationship of vertical parallax, and by which to rectify the original image. Experimental results show that the vertical mismatch between rectified images is greatly reduced, and the rectified images can provide good stereo vision of the subjective test.

Key words: 3DTV; image rectification; un-calibrated image; feature points matching

1 引 言

3DTV(Three Dimension Television)^[1]是继数字高清晰度电视之后的下一代电视技术,其特点是使人们在观看视频节目时具有立体感和沉浸感。3DTV 还能够使得观看者在一定范围内自由

地切换观看的视点和视向。这种多视点功能主要源于其采集端应用多相机阵列,能采集到多个视点的视频序列。

相机阵列的形式可以分为多种,应用较多的是平行相机阵列和弧形相机阵列。理想状态下,这两类阵列所拍摄的图像,各个视点之间应该只

收稿日期: 2010-02-06; 修订日期: 2010-06-24

基金项目: 国家自然科学基金(No. 60832003);上海市教委重点项目(No. 09ZZ90)

作者简介:程明明(1984—),男,湖北黄梅人,硕士研究生,主要研究方向为多视点视频编码及图像校正。

* 通讯联系人, E-mail: anping@shu.edu.cn

有均等的水平视差,而没有垂直视差。但是在实景拍摄时,由于相机物理特性的差异以及手动设置相机位置的误差,使得采集到的图像仍然存在垂直视差,严重影响了系统后端合成的立体图像的质量^[2]。因此,对采集到的多视点图像还须进行图像校正,以减少垂直视差,为后端合成提供良好条件。

根据采集的视点数目可以将图像校正分为双视几何校正、三视几何校正和多视几何校正 3 类。近年来多视点图像校正技术得到了快速的发展,已有多种校正方法,但大多是双目或三目校正方法。如利用极线或单应性矩阵的双目校正方法^[2-4]; Sun Chang-ming 在文献^[5]中以三视张量为基础,利用投影不变性提出的校正三目相机的算法等。

对于由摄像机阵列拍摄的多视图图像校正方面的研究还比较少。如果将双目立体校正方法直接扩展到多视点图像,因为没有同时考虑多个视点的图像,没有统一的基准图像,所以校正效果不佳,韩国的 Yun-Suk Kang 和 Cheon Lee 介绍了一种平行相机多视点图像的校正方法^[6],通过迭代计算公共基线,将原本非共线分布的多相机光心统一投影到一条公共基线上,解决了多视点图像同时校正时基线不一致问题,整体上消除了多视点图像间的垂直失配。相机阵列可以水平或弧形放置,经研究实践证明,弧形摆放相机阵列拍摄的图像立体视觉效果更好。因此,同年他们又提出了一种几何补偿的弧形多相机阵列图像校正方法^[7],通过计算出理想相机的标定参数,然后利用基于平面的单应性变换获得新相机图像。但是该方法在相机间距较大的情况下,校正效果不佳;且其校正准确度受标定参数影响较大,不能完全拟合理想弧形。

为解决上述问题,本文提出了一种利用特征点匹配的无需事先相机标定的校正方法,依据匹配点垂直视差的关系计算校正变换矩阵,用以对原图像进行校正变换,从而减少图像间的垂直失配,校正后的图像可合成质量良好的立体视觉图。

2 未标定多视图图像校正方法

2.1 水平视差与垂直视差

水平视差由相机间距及夹角决定,垂直视差由光心位置的高低决定,因此理想情况下图像间

无垂直视差且水平视差均等。但在实际情况下,人工摆放的相机阵列并不满足理想条件,存在相机位置和方向的失准;此外,由于相机的物理特性各异,所以即使正确摆放,也不能得到理想的图像。以上这些原因导致了多视点图像间对应点位置的垂直失配以及水平视差的不均匀,使得合成的三维图像在显示时物体边缘模糊、立体效果差并会使观看者有眩晕感。因此,在对多视点图像进行立体视觉合成之前必须首先对其进行校正。图像校正的一个重要目的就是减少图像间对应匹配点的垂直视差,并使得图像间水平视差均匀分布。

2.2 校正变换矩阵

对图像进行校正,就是计算图像中每个像素在校正后图像中的新坐标的过程,可通过变换矩阵来实现。假设原图像中的一点 m ,经过校正后为 m' ,满足 $m' = Hm$,则 H 为原图像的校正变换矩阵。如果已知两幅图像一系列对应匹配点的坐标,那么由线性或 MLE(Maximum likelihood estimate)非线性算法可得到两幅图像间的单应性矩阵^[8]。

