

基于 FPGA 的双目实时测距算法设计

陈明明^{1,2,3}, 祝永新¹, 田 犁¹, 封松林¹, 汪 辉¹

(1 中国科学院 上海高等研究院, 上海 201210; 2 中国科学院大学, 北京 100049;

3 上海科技大学 信息科学与技术学院, 上海 201210)

摘 要: 为了克服软件层面实现双目测距算法实时性差的缺点, 提出了一种基于 FPGA 的高速的双目测距算法的硬件架构, 该架构使用并行流水线的结构来实现兴趣点的检测与描述, 通过乒乓操作来实现描述向量的缓存与处理, 同时通过 WTA(Winner Takes All)电路和三角测量法来实现视差的搜索和距离的测量. 实验结果表明: 在 25 MHz 的时钟频率下, 处理 640×480 的图像对只需 12.5 ms, 处理速度约为软件上的 68 倍, 满足了高速场景下实时性的需求. 同时, 处理速度和资源消耗也优于一些相关硬件实现的文献.

关键词: 兴趣点检测与描述; 乒乓操作; 视差搜索; 三角测量; FPGA

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2018)10-0067-05

Design of Binocular Real-Time Distance Measurement Algorithm Based on FPGA

CHEN Ming-ming^{1,2,3}, ZHU Yong-xin¹, TIAN Li¹, FENG Song-lin¹, WANG Hui¹

(1 Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China;

2 University of Chinese Academy of Sciences, Beijing 100049, China;

3 School of Information Science & Technology, Shanghai Tech University, Shanghai 201210, China)

Abstract: In order to overcome the shortcomings of poor real-time performance of the binocular distance measurement algorithm on the software level, we present a high-speed hardware architecture of binocular distance measurement algorithm based on FPGA, the architecture uses a parallel pipeline structure to detect and describe interest points, and description vectors caching and processing is achieved by the ping-pong operation, disparity search and distance measurement are realized by WTA (Winner takes All) circuit and triangulation method simultaneously. The experimental results show that it only takes 12.5ms to deal with a pair of 640×480 images at the clock frequency of 25 MHz, and the processing speed is about 68 times of the software, it meets the real-time needs of some high-speed scenes. Simultaneously, the processing speed and resource consumption are also better than some related hardware implementations.

Key words: detection and description of interest points; ping-pong operation; disparity search; triangulation method; FPGA

1 引言

近年来, 双目视觉算法已经应用于汽车碰撞预警, 无人机和移动机器人避障等多个领域. 由于双目

视觉算法具有计算复杂的特点, 软件上实现难以满足实时性, 采用 GPU 加速来提高实时性, 不适用于低功耗、独立的双目视觉系统实现. FPGA 作为一种具有可重构、资源丰富、实现灵活的可编程系统, 是

收稿日期: 2018-01-13; 修回日期: 2018-01-30

基金项目: 国家重点研发计划(2017YFA0206104); 上海市科学技术委员会科研计划项目(16511108701); 张江管委会公共服务平台项目(2016-14)

视觉类算法实现高速实时性的有效平台。

立体匹配是双目视觉算法的核心,立体匹配主要分为基于灰度的匹配和基于特征的匹配两类。基于灰度的匹配充分利用图像的原始信息,但是对光照,噪声等非常敏感,不够稳定,而基于特征的匹配可有效克服这些缺点^[1-3]。其中 SURF^[2] 算法较其他算法准确率相当,但计算复杂度低,而且许多模块,如盒子滤波器,哈儿小波等非常适合在硬件上进行加速处理。在硬件上实现 SURF 算法的文献中,丁帅帅^[4]提出的方法可以实现 50 fps 的帧率,但是他们没有实现兴趣点的描述和匹配;X. Fan^[5] 等在 FPGA 上完整的实现了 SURF 算法,在时钟频率为 156 MHz 时,最高帧率可达到 356 fps,但是没有进一步地进行测距操作。本文提出了一种基于 FPGA 的双目测距算法的高速硬件架构。

