

单目多视点立体图像提取及应用

宋振东¹ 姜伟^{2*}

(¹ 黑龙江大学信息科学与技术学院, 黑龙江 哈尔滨 150080)
(² 浙江大学工业控制技术国家重点实验室, 浙江 杭州 310027)

摘要 设计并实现了一种基于广角相机和平面镜的单目多视点立体图像摄像系统,给出了硬件装置的设计指标和优化方法;同时,在研究了硬件系统的标定方法基础上,实现了其在三维测距方面的应用。多枚平面镜构成的对称斗型腔体被放置在广角相机前面,物体光线经过不同平面镜反射后,投影到相机图像平面的不同区域,在相机投影平面上生成物体多个影像,形成单目多视点投影图像。该类图像等价于视点不同的多幅图像,可以使用多视点立体视觉算法实现三维测量。

关键词 机器视觉;成像系统;单目立体成像;系统标定;视差估计

中图分类号 TP391.4; TP242.6⁺2 **文献标识码** A **doi**: 10.3788/AOS201232.0515001

Monocular Multi-Viewpoints Stereo Imaging and Application

Song Zhendong¹ Jiang Wei²

(¹ College of Information Sciences and Technology, Heilongjiang University, Harbin, Heilongjiang 150080, China)
(² State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, Zhejiang 310027, China)

Abstract A monocular imaging system is designed to capture the multi-viewpoints stereo images. Several identical isosceles trapezoidal planar mirrors are placed around a wide view-angle camera. The light rays from an object projected onto the camera include direct light and the light reflected by the mirrors, and divide the camera's image into several regions; thus a multi-viewpoints image is captured. These regions correspond to observations with multiple cameras at different positions and orientations, and can be used for multi-views stereo images to reconstruct three-dimensional (3D) information. The system design guideline and calibration method are described, and an application of 3D shape measurement is proposed. Experimental results indicate that the proposed system can capture monocular multi-viewpoints stereo images, and achieve highly robust 3D measurement.

Key words machine vision; imaging system; monocular stereo imaging; system calibration; parallax estimation

OCIS codes 150.6910; 110.2970; 110.4190

1 引 言

基于机器视觉的三维信息提取技术是人工智能以及模式识别领域的研究热点,研究成果已经在品质检测、安全监控以及自律移动体导航等许多领域得到广泛应用。

模仿人类的视觉系统,使用两台相机(两幅图像)的双目立体视觉是一种常用的三维信息测量方法。在已知相机内部参数和外部参数(可以通过系统标定获得)的条件下,通过估计物体特征点在两幅

图像上的投影视差,利用三角测量原理可以计算出物体点的三维空间距离。该方法系统构造简单、理论清晰,是该类研究的经典配置。但是,由于噪声、深度遮挡、重复图案以及相机间色差等影响因素的存在,导致双目立体视觉系统的视差估计结果不稳定,无法得到高精度三维信息^[1]。Okutomi等^[2]提出的基于多相机配置的多基线方法可以有效的抑制噪声等对视差估计的影响,提高结果精度。但是多相机配置增加了系统造价,同时增加了系统标定(尤

