

采用视频拼图方法构建高分辨率 全景视频监控系统

马力¹⁾ 张茂军¹⁾ 徐玮¹⁾ 熊志辉¹⁾ 王瑜²⁾

¹⁾(国防科学技术大学信息系统与管理学院,长沙 410073) ²⁾(海军蚌埠士官学校,蚌埠 233012)

摘要 与普通视频监控系统只能实现单向监控不同,全景视频监控系统可以实现360°全向监控。设计并实现了一种嵌入式高分辨率全景视频监控系统KD-PVS。重点介绍了KD-PVS中多个摄像头的空间位置设计、视频图像变换与拼接算法。KD-PVS通过对多个摄像头采集的视频进行实时变换与拼接以生成全景视频。该系统可方便应用于金融系统、仓库、监狱和移动监控等多种场合,尤其适用于室内监控。

关键词 全景视频 嵌入式系统 视频拼图 视频监控

中图法分类号:TN911.73 文献标识码:A 文章编号:1006-8961(2008)12-2291-06

A High Resolution Panoramic Video Monitoring System Based on Video Mosaicing

MA Li¹⁾, ZHANG Mao-jun¹⁾, XU Wei¹⁾, XIONG Zhi-hui¹⁾, WANG Yu²⁾

¹⁾(College of Information System and Management, National University of Defense Technology, Changsha 410073)

²⁾(Naval Petty Officer Academy, Bengbu 233012)

Abstract Difference from traditional video surveillance system, panoramic video surveillance system can provide viewers with a complete 360° degrees view. One embedded high resolution panoramic video surveillance system called KD-PVS is designed and realized. The paper introduces the multiple-camera configuration and video mosaicing algorithm. KD-PVS stitches the video data from multiple camera sources into the panoramic video. The system is very convenient for various situations, such as finance system, warehouses, prisons, mobile monitoring, etc, especially useful for indoor monitoring.

Keywords panoramic video, embedded system, video mosaicing, video surveillance

1 引言

传统视频监控系统一般采用固定安装的普通摄像机,其视野范围有限,只能监控摄像机前方某一固定角度空间内的场景,无法顾及周围360°范围内发生的所有事件。即使采用带旋转云台的摄像机,同一时刻也只能看到某一角度的画面,监控时不可避免会出现盲区。近几年,全景视频监控系统引起了人们的极大关注^[1-5]。全景视频监控系统能够在任

一时间点同时监控360°范围内的所有目标,真正消除了监控死角与盲区。

设计全景视频监控系统的关键技术之一是如何生成全景视频^[6]。目前生成全景视频的方法可分为两类:折射法和折反射法^[7]。折射法指仅采用具有折射效果的光学元件获取周围场景,其中又分为使用鱼眼镜头和使用多摄像机两种方式。利用鱼眼镜头的宽广视角可直接获取360°场景,经过变换可得到全景视频^[8],目前市场上典型的产品如美国IPIX公司的CommandView系列产品。也可以采用

基金项目:国家自然科学基金项目(K60705013, K60773023)

收稿日期:2007-04-11;改回日期:2007-06-12

第一作者简介:马力(1983~),男,2007年于国防科技大学获系统工程专业硕士学位。研究方向为视频图像处理、多媒体信号处理等。E-mail:nudtml@nudt.edu.cn

多个位置固定的摄像头捕捉不同角度空间的场景,然后将视频图像进行拼接得到全景视频^[4,5,7,9],典型的产品如美国 Honeywell 公司推出的 PARASCAN 全景摄像机。折反射法指采用折射元件和反射元件相结合的方式,主要是利用具有镜面反射效果的空间曲面获取周围 360°场景,然后对摄像头捕获的反射面图像进行实时处理以得到全景视频^[1,2,3,10]。基于该原理的典型产品有日本 SONY 公司的 360°全景摄像机。利用多摄像机进行视频拼图得到的全景视频能够提供较高的分辨率,但由于拼接的原因,全景视频不可避免会存在拼接裂痕和明暗差异。而利用鱼镜头和空间镜面获取全局场景的方法可避免因图像拼接而造成的裂痕,但是生成的全景视频分辨率较低,清晰度较差。

