

双目立体相机中实时匹配算法的 FPGA 实现

单洁, 唐焱, 姜晖

(西安邮电大学通信与信息工程学院, 陕西 西安 710121)

摘要:在双目立体相机中,利用图像处理计算场景深度信息是一项关键技术。通过研究立体视觉图像匹配原理,提出一种基于 FPGA 的立体图像实时匹配算法的实现方法。该算法以 Census 变换为基础借助于像素在邻域中灰度相对值的排序进行相似度比较,来实现区域立体匹配;在左右一致性约束下采用多窗口相关匹配方法改善深度不连续图像的匹配质量,提高匹配准确度。利用 FPGA 流水线和并行处理技术实现了双目立体相机的实时图像匹配。结果表明,该图像匹配结构具有较高的吞吐率和处理速度,可以工作在 97.3MHz 频率下实现 1024 × 1024 灰度图像 30f/s 的实时处理。

关键词:双目立体相机;实时图像匹配;多窗口相关匹配;FPGA

中图分类号:TP391.41; TN911.73 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2013.12.20

Real-time matching algorithm in binocular stereo camera and its implementation with FPGA

SHAN Jie, TANG Yao, JIANG Hui

(School of Communication and Information Engineering, Xi'an University of Post and Telecommunications, Xi'an 710121, China)

Abstract:To compute the depth information using image processing is one of key techniques in the binocular stereo camera. Stereo vision image matching principle is analyzed and a new method of realizing real-time stereo image matching algorithm based on FPGA is proposed to decrease the matching time. In order to achieve regional stereo matching, based on the Census transform, the similarity are compared through the gray level of the pixel in the neighborhood; Multiple window correlation matching method are used to improve the quality of image when depth does not match on the left-right consistency constraints, thus, the matching accuracy is improved. Real-time image matching is implemented by using FPGA pipelining and parallel processing technology. The experimental results demonstrate that the method has higher throughput rate and processing speed, the real-time processing of 1024 × 1024 gray image with 30fps are realized at the frequency of 97.3MHz.

Key words:binocular stereo camera; real-time image matching; multiple window correlation matching; FPGA

1 引言

随着机器视觉的发展,立体视觉系统得到越来越广泛的应用^[1-2]。双目立体相机采用两台相机,模拟人眼成像原理,先通过寻找两个相机中图像对应点的光学,得到各个点处的视差,再根据视差信息和相机的投影模型恢复出原始景物的深度信息。所

以,在通过立体相机计算场景信息的深度过程中,图

基金项目:“973”、“863”项目、国家重大科技专项(No. 61271292);陕西省科技厅工业攻关计划(No. 2012K06-27);西安邮电大学中青年基金(No. 109-0425;109-04367)资助。

作者简介:单洁(1978-)女,讲师,硕士研究生,主要研究方向图像及信号处理。E-mail: shanjie@xupt.edu.cn

收稿日期:2013-04-26

像获取、相机标定、图像匹配和三维重建成为了关键技术。

根据匹配基元的不同,图像匹配可分为区域匹配^[1-2]、特征匹配^[1]和相位匹配^[2-4]。其中,区域匹配以区域相似度为匹配基元,基本原理是在基准图上选取一个模板窗口,在配准图上选取同样大小的一组窗口,再将这些窗口与基准图中的模板窗口进行相似度比较,找出与模板窗口最相似的窗口,那么该窗口的中心点就是模板窗口中心点的匹配点,两点的横坐标差即为视差^[1]。