收稿日期: 2011-10-31; 收到修改稿日期: 2011-12-16

基金项目: 浙江省自然科学基金(Y1110067)和浙江省钱江人才计划(2010R10003)资助课题。

作者简介: 宋振东(1964—),男,副教授,主要从事数值计算及模式识别等方面的研究。E-mail: songzd@hju.edu.cn

* 通信联系人。E-mail: jiangwei@csc.zju.edu.cn

其是色差标定)的难度和工作量。

使用一台相机(一幅图像)的单目立体视觉系统可以有效避免上述多相机系统的诸多缺点。基于一台相机和结构光的主动视觉方法^[3,4]可以高精度地计算距离信息,但是系统价格昂贵,且无法实现实时系统。多重曝光方法可以实现单目三维信息提取^[5,6]。该方法原理类似于共焦点显微镜,利用不同景深的物体合焦距离不同的特点,微小间隔改变相机焦距的同时连续拍摄物体图像,通过检测对应于不同焦点距离的图像上合焦部分,可以逆推出场景的三维距离信息。该方法有效测量范围窄,且需要精密控制焦距变化,系统非常复杂。用镜子反射光线、分割相机视野是一种常用的被动式单目立体视觉方法^[7~9]。它通过同时拍摄物体和物体在镜子中的虚像,在一幅图像上获取物体的两个映像,且这两个映像来自于不同的视点,使用三角测量原理,可以实现三维测距。但是,它同样存在双目立体视觉方法的缺点。

本文设计并实现了一种基于广角相机和平面镜的单目多视点立体图像提取系统:广角相机前面放置由多枚平面镜构成的对称斗型腔,镜面朝向腔体内侧;物体光线经过不同平面镜反射后,投影到相机图像平面的不同区域,在相机图像平面上投影多个影像,生成单目多视点立体图像。该图像等价于视点不同的多幅图像,等价于一个单目多视点立体视觉系统,可以实现高精度三维测量。

2 系统构成及其设计原理

2.1 系统构成

图1为单目多视点立体视觉系统原理图。它主要由 CCD 相机、广角镜头和平面镜斗型腔组成。斗型腔由尺寸相同的等腰梯形平面镜组成,镜面朝向腔体内侧,且腔体的中轴线重合于相机的主光轴。广角镜头由斗型腔小口端探入腔体内,镜头入射光线由斗型腔端口的直射光线和镜面反射光线组成。通过腔体的直射光线投影在相机投影面的中央区域,镜面反射光线投射到相机投影面的周边区域,且不同镜面的投影区域各不相同。因此,该系统拍摄的图像包含目标物体的多个影像,且这些影像来自于不同透视投影点,被称为“多视点投影图像”。

实验系统使用四枚平面镜,因此该图像被分割为5部分[参见图8(b)],分别对应于中央区域的直射光线投影和周边区域的镜面反射光线投影。中央区域影像可以作为基准图像,与周边区域影像构成

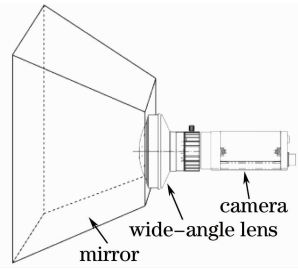


图1 摄像系统原理图

Fig.1 Block diagram of imaging system
立体图像对,实现三维测量。

2.2 等价系统

根据平面镜成像原理,一面平面镜可以生成物体的一个虚像。图2是提案摄像系统的侧视断面图。设定相机投影中心为坐标原点 O ,相机光轴为 Z 轴, Y 轴垂直于镜面端线。上部镜子对应的虚拟相机用 O_1 表示,镜子位置、高度和夹角分别用 b 、 m 和 γ 表示,则根据平面几何关系,中央区域对应相机视角 α 、虚拟相机 O_1 的视角 β 和位置分别为

$$\alpha = \pi - 2\arctan \frac{(m/b)\sin \gamma}{1 + (m/b)\cos \gamma}, \quad (1)$$

