

文章编号:1007-2780(2014)06-1144-07

基于非标定双目网络摄像头的 实时立体显示系统设计

姜瑞凯*, 哈清华

(中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033)

摘要: 立体显示可以为观察者提供场景中物体的深度信息, 并使观察者有身临其境的感觉。为了搭建低成本可应用于辅助远程操纵的实时画面立体显示系统原型, 利用双目网络摄像头和立体显示设备, 在未对相机标定情况下, 采用非标定立体校正的方法, 设计了一套实时立体显示系统, 提供了色差式和快门式 2 种立体显示方式, 并通过 MOS 方法对系统的立体显示效果进行了主观评价。实验结果表明, 经过立体校正后的左右摄像头图像行是对准的。在使用了多线程技术对左右图像并行处理的情况下, 系统输出立体图像的帧频从 21 Hz 提高到了 30 Hz, 满足实时性的要求。在立体显示效果评价中, 色差式和快门式立体显示分别得到了 7.67 分和 9.08 分, 说明立体效果基本令人满意。

关键词: 立体显示; 非标定; 立体校正; 网络摄像头; 双目

中图分类号: TP37 **文献标识码:** A **doi:** 10.3788/YJYXS20142906.1144

Design of real time stereoscopic display system based on binocular web cameras uncalibrated

JIANG Rui-kai*, HA Qing-hua

(Changchun Institute of Optics, Fine Mechanics and Physics China,
Chinese Academy of Sciences, Changchun 130033, China)

Abstract: Stereoscopic display can provide depth information and immersive knowledge about the scenario for the observers. In order to build the low-cost prototype of stereoscopic display system which can aid remote operation, a real-time stereoscopic display system without the calibration to the camera was designed. The system utilized binocular web cameras and stereo display device to provide two stereoscopic display modes: anaglyphic mode and active shutter mode, the performance of which are evaluated by MOS method. Experimental results show that the two images of the binocular web cameras are row-aligned rectified. In the case of parallel processing of images with multi-thread technique, the frame rate of the stereo image output is upgraded to 30 Hz from 21 Hz, which meets the real time requirements. Evaluation proves that the score of anaglyphic mode and active shutter mode are 7.67 and 9.08, respectively. The performance of stereoscopic display is basically satisfactory.

Key words: stereoscopic display; uncalibrated; stereo rectification; web camera; binocular

收稿日期:2014-06-12; 修订日期:2014-07-22.

基金项目:国家 863 计划基金资助项目(No. 2012AA7031010B)

* 通信联系人, E-mail: jrkwood@gmail.com

1 引言

双目立体视觉^[1]是机器视觉中的一个重要分支,从仿生学的角度来讲,双目视觉更符合人类双眼的构造。它利用双目产生视差的原理能够提供景物的深度信息,而在单纯使用单目视觉的情况下却不能提供这个信息。因此在视觉测量应用中,双目视觉更具有优势。在计算出图像中相应的点的视差后,采用基于三角的测量方法可以获得空间点的三维坐标^[2]。在机器人系统应用中,双目视觉可以测量出目标的位置和姿态信息^[3],以为机器人控制系统提供观测值。在人工智能技术并不完善的今天,还有许多机器人控制系统仍然是人在回路的,操作员通过机器人传回的图像或其他观测值,对机器人远程操作或发送指令。立体显示技术可以给操作员更加逼真的距离感和身临其境的感觉。近些年来,立体显示技术得到了迅速的发展和推广,3D 电视和 3D 电影已经走进了大众的生活^[4]。立体显示技术的主要原理是采用的一定的方法,使人的左右眼分别看到具有视差的左右两幅图像。目前主流的立体显示技术分为眼镜式和裸眼式两类^[5-8]。眼镜式 3D 主要包括色差式,快门式和偏光式;裸眼式 3D 主要包括柱状透镜式,光屏障式和指向光源式。虽然裸眼式 3D 技术摆脱了眼镜的束缚,但是在分辨率、可视角度和可视距离等方面还存在很多不足,与眼镜式 3D 技术尚有一定的差距。因此本文采用的立体显示方式主要是基于眼镜式 3D 技术。另外,还有一种头盔式的立体显示技术^[9],在头盔上配有左右两个显示设备和光学系统,使用者戴上头盔后即可观察到左右两幅图像,这种方式比较适合在机器人控制系统中操作员佩戴。本文利用双目网络摄像头和立体显示终端搭建了一套实时立体显示系统,使用非标定校正算法对左右图像校正,介绍了色差式和快门式 3D 显示的图像处理的实现方法,设计了一个时延较小的多线程立体显示图像处理软件。