区域匹配直接利用图像的灰度信息,实质是利用局部窗口之间灰度信息的相关程度,并能取得致密的视差场。区域匹配适用于具有明显纹理特征的图像,可以直接获得稠密的深度图,但区域匹配选取一定大小窗口内的灰度分布特性作为匹配基元,决定了该算法具有以下局限性^[5]:(1)计算量大;(2)对左右两幅图像中光照不均匀较敏感,在这种情况下容易造成匹配困难;(3)匹配窗大小难选择。窗口选择过大,在高度间断处会出现误匹配,选择过小,区域内的灰度分布特性未得以充分展现;(4)对于视差不连续处的情况,如物体边缘附近的区域,适应性较差。文献[6]、[7]提出了在 FPGA 上基于 SAD 匹配准则的算法实现。基于 FPGA 实现图像匹配算法,能够充分发挥 FPGA 的特点,利用流水线 and 并行计算技术,提高运算效率和系统的吞吐率,解决了计算量大的问题,达到实时处理的要求。针对区域匹配对因亮度变化、噪声等因素引起的幅度失真非常敏感的问题,先后有文献提出了对幅度失真不敏感的匹配代价,其中主要有归一化互相关(NCC)、互信息和 Rank、Census 非参数变换^[8]等。Hirschmulle 通过大量实验发现:基于 Census 变换的匹配代价对图像中的噪声和亮度不一致性有很高的鲁棒性,能获得更优的视差结果^[9]。文献[7]采用基于 Census 变换的匹配算法实现了图像匹配的实时处理。此外,传统的固定窗口存在变换窗口大小的选择问题,若窗口太小,则匹配代价区分度过低,在低纹理区域容易出现误匹配;若窗口过大,又会在深度不连续区域出现误匹配。文献[10]提出了基于 Census 变换与自适应窗口的局部匹配算法,对幅度失真具有良好的鲁棒性。但是,经过研究表明,Census 变换由于丢弃了图像色彩信息而仅保留了

像素间大小关系,容易造成误匹配^[11],文献[12]提出了多相关窗口的匹配方法。

本课题主要应用双目立体相机进行障碍物检测,不仅要求对数据的正确性,还要求算法需要较低的复杂度,同时还要具有较好的鲁棒性,处理好噪声、目标区域低纹理、遮挡和图像边界不连续等问题,实现最佳的立体匹配效果。为了解决区域匹配的局限性,对立体匹配算法的进展进行了研究,利用 Census 度量算法的特点,结合文中具体立体相机的硬件结构,提出了一种基于 FPGA 的并行实时 Census 变换立体匹配算法。基于 FPGA 的算法实现结合了多窗口匹配的方法,通过左右一致性约束,实现了基于 FPGA 的立体图像匹配硬件结构,不仅保证了匹配精度,而且对幅度失真具有较好的鲁棒性,为下一步的三维重建及深度测量打下了良好的基础。

2 算法描述

2.1 Census 变换

非参数变换^[8],如 Rank 或 Census,是首先对图像进行非参数化变换,提取包含局部纹理信息的描述符。Census 变换通过计算点在窗口中的灰度顺序。变换以像素为中心选择一个窗口,将窗口中心像素以外的像素变换为一个比特串。如果窗口中一个像素的灰度值比中心像素大,则相应位置为 1,反之,即为 0。Census 变换公式如式(1)(2)所示:

$$\xi[I(m,n),I(x,y)] = \begin{cases} 0, & I(m,n) < I(x,y) \\ 1, & I(m,n) \geq I(x,y) \end{cases} \quad (1)$$

$$\text{code}(m,n) = \bigotimes_{[x,y] \in N} \xi[I(m,n),I(x,y)] \quad (2)$$

式中, $I(m,n)$ 为变换窗口 N 对应的中心像素灰度值; $I(x,y)$ 为窗口内其余像素的灰度值,函数通过对这些像素值的比较,将结果串连(\otimes)成比一个特向量 $\text{code}(m,n)$,它表征了中心像素(m,n)与周围像素的灰度值的比较关系。然后再利用 Hamming 距计算 Census 变换后比特串之间的差异程度,Hamming 距越小,说明两点的匹配度越高。

经过 Census 变换后,中心像素的灰度值被替换为一个 0 和 1 的比特串,该比特串仅与变换窗口中心像素和其他各个像素灰度大小有关,因此,Census 变换对于幅度失真有着很好的抑制能力。

2.2 多相关窗口匹配

由于匹配窗口位于深度不连续的区域时会引入匹配误差,本文中相关窗口的选取采用了文献[10]提出的多相关窗口的匹配方法。当匹配窗口位于深度不连续的区域时,可以通过只采用深度连续区域的相关窗口来减小误差。如图 2 中(b)所示。

