

# 采用桌面视觉系统实现深度测量的方法

刘曦<sup>1</sup> 李大海<sup>1</sup> 王琼华<sup>1,2</sup> 刘效勇<sup>1</sup>

(<sup>1</sup> 四川大学电子信息学院, 四川 成都 610065  
<sup>2</sup> 视觉合成图形图像技术国家重点学科实验室, 四川 成都 610064)

**摘要** 提出了一种基于桌面视觉系统(DVS)的深度测量方法,并推导了相关的立体图像对校正方法和计算公式,主要包括:几何畸变校正、中心投影的重投影变换、缩放变换、竖直平移变换以及亮度和颜色校正等。通过重投影变换可以使内参数彼此不一致的两个会聚式相机拍摄得到的立体图像对,变换为平行式摆放且具有相同内参数(除几何畸变系数外)的两相机拍摄得到的立体图像对。该方法首先对桌面视觉系统拍摄的立体图像对进行校正处理,使其变换为理想的立体图像对;然后,对校正后的立体图像对进行立体匹配,得到各对同名点的水平视差;最后,基于水平视差和摄像机标定的参数,应用三角测量法求空间点的深度信息。实验结果表明,测量方法的相对误差在0.31%以内。

**关键词** 三维测量;立体图像对校正;重投影变换;缩放变换;竖直平移变换;颜色校正

**中图分类号** TP391.4; P234.1 **文献标识码** A **doi**: 10.3788/LOP48.121501

## Depth Measurement Based on Desktop Vision System

Liu Xi<sup>1</sup> Li Dahai<sup>1</sup> Wang Qionghua<sup>1,2</sup> Liu Xiaoyong<sup>1</sup>

<sup>1</sup> School of Electronics and Information Engineering, Sichuan University,  
Chengdu, Sichuan 610065, China

<sup>2</sup> State Key Laboratory of Fundamental Science on Synthetic Vision, Chengdu, Sichuan 610064, China

**Abstract** An effective depth measurement method based on desktop vision system (DVS) with high precision is proposed. For this purpose, rectification techniques for stereo image pairs are proposed and derived, which mainly consist of geometric distortion correction, reprojection transformation by central projection, zooming transformation, vertical translation transformation and brightness and color correction. The reprojection transformation can change the stereo image pairs captured by a binocular converged camera system with different camera intrinsic parameters into those captured by binocular parallel camera system with the same camera intrinsic parameters (except geometric distortion coefficient). For depth measurement, firstly, the stereo image pairs captured by a DVS are rectified and corrected into perfect ones. And then, horizontal parallax of the corresponding points of interest is gotten by stereo matching. Finally, the depth information of the spatial point of interest is calculated based on horizontal parallax and the parameters obtained by camera calibration. Experimental results show that the proposed depth measurement method is simple and practical, and the relative error is demonstrated for depth measurement of the near distant objects to be no more than 0.31%.

**Key words** three-dimensional measurement; stereo image pair rectification; reprojection transformation; zooming transformation; vertical translation transformation; color correction

**OCIS codes** 150.6910; 150.5670; 150.1135

收稿日期: 2011-06-22; 收到修改稿日期: 2011-08-17; 网络出版日期: 2011-10-17

基金项目: 四川省科技厅支撑计划(2008FZ0003)和国家自然科学基金(60877004)资助课题。

作者简介: 刘曦(1988—),男,硕士研究生,主要从事立体视觉和三维立体显示等方面的研究。

E-mail: liuxi\_bazhong@163.com

导师简介: 李大海(1968—),男,博士,教授,主要从事光学信息处理、波前传感、三维立体显示及测量等方面的研究。

E-mail: lidahai@scu.edu.cn(通信联系人)

本文电子版彩色效果请详见中国光学期刊网(<http://www.opticsjournal.net>)

## 1 引言

对物体进行非接触式的深度测量一直是人们研究的热点,而采用立体相机系统实现物体深度的被动式测量是实现非接触式深度测量的一种重要方法。由于桌面视觉系统(DVS)具有巨大的应用前景<sup>[1,2]</sup>,本文的研究重点集中于桌面视觉系统。本文采用的桌面视觉系统是由两个或多个普通廉价相机随意排列并配置相应的图像采集卡来实现的立体相机系统,因此可用于深度测量。但其主要缺点是:1)镜头的几何畸变一般较大(广角镜头尤其);2)两个相机的内参数彼此不一致,相机的空间摆放关系不是严格意义上的平行式,拍摄的立体图像对之间存在亮度和颜色差异。

基于立体相机系统的深度测量一般分为5步<sup>[3]</sup>:摄像机标定、图像获取、特征提取、立体匹配和深度信息重建。其中为了从二维图像对中提取出度量信息,摄像机标定是必须的步骤<sup>[4]</sup>。由于在立体图像对之间很难准确地搜索到对应点,立体匹配成为这种深度测量方法最薄弱的环节<sup>[5]</sup>。相机光轴的不平行导致垂直视差,于是立体匹配成为二维的搜索过程,立体图像对的校正能够消除垂直视差,使立体匹配由二维搜索变为一维<sup>[4,6]</sup>,能够消除匹配的多义性<sup>[7]</sup>,从而提高了立体匹配的精度和效率。因此,深度测量的步骤可以改进为:摄像机标定、图像获取、立体图像对的校正、立体匹配和深度信息重建。