2 算法描述

2.1 适用于硬件的兴趣点检测模块—Hessian 矩阵检测子

Hessian 矩阵检测子用于检测图像中类似斑块的结构,在稳定性和可重复性上都优于最常用的 Harris 角点检测。对于图像中的一个点 (x,y) 处的灰度值用 $X(x,y)$ 表示,尺度为 σ 的 Hessian 矩阵 $H(X,\sigma)$ 定义为:

$$H(X,\sigma) = \begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \tag{1}$$

式中, $L_{xx}(X,\sigma)$ 为点 (x,y) 处 $X(x,y)$ 与高斯二阶导数的卷积, $L_{xy}(X,\sigma)$ 和 $L_{yy}(X,\sigma)$ 与之类似。

通过 Hessian 矩阵的行列式大小来判断图像中的点 (x,y) 是否为兴趣点。在实际运算中,为了加快运算速度,可通过盒子滤波器来近似高斯二阶导数^[2],盒子滤波器与原始图像卷积来实现 Hessian 矩阵中的各元素值。 $\sigma=1.2$ 的盒子滤波器,如图 1 所示,对于 X 和 Y 方向,黑色区域的权重为 -2,白色区域权重为 1,灰色区域的权重为 0;对于 XY 方向,黑色区域的权重为 -1,白色区域的权重为 1,灰色区域权重为 0。分别定义以下盒子滤波器为 D_{xx} , D_{yy} , D_{xy} 。

则 Hessian 矩阵的行列式可近似为^[2]:

$$\text{Det}(H) = D_{xx}D_{yy} - (0.9D_{xy})^2 \tag{2}$$

由于盒子滤波器中的权重值只有 0, 1, -1, -2,在硬件上仅需要加、减、移位操作便可完成卷积运算,用移位操作代替乘法,减少了逻辑资源的使用。

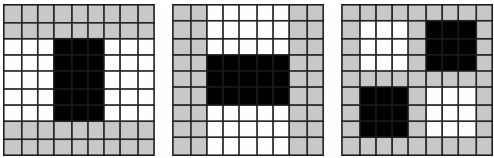


图 1 X,Y,XY 方向盒子滤波器 D_{xx} , D_{yy} , D_{xy}

2.2 适用于硬件的特征描述模块—SURF 描述子

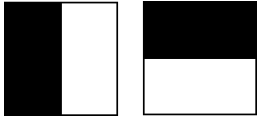
在 SURF 算法中,通过哈尔小波,如图 2(a)所示的计算来实现 SURF 描述子。SURF 描述子不仅比其他描述子速度更快,而且重复性更好,可以在多兴趣点匹配中减小复杂度和提高计算性能。

对于描述子的计算,以兴趣点为中心构建方形描述区域(region, 8×8 像素),同时将此描述区域划分为 2×2 的亚区域(sub-region, 4×4 像素),再讲亚区域划分为 (2×2) 的块区域(block, 2×2 像素)。在块区域内, X 方向的哈尔小波响应为 D_x , Y 方向的为 D_y , 则:

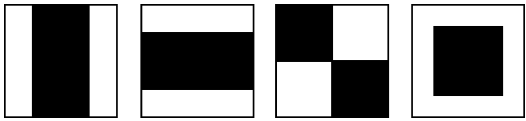
$$D_x = \sum_{x \in \text{left_block}} p(x) - \sum_{x \in \text{right_block}} p(x) \tag{3}$$

$$D_y = \sum_{y \in \text{up_block}} p(x) - \sum_{y \in \text{down_block}} p(x) \tag{4}$$

式中, $p(x)$, $p(y)$ 代表像素的灰度值。在每个亚区域的小波响应和相加,构成了描述向量中的元素。为了引入强度变化的极性信息,提取出绝对值 $|D_x|$, $|D_y|$ 的和。因此,每个亚区域有了 4 维的强度结构描述向量 $v = (\sum D_x, \sum D_y, \sum |D_x|, \sum |D_y|)$, 如此继续计算所有 2×2 亚区域,整个描述区域形成 16 维的描述向量。为了增加描述子的鉴别力,在原有 SURF 描述子的基础上,新加入几种小波来增加描述向量的维度,进而提高匹配准确率。图 2(b) 为所增加的几种类似哈尔小波的小波,称之为类哈尔小波,与哈尔小波共同组成 32 维的描述子。