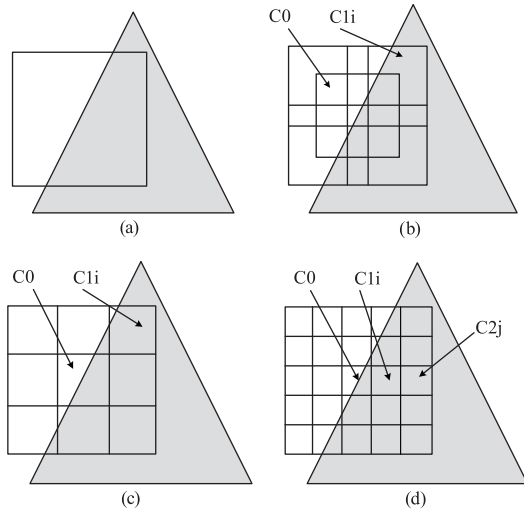


图 1 多窗口匹配示意图

图 2 中(c)所示为 9 窗口匹配方式,其中当中心窗口位于深度不连续的区域时,左右两幅图像的相关过程会存在歧义性。在这种情况下,假设中心窗口的代价值为 C_0 ,在周围 8 个窗口中选取的代价值最小的 4 个窗口,设其代价值分别为 C_{1i_1} 、 C_{1i_2} 、 C_{1i_3} 和 C_{1i_4} ,那么最终的代价值 C 的计算入式(3)所示。

$$C = C_0 + C_{1i_1} + C_{1i_2} + C_{1i_3} + C_{1i_4} \quad (3)$$

这种方法还可以扩展为 25 窗口处理的模式,如图 2(d)所示。可以看出,多匹配窗口可以通过通过小窗口与大窗口结合的方法以适应局部环境。采用周围小窗口可以避免在图像深度不连续的区域引入匹配误差;另一方面,选用较大尺寸的相关窗口可以减少在弱纹理区域的错误匹配。

3 硬件实现

3.1 硬件整体结构

根据任务需求,双目立体相机中采用了一片 FPGA 进行左右两路视频数据的同步采集,预处理,在片上经过立体图像匹配算法后,通过电路中预留的 USB2.0 接口,将得到的视差图传输给 PC 机。整个相机的硬件系统包括视频接口模块,USB2.0 数据传输模块和视频处理模块。整个硬件结构原理图如图 2 所示。

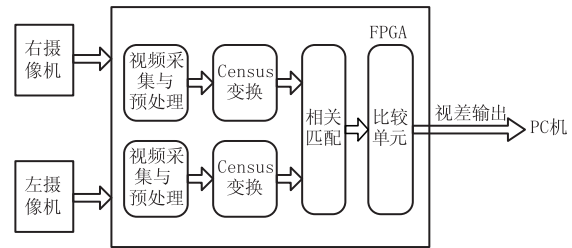


图 2 基于 FPGA 的硬件结构原理图

在立体图像匹配时,要对两组图像做 Census 变换和相似度比较,在这几个步骤中,采用流水线计算方式并对左右两路图像采用并行技术同时计算,大大提高了系统的运算速度。

3.2 多窗口相关匹配法在 FPGA 中的实现

对于硬件实现来说,Census 变换和多窗口代价值计算是基于窗口的操作,在计算之前先要缓存与窗口高度相同的图像行数。本文实现结构中采用了 3×3 子窗口和 9 个子窗口进行代价值计算的方式。首先,需要读入并缓存 3 行数据,然后对中间的 1 行数据进行 Census 变换,变换完毕后将 Census 值缓存到另一个具有 3 行数据大小的缓存器中,为下一步的相关匹配窗口代价值计算做准备。

进行数据缓存时,假设计算的是第 2 行数据,在计算完当前 y 列的数据进行第 $(y+1)$ 列数据的计算时,就可以对第 1 行第 y 列数据进行更新。同样,当计算第 3 行数据时,就可以对第 2 行数据进行同步更新。这样,Census 变换的计算与图像像素采集形成了流水线结构,将以与图像输出同样的速度进行运算,达到实时性。并且在这种情况下,对于宽度为 1024 的 8bit 灰度图像来说,需要耗费的片上存储资源为 $1K \times 3 = 3K\text{byte}$ 。