Fusiello等<sup>[4]</sup>提出了一种基于摄像机标定的立体图像对校正算法,并在此基础上实现了深度测量。然而,该算法仅考虑线性摄像机模型,未考虑非线性模型下的几何畸变对深度测量精度的影响;此外,该算法保持两个原始相机的光心不变,仅将原始立体图像对绕其光心旋转,校正后的图片存在一定的扭曲变形。本文提出了一种有效的立体图像对的校正技术,不但考虑了几何畸变的因素,而且同时执行了绕光心的旋转和对光心的平移,主要包括几何畸变校正、中心投影的重投影变换、缩放变换和竖直平移变换等。实验结果表明,应用本文提出的技术校正得到的立体图像对既适合深度测量又避免了视觉上的扭曲变形现象。

场景光照、遮挡、相机光圈、曝光时间和拍摄视点等因素的差异影响拍摄的立体图像对之间的颜色一致性<sup>[8]</sup>,而照度、相机的灵敏度和非线性自动增益控制等因素的变化导致亮度差异<sup>[9]</sup>。左右视差图像的亮度和颜色差异都会直接影响立体匹配的精度。文献[8]和[9]分别针对颜色差异和亮度差异提出了校正策略。本文基于尺寸不变特征变换(SIFT)特征匹配,提出了一种基于特征匹配点集求校正系数的亮度和颜色校正方法,同时实现了亮度校正和颜色校正。

基于立体图像对的校正技术以及亮度和颜色校正方法,本文提出了一种采用桌面视觉系统实现物体深度测量的方法。首先,对拍摄的立体图像对进行立体图像对的校正以及亮度和颜色校正处理,使其变换成理想的立体图像对。然后,进行立体匹配,计算各对兴趣点的水平视差。最后,基于水平视差和摄像机标定的结果,应用三角测量法求空间点的深度。实验研究表明,本文所提出的深度测量方法简单易行,测量精度高。

## 2 理想系统的空间点深度测量原理

考察如图1所示的理想拍摄系统。该系统中,相机的镜头不存在几何畸变,左右两个相机具有相同的内参数(焦距、主点坐标、像素大小等),两者的光轴相互平行且垂直于基线。

两个相机的镜头置于空气中,其节点与主点重合,分别用 $J_1, J'_1$ 与 $J_2, J'_2$ 表示。在平面上,两个相机的CCD的图像中心分别用 $C_1, C_2$ 表示。空间点 $P$ 在两个相机中

分别成像于像面上 $p_1$ 和 $p_2$ 点。设两个相机的有效焦距均为 $f$ ,两相机之间的基线长度为 $b$ , $P$ 点距离相机物方节点 $J_2$ 的轴向距离(物距)为 $z$ ,像方节点 $J'_2$ 与CCD的垂直距离(像距)为 $l$ , $p_1$ 与 $C_1$ 间的水平距离为 $u_1$ ,

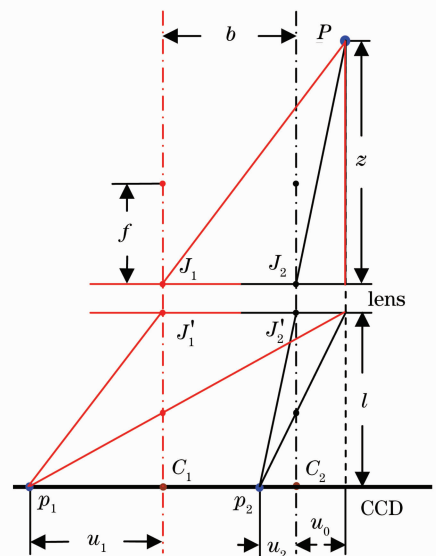


图1 平行式双目视觉测量原理

Fig. 1 Measurement principle of binocular parallel camera system

$p_2$  与  $C_2$  间的水平距离为  $u_2$ ,  $P$  在 CCD 上的垂直投影点与  $C_2$  间的水平距离为  $u_0$ 。

由相似三角形的关系,并结合透镜成像的高斯公式  $1/z + 1/l = 1/f$ ,可得

$$z = \left( \frac{b}{u_1 - u_2} + 1 \right) \cdot f. \quad (1)$$

当空间点  $P$  距离相机较远,使得  $z \gg f$  时,(1)式可化简为

$$z = \frac{b}{u_1 - u_2} \cdot f, \quad (2)$$

式中  $u_1 - u_2$  为水平视差, $z$  为空间点的深度。通过立体匹配得到水平视差,然后结合摄像机标定得到的参数,利用(2)式即可得到空间点的深度信息<sup>[3]</sup>。

### 3 桌面视觉系统的空间点深度测量原理

实际用于深度测量的双目立体相机系统(例如桌面视觉系统)的两个相机在空间中摆放不能达到严格意义上的平行,如图 2 所示。于是,提出对该系统拍摄得到的立体图像对做以下处理。