(a) X 方向和 Y 方向的哈尔小波



(b) 几种类哈尔小波

图 2 哈尔小波与类哈尔小波

(黑色区域代表的权重为 1, 白色区域代表的权重为 -1)

硬件上实现 SURF 描述子时,仅需对描述区域内的像素进行加、减运算便可完成描述子的构建。

2.3 基于 WTA(Winner Takes All)的匹配与三角测量模块

匹配或视差搜索用于找到左右图像上同源的感兴趣点,进而得到准确的视差。左右图像上兴趣点的匹配本质上是找出参考兴趣点与目标兴趣点之间的几何变换关系,本文采用 SURF 描述子作为兴趣点的特征进行匹配,以提高匹配的稳定性与准确率。用 $I_R(x, y)$ 代表左图上参考兴趣点 (x, y) 的描述子, $I_T(x+d, y)$ 表示右图上与左图上参考兴趣点视差为 d 的目标兴趣点的描述子,采用 SED(Squared Euclidean Distance, 欧式距离平方)进行相似度度量,如式(5)所示,使用 WAT 进行最近邻搜索,在右图像上找出与参考描述子最接近的目标描述子,此时对应的视差即为搜索到的视差,如式(6)所示,其中 e 代表匹配代价。

$$SED(x, y, d) = || I_R(x, y) - I_T(x+d, y) ||_2^2 \quad (5)$$

$$e(x, y, d) = \min\{SED(x, y, d)\} \quad (6)$$

为了排除左右图像上的非公共区域、遮挡等带来的误匹配, Lowe^[1] 提出了比较最近邻距离与次近邻距离的方法,最近邻与次近邻的距离比率小于一定的阈值,则认为是真实有效的匹配。本文在此基础上再对匹配代价 e 进行约束,匹配代价小于设定阈值方可认为是真实有效的匹配。

匹配模块得到视差后,通过三角测量法,完成视差到距离的转换,如图 3 所示。

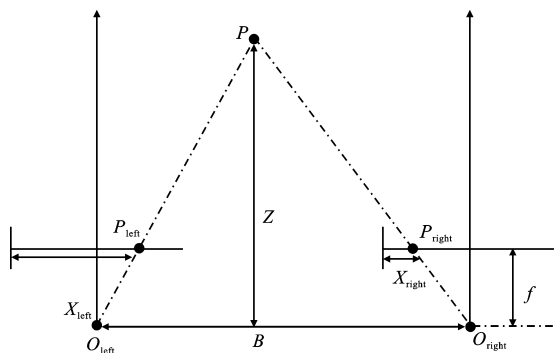


图3 三角测量法

其中 Z 为点 P 到摄像头的距离, B 为基线长度(即摄像头间距), f 为焦距, x_{left}, x_{right} 分别为点 P 在左右图像 X 方向上的坐标, 其中视差 $d = x_{left} - x_{right}$, 则:

$$Z = \frac{Bf}{x_{left} - x_{right}} = \frac{Bf}{d} \quad (7)$$

3 硬件设计

本文整体硬件架构如图 4 所示, 主要包括基于 Linebuffer 的兴趣点检测与描述子提取模块, 4 个 RAM 模块, SED 和 WTA(Winner Takes All) 计算模块。左右图像传感器传输的图像在 FPGA 开发板 SDRAM 上帧缓存, 每个时钟周期读出 1 个像素, 将读出的像素由 RGB 值转化为灰度值, 即为 Linebuffer 的输入 $Pixel_1, Pixel_2$ 。像素流满 Linebuffer 后来实现兴趣点的检测与描述子的提取; RAM_1 和 RAM_3 先分别写入左右图像上的描述向量, 写满后开始读出, 读出的同时换成 RAM_2 和 RAM_4 写入, RAM_2 和 RAM_4 写满后开始读出, 同时 RAM_1 和 RAM_3 又开始写入, 如此相互配合的切换, 进行乒乓操作。从 RAM 中读出的描述向量经过 SED, WAT 和三角测量模块, 便可输出图像上兴趣点的距离与坐标信息。