2 非标定立体校正

在立体显示中,左右摄像机理想的成像平面应该是共面和行对准的,因为这样更符合人眼的

观察习惯。但是,完美的对准结构在真实的立体视觉系统中是几乎不存在的,尤其对于网络摄像头,左右图像存在较大行错位和非共面的问题。如果不对图像进行一定处理,而是通过立体显示设备直接给人眼观察,会造成观察者有眩晕等不适的感觉。通过数学方法将左右摄像机的图像平面重新投影,使它们落在同一个平面上,而且行对准的过程称为立体校正。立体校正可分为两类^[10]:一类是标定立体校正,先通过标定的方法得到左右摄像机的内外参数,然后使用内外参数进行立体校正,这种方法有对图像校正比较准确和产生畸变较小的优点,但是,由于标定需要标定参考物和人工干预,因此不太适合摄像机参数易发生变化(如根据场景调焦)的情况;另一类是非标定立体校正,这类方法无需对摄像机标定,利用场景中左右图像的一些匹配点,即可完成对图像的校正,因此比较灵活。在非标定立体校正中比较典型的是 Hartley 算法^[11]。

Hartley 算法的主旨是寻找使两幅图像对应点总的视差最小化,并且将极点映射到无穷处的投影变换矩阵 H_l 和 H_r ,其中极点是指左右相机光心连线与左右图像的交点。Hartley 算法的基本步骤如下:

- 1. 识别左右图像中的匹配点 $u_i \leftrightarrow u'_i$,至少需要找出 8 对匹配点,而且越多越好。
- 2. 计算基础矩阵 F ,基础矩阵是秩为 2 的 3×3 矩阵,对于图像中的匹配点 $u_i \leftrightarrow u'_i$,它们满足:

$$u_i'^T F u_i = 0. \tag{1}$$

如果给出至少 8 对匹配点,可以求解由(1)式组成的线性方程组得到 F 。

- 3. 计算左右极点在图像中的坐标,设左右极点分别为 e_l 和 e_r ,则它们满足:

$$F e_l = 0, F^T e_r = 0. \tag{2}$$

根据(2)式可以求得左右极点的位置。

- 4. 构造投影变换矩阵 H_r 将右极点映射到水平方向的无穷远处。平移变换矩阵 T 将右图像坐标原点移到图像中心,其形式如下:

$$T = \begin{pmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3}$$

(u_0, v_0) 是右图像中心像素坐标。经过平移变换后,右极点变换到位置 $T e_r$,假设此时右极点

坐标为 $(u, v, 1)^T$,再经过旋转变换,使右极点变换到 X 轴上,旋转变换矩阵 R 的形式为:

$$R = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

其中: $\theta = \arctan(v/u)$,经过旋转变换后,右极点变换为 $R\mathbf{T}e_r$,已经位于 X 轴上,假设此时右极点坐标为 $(k, 0, 1)^T$,还需要一个变换将右极点移至水平无穷远处,即 $(1, 0, 0)^T$,这个变换的形式为:

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/k & 0 & 1 \end{pmatrix}. \tag{5}$$

经过以上三个变换,右极点被映射到水平无穷远处,所以投影矩阵 H_r 为:

$$H_r = GRT. \tag{6}$$

5. 查找匹配的投影矩阵 H_l ,它将左极点映射到无穷远,并保证两幅图像的行对准,同时使得左右匹配点总的视差最小,即平方和距离最小化:

$$\sum_i d(H_l \mathbf{u}_i, H_r \mathbf{u}'_i)^2. \tag{7}$$

由文献[11]可知,满足上述条件的 H_l 应具有如下形式:

$$H_l = A H_r M. \tag{8}$$

其中: M 是一个非奇异矩阵,由基础矩阵 F 因数分解得到:

$$F = [e_r] \times M. \tag{9}$$

A 是一个仿射变换矩阵:

$$A = \begin{pmatrix} a & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{10}$$

