

Harris 角点结合金字塔光流法的目标跟踪算法设计研究

徐里萍¹, 耿斌¹, 李小龙¹, 赵丽²

(1. 川北幼儿师范高等专科学校, 四川 广元 628017; 2. 山西大学 软件学院, 太原 030013)

摘要: 针对现存很多跟踪算法在速度和准确度方面很难满足嵌入式跟踪开发的需要, 提出一种基于 Harris 角点和金字塔光流法的目标跟踪算法, 并详细给出了 DSP-FPGA 的硬件设计; 首先, 使用 Harris 角点提取目标角点特征; 然后, 使用金字塔光流法为后续视频帧匹配角点; 最后, 基于角点的质心跟踪算法用于匹配目标的重心, 确定目标的位置, 重心跟踪算法可以较好地抵消由于旋转或扭曲带来的形变问题; 在硬件实现过程中, FPGA 方便电路设计, 使用硬件描述程序语言实现硬件算法、逻辑控制和外部接口, DSP 则运行目标跟踪算法; 实验结果验证了本文硬件实现算法的有效性, 相比于 AVT21 开发板的质心跟踪算法、相位相关跟踪算法和金字塔相关性跟踪算法相比, 文章算法在平均重叠和平均中心误差方面具有一定优势, 在 720p 的视频流上可以满足 25fps。

关键词: 目标跟踪; Harris 角点; 金字塔光流; AVT21; 数字信号处理器 (DSP); 现场可编程门阵列 (FPGA)

Design and Research of Objective Tracking Algorithm Using Harris Corner and Pyramid Optical Flow Method

Xu Liping¹, Geng Bin¹, Li Xiaolong¹, Zhao Li²

(1. North Sichuan College of Preschool Teacher Education, Guangyuan 628017, China;

2. School of Software, Shanxi University, Taiyuan 030013, China)

Abstract: For many existing tracking algorithm is difficult to meet the development needs of embedded tracking in terms of speed and accuracy, fast tracking algorithm is proposed based on Harris corner and Pyramid optical flow, with details on DSP-FPGA hardware design given. Firstly, Harris corner is used to extract feature points of targets. Then, Pyramid optical flow is used to match corner for subsequent video frame. Finally, centroid tracking algorithm based on corner is used to match the center of gravity, determining the location of the target, better counteracting distortions caused by the rotation or deformation. In the hardware implementation process, FPGA convenient circuit design using a hardware description language program hardware algorithms, logic control and external interface, and DSP runs target tracking algorithm. The effectiveness of the proposed algorithm implemented in hardware is verified by experimental results. Compared to trackers in AVT21 development board, centroid algorithm, phase correlation tracking algorithm and pyramid correlation tracking algorithm, the proposed algorithm has more advantages in means of overlapping and means of center error. In 720p video stream, the video can satisfy 25fps.

Keywords: target tracking; Harris corner; pyramid optical flow; AVT21; digital signal processor (DSP); field programmable gate array (FPGA)

0 引言

目标跟踪技术^[1]在机器视觉、模式识别等诸多方面均有重要的应用, 它以图像处理技术为主, 将光学、计算机控制和几何运算等技术融为一体, 可通过图像处理程序的串行实现目标捕获, 识别和跟踪^[2]。目前的跟踪技术已广泛应用于多个领域: 如气象云图分析^[3]、智能交通视频监控系统^[4]等。在这些系统中, 如何获得鲁棒快速的跟踪, 是该领域的研究关键。

一般跟踪算法中最重要的是目标表示形式, 如: 颜色直方

图特征^[5]、直方图方向梯度^[6]、Harr 特征描述符^[7]和角点特征^[8]等, 已经有很多研究者对其进行了大量研究, 下面介绍几个具有代表性的目标跟踪算法。

文献 [9] 提出一种二次约束的目标跟踪算法, 在目标定位的搜索机制中, 使用二次等式约束的 Kalman 滤波器提高跟踪精度, 同时匹配和更新目标外观模型。但这种方法容易陷入局部最优, Kalman 滤波的假设前提条件也比较多。

