

一种新的用于立体图像编码的视差估计算法

韩军功 卢朝阳

(西安电子科技大学通信工程学院 ISN 国家重点实验室 西安 710071)

摘 要 视差估计是立体视觉和立体图像编码领域的研究重点,其估计的准确程度直接影响编码的精度和效率.该文在研究传统估计算法的基础上,提出了一种新的带极线约束的 DT 基立体图像视差估计算法.该算法首先把立体图像中的左图像作为参考图像,用 DT 网格进行描述.然后,利用极线约束条件在右图像中估计出各三角形顶点的对应点,整个估计过程要求 DT 结构保持不变.通过估计出来的对应点计算出这些顶点的视差矢量,再用六参数仿射变换估计出三角形内部各点的视差矢量.编码时,只需编码各顶点的视差矢量.从实验结果看,这种新算法同传统的“块”匹配视差估计算法相比,视差矢量的估计更加准确,视差图中主要物体的轮廓更加清晰,有利于进一步的物体分割.

关键词 立体图像;极线约束;视差估计;Delaunay 三角形

中图法分类号 TP391

A Novel Disparity Estimation Algorithm for Stereo Image Encoding

HAN Jun-Gong LU Zhao-Yang

(ISN National Key Laboratory, School of Telecommunication Engineering, Xidian University, Xi'an 710071)

Abstract Disparity estimation is important in the stereo vision and stereo image coding whose results influence the precision and efficiency of the coding system. This paper introduces a novel Epipolar line-Constraint DT-based Stereo Image Disparity Estimation Algorithm after studying the traditional disparity estimation algorithms. This algorithm uses the DT mesh and epipolar line constraint between two corresponding points. The first step of this algorithm is to describe the left image by DT mesh, which is regarded as the reference image. Secondly, the corresponding vertexes of triangles are estimated by using the theory of epipolar line-constraint within the right image. In the course of estimation, the DT structure must keep changeless. Finally, the disparity vectors of vertexes are calculated and the disparity vectors inner the triangles are estimated by six-parameters affine transformation. While encoding, only the disparity vectors of vertexes are coded, which carry compression's point. Some real stereo images are used in the experiment, and the results show that the estimation veracity of this algorithm is higher than the traditional algorithms, which are based on “block” matching, and the object's contour in the disparity map is clearer than before, which means that the disparity map produced by this algorithm is more beneficial for object's segmentation. As a result, this algorithm is adaptive to coding system.

Keywords stereo image; epipolar line-constraint; disparity estimation; Delaunay triangulation

1 引言

立体视觉是人脑的一种能力,它利用左、右眼所看到的图像之间微弱的差异来感知外部的 3D 世界. 立体图像正是利用人类这种能力,使用左、右两组图像来描述一幅场景,人眼通过处理同一时刻的两副图像之间的相对位置来感知场景的 3D 深度信息. 立体图像和传统的二维平面图像相比,具有更强的逼真度. 目前,它已经被广泛应用到立体电视、地形地理、航空导航、网上参观、远程医疗诊断^[1~3]等领域. 但在实际的应用中,由于立体图像描述一组场景需要左右两组序列,同传统的平面图像相比,数据量更大,所以传输时必须使用立体图像编码技术.

立体图像编码技术利用立体图像对之间的冗余信息,达到信息压缩的目的. 视差矢量 (disparity vector) 是去除立体图像中冗余信息的主要手段,它是同一个点在两副图像中投影点之间的位置偏差. 传统的用于编码的视差矢量估计方法是基于“块”匹配的估计算法^[4~6],它首先把右图像分块,然后对每一块在左图像中用匹配算法做匹配,计算对应块之间的位置差异,形成块的视差矢量. 对于立体静止图像编码,利用视差矢量和左图像就可以恢复出右图像;对于立体视频图像编码,视差矢量经过平滑处理可以粗略地分割物体轮廓^[7],计算出物体的 3D 深度数据,进行物体的 3D 建模,然后对这些主要物体编码. 因此,视差估计是立体图像编码中的重要环节.

基于块的视差估计算法没有充分利用立体图像对的几何特性,对于点的视差估计不是十分准确,解码重建图像存在明显的块效应,并在低比特率时主观质量明显下降. 本文提出的带极线约束的 DT 基立体图像视差估计算法,利用了空间中同一个点在左、右两副图上的投影点之间的极线约束关系,还利用了 DT 网格良好的几何特性和网格的跟踪保持特

性,对立体对进行视差估计. 从实验结果看,这种新的算法较传统算法估计的准确率更高,这样更有利于准确地计算物体的 3D 深度信息;所形成的视差图像也更利于物体分割.