为消除垂直视差,本文提出了一种计算图像校正变换矩阵的方法。首先,选取两个邻近相机的图像(图像 1 和 2),以其中一相机的图像(图像 1)为基准(该图像不做校正),寻找其与邻近相机图像间的匹配特征点,如基准图像 1 的点 (x_1, y_1) 对应邻近图像 2 的点 (x_2, y_2) 。理想情况下,该两点应无垂直视差,因此假设图像 2 对应图像 1 的理想匹配点坐标为 (x_2, y_1) 。以此类推寻找一系列匹配点,从而得到一系列理想匹配点。然后根据图像 2 的特征点坐标和理想匹配点坐标可得到图像 2 校正前后图像之间的单应性矩阵,即为图像 2 的校正变换矩阵。

假设两幅图像间的匹配点坐标分别为:
$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{bmatrix} \text{ 和 } \begin{bmatrix} x'_1 & x'_2 & \cdots & x'_n \\ y'_1 & y'_2 & \cdots & y'_n \end{bmatrix},$$
前者为基准图像的点,则得到后者校正前后理想的匹配关系为:
$$\begin{bmatrix} x'_1 & x'_2 & \cdots & x'_n \\ y'_1 & y'_2 & \cdots & y'_n \end{bmatrix} \text{ 和 } \begin{bmatrix} x'_1 & x'_2 & \cdots & x'_n \\ y_1 & y_2 & \cdots & y_n \end{bmatrix},$$
由这一组理想匹配点,通过线性算法即可计算出校正前后图像间点的匹配关系,即单应性矩阵 H ,

满足:
$$\begin{bmatrix} x'_n \\ y_n \\ 1 \end{bmatrix} = \begin{bmatrix} x'_n \\ y'_n \\ 1 \end{bmatrix} \cdot H, \text{ 其中 } H \text{ 为 } 3 \times 3 \text{ 的方阵。}$$

利用求得的 \mathbf{H} ,对原图像每点的坐标进行变换,得到其在校正后图像中对应的坐标,即完成图像的校正。假设原图像坐标为 $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$,则校正后的坐标为 $\begin{bmatrix} x_r \\ y_r \end{bmatrix}$,且满足 $\begin{bmatrix} x_r \\ y_r \\ 1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} \cdot \mathbf{H}$ 。校正后图像中没有坐标对应的像素值,可由其邻近像素值插值产生。

2.3 校正过程

由于一次只能实现两幅图像的校正,因此对于多视图图像,首先必须确定校正的顺序。本文的测试序列为弧形摆放的 8 相机阵列获取。先以相机 5 为基准,因为该相机位于场景中间,包括场景大部分内容,由相机 5 得到相机 4 和 6 的校正变换矩阵;再以校正后的相机 4 和 6 为基准,分别得到相机 3、2、7 和 8 的校正变换矩阵;最后再以校正后的相机 2 为基准得到相机 1 的校正变换矩阵。这样就保证了所有的相机都是以相机 5 为基准,光心高度都与相机 5 的一致,但间距仍保持不变。如图 1 所示。

校正顺序确定后,就开始进行两两校正,首先需要寻找两幅图像的匹配点。我们采用计算机视觉领域中广泛使用的 Harris 算法来提取匹配点坐标,因其计算简单、稳定性强^[9]。采用此方法寻找图 1 所示各相机对的匹配点(其中相机 2、4 和 6 为其校正后的图像)。

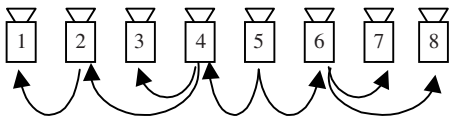


图 1 校正顺序示意图
Fig. 1 Rectification order diagram

得到相机图像对间的匹配点坐标后(本文提取了 8 对匹配特征点坐标),由这些坐标进而确定 8 对理想匹配点坐标,得到 8 对匹配关系。然后采用文献^[8]中的线性方法计算出单应性矩阵,也即为待校正图像的校正变换矩阵。

最后用得到的校正变换矩阵,对原图像进行图像变换就可获得校正后的图像。依据前面确定的校正顺序依次对各图像进行校正,最终实现 8 幅图像的校正。

3 实验结果分析

对校正效果的评价一般可采用下面几种方法:一是极线几何关系,校正后图像间的极线应相互平行;二是匹配点之间的垂直视差;三是评价函数^[10]。后两种越小越好。其中评价函数定义如下:

$$f(\mathbf{H}_1, \mathbf{H}_2) = \frac{1}{n} \sum_{i=1}^n |[(\mathbf{H}_2 \mathbf{m}_2) \mathbf{F} (\mathbf{H}_1 \mathbf{m}_1)]| \quad (1)$$

式中 \mathbf{F} 为理想基本矩阵, \mathbf{m}_1 和 \mathbf{m}_2 为两图像的匹配点, \mathbf{H}_1 和 \mathbf{H}_2 分别为两图像的校正变换矩阵, n 为匹配点数目。