用于大范围监控目的的全景视频(如监控广场、停车场等)必然要求提供较高的视频图像分辨率。该文采用视频拼图的方法设计并实现了一种高分辨率的全景视频监控系统 KD-PVS。该系统基于嵌入式系统设计与开发,集成度高,无需安装其他的软硬件,可直接输出全景视频信号。系统体积小、便于安装和维护,并具备较强的可扩展性。

2 KD-PVS 的系统结构

KD-PVS 系统分为两大模块:全局场景获取模块和视频信号处理模块,其结构如图 1 所示。全局场景获取模块包含 4 个广角摄像头及其固定装置,其功能是捕获 360°全局场景。视频信号处理模块负责 4 路视频的采集、视频图像变换与拼接以及全景视频信号的输出。

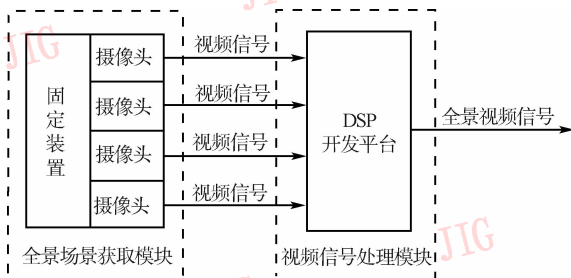


图 1 系统总体结构图
Fig. 1 Structure of KD-PVS

要获取全局场景,所需摄像头的数量与摄像头本身的视角大小有关。考虑到成本与拼接复杂性问

题,选用相同的 4 个视角约为 90°的微型摄像头共圆放置。4 个摄像头保持在同一水平面上,否则相邻图像会在垂直方向上产生位移,给拼接造成困难。摄像头的光心位于同一个圆上,相邻两个摄像头光心与该圆心的连线互成 90°。

如果将摄像头水平放置,即摄像头的光轴位于由 4 个摄像头光心所形成的平面内,这样从理论上说,4 个视角小于 90°的摄像头便无法拍摄到周围 360°的全部场景。因此考虑将每个摄像头在垂直方向上向下倾斜一定角度进行放置,使每个摄像头的光轴与所在的水平面成一个固定的角度,以增加视频叠加的范围。基于上述想法,设计与实现了一个形状类似倒置的正四棱台的固定装置,如图 2 所示。摄像头固定在棱台 4 个侧面上的同一位置处,镜头光轴与侧面垂直。若将该装置底面朝上水平放置,则每个摄像头斜向下与竖直方向成一固定角度,可捕获 360°的全局场景。这种设计不仅使得利用 4 个视角小于 90°的摄像头也能成功获取 360°全局场景,并且由于在大多数的应用场合,监控摄像头一般安装在位置较高的地方(如天花板),摄像头斜下方的区域正是需要监控的重点区域,因而这种设计更加适应大部分应用场合的需要,尤其适合于室内监控。

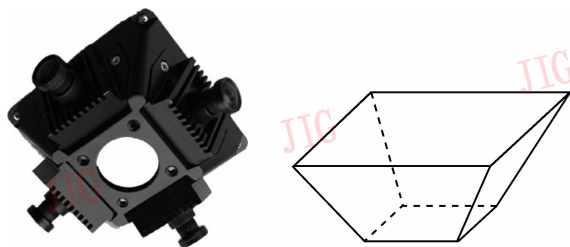


图 2 摄像头固定装置外观及形状示意图
Fig. 2 Appearance and shape of the device fixed cameras

3 全景视频生成算法

全景视频生成算法是 KD-PVS 系统的核心。该算法主要分为两个步骤:第 1 步将每一路视频图像统一投影到一个圆柱面上,目的是使现实世界中相同的景物在不同的局部图像中也是相同的,以便于拼接;第 2 步将 4 幅投影图像进行拼接,生成一张全景图像。对实时采集的每一帧视频图像均按照上述步骤进行处理,即生成连贯的全景视频。