$$\beta = \frac{1}{2}(\pi - \alpha), \quad (2)$$

$$O_1(y, z) = [b(1 - \cos 2\gamma), -b\sin 2\gamma], \quad (3)$$

由图2可以看出,虚拟相机和真实相机具有不同的位置和空间姿态,构成了双目立体视觉配置,可以实现三维测量。

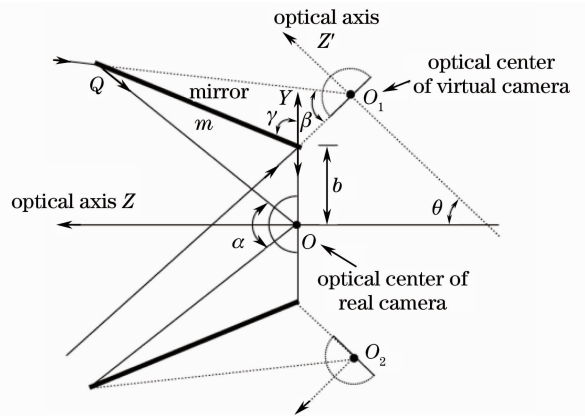


图2 系统的侧视断面图

Fig.2 Cross view of the proposed system

同样,其他平面镜对应的虚拟相机的位置和姿态也可以用类似方法计算。因此,提案系统的硬件装置等价于一个使用多台相同相机构成的多机立体视觉系统。与多机系统相比较,提案系统性价比高,并且所有“相机”具有严格一致的内部参数和颜色系统,有利于特征匹配,实现高精度视差估计。

2.3 系统设计优化

如 2.2 节所述, 视角 α 、 β 、虚拟相机的位置和姿态信息均可通过平面镜尺寸 m 、位置 b 和角度 γ 计算。本节通过评估相机间的共同视角(关联到三维测量视野范围) 阐述摄像装置的设计指针, 目的在于最大化相机间的共同视角, 从而提高立体视觉的测量范围。图 2 中相机 O 、 O_1 和 O_2 的共同视角用 Ω 表示, 则摄像装置的设计标准就是确定参量 m 、 b 和 γ 的取值, 最大化共同视角 Ω 。

图 3 显示共同视角 Ω 随着 m 、 b 和 γ 的变化情况。图中 m' 表示 m/b , 即正规化装置尺寸。该图可以看出, 共同视角 Ω 的最大值依存于 m' 的取值, 大的 m' 可以得到较大视角 Ω 。但是, 大的 m' 造成装置尺寸过大, 不利于移动。因此, 实验中 m' 取值 3, 由图 3 可见, 当 γ 取值为 56° , Ω 有最大值, 约为 43° 。

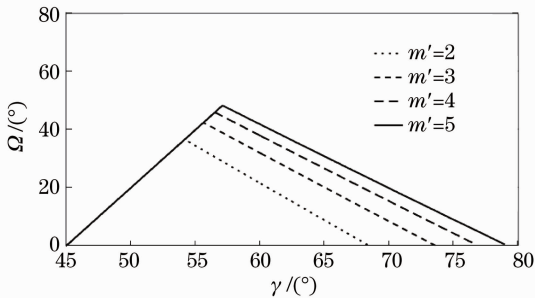


图 3 共同视角 Ω

Fig. 3 Common view angle Ω

3 系统标定

机器视觉系统的标定是获得高精度测量结果的基础和保证^[10]。使用一幅标定板的多视点投影图像计算相机内部参数、镜头畸变参数和镜面外部参数(镜面位置和姿态)三方面内容。

3.1 相机及镜头内部参数标定

视觉系统的相机和镜头使用前必须进行标定。为了增加实体相机的视角, 选用了 $360^\circ \times 185^\circ$ 的鱼眼广角镜头。因此, 内部参数的标定包括鱼眼广角镜头投影模型参数、镜头半径方向和切线方向畸变参数、相机的焦点距离和图像主点(即光轴与投影面的交点)位置参数。

鱼眼镜头有多种投影模型^[11], 本研究使用效果较好的统一投影模型(Unified projection model)^[12,13]。在该模型中, 一个三维空间点投影到相机投影面需要经过如下四个步骤: 投影到单位球、投影到单位平面、镜头畸变和透视投影变换。基于文献^[14]的方法, 使用已知尺寸的棋盘格子标定板,

同时计算出所有标定变量($\zeta, k_1, k_2, k_3, k_4, k_5, f_u, f_v, u_0, v_0$)。其中 ζ 关联于单位投影平面的位置, k_1, k_2, k_3, k_4, k_5 是镜头畸变参数, f_u, f_v, u_0, v_0 是镜头焦距和图像主点位置。