在下一个相关匹配的计算单元中,因为采用 3×3 的子窗口,也要首先缓存 3 行 Census 值,再进行窗口代价值的计算。在这一步中,需要计算不同视差情况下,匹配子窗口中中心窗口的代价值和周围 8 个子窗口的代价值,并且找出其中代价值最小的 4 个,与中心窗口代价值进行累加。

通过以上的判决和累加后,找到代价值最优的情况,就完成了多窗口的相关匹配,也就是得到了视差值。在本文中,立体视觉试验系统将视差输出给上位机,由上位机根据此视差值和相机的标定结果,计算出实际场景中物体的景深。图 3 所示为 FPGA 中匹配算法的结构流程图。

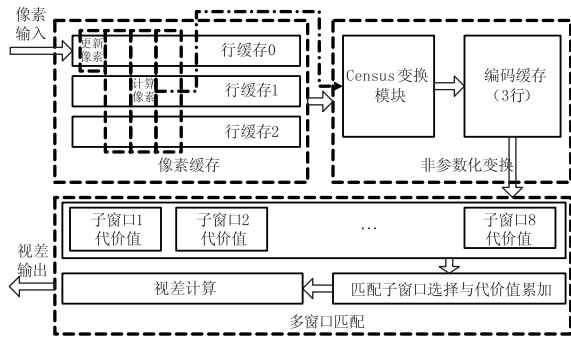


图3 匹配算法结构流程图

4 结果分析

通过流水线技术,采用循环数据更新结构,不仅使得计算可以实时进行,还可以充分利用缓存空间,大大降低了缓存的使用量。对于一幅尺寸为 $N \times N$ 个像素的图像,假设相关窗口尺寸 $n \times n$,搜索视差的范围是 d ,直接进行块匹配计算时的复杂度为 $O(N^2n^2d)$,采用本文中的方法完成一幅灰度图像的匹配所需的计算复杂度为 $O(N^2d)$,与窗口尺寸无关。

在 Stratix II FPGA EP2S30 实验板上实现了的文中提出的硬件结构,如图4所示。



图4 立体图像处理实验板

左右两路输入图像通过视频采集接口传输到实验板,经过预处理后送入 FPGA 进行立体视觉算法的处理。匹配的视差图像结果再通过 USB2.0 送给上位机,与 Matlab 变换的结果进行比较,实验验证了设计的正确性。图5中,图5(c)、图5(d)所示为通过硬件仿真,分别将图5(a)、图5(b)两帧图像送入实验板,经过 FPGA 变换后的结果。

通过与 Matlab 运算结果相比较,变换后的结果完全一致,验证了硬件运算的正确性。FPGA 硬件结构的描述全部采用 VHDL 语言,可以较方便的在不同FPGA 平台上进行移植。综合和分析采用了 QuartusII 软件,采用 Modelsim 进行功能仿真和时序仿真。实现 $1024 \times 1024 @ 30\text{f/s}$ 灰度视频图像实时匹配功能,经过综合,在 Stratix II FPGA EP2S30FPGA 平台中布局布线后的结果如表1所示。

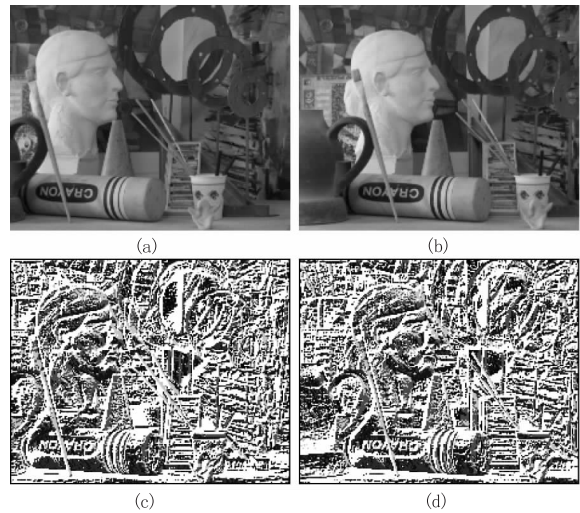


图5 FPGA 实现 Census 变换结果

表1 FPGA 综合结果

器件	Stratix II EP2S30F67214
Logic Element	9486 (28%)
Memory	907520bit (66%)
Maximum Frequency	97.3MHz