3.1 视频图像畸变校正与柱面投影变换

由于系统采用的是广角摄像头,其光学镜头与理想的小孔成像模型不一致,拍摄的图像存在较严重的径向畸变,因而在实现柱面投影变换之前需要对原始视频图像进行畸变校正。由光学系统的成像原理可知,视场内每个像点对应一个不同的畸变量,离光轴越远,畸变量越大。该畸变量与像点距图像中心的距离以及镜头焦距有关,因而可根据其映射关系计算出原始图像中每个像素点的理想坐标^[11]。系统采用文献[11]中提出的方法对视频图像进行畸变校正。

本系统将经畸变校正后的 4 幅视频图像均投影到同一圆柱面上,以便拼接成全景图像。柱面投影变换分为两个步骤:(1)利用小孔成像模型将摄像机所拍摄的图像平面上的像素点映射到圆柱侧表面上;(2)将圆柱侧表面的投影区域展开,使圆柱面上的像素点以 2 维坐标的形式存储。下面以 1 幅图像为例,采用逆向推算的方法示范投影变换的具体步骤和公式。

如图 3 所示,设相机的透视投影中心为 O , OP 为相机镜头光轴, $ABCD$ 为焦平面。 OP 与 z 轴的夹角为 α , 相机水平视角为 2β , 垂直视角为 2γ , 相机焦距为 f , 投影圆柱面底面半径为 R 。

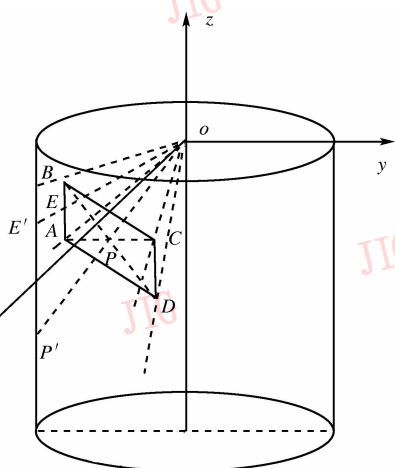


图 3 视频图像柱面投影示意图

Fig. 3 Cylindrical projection of video images

(1)构造投影平面 2 维坐标系,并计算投影平面上点的 2 维坐标与其在图 3 所示 $oxyz$ 坐标系中 3 维坐标的对应关系。以原始图像上边缘 AB 的中点 E 在柱面上的投影 E' 为投影平面顶部的中心点,设宽为 W ,高为 H ,如图 4 构造直角坐标系。

设投影平面上任意点 $M'(x_0, y_0)$, 其对应的 3 维坐标为 (x_1, y_1, z_1) , 由空间几何关系可知:

$$\begin{cases} x_1 = R \sin \frac{W/2 - x_0}{R} \\ y_1 = -R \cos \frac{W/2 - x_0}{R} \\ z_1 = -R \cot(\alpha + \gamma) - y_0 \end{cases} \quad (1)$$

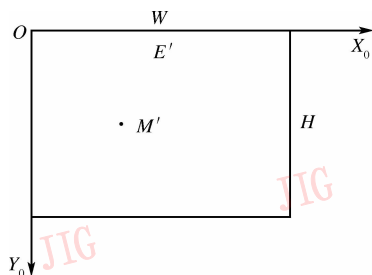


图 4 投影平面 2 维坐标系

Fig. 4 Planimetric rectangular coordinates of projected image

(2)计算点 M' 所对应的原始图像上某点 M 的 3 维坐标 (x_2, y_2, z_2) 。由直线 OM' 的方程和焦平面 $ABCD$ 的方程可求得

$$\begin{cases} x_2 = \frac{fR \sin \frac{W/2 - x_0}{R}}{(R \cot(\alpha + \gamma) + y_0) \cos \alpha + R \sin \alpha \cos \frac{W/2 - x_0}{R}} \\ y_2 = \frac{-fR \cos \frac{W/2 - x_0}{R}}{(R \cot(\alpha + \gamma) + y_0) \cos \alpha + R \sin \alpha \cos \frac{W/2 - x_0}{R}} \\ z_2 = \frac{-f(R \cot(\alpha + \gamma) + y_0)}{R \sin \alpha \cos \frac{W/2 - x_0}{R} + (R \cot(\alpha + \gamma) + y_0) \cos \alpha} \end{cases} \quad (2)$$