一般方法^[14], 特别是鱼眼镜头的标定, 使用一幅图像无法实现全部模型参数的精确标定。如图 4 所示, 本研究提案摄像装置使用平面镜反射光线, 拍摄的一幅多视点投影图像上包含标定板的多个影像。由平面镜反射原理可知, 这些标定板影像等价于同一标定板在不同位置上分别投影到相机投影平面上所获得的多幅标定板图像(参见 3.2 节及图 7)。所以, 本研究使用一幅标定板的多视点投影图像标定相机内部参数, 等价于使用多幅标定板图像的标定操作, 能够获得高精度的标定结果。

图 4 是本研究实验用装置拍摄的一幅标定板多透视投影图像, 图像中包含标定板的 5 个独立影像。从该图像可以看出, 鱼眼镜头畸变很大, 边缘区域尤为严重(标定板的直线边变成曲线)。使用该图像对相机和鱼眼镜头进行了标定, 结果如下: ($\zeta, k_1, k_2, k_3, k_4, k_5, f_u, f_v, u_0, v_0$) = (1.6988, -0.06093, 0.18404, -0.00015, -0.00017, 0, 871.54278, 868.49105, 791.49429, 595.47177)。使用标定结果, 对图 4 所示图像进行了畸变矫正, 并进行透视投影变换, 变换结果如图 5 所示。比较图 4 和图 5 可

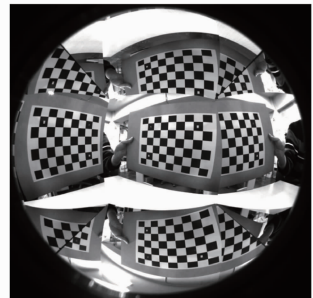


图 4 标定图像

Fig. 4 Calibration image

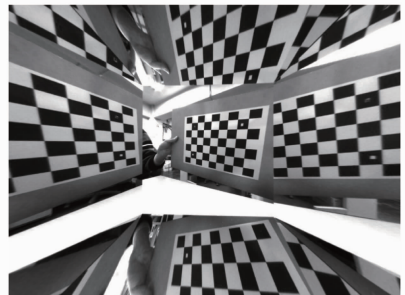


图 5 透视投影图像

Fig. 5 Perspective projection image

以看出,镜头畸变导致的歪扭曲线边已经被矫正为直线,得到了标准的透视投影图像。

3.2 虚拟相机外部参数标定

首先计算各标定板位置信息,然后计算平面镜的位置和姿态信息,最后估计出各虚拟相机的外部参数。文献[14]方法在标定相机内部参数的同时可以计算出各标定板的空间位置和姿态信息。图6是基于图4的标定板多视点投影图像计算的标定板三维空间位置和姿态。其中0号所示为真实标定板、其他为镜面反射的虚拟标定板位置和姿态。

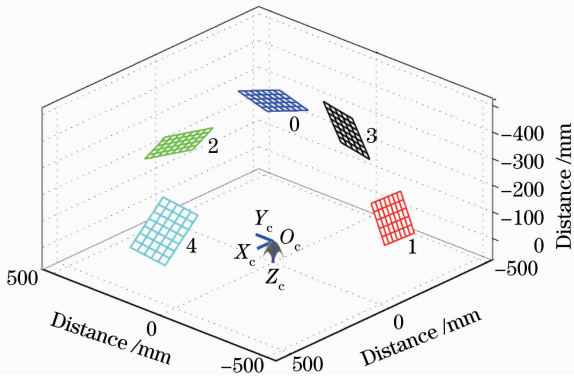


图6 标定板位置

Fig. 6 Position of target

图7为真实物体(包括相机)和被镜面反射的虚拟物体之间的关系。如果用 \mathbf{n}'_0 和 d'_0 表示真实标定板的法向量和距离, \mathbf{n}'_i 和 d'_i 表示第 i 枚镜子反射的虚拟标定板的法向量和距离,则根据平面镜反射原理,第 i 枚平面镜的法向量值和距离 $[\mathbf{n}^m_i, d^m_i]$ 可以利用下面公式计算:

$$[\mathbf{n}^m_i, d^m_i] = \left[\frac{\mathbf{n}'_0 - \mathbf{n}'_i}{|\mathbf{n}'_0 - \mathbf{n}'_i|}, \frac{(\mathbf{n}'_0 - \mathbf{n}'_i) \cdot \mathbf{n}'_i}{|\mathbf{n}'_0 - \mathbf{n}'_i| (1 - \mathbf{n}'_0 \mathbf{n}'_i)} (d'_0 - d'_i) \right].$$

同时,利用上述求解的平面镜的位置和姿态信息 (\mathbf{n}^m_i, d^m_i) ,图7中镜面 M_i 关联的虚拟相机 O_i 的外部参数可以由下式计算:

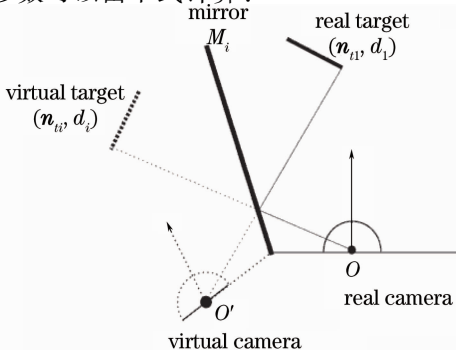


图7 真实物体和虚拟物体

Fig. 7 Real object and mirrored object

$$\mathbf{O}_i = -2d^m_i \mathbf{n}^m_i, \quad (4)$$

$$\mathbf{R}_i = \mathbf{I} - 2\mathbf{n}^m_i \mathbf{n}^m_i{}^T, \quad (5)$$

式中 \mathbf{O}_i 表示位置, \mathbf{R}_i 为表示姿态的旋转矩阵, \mathbf{I} 为 3×3 的单位矩阵。

本研究基于平面镜反射原理,利用上述方法,实现了 CCD 相机内部参数、鱼眼广角镜头投影参数和畸变参数、平面镜外部参数和相应虚拟相机外部参数的标定。

4 基于多视点图像的视差估计算法

作为提案摄像装置的一种应用,设计了多视点图像视差估计算法,实现三维测量。

3.2 节实现了镜面反射虚拟相机的外部参数标定,即知道了中央的真实相机和周围的虚拟相机间的位置关系。因此,真实相机和任何一台虚拟相机可构成一个双目立体视觉系统,实现三维测量。然而,双目系统对噪声敏感,且存在“深度遮挡”弱点,无法获得稳定结果^[2]。本文将中央真实相机作为基准相机,周围虚拟相机作为参考相机,构成多机系统,通过多图像间特征点匹配,抑制图像噪声、深度遮挡等因素的影响,提高视差估计的稳健性,实现稳定三维重建。

首先,使用 3.1 节标定的鱼眼广角镜头模型参数、镜头畸变参数和相机内部参数变换多透视投影图像为透视投影图像。由于鱼眼广角镜头垂直视角超过 180° ,导致一张多透视投影图像无法变换为一张透视投影图像。根据各虚拟相机的位置和姿态信息,可将多透视投影图像的不同部分投影到各相机的投影平面上,分别获得多枚对应于不同相机的标准透视投影图像。

然后,根据相机间的位置关系,计算真实相机和各虚拟相机之间的外极约束关系,确定各参照图像(即各虚拟相机对应的图像)上的外极线。采用类似于基滑动窗口的对应点匹配方式,对基准图像上的每一个像素点,在参照图像上沿着外极线滑动窗口,相似度关系可表示为

$$S_{I_c, I_{mi}}(l; x_0, y_0) = \sum_w [I_c(x_0, y_0) - I_{mi}(l; x_0, y_0)]^2, \quad (6)$$