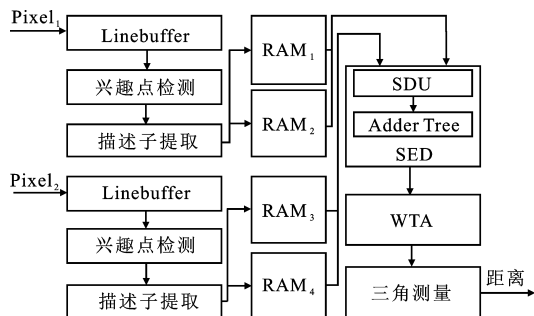


图4 本文算法整体硬件架

本文整体采用流水线式的架构, 同时各个模块中大量的运算采用并行结构, 大大提升了处理速度。

3.1 基于 Linebuffer 的兴趣点检测与描述模块

在硬件设计中, 如果像在软件中那样将整幅图像存储进而来进行运算, 在硬件中是无法接受的。为了解决这个问题, 本文采用流水线式设计, 使用基于 RAM 的移位寄存器(Shift Register)来实现 Linebuffer 的设计。对于 640×480 的灰度图像, 要和 9×9 的盒子滤波器进行卷积运算, 移位寄存器的输入位宽设为 8 bits, 抽头(taps)设置为 9, 抽头间距设为 640(图像宽度), 即只需缓存图像的 9 行。

如图 5 所示, 图中小方框代表存储像素灰度值的寄存器。虚线方框内 9×9 像素大小的区域(以寄存器 i_1 中的像素为中心)与盒子滤波器进行卷积运算, 通过 Hessian 矩阵的行列式与设定的阈值比较, 来确定 i_1 中的像素是否为兴趣点。由于从像素流入 i_1 , 到判断出像素是否为兴趣点需要一定时间(此处

为 5 个时钟周期), 若 i_1 中的像素为兴趣点, 则经过 5 个时钟周期, i_1 处的像素流入到 i_2 中, 即此时寄存器 i_2 中的像素为兴趣点. 此时, 以 i_2 为中心构建 $8 \times$

8 像素大小的描述区域, 如图中实线框所示, 在此描述区域内进行哈儿小波响应与类哈儿小波响应的累加, 来构建描述向量.

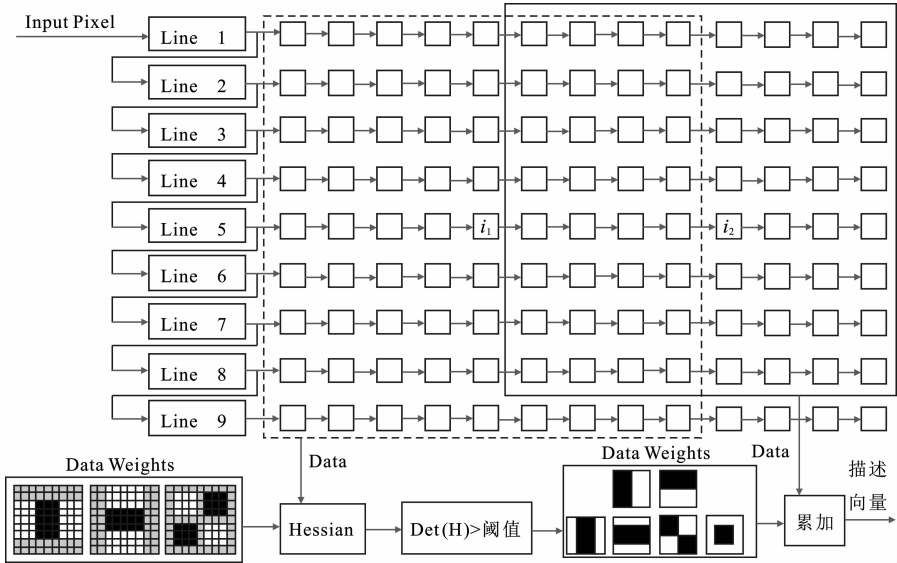


图 5 Linebuffer 实现兴趣点检测与描述模块

3.2 匹配(视差搜索)模块

匹配模块主要包含两个子模块, SED(Squared Euclidean Distance) 计算模块和 WTA (Winner Takes All) 电路模块.