(3)计算焦平面上的点 $M(x_2, y_2, z_2)$ 在以 A 为原点的平面直角坐标系(如图 5 所示)下的 2 维坐标 (u, v) 。易知 u 即为点 M 到直线 AD 的距离, v

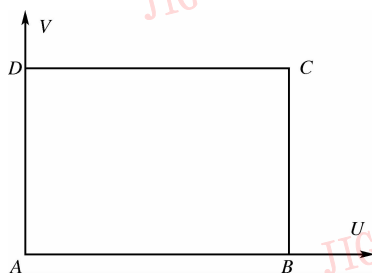


图 5 原始图像平面坐标系

Fig. 5 Planimetric rectangular coordinates of original image

为点 M 到直线 AB 的距离。可得：

$$u = f \tan \beta - x_2 \quad (3)$$

$$v = \frac{y_2}{\cos \alpha} + f \tan \alpha + f \tan \gamma \quad (4)$$

将式(2)代入式(3)、式(4)式可得：

$$\begin{cases} u = f \tan \beta - \frac{f R \sin \frac{W/2 - x_0}{R}}{(R \cot(\alpha + \gamma) + y_0) \cos \alpha + R \sin \alpha \cos \frac{W/2 - x_0}{R}} \\ v = f(\tan \alpha + \tan \gamma) - \frac{f R \cos \frac{W/2 - x_0}{R}}{(R \cot(\alpha + \gamma) + y_0) \cos^2 \alpha + R \sin \alpha \cos \alpha \cos \frac{W/2 - x_0}{R}} \end{cases} \quad (5)$$

式(5)表示了相机所拍摄的原始图像上的点 (u, v) 与其在展开的投影柱面上的投影点 (x_0, y_0) 之间的映射关系。

在算法的实现过程中,由于根据式(5)计算所得的位置坐标 (u, v) 可能是非整数值,我们采用了双线性插值算法对原始图像进行重采样。图 6(a)是采用水平视角为 120° 的摄像头拍摄的图像,对其进行畸变校正后再进行柱面投影变换,所得图像如图 6(b)所示。在系统实现过程中,4 个摄像机的光心并不是同一个空间点,即不满足单视点约束,但可近似等效为同一视点(即图 3 中的原点 O)。故每个摄像机所拍摄的视频图像均可通过式(5)所描述的映射关系实现柱面投影变换。

柱面的底面半径为 R ,取展开的柱面投影图像宽为 W ,则投影图像两边 $\frac{W}{2} - \frac{1}{4}\pi R$ 的区域是重叠区域。

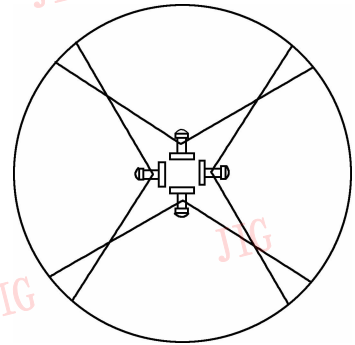


图 7 摄像头重叠视野范围示意图

Fig. 7 Overlap between adjacent cameras

考虑到上述算法需基于嵌入式系统实现,为确保系统的实时性,简化拼接算法如下:将投影图像剪除重叠区域后直接拼接以生成全景图像。基于此想法,设定生成的柱面投影图像宽 $W = \frac{1}{2}\pi R$,这样相邻图像刚好无重叠区域,可直接进行拼接。图 8 中