文献 [10] 提出了很有名的 TLD 算法, 即跟踪-学习-检测, 将学习方法融入到目标跟踪系统中, 利用 N-P 专家的学习对跟踪和检测出现的正负样本进行分类, 同时, 跟踪模板和检测模板也相应进行更新。虽然 TLD 跟踪单个目标准确且稳定, 但 NP 学习比较耗时, 而且这种算法仅限于单目标跟踪, 因为 NP 学习无法扩展到多目标的学习中。

文献 [11] 提出了粒子滤波的随机搜索算法, 该算法可以消除局部最小化问题, 一般目标的外观模型随时间变化而变化, 例如光照变化, 温度湿度变化, 这时目标需要正则化更新。但是, 不恰当的更新操作会导致目标漂移。

收稿日期: 2017-09-18; 修回日期: 2017-10-27。

作者简介: 徐里萍 (1981-), 女, 四川广元人, 讲师, 硕士, 主要从事图形图像处理等方向的研究。

耿斌 (1969-), 男, 四川广元人, 副教授, 硕士, 主要从事图像处理等方向的研究。

赵丽 (1980-), 女, 山西长治人, 副教授, 硕士, 主要从事计算机应用方向的研究。

文献 [12] 利用在线更新外观模型, 充分挖掘目标周围局部的可视数据, 在背景粒子的协助下选择判别特征, 但该算法的更新过程计算量较大, 实时性比较弱。

以上方法均在实时性方面具有较大的缺点, 这可能有两方面原因, 一方面是算法本身比较复杂, 另一方面在实现过程中实现的方法造成大量的运算时间。本文是一种嵌入式的算法设计, 本文围绕目标的角点特征展开, 这是因为角点特征集中了大量的图像信息, 且很少受光照变化影响^[8]。且它具有旋转不变性, 仅含约 0.05% 的图像像素, 需要处理的数据非常小, 非常适合嵌入式开发的实时要求, 本文的主要贡献如下:

1) 提出一种围绕角点的多级跟踪算法, 这些算法没有复杂华丽的推导和理论, 但非常实用, 这种设计思想很大程度上启发于 AVT21 开发板 (一种比较通用的跟踪开发板, 可以在网上搜索相关介绍) 的设计思想。

2) 设计一种 DSP-FPGA 硬件实现框架, 这种硬件设计使算法、控制、接口等处理更加优化。

1 角点特征提取与金字塔 LK

1.1 角点特征提取

目标跟踪过程中, 选择一种好的特征提取算法非常重要, 图像中, 目标角点是重要的局部特征, 它集中了大量的重要图像信息, 具有旋转不变性和对光照的鲁棒性, 且需要处理的数据非常少。因此, 角点检测已应用于许多重要的实践中, 特别是, 实时目标跟踪领域。目前, 很多采用角点检测的目标跟踪算法主要基于 SUSAN 和 Harris 算法^[13]。提出的算法使用 Harris 算法提取目标角点特征。

1.2 光流跟踪器

光流 (Lucas Kanade, LK) 也是一种完善且广泛使用的算法, 由于该算法仅依赖局部信息, 而局部信息来源于提取出的角点周围区域, 因此可以用于处理稀疏纹理图像。然而, LK 光流算法的缺点是大幅度运动会将角点移出局部窗口, 从而使 LK 算法难以发挥作用。

金字塔 LK 算法可削弱这种缺陷, 即首先从图像金字塔的最高层开始处理, 逐步处理到金字塔的最低层, 图像金字塔的最高层拥有最少简化的细节信息, 具体过程如图 1 所示。