其中 SED 计算模块(如图 6 所示, 其中 l_i, r_i 分别为左右描述向量中的元素)用来计算 RAM 读出的左右图像上兴趣点描述向量之间的 SED. 主要包含两个子模块, SDU(Square Difference Unit)和 加法树(Adder Tree). 由于 SED 计算存在大量内在的并行度, 我们将描述向量中的各个元素同时进行差的平方运算, 再进行求和, 得到两个向量间的 SED, 与软件上的循环计算相比, 很大程度的节约了时间.

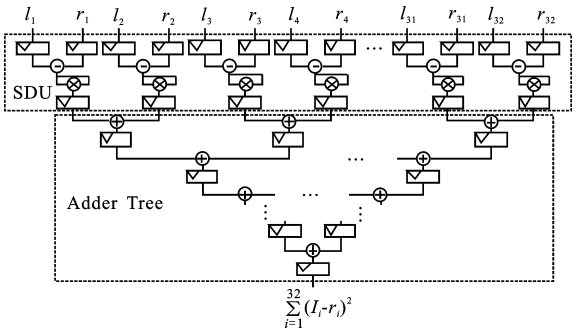


图 6 SED(Squared Euclidean Distance)计算模块

传统的 WTA 电路用来进行最近邻搜索, 本文在此基础上加入次近邻的搜索的功能. 修改后的 WTA 电路模块用来将左图像上的基准兴趣点的描述向量与右图像上多个目标兴趣点的描述向量进行

最近邻与次近邻的搜索, 主要包含 2 个比较器, 一个选择器和一些寄存器, 来实现流水线式最小以及第二小距离的搜索, 如图 7 所示.

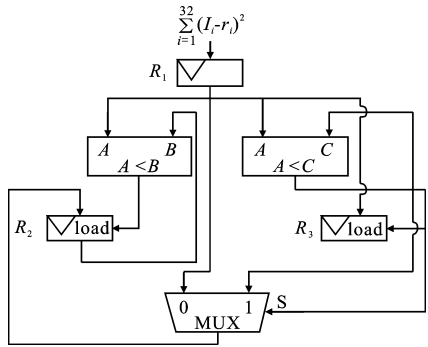


图 7 WTA 电路模块

若用 $I_R(x, y)$ 代表左图上参考兴趣点 (x, y) 的描述向量, $I_T(x+d, y)$ 表示右图上与左图上参考兴趣点视差为 d 的目标兴趣点的描述向量. 在一定时期内, 维持 $I_R(x, y)$ 不变, 不断更新 $I_T(x+d, y)$, 这个时期结束时, R_3 中存储的即为与参考兴趣点 (x, y) 描述向量最相似的目标兴趣点描述向量间的 SED(与坐标信息), 称为 SED_1 , 同样 R_2 中存储的为与参考兴趣点 (x, y) 描述向量第二相似的目标兴趣点描述向量间的 SED(与坐标信息), 称为 SED_2 . 若 SED_1 、 SED_2 的比率小于一定阈值, 同时 SED_1 小于设定的阈值, 则认为 R_3 中存储的目标兴趣点与参考兴趣点 (x, y) 是真实有效的匹配, 通过公式(7)便

可得出左图上点 (x,y) 的距离信息。

如此再不断更新左图上参考兴趣点的描述向量(与坐标信息),便可得到左图上所有获得真实有效匹配的兴趣点的距离信息。

4 实验结果及分析

本文算法采用 Altera DE1-SoC FPGA 开发板 5CSEMA5F31C6 器件仿真,系统时钟频率为 25 MHz 时,对于大小为 640×480 的双目图像,处理一幅图像的时间可以达到 12.5 ms. 仿真产生的结果通过 OpenCV 显示,通过计算得出,本文算法测量出的距离与真实距离的平均误差为 3.67%。