本文先采用微软 breakdancers 序列的一组图像作为测试图像,为 2D 弧形摆放的 8 个间距为 20 cm 的相机阵列。图 2 示出了校正后的外极线,可见校正后图像的外极线平行,且对应点都在极线附近。图中假设校正后的基本矩阵为理想值。

对每对图像都重新选取 8 对匹配点,校正前后匹配点的平均垂直视差和评价函数变化如表 1

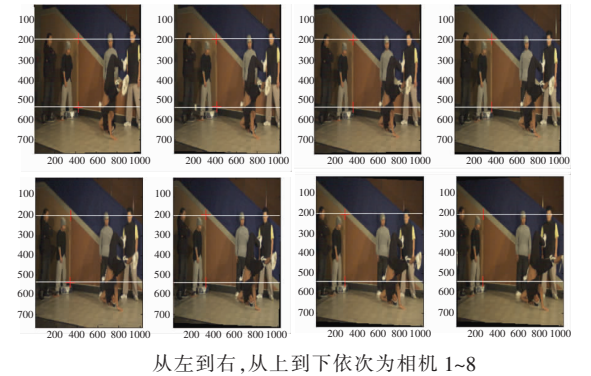
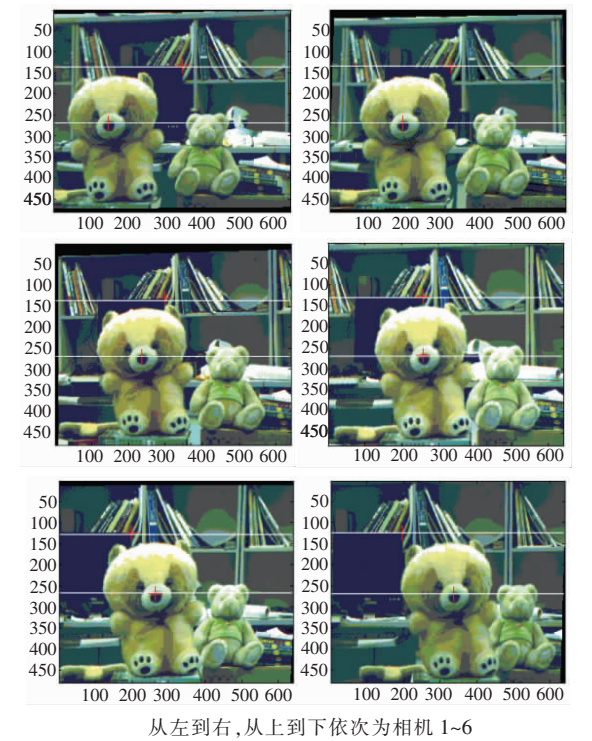


图 2 Breakdancers 序列校正后的外极线
Fig. 2 Epipolar lines after rectification for test sequences

表 1 测试序列校正前后垂直视差和评价函数的变化
Table 1 Vertical mismatch between before and after rectification for test sequences

相机对	垂直视差(像素)		评价函数	
	原序列	校正后	原序列	校正后
1 和 2	8.125	0.911 1	7.375	1.793
3 和 4	3.125	1.516 0	2.125	1.992
5 和 6	15.370	1.063 0	14.375	0.918
7 和 8	4.000	2.881 0	4.500	2.891
2 和 3	3.875	2.476 0	3.875	3.051
4 和 5	7.750	0.331 0	8.750	1.013
6 和 7	4.375	2.497 0	5.375	3.186



从左到右,从上到下依次为相机 1~6

图 3 自拍序列校正后外极线

Fig. 3 Epipolar lines after rectification for self sequences

所示,可以看出校正后图像间垂直视差约平均减少 1/2。

根据校正前后合成的立体视觉效果,可以看出校正后合成图像的画面更加清晰、立体感更强,主观评价三维效果明显好于原图像。实验中也发现有少量局部区域改善不明显,如最右边的人头部分。分析其原因主要是受以下两方面因素的影响:一是采用整像素匹配的精度不够高和匹配点的数量不够多,以及点位置选取不够均匀;二是由

表 2 自拍序列校正前后垂直视差和评价函数的变化
Table 2 Vertical mismatch between before and after rectification for self sequences

相机对	垂直视差(像素)		评价函数	
	原序列	校正后	原序列	校正后
1 和 2	0.875 0	1.467 3	0.625 0	0.865 5
3 和 4	14.625 0	0.930 9	13.625 0	1.203 8
5 和 6	15.500 0	3.605 8	14.500 0	3.587 5
2 和 3	16.500 0	2.881 0	17.500 0	1.131 5
4 和 5	22.250 0	2.476 0	23.250 0	1.502 7