式中 I_c, I_{mi} 表示中央基准图像和第 i 个虚拟相机对应的参照图像, l 为距离, w 表示窗口尺寸。由(6)式可以看出,本文方法虽然是使用滑动窗口方式匹配对应点,但是参照图像上的窗口变化不是逐像素变化,而是通过以某一间隔变化待测量的距离 l , 通过

相机间的位置关系计算参照图像上对应点的位置，进而使用上式计算窗口间相似度。

最后，通过整合所有立体图像对之间的对应点匹配信息，计算三维距离。对于一个距离值 l ，在所有参照图像上均可以计算出对应点位置，并且可以通过上式计算出对应点间相似度。在一般情况下，真实的距离值将使所有立体图像间对应点相似度的评价值最小(相似度高)，所有立体图像间对应点的相似度评价值的和也将取得最小值。因此，本研究将不同图像间的相似度评价量 $S_{I_c, I_m i}(l; x_0, y_0)$ 相加 S_{Σ} ，其和作为新的评价量，估计三维距离信息。具体如下式所示：

$$S_{\Sigma}(l; x_0, y_0) = \sum_{i \in S(x_0, y_0)} S_{I_c, I_m i}(l; x_0, y_0). \quad (7)$$

因为噪声具有随机性，且深度遮挡往往发生在单侧，只影响某一对立体图像间的相似度评价量 $S_{I_c, I_m i}(l; x_0, y_0)$ ，对于其它图像对则无影响，即对于所有评价量之和影响较小，应用(7)式可以得到正确的距离值 l 。

5 实验结果

为了检验单目立体视觉摄像系统的正确性和有效性，试制了摄像系统硬件装置，并针对视觉系统的三维复原应用，进行了真实场景实验。图 8 是本研究试制的实验装置和拍摄的单目多视点投影图像。考虑到虚拟相机的水平视角，该实验装置使用四枚平面镜[见图 8(a)]，因此等价于一个使用五台相机的系统，拍摄的图像如图 8(b)所示。实验装置所用相机的内部参数标定结果如下： $(\zeta, k_1, k_2, k_3, k_4, k_5, f_u, f_v, u_0, v_0) = (1.6988, -0.06093, 0.18404, -0.00015, -0.00017, 0.87154278, 868.49105, 791.49429, 595.47177)$ 。

根据装置标定的结果，对图 8(b)图像进行透视

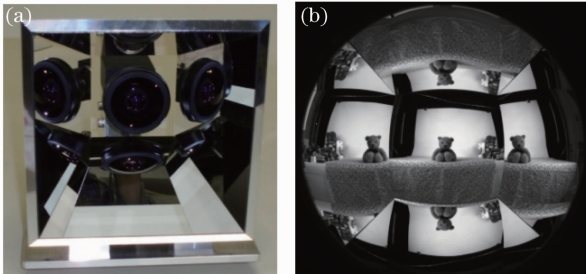


图 8 实验装置和原始图像。(a)实验装置；
(b)单目多视点图像

Fig. 8 Experiment device and original image.

(a) Experimental device; (b) monocular multi-view image

投影再变换，获得 5 幅标准的目标物体透视投影图像，如图 9 所示。使用第 4 节的算法进行视差估计，得到目标物体的视差图，如图 10 所示。结果显示本研究方法可以获得良好的视差估计结果。

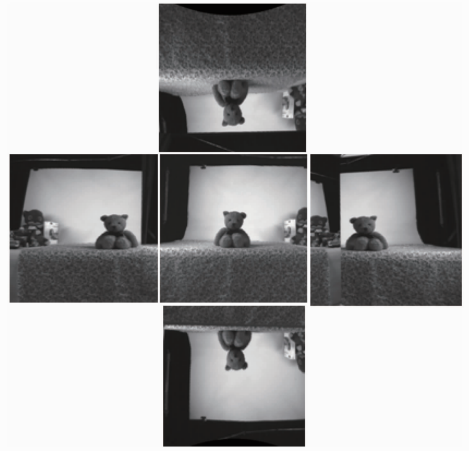


图 9 图 8(b)的透视投影变换图像

Fig. 9 Perspective re-projected images from Fig. 8 (b)