为了对比,分别在 PC 和 FPGA 上实现本文算法,计算相同的图像,硬件平台处理速度约为软件平台的 68 倍(软件平台的处理速度为 847 ms). 表 1 为本文算法在 FPGA 上的资源使用情况. 表 2 为本文算法与一些已有相近算法的比较,对比结果显示,本文算法在测量误差上接近软件平台实现的结果^[6],但处理速度是它们的 16 到 40 倍;同时处理速度和 BRAM 消耗也优于相关硬件平台上实现的结果^[7]。

表 1 FPGA 资源使用情况

BRAM	LUT	Register	DSP
28.6 kB	15 096	8 862	144

表 2 本文算法与相关文献比较

文献	图像大小	处理平台	时钟频率 /MHz	处理速度 /ms	BRAM /kB	测量误差/%
[6]	1 008 * 1 018	软件	/	500	/	<4
[7]	未知	软件	/	200	/	1~6
[8]	1 024 * 1 024	硬件	97.3	33.3	113.4	/
[9]	320 * 240	硬件	50	15	90.2	/
本文	640 * 480	硬件	25	12.5	28.6	0.8~8.5

5 结束语

本文设计并实现了一种基于 FPGA 的双目实时测距算法的硬件架构,引入了类哈儿小波来加强描述子鉴别力,同时加入了剔除机制以提高视差搜索的准确率,流水线的结构来实现兴趣点的检测与描述,使用乒乓操作实现描述向量无缝地缓存与处理,节省了硬件资源,加快了处理速度. 实验结果表明,本文的算法在测量误差上接近软件平台实现结果的同时,处理速度大幅提升,满足了一些高速场景(如汽车碰撞预警,无人机与移动机器人避障等)下实时测距的需求. 下一步将在原有硬件架构的基础

上,采用一些拟合的方法来实现图像亚像素级别的匹配,进一步提高测量精度。

参考文献:

[1] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.

[2] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF) [J]. Computer Vision and Image Understanding, 2008, 110(3): 346-59.

[3] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]//proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005. San Diego, CA, 2005.

[4] 丁帅帅,柴志雷. 基于 HLS 的 SURF 特征提取硬件加速单元设计与实现 [J]. 微电子学与计算机, 2015, 33(9): 133-140.

[5] Fan X, Wu C, Cao W, et al. Implementation of high performance hardware architecture of OpenSURF algorithm on FPGA[C]//proceedings of the 2013 International Conference on Field-Programmable Technology (FPT). Kyoto, Japan, 2013.

[6] 赵勋杰,李成金. 双目立体实时测距系统的关键技术研究 [J]. 激光与红外, 2006, 35(9): 874-881.

[7] Wang C, Sun Y. A new method of depth measurement with binocular vision based on SURF[C]// proceedings of the International Workshop on Computer Science and Engineering, 2010.

[8] 单洁,唐垚,姜晖. 双目立体相机中实时匹配算法的 FPGA 实现 [J]. 激光与红外, 2013, 43(12): 1406-1416.

[9] 徐彦泓,朱云芳,杜歆,等. 非参数化立体匹配算法的 FPGA 实现[J]. 浙江大学学报(工学版), 2009, 53(2): 250-254.

作者简介:

陈明明 男,(1991-),硕士研究生. 研究方向为图像处理算法与硬件设计. E-mail: 205528753@qq.com.

祝永新 男,(1969-),博士,研究员. 研究方向为计算机体系结构、嵌入式系统和大数据处理.

田犁 男,(1985-),博士,副研究员. 研究方向为 CMOS 图像传感器芯片、智能车辆视觉系统以及数字图像信号处理电路和算法.

封松林 男,(1964-),博士,研究员. 研究方向为无线传感器网络和微系统技术.

汪辉 男,(1975-),博士,研究员. 研究方向为高性能的成像和显示面板驱动.