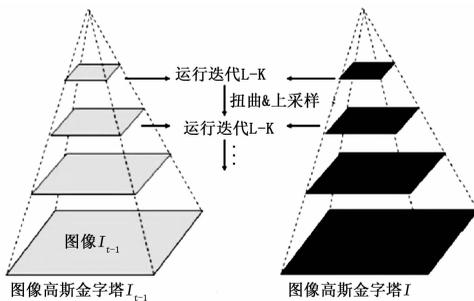


图 1 金字塔 LK 光流算法

目标跟踪误差是在参考图像块以及第二帧变换图像块之间的图像网格上测量, 而图像 I_1 和 I_2 之间的映射误差是相应像素之间的色差, 表示如下:

$$e_{of}(x_i, \hat{\theta}) = I_2[m(x_i, \hat{\theta})] - I_1(x_i) \quad (1)$$

式中, $\hat{\theta}$ 是一个八维向量, 即空间坐标, $m_j = m(x_i, \hat{\theta})$, x_i 和 m_j 分别是图像 I_1 和 I_2 的像素, 假设 $\hat{\theta} = \hat{\theta} + \delta\theta$, 考虑式 (1) 的一阶泰勒展开形式, 有:

$$e_{of}(x_i, \hat{\theta}) \approx I_2(m_i) + J_{I_2|_{\hat{\theta}}}^T \delta\theta - I_1(x_i) \quad (2)$$

式中, $J_{I_2|_{\hat{\theta}}}^T$ 是 $\hat{\theta}$ 中计算的 8×3 雅可比矩阵, 即相对于每一个变换参数均有颜色分量的导数。

为了估计 $\hat{\theta}$, 在 x_0 处 N_{x_0} 中心邻域内所有像素 $x_i, i = 1, \dots, n_x$ 上最小化式 (2) 的误差。为了评估计算过程的鲁棒性, 使用双权重 M 估计, 像素的贡献也根据其具邻域中心 x_0 的距离用空间核进行加权, 即为广义 M-估计算法, 从式 (1) 和式 (2) 得出优化准则为:

$$J_{of} = \sum_{i=1}^{n_x} K_e \left(\frac{x_i - x_0}{h} \right) \rho \left(\left\| \frac{J_{I_2|_{\hat{\theta}}}^T \delta\theta + e_{of}(x_i, \hat{\theta})}{\sigma} \right\| \right) \quad (3)$$

式中, h 表示空间带宽, σ 为颜色误差的比例, $\rho(u)$ 为双加权损失函数。对于多变化情况, 该损失函数通过论证准则来计算。邻域应包含足够多局部信息, 确保一阶近似中较大误差不会影响估计。 K_e 为完全对称的内核函数, 本文选取 Bhattacharyya 核函数。

统计学中, Bhattacharyya 距离是一种广泛应用的距离测量算法, 即对于给定高斯分布 g_i, g_j , 其计算如下:

$$BD(g_i, g_j) = \frac{1}{8} (\mu_i - \mu_j)^T \Sigma^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left[\frac{\det \Sigma}{\sqrt{\det \Sigma_i \det \Sigma_j}} \right] \quad (4)$$

式中, $\Sigma = \frac{\Sigma_i + \Sigma_j}{2}$, 定义高斯分布的 Bhattacharyya 核函数如下:

$$K_{BD}(g_i, g_j) = \exp \left(- \frac{BD(g_i, g_j)}{2t^2} \right) \quad (5)$$

2 目标跟踪算法及其硬件设计

2.1 算法步骤描述

提出的目标跟踪算法主要包含使用基于光流的特征点匹配算法, 通过目标角点检测和匹配, 描述如下。

步骤 1: 选择目标区域后预处理图像。为了消除噪声, 采用中值滤波算法。 3×3 模板的滤波器在具有很高的并行处理能力的现场可编程门阵列 (Field Programmable Gate Array, FPGA)^[14] 中编程时, 处理速度更快, 数值结果更好。

步骤 2: 使用 Harris 算法在目标区域中提取角点特征, 同时删除预选目标区域内背景的特征点。在该目标跟踪系统中, 首先手动选择目标图像区域, 由于目标区域比整个图像尺寸小, 此过程需要较少时间。由于两个相邻图像帧之间的目标位移非常小, 提出的算法仅在目标周围较小区域中计算目标位移。