其中 a, b, c 使得下式最小化:

$$\sum_i (a\hat{u}_i + b\hat{v}_i + c - \hat{u}'_i)^2. \tag{11}$$

上式中 $(\hat{u}_i, \hat{v}_i, 1) = H_l \mathbf{u}_i, (\hat{u}'_i, \hat{v}'_i, 1) = H_r \mathbf{u}'_i$,可以用线性最小二乘的方法求解 a, b, c ,进而得到 H_l 。

6. 根据投影矩阵 H_l 和 H_r ,对左右图像分别做投影变换,得到校正后的图像。

3 系统设计与实现

3.1 系统环境搭建

本文中的立体显示系统组成框图如图 1 所

示,主要包括以下 4 个部分:

1. 图像获取设备 选用的是 2 个同样的 ANC720P 网络高清摄像头,输出图像分辨率为 1280×720 ,帧频为 30 帧/秒。

2. 图像处理器 选用的是 DELL 通用 PC 机,搭载了酷睿 i3-3240@3.40 GHz 处理器和 4G DDR3 内存。操作系统为 Windows XP,开发工具使用 VC++6.0 和 OpenCV1.1。

3. 立体显示设备 对于色差式的立体显示方式,使用普通显示器即可。对于快门式或偏振式立体显示方式,需要使用快门式或偏振式立体显示器。本系统选用的是海信的一款 3D 电视作为立体显示设备,它支持快门式立体显示方式,也可以作为普通显示器使用。

4. 立体眼镜 对于色差式的立体显示,配合使用红蓝立体眼镜。对于快门式立体显示,系统使用的是海信 3D 电视配套的快门式立体眼镜。

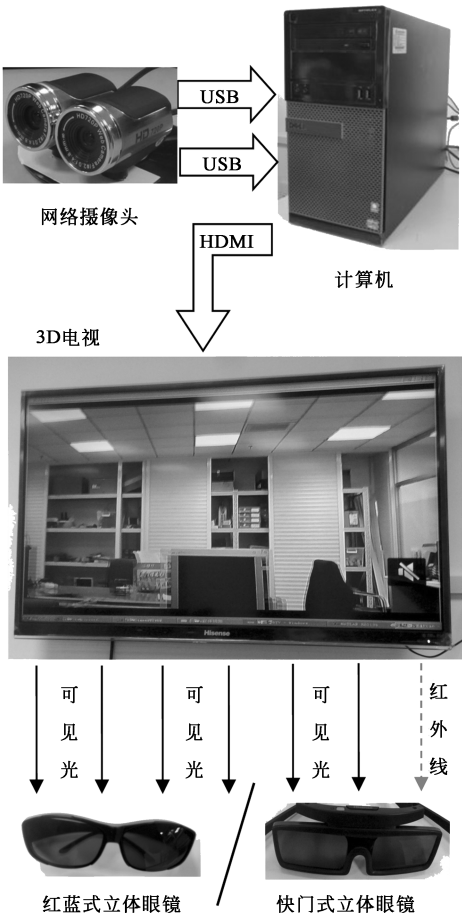


图 1 立体显示系统框图

Fig. 1 Flowchart block diagram of stereoscopic display system

3.2 系统软件设计

本文中设计的立体显示系统软件部分需要完成的功能是:根据 Hartley 算法计算投影变换矩阵 H_l 和 H_r ;对实时采集的图像进行校正,使校正后左右图像共面并且行对准,再将左右图像的公共区域合成为立体图像输出显示。首先按照 Hartley 算法的基本步骤计算投影变换矩阵 H_l 和 H_r ,如果摄像头的参数和相对位置不发生变化,此过程只需要进行一次;如果改变了摄像头的参数和相对位置,则需要重新计算投影变换矩阵 H_l 和 H_r 。使用 SURF 特征点匹配可以完成在当前的场景中寻找左右图像的匹配点,此过程自动完成,不需要人工干预。在计算得到投影矩阵 H_l 和 H_r 后,对实时采集的图像进行校正。对于高分辨率的图像,投影变换是一种较为耗时的处理,由于对左右图像的校正可以同时进行,因此系统采用多线程的方法对左右图像并行处理,以减少总的处理时间。立体显示系统软件的处理流程如图 2 所示。