1) 对两个相机拍摄得到的图像分别进行几何畸变校正,以降低其对水平视差的影响。

2) 以立体图像对中一个图像为参考,对另一个图像进行中心投影的重投影变换,将会聚式得到的立体图像对变换为平行式得到的立体图像对,从而消除梯形失真的不一致。

3) 对重投影变换处理后的立体图像对进行缩放变换,以减小因两个相机镜头的像方主平面不共面而导致的立体图像对水平视差的失真。

对于桌面视觉系统拍摄得到的立体图像对,按以上 3 个步骤校正处理后,利用(2)式即可求得拍摄场景中待测空间点的深度信息。但是,经过以上 3 步处理后的立体图像对仍然存在垂直视差以及亮度和颜色不一致问题,因此还需要做以下的处理:

4) 进行竖直平移变换,消除来自相机镜头的几何畸变<sup>[10]</sup>、梯形失真的不一致<sup>[10]</sup>和图像相对缩放之外两个相机光心在垂直方向的相对平移带来的立体图像对的垂直视差。

5) 对立体图像对进行亮度和颜色不一致的校正处理,消除它们对立体匹配的影响。

对于采用立体视觉系统的深度测量,摄像机标定是必须的步骤<sup>[4]</sup>。鉴于此,本文提出了一种基于摄像机标定的立体图像对的校正方法。

#### 3.1 双目立体相机系统的标定

通过对左右相机分别进行摄像机标定<sup>[1]</sup>可以得到:左(右)相机拍摄到的图像对的主点坐标  $(u_{10}, v_{10})$  [ $(u_{20}, v_{20})$ ]、左(右)相机镜头的几何畸变系数  $k_1$  ( $k_2$ )、左(右)相机焦距  $f_1$  ( $f_2$ ) 分别与像面上单一像素横向尺寸  $d_{x_1}$  ( $d_{x_2}$ ) 和纵向尺寸  $d_{y_1}$  ( $d_{y_2}$ ) 的比值  $(\alpha_1, \beta_1)$  [ $(\alpha_2, \beta_2)$ ]、左(右)相机坐标系相对于世界坐标系的旋转矩阵  $\mathbf{R}_1$  ( $\mathbf{R}_2$ ) 和平移向量  $\mathbf{t}_1$  ( $\mathbf{t}_2$ )。其中  $(u_{10}, v_{10})$ ,  $k_1$ ,  $(\alpha_1, \beta_1)$  称为左相机的内参数,  $\mathbf{R}_1$ ,  $\mathbf{t}_1$  称为外参数;对于右相机同样有内、外参数如上。

基于  $\mathbf{R}_1$ ,  $\mathbf{t}_1$  和  $\mathbf{R}_2$ ,  $\mathbf{t}_2$ , 可以得到右相机坐标系相对于左相机坐标系的空间位置参数:

$$\mathbf{R} = \mathbf{R}_2 \mathbf{R}_1^{-1}, \quad \mathbf{t} = \mathbf{t}_2 - \mathbf{R}_2 \mathbf{R}_1^{-1} \mathbf{t}_1, \quad (3)$$

式中  $\mathbf{R}$  和  $\mathbf{t}$  分别表示右相机坐标系相对于左相机坐标系的旋转矩阵和平移向量。且有

$$\mathbf{X}_R = \mathbf{R} \cdot \mathbf{X}_L + \mathbf{t}, \quad (4)$$

式中  $\mathbf{X}_R = [X_{C_2}, Y_{C_2}, Z_{C_2}]^T$  和  $\mathbf{X}_L = [X_{C_1}, Y_{C_1}, Z_{C_1}]^T$  分别表示空间点在右相机坐标系  $O_2 X_{C_2} Y_{C_2} Z_{C_2}$  和左相机坐标系  $O_1 X_{C_1} Y_{C_1} Z_{C_1}$  下的坐标。

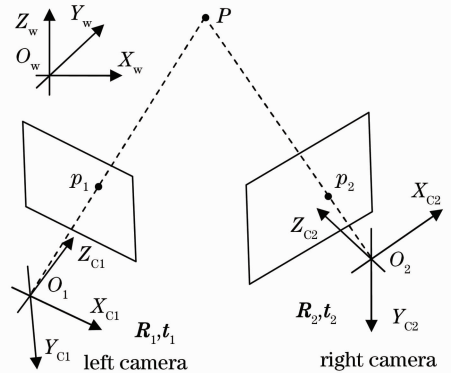


图 2 双目桌面视觉系统示意图

Fig. 2 Schematic of binocular DVS

### 3.2 立体图像对的校正

本文以右图为参考对左图进行立体图像对的校正,下面是对该方法的具体描述。

#### 3.2.1 几何畸变校正

这里主要考虑径向畸变和切向畸变。径向畸变是引起摄像机镜头畸变最主要的因素<sup>[1]</sup>,一般情况下,对于摄像机的非线性模型,一阶径向畸变已能足够描述非线性畸变<sup>[2]</sup>。设通过摄像机标定得到的一阶径向畸变系数为  $k$ ,  $(x, y)$  表示针孔线性成像模型下图像点坐标的理想值,  $(x_d, y_d)$  表示实际的考虑到镜头几何畸变的图像点坐标。径向畸变的中心与图像的主点重合,可得<sup>[2,11,12]</sup>

$$\begin{cases} x_d = [1 + k(x^2 + y^2)]x \\ y_d = [1 + k(x^2 + y^2)]y \end{cases}, \quad \begin{cases} x = (u - u_0)d_x \\ y = (v - v_0)d_y \end{cases}, \quad \begin{cases} x_d = (u_d - u_0)d_x \\ y_d = (v_d - v_0)d_y \end{cases}, \quad (5)$$