步骤 3: 通过金字塔 LK 光流法为后续图像帧匹配图像角点, 该算法虽然精度高, 但它需要较长时间, 使得当目标和其图像都很大时不能满足实时性要求。为了克服该缺点, 提出的算法中仅使用合理角点特征, 以减少计算时间, 借鉴 Opencv 里源码处理角点的方式。

步骤 4: 计算匹配点, 并通过质心算法得到目标位置。当使用金字塔 LK 法在后续帧中计算特征角点的新位置时, 可能与之前步骤中计算的之前的位置有一些小的匹配点偏差, 由于旋转和被跟踪目标的失真。实验结果还表明, 大多数匹配点接近目标, 仅几个点由于目标旋转、扭曲和/或屏蔽偏离目标。

为了得到正确的结果,提出的算法必须考虑所有匹配点,采用质心法来计算所有匹配角点的质心。

关于质心算法,一般是通过黑热化或白热化操作,对原有图像帧进行二值化处理,这个处理直接关系到最终的质心运算的准确度^[15]。这里要着重说明下,本文直接获得角点位置,利用角点位置计算质心位置,不需要进行二值化处理。

2.2 实时目标跟踪系统的硬件设计

为了高速处理整个跟踪算法,设计了一种硬件系统,具有高速处理器数字信号处理器 (Digital Signal Processor, DSP) 和 FPGA。

提出的实时目标跟踪的硬件系统主要构成有:核心芯片 TMS320C6416T^[13]作为数字信号处理器、可编程逻辑芯片组 CPLD 和现场可编程阵列 FPGA,硬件设计如图 2 所示。

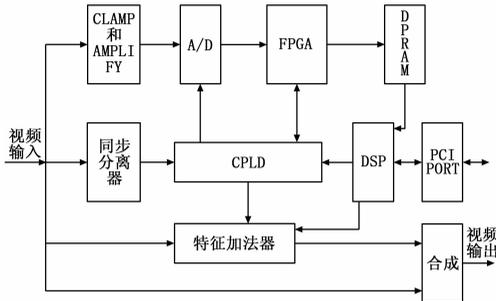


图 2 硬件结构

硬件的主要部分包括视频信号处理器、高速率模数 (A/D)、数字图像存储器、转换器数字图像处理器、数据分析器、计算机数据通信接口 RS232 以及同步和加法器显示。图 2 中的流程图可以大致描述如下:预处理图像传感器中的模拟视频信号经过固定、滤波和放大,可以满足 A/D 输入视频信号的需求。接着,此信号由具有 8 bit 的高速率 A/D 转换器采样,输出的信号直接发送到 FPGA,该系统在运行预处理算法的同时,还执行其它函数。DSP 芯片读取由 FPGA 处理过的数据,通过 FPGA 实现的双口 RAM 内存,并行运行跟踪算法。数据包括由 DSP 计算的目标位置,经过 PCI 端口发送到伺服系统。DSP 的外部同步信号作为控制信号,CPLD 的外部同步信号作为参考信号。特征发生芯片组 (UPD6464) 由 DSP 的多信道缓冲串行端口 (Multi-channel Buffer Serial Port, McBSP) 控制。然后添加 UPD6464 特征信号到该模拟信号,这样,所计算的数据信号最终显示在显示器上。

2.3 实时目标跟踪系统的硬件设计

作为核心处理器, TMS320C6416T 运行目标跟踪算法,并计算所述目标实时位置,处理器以 EDMA 模式,通过 DSP 的 EMIFB 端口,直接从双端口的 RAM (即 DPRAM) 上读取图像数据。

尽管使用德州仪器公司制造的高性能 DSP 处理器,该系统仍很难完成较大图像数据的实时处理任务。因此,可以看出,结合了 FPGA 和 DSP 芯片的快速目标跟踪系统可以实现软件硬化和硬件软化。其中,中值滤波、双阈值处理和边缘提取操作这些耗时处理由高度并行化的 FPGA 完成。同时, FPGA 方便了电路设计和算法修改,使用硬件描述程序语言 (Hardware description Language, HDL) 来实现硬件算法、逻辑控制和外部接口等,而 DSP 则运行复杂的目标跟踪算法。