5 结论

本文提出了在双目立体相机中基于 FPGA 实现立体图像匹配算法的一种设计。采用了改进的基于 Census 变换的方法,结合多窗口相关匹配和左右一致性原理,实现立体图像匹配的实时运算。设计中结合 FPGA 结构的特点,充分利用流水线结构和并行执行的思想,减少运算量和存储器的使用,降低了资源的使用量,提高运算速度。实验结果证明,该匹配算法结构具有较高的吞吐率和变换速度,可以工作在 97.3MHz 的频率下,能够实现 1024×1024 灰度图像 30f/s 的图像实时匹配。

参考文献:

[1] Zhou Xing, Gao Zhi-jun. Application and Future Development of Stereo Vision Technology[J]. Journal of Graphics, 2010, 04: 50 - 55.
 周星,高志军. 立体视觉技术的应用与发展[J]. 工程图学学报, 2010, 04: 50 - 55.

[2] You Suya. State of the art and Future the development of stereo vision [J]. Journal of Image and Graphics, 1997, 01: 17 - 22.
 游素亚. 立体视觉研究的现状与进展[J]. 中国图象图形学报, 1997, 01: 17 - 22.

[3] Xu Yanjun, Du Limin, Hou Ziqiang. A Scale-Adaptive

- Phase-Based Stereo Matching Method[J]. *Acta Electronica Sinica*, 1997, 07: 38 - 41.
- 徐彦君, 杜利民, 侯自强. 基于相位的尺度自适应立体匹配方法[J], *电子学报*, 1997, 07: 38 - 41.
- [4] Zitova B, Flusser J. Image registration methods: a survey [J]. *Image and Vision Computing*, 2003, 21 (11): 977 - 1000.
- [5] Bai Ming, Zhuang Yan, Wang Wei. Progress in binocular stereo matching algorithms [J]. *Control and Decision*, 2008, 23(7) : 721 - 729. (in Chinese)
- 白明, 庄严, 王伟. 双目立体匹配算法的研究与进展 [J]. *控制与决策*, 2008, 23(7) : 721 - 729.
- [6] Ambrosch K, Humenberger M, Kubinger W, et al. Hardware implementation of an SAD based stereo vision algorithm [C]. *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshops*, Minneapolis, Jun. 17, 2007: 1 - 6.
- [7] Chen L, Jia Yunde. A parallel reconfigurable architecture for real-time stereo vision [C]. *Proceedings of the International Conference on Embedded Software and Systems*, Hangzhou, May 25, 2009: 32 - 39.
- [8] Zabih R, Woodfill J. Non-parametric local transforms for computing visual correspondence [C]//*European Conference on Computer Vision*, Stockholm, 1994: 151 - 158.
- [9] Jin S, Cho J, Phan X, et al. FPGA design and implementation of a Real-Time stereo vision system [J]. *IEEE Transaction on Circuits and Systems for Video Technology*, 2010, 20(1) : 15 - 26.
- [10] Zhou L, Xu G L, Li K Y. Stereo matching algorithm based on Census transform and modified adaptive windows [J]. *ACTA*, 2012, 33: 886 - 892. (in Chinese)
- 周龙, 徐贵力, 李开宇, 等. 基于 Census 变换和改进自适应窗口的立体匹配算法 [J]. *航空学报*, 2012, 33: 886 - 892.
- [11] Xiong G M, Li X. Color rank and census transforms using perceptual color contrast [C]//*the 11th International Conference on Control, Automation, Robotics and Vision*, Singapore, 2010: 1225 - 1230.
- [12] Hirschmüller H, Innocent P R, Garibaldi J. Real-time correlation-based stereo vision with reduced border errors [J]. *International Journal of Computer Vision*, 2004, 47 (1 - 3): 229 - 246.