匹配点得到单应性矩阵采用的是线性算法,其准确度要低于非线性算法。这些也是我们今后需要进一步研究的内容。

此外,实验中还用 6 相机的弧形阵列实拍了一组实验图像(以相机 4 为基准,图像分辨率为 640×480)。实验结果如图 3 和表 2 所示。由实验结果可知,虽然相机 1 和 2 间的结果不是很理想,但整体效果明显改善。

4 结 论

提出了一种无相机标定的多视图像校正方法,首先准确地寻找两幅图像间的匹配特征点,然后由匹配点坐标准确地计算出两幅图像间的单应性矩阵,从而有效减少了校正后匹配点间的垂直视差,校正后图像客观评价和主观合成效果都比校正前有明显改善。如何进一步提高匹配点和单应性矩阵的精确度、通过寻找更多的匹配特征点以及采用更优的算法计算单应性矩阵是今后需要进一步研究的问题。

参 考 文 献:

[1] 张兆杨, 安平, 刘苏醒. 3D 多视点立体显示及其关键技术 [J]. 电子器件, 2008, 31(1): 302-307.
[2] Lin Guoyu, Chen Xu, Zhang Weigong. A robust epipolar rectification method of stereo pairs[C]//2010 International Conference on Measuring Technology and Mechatronics Automation, Changsha, China: IEEE, 2010: 322-326.
[3] Fusiello A, Trucco E, Verri A. A compact algorithm for rectification of stereo pairs [J]. Machine Vision and Application, 2000, 12(1): 16-22.
[4] Zhang Z. A flexible new technique for camera calibration [J]. Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
[5] Sun Chang-ming. Uncalibrated three-view image rectification [J]. Image and Vision Computing, 2003, 21(3): 259-269.
[6] Kang Y S, Lee C, Ho Y S. An efficient rectification algorithm for multi-view images in parallel camera array[C]//3D Conference: The True Vision - Capture, Transmission and Display of 3D Video, Istanbul, Turkey: IEEE,

2008:61-64.

- [7] Kang Y S, Ho Y S. Geometrical compensation algorithm of multi-view image for arc multi-camera arrays [C]//*Proceedings of the 9th Pacific Rim Conference on Multimedia*, Tainan: LNCS, 2008; 543-552.
- [8] Hartley R, Zisserman A. *Multiple View Geometry in Computer Vision* [M]. London: Cambridge University Press, 2003; 237-360.
- [9] Guo Chenguang, Li Xianglong, Zhong Linfeng, *et al.* A fast and accurate corner detector based on Harris algorithm [C]//*Proceedings of the 3rd International Conference on Intelligent Information Technology Application*, Nanchang China: IEEE, 2009; 49-52.
- [10] 文福林, 李大海, 王琼华, 等. 立体图像对校正技术在三维显示中的应用 [J]. 激光杂志, 2009, 130(1): 38-40.

(上接 P592)

例:[12]西安电子科技大学. 光折变自适应光外差探测方法:中国,01128777.2[P/OL].2002-03-06[2002-05-28].<http://211.152.9.47/sipoasp/zljs/hyjs-yx-new.asp?recid=01128777.2&leixin=0>.

电子文献:[序号]主要责任者. 题名:其他题名信息[文献类型标志]. 出版地:出版者,出版年(更新或修改日期)[引用日期]. 获取和访问路径.

例:[13]PACS-L: the public-access computer systems forum[EB/OL]. Houston,Tex: University of Houston Libraries, 1989[1995-05-17]. <http://info.lib.edu/pacsl.html>.

例:[14]Online Computer Library Center, Inc. History of OCLC [EB/OL]. [2000-01-08]. <http://www.oclc.org/about/history/default.htm>.

2. 文后参考文献应在正文中引用该文献处进行标注。
3. 关于作者姓名,姓在前,名在后(拉丁文只用缩写);作者 3 人以下应全部列出,4 人以上仅列出前 3 人,其后加“等”(外文加“et al”)。
4. 注意参考文献标注的标点符号、次序,并不得缺项。
5. 参考文献著录不规范的有可能影响文章的录用。

注:文献类型和电子文献载体标志代码:

文献类型和标志代码

文献类型	标志代码	文献类型	标志代码
普通图书	M	报告	R
会议录	C	标准	S
汇编	G	专利	P
报纸	N	数据库	DB
期刊	J	计算机程序	CP
学位论文	D	电子公告	EB

电子文献载体和标志代码

载体类型	标志代码
磁带(magnetic tape)	MT
磁盘(disk)	DK
光盘(CD-ROM)	CD
联机网络(online)	OL