图 3 是本文算法所用的开发板,基本上与 AVT21 的尺寸大小差不多。

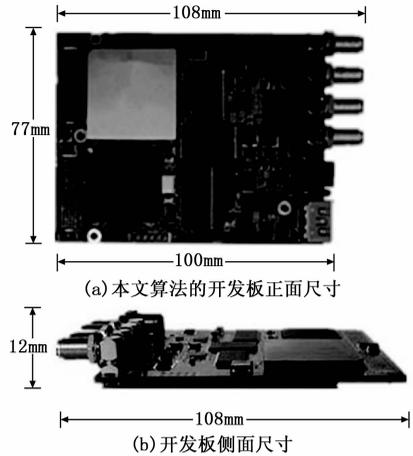


图 3 算法实现的开发板

3 实验结果与分析

实验比较了 4 个算法,其中算法 1 是 AVT21 开发板的质心跟踪算法,该算法具有较高速度,适用于当背景很简单,如黑色或白色明显的目标。算法 2 是基于金字塔相关性^[14]的跟踪算法,其中目标模板的大小是目标跟踪区的四分之一,算法 3 是 AVT21 开发板上的相位相关跟踪算法,算法 4 是本文提出的算法,算法 1 和 3 是基于 AVT21 开发板的,算法 2 和 4 是基于本文的高速目标跟踪系统。

3.1 评估指标

本文利用两个评估指标评价嵌入式跟踪器的性能:重叠和均方差 (Root Mean Square Estimation, RMSE),前者量化为:

$$A_k = \frac{TT_k \cap GT_k}{TT_k \cup GT_k} \quad (6)$$

式中, TT_k 是跟踪器的边界框, GT_k 是实际边界框, RMSE 量化所有帧上跟踪器目标估计中心 TT_k^c 和实际中心 GT_k^c 之间的总体边界框中心误差:

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N \|TT_k^c - GT_k^c\|^2} \quad (7)$$

式中, N 是总帧数。

3.2 实验结果

图 4 所示为使用提出算法的目标跟踪测试结果。图 4 (a) 是初始图像,手动选择目标区域,预处理后由 Harris 算法提取角特性。七个角特性由符号“×”表示,这也是光流算法匹配的对象。图 4 (b) ~ (e) 为单独匹配和跟踪测试中第 5 帧、第 158 帧、第 252 帧、第 288 帧的结果。图 3 中符号“+”表示光流算法的匹配点。虽然目标在图 4 (b) 几乎无变化、在图 4 (c) 中有小旋转、在图 4 (d) 比图 4 (a) 中存在模糊现象,提出的算法能够正确地得到匹配角特征,并成功地跟踪目标。如图 4 (e) 所示,由于目标图像很大且旋转,一些特征角点偏离初始位置。对于图 4 (e) 的情况,应使用质心算法来计算特征角点的质心,以确定目标位移和位置。

图 5 是使用遮蔽直升机的另一测试。其中,图 5 (a) 表示手动选择目标区域的初始图像,如图 5 (b) ~ (f) 分别呈

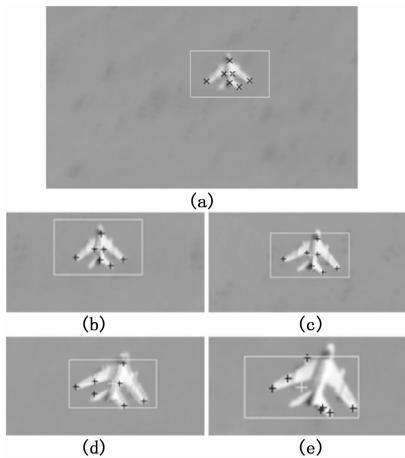


图 4 飞机的跟踪分析和结果

现了第 70 帧、第 75 帧、第 77 帧、第 85 帧、第 110 帧的跟踪结果, 从图 5 中可以看到, 当高塔轻微遮蔽直升机时, 提出的算法可以准确地跟踪目标。