图 10 视差图

Fig. 10 Disparity map

6 结 论

提出了一种基于广角相机和平面镜的单目多视点立体视觉系统，给出了系统设计方法、标定方法、多视点图像视差估计算法，并进行了系统正确性验证试验。实验结果表明，与双目立体视觉系统相比较，它能够改善噪声、深度遮挡等因素对视差估计的影响，有效提高视差估计结果的质量。

今后，将在硬件装置一体化、小型化方面作进一步的研究；同时，将进行系统在三维视频制作等方面的应用推广。

参 考 文 献

- 1 M. Okutomi, Y. Katayama, S. Oka. A simple stereo algorithm to recover precise object boundaries and smooth surfaces[J]. *International J. Computer Vision*, 2002, **47**(1-3): 261~273
- 2 M. Okutomi, T. Kanade. A multiple-baseline stereo[J]. *IEEE Trans. Pattern Anal.*, 1993, **15**(4): 353~363
- 3 J. Diebel, K. Reutersward, S. Thrun *et al.*. Simultaneous

- localization and mapping with active stereo vision [C]. Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan; IEEE, 2004; 3436~3443
- 4 Ding Yabin, Peng Xiang, Tian Jindong *et al.*. Pose estimation of multiple viewpoints for three-dimensional digital imaging system [J]. *Acta Optica Sinica*, 2007, **27**(3): 451~456
丁雅斌, 彭翔, 田劲东等. 一种三维数字成像系统的多视点姿态估计方法[J]. *光学学报*, 2007, **27**(3): 451~456
- 5 E. H. Adelson, J. Y. A. Wang. Single lens stereo with a plenoptic camera[J]. *IEEE Trans. Pattern Anal.*, 1992, **14**(2): 99~106
- 6 S. Hiura, T. Matsuyama. Depth measurement by the multi-focus camera[C]. Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. California, USA; IEEE, 1998: 953~959
- 7 J. Gluckman, S. K. Nayar. Catadioptric stereo using planar mirrors[J]. *Int. J. Comput. Vision*, 2001, **44**(1): 6~79
- 8 J. Gluckman, S. K. Nayar. Rectified catadioptric stereo sensors [J]. *IEEE Trans. Pattern Anal.*, 2002, **24**(2): 224~236
- 9 O. Avni, T. Baum, G. Katzir *et al.*. Recovery of 3D animal motions using cameras and mirrors[J]. *Mach. Vision. Appl.*, 2009, **21**(6): 879~888
- 10 Huang Tao, Zhu Qiudong, Hao Qun *et al.*. Study on distortion elimination in stereoscopic display systems [J]. *Acta Optica Sinica*, 2010, **30**(2): 405~410
黄涛, 朱秋东, 郝群等. 立体显示系统的消畸变研究[J]. *光学学报*, 2010, **30**(2): 405~410
- 11 M. D. Grossberg, S. K. Nayar. A general imaging model and a method for finding its parameters[C]. Proceedings of the Eighth International Conference on Computer Vision, Vancouver, British Columbia, Canada; IEEE, 2001; 108~115
- 12 C. Geyer, K. Daniilidis. A unifying theory for central panoramic systems and practical implications [C]. Proceedings of the European Conference on Computer Vision, Dublin, Ireland, 2000; 445~461
- 13 J. Courbon, Y. Mezouar, L. Eck *et al.*. A generic fisheye camera model for robotic applications[C]. Proceedings of IEEE / RSJ International Conference on Intelligent Robots and Systems. San Diego, California, USA; IEEE, 2007; 1683~1688
- 14 J. Y. Bouguet. Camera calibration toolbox for Matlab [OL] http://www.vision.caltech.edu/bouguetj/calib_doc/index.html (2005)

栏目编辑: 李文喆