式中  $(u, v)$  和  $(u_d, v_d)$  表示以像素为单位的图像点坐标。于是,基于标定得到的一阶径向畸变系数  $k$ 、图像的主点坐标  $(u_0, v_0)$ 、图像中单一像素横向尺寸  $d_x$  和纵向尺寸  $d_y$ , 应用(5)式即可实现几何畸变校正。

#### 3.2.2 中心投影的重投影变换

为了消除梯形失真不一致的影响,提出了中心投影的图像纠正重投影变换。重投影变换是以一个相机拍摄的图像为参考,对另一个相机拍摄的图像以相机的像方节点为投影中心进行变换重采样,成为一个光轴与参考图像对应相机的光轴平行的虚拟相机拍摄的图像。于是,重投影得到的图像与参考图像组成的立体图像对变换成了平行式双目相机系统拍摄得到的立体图像对。

对于(4)式,当不考虑平移向量,即令  $t=0$  时,有

$$[X_{C2}, Y_{C2}, Z_{C2}]^T = \mathbf{R} \cdot [X_{C1}, Y_{C1}, Z_{C1}]^T. \quad (6)$$

根据相机成像的透视投影关系,有<sup>[11]</sup>

$$x_i = f_i \cdot X_{Ci}/Z_{Ci}, \quad y_i = f_i \cdot Y_{Ci}/Z_{Ci}, \quad (i = 1, 2) \quad (7)$$

将(7)式代入(6)式并考虑(5)式中  $(x, y)$  和  $(u, v)$  之间的关系,可以得到

$$\begin{cases} u_2 = u_{20} + \alpha_2 \cdot \frac{a_1(u_1 - u_{10})/\alpha_1 + a_2(v_1 - v_{10})/\beta_1 + a_3}{c_1(u_1 - u_{10})/\alpha_1 + c_2(v_1 - v_{10})/\beta_1 + c_3} \\ v_2 = v_{20} + \beta_2 \cdot \frac{b_1(u_1 - u_{10})/\alpha_1 + b_2(v_1 - v_{10})/\beta_1 + b_3}{c_1(u_1 - u_{10})/\alpha_1 + c_2(v_1 - v_{10})/\beta_1 + c_3} \end{cases} \quad \text{with } \mathbf{R} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}. \quad (8)$$

根据前述分析,(8)式就是以右图为参考对左图进行中心投影的重投影变换的结果,将结果表示为  $(u'_1, v'_1)$ 。

设重投影变换后的左相机坐标系为  $O'_1 X'_{C1} Y'_{C1} Z'_{C1}$ , 相机的焦距为  $f'_1$ 。即使  $f_1 \neq f_2$ , 但是当两个相机 CCD 阵列结构一样时,由于  $d_{x'_1} = d_{x_1} = d_{x_2}$ ,  $d_{y'_1} = d_{y_1} = d_{y_2}$  以及  $(\alpha'_1, \beta'_1) = (\alpha_2, \beta_2)$ , 于是可得  $f'_1 = f_2 = f$ 。因此,通过重投影变换可以使内参数彼此不一致的两个会聚式相机拍摄得到的立体图像对,变换为平行式摆放且具有相同内参数(除几何畸变系数外)的两相机拍摄得到的立体图像对。

#### 3.2.3 缩放变换

利用重投影变换处理后的立体图像对的水平方向坐标信息求缩放因子,进行缩放变换。

重投影变换使得  $\mathbf{R} = \mathbf{I}$ (单位矩阵),于是得到  $\mathbf{X}_R = \mathbf{X}'_L + \mathbf{t}$ , 其中  $\mathbf{X}'_L = [X'_{C1}, Y'_{C1}, Z'_{C1}]^T$ 。令  $\mathbf{t} = [\Delta X, \Delta Y, \Delta Z]^T$  (其中  $\Delta X, \Delta Y$  和  $\Delta Z$  分别表示右相机坐标系相对于左相机坐标系在水平、竖直和光轴方向的平移量),  $(x'_1, y'_1)$  表示重投影变换后左图的图像坐标,考虑相机成像的透视投影关系,可得

$$Z_{C2} = \frac{x'_1 \cdot \Delta Z - f \cdot \Delta X}{x'_1 - x_2}. \quad (9)$$

设缩放变换后的图像坐标为  $(x''_1, y''_1)$ , 则横向和纵向的缩放因子分别为  $\eta_x = x''_1/x'_1$ ,  $\eta_y = y''_1/y'_1$ 。缩放前后满足关系  $\mathbf{X}''_L = \mathbf{X}'_L + [0, 0, \Delta Z]^T$ , 其中  $\mathbf{X}''_L = [X''_{C1}, Y''_{C1}, Z''_{C1}]^T$ 。进一步考虑相机成像的透视投影关系,可得

$$\eta = \eta_x = \eta_y = (Z_{C2} - \Delta Z)/Z_{C2}. \quad (10)$$

由(10)式可以看出,缩放因子  $\eta$  的大小取决于空间点的深度  $Z_{C2}$ 。本文通过选择左右视差图像中的特征点来求缩放因子,为简单起见,假定检测的特征点具有相似的深度<sup>[13]</sup>。由于对图像旋转、尺度缩放、亮度和颜色变化、视角变化、噪声等保持不变性和稳定性<sup>[14]</sup>, 为了实现差异较大的两幅图像的特征点匹配,采用 SIFT 特征匹配算法<sup>[15]</sup>进行特征点的匹配。通过 SIFT 特征匹配,假设获得了  $N$  对匹配点,通过(9)式计算