(a)、(b)、(c)、(d)分别为实验中 4 个摄像机在同一时刻不同方向拍摄到的 4 个场景的图像,图 8 (e)、(f)、(g)、(h)分别是将上述 4 幅图像进行畸变校正和柱面投影变换后得到的图像,将图 8 (e)、(f)、(g)、(h)4 幅图像直接进行拼接,即得到如图 8 (i)所示的一张全景图像。

由于工艺的限制,系统的加工与安装都会有一定的误差。忽略摄像头固有参数的差异,系统加工时造成摄像头水平方向和竖直方向的角度偏差是导致图像拼接误差的主要因素。假设在水平方向上,某摄像头的光轴偏离预定位置的角度为 $\Delta\lambda$,竖直方向上偏离的角度为 $\Delta\alpha$,易知造成的图像重叠范围

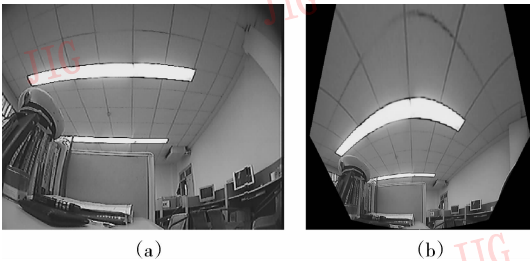


图 6 摄像机所拍摄的原始图像及实现畸变校正和柱面投影变换后的图像

Fig. 6 Original image and transformed image

3.2 图像拼接

图像拼接的难点在于找到并处理两个相邻图像样本之间的重叠区域。本系统利用其物理结构的特殊性,可快速确定相邻图像的重叠区域,然后剪除其重叠区域后直接将图像进行拼接,可快速生成柱面全景图像。

由于系统结构的轴对称特性,且采用 4 个相同的摄像头,故相邻两个摄像头所拍摄的视频图像重叠范围必然近似相同(如图 7 所示)。如果设投影

的误差为 $\frac{R\Delta\lambda}{\sin\alpha}$, 图像拼接时所产生的竖直方向上的偏差为 $R\cot(\alpha + \Delta\alpha) - R\cot\alpha$ 。实验中设定全景视频的分辨率为 $1\,280 \times 260$, 此时投影圆柱面底面半径约为 204 个像素值, 设定摄像头倾斜角 $\alpha = 60^\circ$, 若要将图像重叠范围误差和图像拼接时竖直方向上的误差均控制在 2 个像素点以内, 则需 $\Delta\lambda < 0.5^\circ$, $\Delta\alpha < 0.45^\circ$ 。