2 立体视觉中的极线约束

图 1 是一个立体图像的基本结构图,为了计算简便,我们假设两个摄像机的光轴平行且具有相同的内外参数. 图中, P 表示空间中一点,它的空间三维坐标为 $P(x, y, z)$, F_l, F_r 分别表示左、右两个摄像机的位置,而点 $P_l(x_l, y_l)$, $P_r(x_r, y_r)$ 则分别表示 P 点在左、右两个图像平面中的投影坐标. 利用几何计算可以得到, P 点的深度信息 ($f-z$) 和视差 d 成反比.

$$f - z = \frac{f \times s}{d} \quad (1)$$

式(1)中, f 表示摄像机的焦距, s 表示两个摄像机中心之间的距离,即基线 (baseline) 长度, d 表示视差,它等于 $x_l - x_r$. 从这个立体视觉的基本模型可以看出,立体图像对隐含着物体的三维深度信息,这个三维深度信息和视差 d 有着密切的关系,通过 d 可以计算出物体的三维信息.

在立体视觉系统中,由两个摄像机获得图像 I_l 和 I_r . 如果 P_l, P_r 是空间中同一点 P 在两个图像上的投影点,我们称 P_l, P_r 互为对称点. 视差估计就是已知某图像中的一点,在另一图像中估计它的对称点. 由计算机视觉的知识^[8], P_l 的对应点一定位于 I_r 上由 P_l 与两个摄像机的相对几何位置决定的某一条直线上,该直线称为 P_l 点在图像 I_r 上的极线 (如图 1 中的虚线). 有了这种极线约束关系就可以找到 P_l 的对应点,所以,本文视差估计的一个重要环节就是正确计算极线方程. 文献^[8]中详细推导了立体图像的几何原理以及在已知两摄像机投影矩阵 (由摄像机定时求出) 条件下求解极线方程的方法.

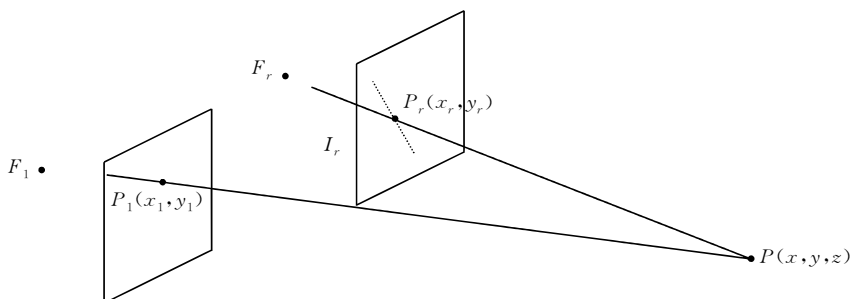


图 1 平行光轴立体图像基本几何结构

3 传统的基于“块”匹配的视差估计算法

文献[4,5]中介绍了传统的块匹配的视差估计算法,文献[6]中的视差估计部分也使用了该算法.块匹配视差估计算法主要分为以下几个步骤:

1. 把左图像作为参考图像.
2. 右图像作为待预测图像,分成 8×8 或者 16×16 的子块.
3. 对于右图像中的每一个子块,在左图像中寻找其最优的匹配块.具体方法是:记录每一个块的位置,在左图像中的相同位置以半径为 R 建立一个搜寻窗口,遍历窗口中的每一个位置,分别计算它们和待匹配块的互相关系数,相关系数最大的为最优匹配.

4. 记录最优匹配的 X 方向、 Y 方向坐标.
5. 用匹配块的 X, Y 坐标减去待匹配块的 X, Y 坐标,分别作为这个块在 X, Y 方向上的视差矢量.

传统的块匹配的视差估计算法类似于基于块的运动估计算法,它的优点是简单、实用、易于硬件实现,是 3D 立体电视拟采用的主要方法之一.但这种算法的视差估计准确性不高.