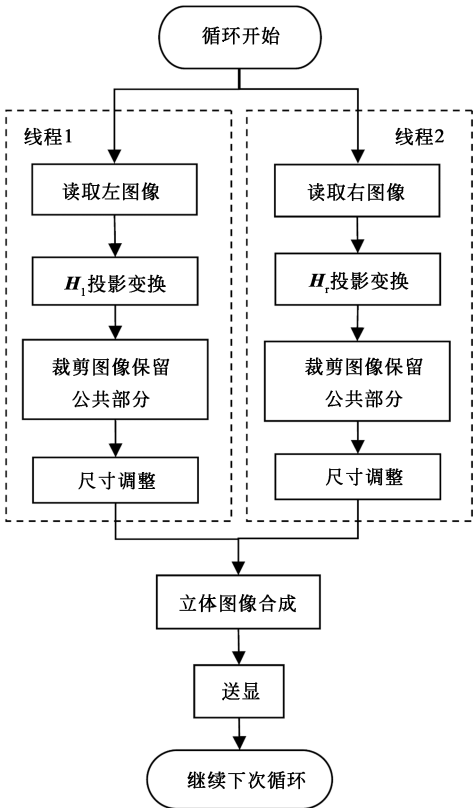


图 2 立体显示系统软件流程图
Fig. 2 Flow chart of stereoscopic display system software

3.2.1 读取图像与投影变换

读取图像完毕后,对图像进行投影变换,这里采用的是后向映射双线性插值的方法计算变换后的每个像素点灰度值。

3.2.2 选取公共区域

因为左侧图像和右侧图像分别采用了不同变换矩阵,所以变换后的两幅图像的有效范围不再重合。为了便于人眼观察,对变换后的图像进行裁剪,只保留矩形公共部分,如图 3 中所示,其中实线框代表左图像有效区域,虚线框代表右图像有效区域,矩形公共区域用灰色部分表示。计算公共区域的具体方法如下:设公共矩形区域的左上角顶点坐标为 (c, r) ,宽度为 w ,高度为 h 。将左右图像有效区域共 8 个顶点横坐标升序排列,设 $x_n (n=1, 2, \dots, 8)$ 为排序后的横坐标序列,则 $r = x_4, w = x_5 - x_4$;将左右图像有效区域的顶点纵坐标升序排列,设 $y_n (n=1, 2, \dots, 8)$ 为排序后的纵坐标序列,则 $c = y_4, h = y_5 - y_4$ 。

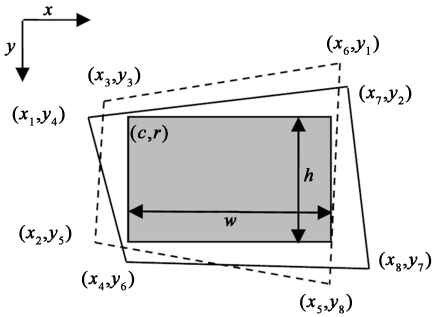


图 3 选取公共区域示意图
Fig. 3 Schematic diagram of selecting common area

3.2.3 尺寸调整

为了与立体显示的分辨率匹配,还需要对裁剪后的图像做尺寸变换。尺寸变换仍然使用后向映射双线性插值的方法计算变换后的每个像素点的灰度值。为了保持原始图像的长宽比,使用零填充的方法,达到与立体显示分辨率匹配的目的,即在图像上下或左右两侧填充一定数量的像素值为 0 的行或列。