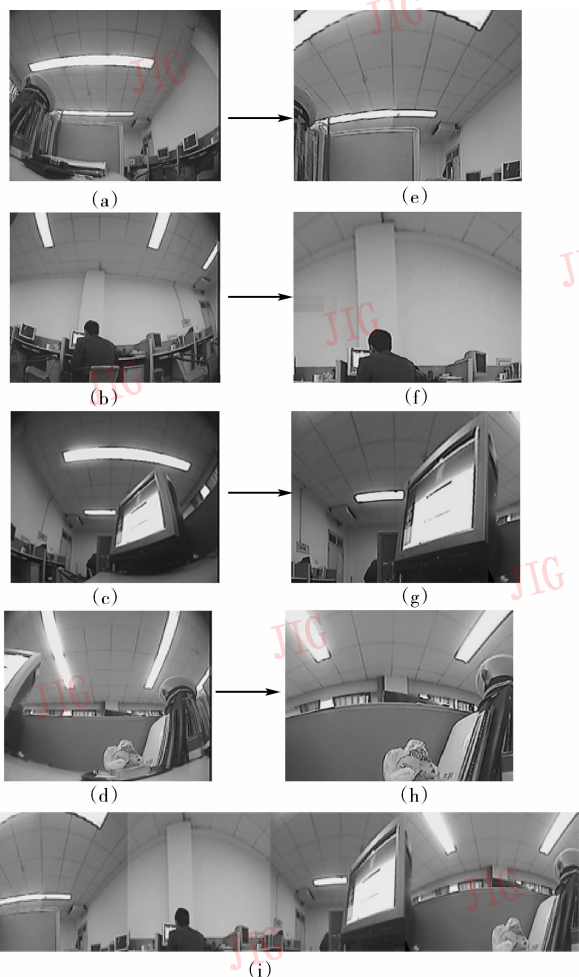


图 8 图像投影变换与拼接示意图

Fig. 8 Projection and mosaicing of video images

3.3 确定摄像机倾斜角

按照前文所述的摄像头摆放方式对 4 个摄像头进行固定放置, 即使摄像头视野范围小于 90° , 也可以捕获 360° 的全局场景。如图 3 所示, 减小摄像头的倾斜角 (即 OP 与 z 轴的夹角 α) 可增大相邻视频图像的重叠范围。然而, 摄像头倾斜角的减小也直接导致了摄像头垂直视野范围的减小。在相邻图像能实现正常拼接的同时, 为确保

生成的全景图像有较大的垂直视野范围, 下面推导全景图像的垂直视野范围与摄像头倾斜角的关系。

如图 3 所示, 考虑两幅相邻的视频图像, 将 $ABCD$ 绕 z 轴逆时针旋转 90° , 得到光轴位于 xoz 平面内的摄像机的焦平面。易知其右边沿和 $ABCD$ 的左边沿 AD 必然相交于一点。而该点在圆柱面上的投影点则位于最终生成的全景图像的上沿, 因而可由该点的空间坐标求得全景图像的垂直视野范围。该点的坐标为

$$\begin{cases} x = f \tan\beta \\ y = -f \tan\beta \\ z = f \left(\tan\alpha \tan\beta - \frac{1}{\cos\alpha} \right) \end{cases} \quad (6)$$

设此时全景图像垂直视角为 θ (指在全景图像中能看到的场景点和摄像头光心的连线与 z 轴负方向的夹角), 则

$$\theta = \arctan \frac{\sqrt{x^2 + y^2}}{|z|} = \arctan \frac{\sqrt{2} \tan\beta \cos\alpha}{1 - \sin\alpha \tan\beta} \quad (7)$$

可求得当 $\sin\alpha = \tan\beta$ 时, θ 可取最大值 $\arctan(\sqrt{2} \tan\alpha)$ 。实验中采用了水平视角约为 85° 的摄像头进行试验, 当将摄像头倾斜角设置为 66° 时, 生成全景视频的垂直视野范围约为 72° , 和理论值完全相符。

若选用水平视野范围大于 90° 的摄像头时, 便不需要考虑摄像头倾斜角的设置, 因为无论倾斜角度为多少, 相邻视频图像总是存在重叠区域。此时可根据实际应用场合的需要, 确定一个合适的倾斜角度。

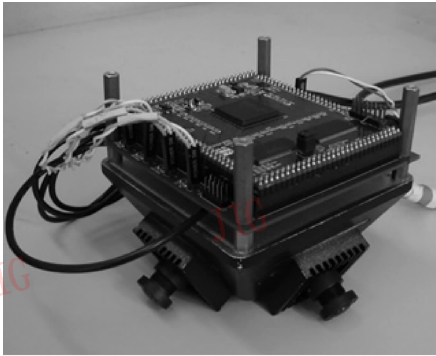
4 实验结果

基于自行研制的嵌入式 DSPIC 开发平台 (DSP for Picture Processing, 图像处理 DSP 平台), 实现了一套原型系统 KD-PVS, 如图 9(a) 所示。DSPIC 开发平台采用 TI 公司 TMS320DM642 DSP 作为嵌入式处理器, 该处理器主频 600MHz, 可支持 8 路指令并行处理。在 DSPIC 平台上集成了 4 个 TVP5150 芯片负责视频信号采集, 一个 SAA7121 芯片负责视频信号输出, 并集成有 32MB 外部存储器, 4MB FLASH 存储器。