4 本文算法

本文提出了一种带极线约束的 DT 基立体图像视差估计算法,这种算法首先对左图像进行 DT 网格描述,DT 描述产生的三角形顶点大多是图像的灰度不连续点或一些物体的边缘点.计算这些“特殊点”在右图像中的极线方程,在极线上寻找其最优的对应点.计算各顶点的视差矢量,三角形内部点的视差矢量用六参数仿射变换求得.具体算法流程如下:

1. 左图像作为参考图像,进行 DT 描述^[9].
2. 选定某一个三角形.
3. 取出它的一个顶点,用文献[8]中的方法求出它的极线方程.
4. 对极线上的点进一步筛选,因为极线上和 P_1 相关值最大的点并非一定是 P_1 的对应点^[7].具体筛选做法是,记录左图像中一个顶点的位置,以右图像中同一位置点为中心,做一个矩形框(对本文立体对图像,经过多次实验,以 30×30 为优),记录经过这个矩形框内的极线上的点.

5. 对记录的每个点,假设它就是该顶点的对应点,结合三角形另外两个顶点的对应点判断是否会引起三角形的过度形变(保证变化后的三角形不重复覆盖图像区域).如果没有,记录下来.

6. 以顶点为中心建立一个 8×8 的匹配块 I_1 ,以第 5 步完成后记录的点为中心也建立一个 8×8 的块 $Q_i (i=1, 2, \dots)$,遍历每一个点,用式(2)计算 Q_i 与 I_1 的互相关系数 r_{xy} ,记录相关系数最大的点的 X, Y 坐标.

$$r_{xy} = \frac{R_{xy}(\tau) - m_x m_y}{\sigma_x \sigma_y} \quad (2)$$

其中, $R_{xy}(\tau)$ 表示两块的相关函数, m_x, m_y 分别表示两块的均值, σ_x, σ_y 分别表示两块的方差.

7. 原始坐标减去估计位置坐标作为该顶点的视差矢量.转向步 3,计算三角形另外两个顶点的视差矢量.对于该三角形内部的点用下面的六参数仿射变换求得其视差矢量.

$$\begin{cases} \Delta d_x = a_{11}x + a_{12}y + a_{13} \\ \Delta d_y = a_{21}x + a_{22}y + a_{23} \end{cases} \quad (3)$$

式中 x, y 表示点的坐标, Δd_x 和 Δd_y 分别表示 X, Y 方向上的视差矢量.其中, a_{11}, a_{12} 等 6 个参数可以利用已获得的三角形 3 个顶点的坐标和视差矢量联立方程组解得.

8. 转向步 2,直到所有三角形被处理完为止.
9. 形成立体图像的视差矢量图.

5 实验结果及分析

本文对两种算法进行了仿真,比较了两种估计算法的性能.图 2 是实验使用的一幅立体图像对,其中(a)是作为参考图像的左图像,(b)是右图像,均为 512×512 的灰度 PGM 格式立体图像.

首先,在立体图像的左图像中选取 30 个点用两种算法分别估计,表 1 是使用包括图 2 在内的 3 幅立体图像分别估计的结果.对于估计的准确率,我们采用主观观察并结合统计计算的方法来衡量.其中,“估计准确”是指主观匹配基本正确;“估计较准确”是指主观匹配有一定偏差,且统计偏差在 10 个像素以内;“估计不准确”是指主观匹配有较大偏差,且统计偏差在 10 个像素以上.

表 1 估算结果
(a) 本文算法

准确性	第 1 幅	第 2 幅	第 3 幅
准确	21 对	20 对	22 对
较准确	7 对	7 对	7 对
不准确	2 对	3 对	1 对

(b) 传统算法

准确性	第 1 幅	第 2 幅	第 3 幅
准确	16 对	15 对	16 对
较准确	11 对	12 对	12 对
不准确	3 对	3 对	2 对



(a) 本文算法处理的结果



(b) 用传统的块匹配方法处理的结果

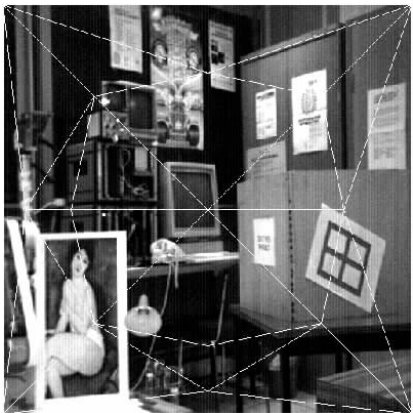
图 2 两种方法处理结果的比较

从结果看,本文算法估计准确的点较传统算法更多,这是因为它利用了极线约束条件,且 DT 剖分后的顶点大多是图像的灰度不连续点和物体的边缘点的原因.两种算法中估计不准确的点数相差不多,它们多位于物体有较大变化的区域.两种算法对其中的第 2 幅图像的处理都稍微差一些,因为第 2 幅图像两个摄像机的位置相差较远,左右图像的场景变化较大.