3.2.4 立体图像合成与送显

待左右图像校正完毕后,即可将左右图像合成为立体图像送出显示。根据立体显示方式的不同,合成立体图像的方法有所不同。色差式立体显示利用了红、绿、蓝为互斥颜色的特性,如果使用红蓝眼镜(左红右蓝)作为观察设备,将左图像

放置在立体图像的 R 通道,将右图像放置在立体图像的 G 和 B 通道,透过红蓝眼镜,利用滤光原理,左眼只会看到左图像,右眼只有看到右图像,经过大脑合成后产生立体效果。另外,色差式也可以用于观察彩色的立体图像,方法是将左图像的 R 通道置于立体图像的 R 通道,将右图像的 G、B 通道分别置于立体图像的 G 通道和 B 通道,再透过红蓝眼镜可以观察到彩色的立体图像,但颜色失真会比较严重。对于快门式的立体图像一般为左右格式,即将原始立体图像一分为二,左侧二分之一放置左图像,右侧二分之一放置右图像。快门式立体显示器再将接收到原始立体图像分为左右幅图像,并以 120 Hz 的频率将左右图像隔帧交替显示。快门式立体眼镜与显示器通过红外线同步,在显示左侧图像时,关闭右侧镜片快门,在显示右侧图像时,关闭左侧镜片快门,利用人眼视觉残留效应原理,经过大脑合成产生立体效果。

4 实验结果与质量评价

4.1 实验结果与分析

应用上述方法,对本文设计的立体显示系统进行了实验验证。首先在当前场景中寻找左右图像匹配点,图 4(a)是使用 SURF 算法自动匹配特征点的结果。从图中可以看出左右图像的对应点不在一条水平线上,说明此时左右图像没有行对准。在图像中通过极点的直线称为极线,图 4(b)所示的是对图像校正前的极线位置情况,左右极线都有相交于一点的趋势,说明此时极点并不在无穷远处。图 4(c)是利用 Hartley 算法对图像进行立体校正的结果,图中的直线是校正后的极线位置情况,可以看出经过立体校正后,左右图像的对应点在一条水平线上,说明已经行已经对准,极线相互平行,说明左右极点已经在无穷远处。



(a)特征点匹配

(a)Correspondence match



(b)立体校正前的极线

(b)Epipolar lines before rectification



(c)立体校正后的极线

(c)Epipolar lines after rectification

图 4 立体校正实验结果

Fig. 4 Results of stereo rectification

将校正后图像合成为立体图像,图 5(a)为合成后色差式立体图像,佩戴红蓝立体眼镜可以观察到立体效果。图 5(b)为合成后的左右格式立体图像,需要将图像送显至立体显示器,佩戴快门式立体眼镜可以观察到立体效果,图 5(c)为快门式立体图像在 3D 电视上裸眼观察的效果。



(a) 色差式
(a)Anaglyphic mode



(b) 左右格式
(b)Left and right format



(c) 快门式
(c)Active shutter mode

图 5 立体图像合成效果
Fig. 5 Results of stereo rectification

在没有使用多线程对左右图像并行处理的情况下,测得系统输出立体图像的帧频为 21 Hz;而在使用了对左右图像并行处理的情况下,系统输出立体图像的帧频提高到了 30 Hz。

4.2 立体显示质量评价

由于立体显示主要是用于人眼的观看,因此

本文采用了具有代表性的 MOS(Mean Opinion Score)方法对立体显示系统的质量进行了主观评价。挑选了 20 位观察者对立体显示效果进行评分,满分为 10 分,分数越高代表显示效果越好。表 1 是 20 位观察者分别色差式和快门式立体显示评分情况,其中色差式立体显示平均得分为 7.67 分,快门式立体显示平均得分 9.08 分。可以看出,由于色差式立体显示技术本身的缺点,所以立体显示效果的得到分数较低。快门式立体显示的效果比较令人满意。

表 1 立体显示评分结果
Tab. 1 Score of stereoscopic display system

序号	分数(分)		序号	分数(分)	
	色差式	快门式		色差式	快门式
1	7	9	11	7	9
2	8	9	12	8	9
3	7	8	13	7	8
4	7	10	14	8	10
5	9	9	15	8	9
6	8	9	16	6	8
7	8	9	17	8	9
8	6	9	18	9	9
9	8	10	19	8	9
10	7	8	20	7	10