对系统加电后, 直接将输出视频线接到普通显示设备 (如彩色电视机) 上, 便可看到实时的全景视

频。该系统最大可输出 2 560 × 600 分辨率的全景视频图像,帧率可达 40fps,垂直视野范围约为 80° (水平向上 10°,向下 70°)。为适应普通电视机显示

的需要,目前将输出的全景视频分辨率设置为 1 280 × 260,并分割成上下两部分显示,即显示分辨率为 640 × 520,实际显示效果如图 9(b)所示。



(a) 原型系统KD-PVS实物图



(b) 全景监控视频

图 9 KD-PVS 实物与监控效果图

Fig. 9 Appearance of KD-PVS and panoramic monitoring video

5 结 论

本文设计和实现了一种基于视频拼图的全景视频监控系 统,可实时输出 360°全景视频信号,图像分辨率高,视觉效果较好。系统轻便健壮,集成度高,可方便应用于金融系统、仓库、监狱和移动监控等多种场合,尤其适用于室内监控。

由于系统采用了单视点近似和简化的拼接算法,所确定的相邻视频图像的重叠区域不十分精确;在图像拼接处,垂直方向上的场景存在微小的偏差。从实验结果看,全景视频的拼接仍有痕迹。因此,如何进一步消除视频图像的拼接痕迹,提高拼接质量,是下一步系统改进的方向。

参考文献 (References)

- 1 Onoe Y, Yamazawa K, Yokoya N, *et al.* Visual surveillance and monitoring system using an omni-directional video camera[A]. In: Proceedings of the 14th International Conference on Pattern Recognition[C], Brisbane, Australia, 1998: 588 ~ 592.
- 2 Chahl J S, Srinivasan M V. A Complete Panoramic Vision System, Incorporating Imaging, Ranging, and Three Dimensional Navigation [A]. In: Proceedings of the IEEE Workshop on Omnidirectional Vision[C], Hilton Head Island, South Carolina, USA, 2000: 104 ~ 111.
- 3 Morita S, Yamazawa K, Yokoya N. Networked video surveillance using multiple omnidirectional cameras[A]. In: Proceedings of IEEE International Symposium on Computational Intelligence in Robotics

- and Automation[C], Kobe, Japan, 2003: 1245 ~ 1250.
- 4 Foote J, Kimber D. FlyCam: Practical panoramic video and automatic camera control[A]. In: Proceedings of IEEE International Conference on Multimedia and Expo[C], Hilton New York & Towers New York City, NY, USA, 2000: 1419 ~ 1422.
- 5 Rybski P E, Torre F de la, Patil R, *et al.* Cameo: Camera assisted meeting event observer [A]. In: Proceedings of the IEEE International Conference on Robotics and Automation[C], New Orleans, USA, 2004: 1634 ~ 1639.
- 6 Benosman R, Kang S B, Faugeras O. Panoramic Vision: Sensors, Theory, and Applications[M]. New York: Springer Verlag, 2001.
- 7 Tan K, Hua H, Ahuja N. Multiview panoramic cameras using mirror pyramids[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, **26**(7): 941 ~ 946.
- 8 Xiong Y, Turkowski K. Creating image-based VR using a self-calibrating Fisheye Lens[A]. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C], Puerto Rico, USA, 1997: 237 ~ 243.
- 9 Dasgupta Sumantra, Banerjee Amarnath. An augmented-reality-based real-time panoramic vision system for autonomous navigation [J]. IEEE Transactions on System, Man, and Cybernetics, Part A: Systems and Humans, 2006, **36**(1): 154 ~ 161.
- 10 Nayar S K, Peri V. Folded Catadioptric Cameras [A]. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C], Corfu, Greece, 1999: 217 ~ 223.
- 11 Wang Wei, Zhong Yan-li, Wu Jian-kun. Real-time system designed for distortion-correction in wide angle lens [J]. Optical Technique, 2004, **30**(4): 489 ~ 494. [王伟, 钟堰利, 吴建坤. 广角镜头图像畸变校正系统的研究 [J]. 光学技术, 2004, **30**(4): 489 ~ 494.]