图 3 是本文算法对三角形顶点进行估计的实验结果.可以看出,除了右下角物体变化比较大的区

域有个别点估计不太准确外,其它各点的估计都非常准确.

图 4 是用两种算法估计后形成的视差图,如前所述,不管对立体静止图像还是立体视频图像,这种视差矢量都非常重要.从图中的结果看,传统算法形成的视差图的方块效应比较明显,且形成的物体轮廓的边缘比较粗糙.比较而言,本文算法形成的视差矢量图中主要物体(如画像、电视机等)的轮廓形状优于传统算法.

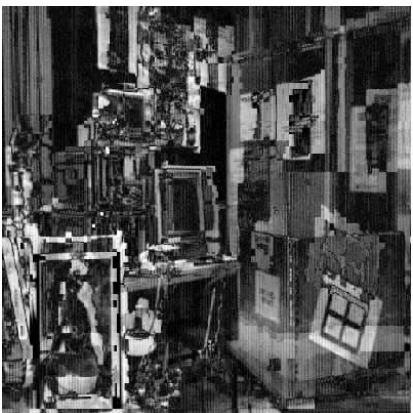


(a) 使用 DT 剖分方法对左图像进行描述的结果(15 个点的描述结果)



(b) 对三角形顶点进行估计的结果

图 3



(a) 用本文算法形成的视差图



(b) 用传统算法形成的视差图

图 4

6 结束语

本文针对立体图像编码中的视差估计问题,提出了一种新的带极线约束的 DT 基立体图像视差估计算法. 这种算法利用了立体图像对应点之间的极线约束关系,还利用了 DT 网格良好的几何特性和网格的跟踪保持特性. 本文算法首先精确估计三角形顶点这些图像的“特殊点”的视差矢量,然后再用仿射变换估计其它点的视差矢量. 这种方法不同于传统以块为单位进行估计的方法. 从实验结果看,本文算法视差估计的准确率优于传统算法,说明这种基于内容的图像剖分方法在图像匹配和图像编码方面有一定优势.

致 谢 本文图像数据由中国科学院马颂德教授和华军老师提供,在此表示感谢.

参 考 文 献

- 1 Waldowski M. A new segmentation algorithm for videophone applications based on stereo image pairs. *IEEE Transactions on Communications*, 1991, 39: 1856~1868
- 2 Kauffman D, Wood S. Digital elevation model extraction from stereo satellite images. In: *Proceedings of International Geosci-*

- ence and Remote Sensing Symposium, 1987, 1: 349~352
- 3 Antonisse H. Active stereo vision routines using PRISM3. In: *Proceedings of SPIE—The International Society for Optical Engineering*, 1993, 1825: 745~756
- 4 Gunatilake P D, Siegel M W, Jordan A G. Compression of stereo video streams. In: *Proceedings of International Workshop on HDTV'93*, Ottawa, Ontario, Canada, 1993
- 5 Tzovaras D, Grammalidis N, Srintzis M G. Depth map coding for stereo and multiview image sequence transmission. In: *Proceedings of Internat Workshop on Stereoscopic and 3D Imaging (IWS3DI95)*, Satorini, Greece, 1995. 75~80
- 6 Aydmoglu H, Hayes M H. Stereo image coding: A projection approach. *IEEE Transactions on Image Processing*, 1998, 7 (4): 506~516
- 7 Ma Song-De, Zhang Zheng-You. *Computer Vision—Computation Theory and Algorithm Base*. Beijing: Science Press, 1998 (in Chinese)
(马颂德,张正友. 计算机视觉——计算理论与算法基础. 北京: 科学出版社, 1998)
- 8 Lu Zhao-Yang, Zhou Xing-Ni. Image coding algorithm based on DT modal. *Chinese Journal of Computers*, 1997, 20(9): 769~774(in Chinese)
(卢朝阳,周幸妮. DT 模型基图像编码方法. 计算机学报, 1997, 20(9): 769~774)
- 9 Malassiotis S, Srintzis M G. Object-based coding of stereo image sequences using three-dimensional models. *IEEE Transactions on Circuits and Systems for Video Technology*, 1997, 7 (6): 892~905



HAN Jun-Gong, born in 1977, Ph. D. candidate. His main research interests include image processing, video coding and stereo image coding.

LU Zhao-Yang, born in 1963, Ph. D., professor and Ph. D. supervisor. His main research interests include image transmission and processing, image coding, image recognition and so on.