5 结 论

本文基于非标定立体校正 Hartley 算法,设计了一套实时立体显示系统,提供了色差式与快门式两种立体显示方式。实验结果表明,经过校正后的左右图像行已对准,经过测试立体图像输出的频率可以达到 30 Hz,与摄像头采集图像的频率相符合,说明系统能够实时处理并合成立体图像。通过 MOS 方法对立体显示效果的进行了评价,其中色差式立体显示得到 7.67 分,快门式立体显示得到 9.08 分,可以说明本系统得到了较好的立体显示效果。如果经过系统移植,本系统还可以应用于机器人远程操控立体显示系统中,为操作员提供更真实、更准确的场景信息。

参 考 文 献:

- [1] 游素亚,徐光祐. 立体视觉研究的现状与发展 [J]. 中国图形图像学报,1997,2(1): 17-24.
You S Y, Xu G Y. State of the art and Future the development of stereo vision [J]. *Journal of Image and Graphics*,1997, 2(1): 17-24. (in Chinese)
- [2] 张广军. 机器视觉 [M]. 北京:科学出版社, 2005.
Zhang G J. *Machine Vision* [M]. Beijing: Science Press, 2005. (in Chinese)
- [3] 宋晓伟. 双目视觉位姿测量算法研究 [D]. 北京:中国地质大学, 2011.
Song X W. Binocular vision position and attitude measurement algorithm [D]. Beijing: China University of Geosciences, 2005. (in Chinese)
- [4] 张兆杨,安平,张之江,等. 二维和三维视频处理及立体显示技术 [M]. 北京:科学出版社,2010.
Zhang Z Y, An P, Zhang Z J, *et al.* *2D and 3D Video Processing and Stereoscopic Display Technologies* [M]. Beijing: Science Press, 2010. (in Chinese)
- [5] 张兴,郑成武,李宁,等. 液晶材料与 3D 显示[J]. 液晶与显示,2012,27(4): 448-454.
Zhang X, Zheng C W, Li N, *et al.* Liquid crystal materials and 3D display [J]. *Chinese Journal of Liquid Crystals and Displays*,2012, 27(4): 448-454. (in Chinese)
- [6] 郭畅. 一种新型偏光快门式 3D 显示系统的设计与实现 [D]. 青岛:中国海洋大学, 2013.
Guo C. The design and implementation of a new patterned retarder/shutter glasses 3D LCD [D]. Qingdao: Ocean University of China, 2013. (in Chinese)
- [7] Dodgson N A. Autostereoscopic 3D Displays [J]. *Computer*,2005, 38(8): 31-36.
- [8] 帖志成,梁发云,黄伟莉,等. 嵌入式设备立体图像显示技术研究[J]. 液晶与显示,2013,28(1): 71-75.
Tie Z C, Liang F Y, Huang W L, *et al.* Stereoscopic display based on embedded system [J]. *Chinese Journal of Liquid Crystals and Displays*,2013, 28(1): 71-75. (in Chinese)
- [9] 丁剑飞,刘永进. 三维立体显示技术综述 [J]. 系统仿真学报,2008,20(suppl): 132-135.
Ding J F, Liu Y J. A Survey on three-dimension display technologies [J]. *Journal of System Simulation Journal*, 2008, 20(suppl):132-135. (in Chinese)
- [10] Bradski G, Kaehler A. *Learning OpenCV* [M]. O'Reilly Media, 2008.
- [11] Hartley R. Theory and practice of projective rectification [J]. *International Journal of Computer Vision*,1999, 35 (2): 115-127.

作者简介:姜瑞凯(1982—),男,吉林长春人,硕士,助理研究员,主要从事嵌入式软件、图像处理等方面的研究。E-mail: jrkwood@gmail.com

哈清华(1983—),男,吉林长春人,博士研究生,助理研究员,主要从事软件测试、信息融合等方面的研究。E-mail: haqinghuaha@163.com