各对匹配点的深度,表示为  $Z_{C2i}$  ( $i = 1, 2, \dots, N$ )。可得深度  $Z_{C2}$  为

$$Z_{C2} = 1/N \sum_{i=1}^N Z_{C2i}. \quad (11)$$

各对 SIFT 特征匹配点的像素图像坐标表示为  $(u'_{1i}, v'_{1i})$  和  $(u_{2i}, v_{2i})$ , 而重投影变换后的左右视差图是具有相同内参数的双目相机拍摄得到的,有

$$x'_{1i} = (u'_{1i} - u_{20})d_{x_2} = (u'_{1i} - u_{20})f/\alpha_2, \quad x_{2i} = (u_{2i} - u_{20})f/\alpha_2. \quad (12)$$

于是,将(9),(12)式代入(11)式,可得

$$Z_{C2} = 1/N \sum_{i=1}^N \frac{(u'_{1i} - u_{20}) \cdot \Delta Z - \alpha_2 \cdot \Delta X}{u'_{1i} - u_{2i}}. \quad (13)$$

因此,利用(13)式求到深度  $Z_{C2}$ ,再由(10)式即可求得缩放因子  $\eta$ 。

### 3.2.4 竖直平移变换

重投影变换和缩放变换后:  $\mathbf{R} = \mathbf{I}, \mathbf{t} = [\Delta X, \Delta Y, 0]^T$ 。于是有,  $\mathbf{X}_R = \mathbf{X}_L'' + [\Delta X, \Delta Y, 0]^T$ 。考虑相机成像的透视投影关系,可得

$$y_2 - y_1'' = f \cdot \Delta Y / Z_{C2}. \quad (14)$$

类似于(12)式,有  $y_1'' = (v'_{1i} - v_{20})f/\beta_2, y_2 = (v_{2i} - v_{20})f/\beta_2$ 。代入(14)式可得竖直平移变换量  $t_y$

$$t_y = v_2 - v_1'' = \beta_2 \cdot \Delta Y / Z_{C2}. \quad (15)$$

因此,利用(15)式即可求得竖直平移变换量  $t_y$ 。

### 3.3 亮度和颜色校正

左右视差图的亮度和颜色不匹配会直接影响立体匹配的精度。本文基于 SIFT 特征匹配,提出了一种基于特征匹配点集求校正系数的亮度和颜色校正方法。它根据获得的  $N$  对匹配点,在 YCbCr 颜色空间(其中 Y 代表亮度,Cb 和 Cr 表示色度)下求各通道的校正系数  $k_Y, k_{Cb}$  和  $k_{Cr}$ ,进而对目标图像左图进行亮度和颜色校正,最后将校正后的目标图像转换到红绿蓝(RGB)颜色空间。

1) 将左右视差图中的各个匹配点分别进行 RGB 颜色空间到 YCbCr 颜色空间的转换。转换公式为

$$\begin{bmatrix} y \\ c_b \\ c_r \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1687 & -0.3313 & 0.5000 \\ 0.5000 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}. \quad (16)$$

2) 分别求参考图像中各匹配点的  $y$  分量、 $c_b$  分量、 $c_r$  分量的总和  $Y_1, C_{b1}, C_{r1}$  以及目标图像的  $Y_2, C_{b2}, C_{r2}$ 。然后可得 Y, Cb, Cr 通道的校正系数  $k_Y, k_{Cb}, k_{Cr}$ :

$$k_Y = Y_1/Y_2, \quad k_{Cb} = C_{b1}/C_{b2}, \quad k_{Cr} = C_{r1}/C_{r2}. \quad (17)$$

3) 将目标图像中各个像素点分别进行 RGB 颜色空间到 YCbCr 颜色空间的转换,转换公式同 1) 中所述。对目标图像中各个像素点的  $y$  分量、 $c_b$  分量、 $c_r$  分量的值分别乘以校正系数  $k_Y, k_{Cb}, k_{Cr}$  进行校正。

4) 将校正后的目标图像进行 YCbCr 颜色空间到 RGB 的颜色空间的转换,转换公式为

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.402 \\ 1 & -0.34414 & -0.71414 \\ 1 & 1.772 & 0 \end{bmatrix} \begin{bmatrix} y \\ c_b - 128 \\ c_r - 128 \end{bmatrix}. \quad (18)$$

### 3.4 深度测量

(2)式中,  $z, b$  和  $f$  以物理单位(例如毫米)进行描述,而  $u_1$  和  $u_2$  以像素为单位,于是,引入单一像素的横向尺寸  $d_x$ 。桌面视觉系统拍摄的立体图像对进行了前述校正处理后,三角测量法的公式可以表示为

$$z = \frac{b \cdot f}{(u_1'' - u_2) \cdot d_{x_2}} = \frac{b}{u_1'' - u_2} \cdot \alpha_2, \quad (19)$$

式中  $\alpha_2$  和基线长度  $b$  ( $b = \Delta X$ ) 可通过双目立体相机系统的标定得到,水平视差  $(u_1'' - u_2)$  通过立体匹配得到。

## 4 桌面视觉系统深度测量的实现步骤

本文实现深度测量的基本步骤如下:

1) 对拍摄系统进行摄像机标定,得到两个相机各自的内参数及相对空间位置参数;

- 2) 利用桌面视觉系统拍摄立体图像对,并采用均值滤波法降低随机噪声的影响<sup>[16]</sup>;
- 3) 利用标定所得几何畸变系数分别对左右视差图像进行几何畸变校正;
- 4) 利用标定所得参数,以校正了几何畸变后的右图作为参考,对校正了几何畸变后的左图依次进行中心投影的重投影变换、缩放变换和垂直平移变换;
- 5) 对校正后的立体图像对,以左图作为参考对右图进行亮度和颜色校正;
- 6) 应用立体匹配算法对校正处理后立体图像对的右图中感兴趣的像素点在左图中分别搜索其对应点,求各对同名点的水平视差。采用基于自适应权值算法、Rank 变换和视差校准的彩色图像立体匹配算法<sup>[17]</sup>。
- 7) 基于水平视差和标定参数,应用(19)式求各兴趣点的深度信息。

## 5 实验结果

实验研究发现,当以视场较小的相机拍摄的图像为参考进行立体图像对的校正时,视场较大的相机拍摄的图像在校正后存在大量的“裂缝”(如图 3 所示),且得到的立体图像对用于立体匹配时的误匹配相当严重。因此,本文以视场较大的相机拍摄的图像为参考。此时如果  $k < 0$  (桶形畸变),几何畸变校正也会带来一些“裂缝”,需要进行插值处理。当两个相机的视场相等时,几何畸变校正、重投影变换和缩放变换都分别可能会带来“裂缝”。为了减少计算量,“裂缝”的插值处理是通过周围像素点的像素值求平均实现的<sup>[13]</sup>。

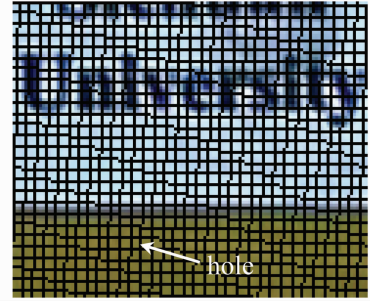


图 3 图像中的“裂缝”

Fig. 3 “Holes” in the sample image

图 4 所示的立体图像对是在室内利用双目桌面视觉系统拍摄得到的。

采用本文提出的立体图像对的校正方法,以图 4(b)为参考,对图 4(a)进行校正处理,校正后的立体图像对如图 5 所示。校正过程中相应的实验参数如下:

$$k_1 = -0.1185, \quad (u_{10}, v_{10}) = (412.6, 278.2), \quad (\alpha_1, \beta_1) = (985.8, 980.4);$$

$$k_2 = -0.1770, \quad (u_{20}, v_{20}) = (398.4, 269.1), \quad (\alpha_2, \beta_2) = (725.1, 721.5);$$

$$\mathbf{R} = \begin{bmatrix} 0.9981 & -0.0132 & 0.0598 \\ 0.0141 & 0.9998 & -0.0149 \\ -0.0596 & 0.0157 & 0.9981 \end{bmatrix}, \quad \mathbf{t} = [-161.7, 4.6, -1.2]^T; \quad \eta = 1.0011, \quad t_y = 3.0 \text{ pixel}.$$

为了直观地评估校正结果,在图 4 和图 5 中分别画出了 7 条白色水平线,可以看出,垂直视差很好地消除了,而且校正后的立体图像对不存在明显的视觉上的扭曲变形现象。此外还对垂直视差的变化做出定量评估。通过 SIFT 特征匹配对原始立体图像对和校正后的图像对分别选取 40 对匹配点,然后计算校正前后各对匹配点纵坐标绝对差异的平均(MAD)<sup>[18]</sup>和垂直视差的均方根(RMS)值<sup>[13]</sup>,结果如表 1 所示。由表 1 可以看出,本文的立体图像对校正方法能有效地消除垂直视差。

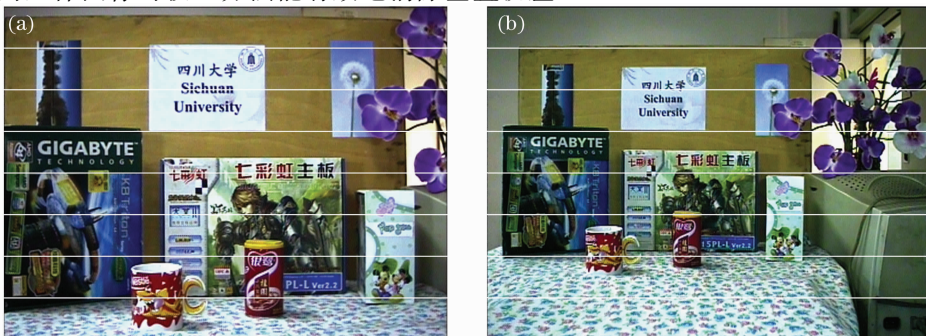


图 4 拍摄得到的立体图像对。(a)左图;(b)右图

Fig. 4 Original stereo images. (a) The left image; (b) the right image

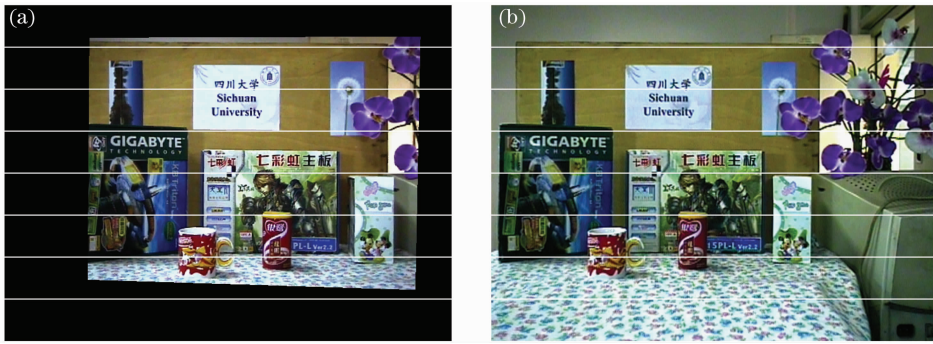


图 5 校正后的立体图像对。(a)左图;(b)右图

Fig. 5 Rectified stereo images. (a) The left image; (b) the right image

表 1 垂直视差的定量评估

Table 1 Quantitative evaluations of vertical parallax

	Original stereo images	Rectified stereo images
MAD /pixel	39.6517	0.3482
RMS /pixel	43.6987	0.4124

从图 5 可以明显看出,校正后的左右视差图像存在亮度和颜色差异。采用本文提出的亮度和颜色校正方法,以图 5(a)为参考对图 5(b)进行校正处理,结果如图 6 所示。校正参数为  $k_Y=1.0821$ ,  $k_{Cb}=1.0001$ ,  $k_{Cr}=1.0002$ 。

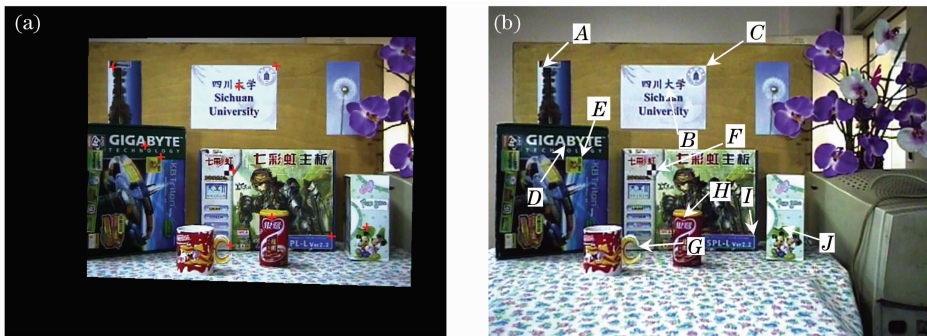


图 6 亮度和颜色校正后的立体图像对。(a)左图;(b)右图

Fig. 6 Generated stereo images after rectification and correction. (a) The left image; (b) the right image

考察图 6(b)中随机选择的 10 个拟被测量点 A~J。右图中这 10 个点的图像坐标如表 2 中“Interested coordinate”栏所示。由于立体图像对的垂直视差已经校正至 0.5 pixel 以下,进行立体匹配时只需要一维搜索。采用文献[17]的立体匹配算法分别得到图 6(b)中 10 个兴趣点在左图中的对应点[如图 6(a)所示,以红色“+”表示],其坐标如表 2 中“Searched coordinate”栏所示。计算各对匹配点的水平视差,利用(19)式求得所考察的 10 个点的空间深度,如表 2 中“Measured depth”栏所示。为了验证测量的结果,实际测量了这 10 个空间点的深度,其结果如表 2 中“Real depth”栏所示。同时,计算了深度测量的相对误差,其结果也表示在表 2 中。

通过表 2 可以看出,本文的深度测量方法的结果与实际的深度信息之间的相对误差非常小。实验结果表明,本文提出的基于立体图像对校正及亮度和颜色校正的桌面视觉系统深度测量的方法简单易行,而且其测量结果具有较高的精度。然而,需要指出的是,与其他所有的被动光学深度测量方法一样,本文的测量结果还依赖于摄像机标定的精度和立体匹配的精度。

表 2 深度测量的结果  
Table 2 Results of depth measurement

Tested points	Interested coordinate /pixel	Searched coordinate /pixel	Measured depth /m	Real depth /m	Relative error /%
A	(95, 107)	(189, 107)	1.2471	1.251	0.31
B	(314, 140)	(408, 140)	1.2471	1.251	0.31
C	(379, 105)	(473, 105)	1.2471	1.251	0.31
D	(134, 243)	(245, 243)	1.0561	1.058	0.18
E	(163, 263)	(272, 263)	1.0755	1.077	0.14
F	(292, 286)	(398, 286)	1.1059	1.109	0.28
G	(261, 417)	(392, 417)	0.8948	0.896	0.13
H	(341, 368)	(465, 368)	0.9454	0.943	0.25
I	(463, 401)	(569, 401)	1.1059	1.109	0.28
J	(512, 385)	(630, 385)	0.9934	0.995	0.16

## 6 结 论

考虑到桌面视觉系统的特点,提出了一种立体图像对的校正技术以及一种亮度和颜色校正方法,基于此,提出了一种有效的采用桌面视觉系统的深度测量方法。首先,对桌面视觉系统拍摄的立体图像对进行立体图像对的校正处理,使其变换为理想的立体图像对;其次,对校正后的立体图像对进行亮度和颜色校正;然后,应用立体匹配,得到各对同名点的水平视差;最后,基于水平视差和摄像机标定的参数,由三角测量法求空间点的深度信息。其中,立体图像对的校正处理技术主要包括几何畸变校正、中心投影的重投影变换、缩放变换和竖直平移变换等。实验结果表明,该方法简单易行,测量结果的相对误差在 0.31% 以内。

## 参 考 文 献

- Zhengyou Zhang. A flexible new technique for camera calibration[J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000, **22**(11): 1330~1334
- Zhengyou Zhang. Flexible camera calibration by viewing a plane from unknown orientations[C]. Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999, **1**: 666~673
- Wang Feng, Dai Shuguang. 3-D reconstruction based on binocular stereo vision[C]. *SPIE*, 2009, **7283**: 728331
- Andrea Fusiello, Emanuele Trucco, Alessandro Verri. A compact algorithm for rectification of stereo pairs[J]. *Machine Vision and Applications*, 2000, **12**(1): 16~22
- Kazuyuki Miyazawa, Takafumi Aoki. A robot-based 3D body scanning system using passive stereo vision[C]. 15th IEEE International Conference on Image Processing, 2008, 305~308
- Jin Zhou, Baoxin Li. Image rectification for stereoscopic visualization[J]. *J. Opt. Soc. Am. A*, 2008, **25**(11): 2721~2733
- Zeng Jiyong, Su Xianyu. A new rectification method of stereo pairs without calibration for cameras[J]. *Acta Optica Sinica*, 2004, **24**(5): 628~632
- 曾吉勇, 苏显渝. 一种无相机标定的立体图像对校正新方法[J]. *光学学报*, 2004, **24**(5): 628~632
- Feng Shao, Mei Yu, Gangyi Jiang. Dominant color extraction based color correction for multi-view images[J]. *Chin. Opt. Lett.*, 2007, **5**(8): 449~451
- Yunfang Zhu. A new method of brightness correction for multi-view images[C]. IEEE International Conference on Intelligent Computing and Intelligent Systems, 2009, **4**: 547~551
- Andrew Woods, Tom Docherty, Rolf Koch. Image distortions in stereoscopic video systems[C]. *SPIE*, 1993, **1915**: 36~48
- Juyang Weng, Paul Cohen, Marc Herniou. Camera calibration with distortion models and accuracy evaluation[J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1992, **14**(10): 965~980
- Ye Nan, Zhang Liyan. Determining plastic strain ratio of sheet metal based on stereo vision[J]. *Acta Optica Sinica*, 2010, **30**(12): 3515~3523
- 叶 南, 张丽艳. 基于立体视觉的板料塑性应变比测量[J]. *光学学报*, 2010, **30**(12): 3515~3523
- H.-M. Wang, C.-W. Chang, J.-F. Yang. An effective calibration procedure for correction of parallax unmatched image



- pairs[J]. *IET Image Processing*, 2009, **3**(2): 63~74
- 14 Zhang Chao, Wang Qionghua, Li Dahai *et al.*. Acquisition of parallax images for multi-view auto-stereoscopic display based on SIFT matching algorithm[J]. *Acta Optica Sinica*, 2010, **30**(7): 1989~1993  
张超, 王琼华, 李大海等. 基于 SIFT 匹配算法的多视点自由立体显示视差图像的生成[J]. *光学学报*, 2010, **30**(7): 1989~1993
- 15 David G. Lowe. Distinctive image features from scale-invariant keypoints[J]. *Internat. J. Computer Vision*, 2004, **60**(2): 91~110
- 16 Ding Xiaohua, Li You, Yu Qifeng *et al.*. CCD noise calibration and its application in edge location[J]. *Acta Optica Sinica*, 2008, **28**(1): 99~104  
丁晓华, 李由, 于起峰等. CCD 噪声标定及其在边缘定位中的应用[J]. *光学学报*, 2008, **28**(1): 99~104
- 17 Zheng Gu, Xianyu Su, Yuankun Liu *et al.*. Local stereo matching with adaptive support-weight, rank transform and disparity calibration[J]. *Pattern Recogn. Lett.*, 2008, **29**(9): 1230~1235
- 18 H.-H. P. Wu, Y.-H. Yu, W.-C. Chen. Projective rectification based on relative modification and size extension for stereo image pairs[J]. *IEE Proceedings-Vision, Image and Signal Processing*, 2005, **152**(5): 623~633