图 6 所示为坦克目标的跟踪实验结果, 其中背景非常复杂。图 6 (a) ~ (d) 分别呈现第 10 帧、第 147 帧、第 192 帧、第 210 帧的跟踪结果。从图 5 (c) 中可以看到, 当坦克几乎由烟遮蔽时目标丢失, 其中无坦克角特征提取。

表 1 和表 2 是各算法的重叠和 RMSE 指标结果, 加粗数据表最优数据, 可以看出本文算法获得了最好的结果, 这表明了嵌入式开发中算法设计并不是越复杂越好, 嵌入式开发的核心原则是在实现的可能性和精度中寻求一个折中方案。

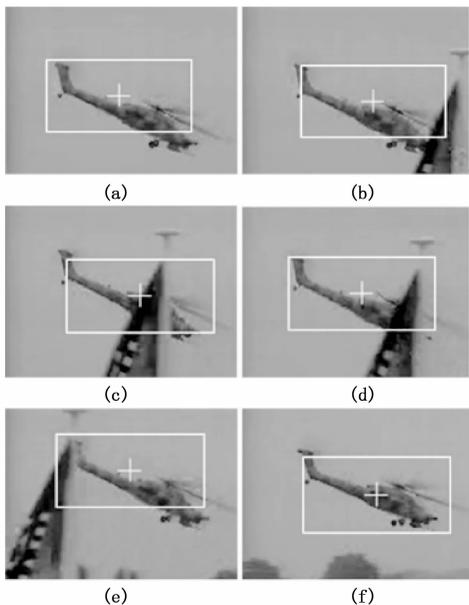


图 5 遮蔽直升机的测试结果

表 1 各方法对每个视频序列的平均重叠

视频	算法 1	算法 2	算法 3	算法 4
直升机	0.1555	0.2914	0.3320	0.4963
坦克	0.2097	0.3021	0.4848	0.4656
飞机	0.3673	0.4362	0.4594	0.6022

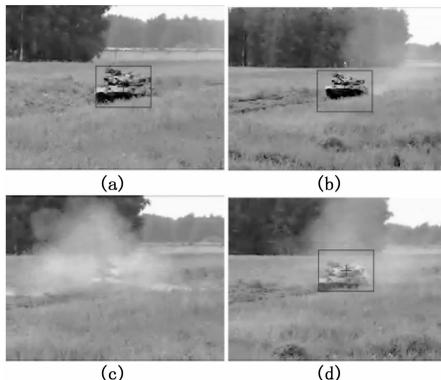


图 6 坦克视频的测试结果

表 2 各方法对每个视频序列的中心误差(RMSE)

视频	算法 1	算法 2	算法 3	算法 4
直升机	49	36	38	12
坦克	39	37	28	9
飞机	16	12	13	9

3.3 运行速度比较分析

表 3 所示为不同算法的处理速度比较, 所有这些实验在高速目标跟踪系统上完成。从表 3 可以看出本文算法只比单纯的质心跟踪算法稍微慢一点, 这充分体现了较好的硬件设计所带来的好处。

对于 720×576 和 1080×720 (720p) 两种视频流, 比较了各自的帧率, 从帧率捕捉器上获得的数据如表 4 所示。可以看出本文算法达到了非常高的实时要求, 与 AVT21 开发板上的跟踪算法基本没有太大区别。

表 3 不同方法的处理速度

跟踪区大小	算法 1	算法 2	算法 3	算法 4
80×60	5 ms	18 ms	10 ms	8 ms
100×100	7 ms	20 ms	15 ms	10 ms
160×120	9 ms	29 ms	24 ms	14 ms

表 4 不同帧率的比较

视频流	算法 1	算法 2	算法 3	算法 4
720×576	30fps	25fps	25fps	30fps
1080×720	28fps	20fps	25fps	25fps

4 结论与展望

本文提出了使用角点检测基于金字塔光流算法的目标跟踪。该算法首先使用 Harris 方法提取角点特征, 然后使用金字塔光流法为后续视频帧提取角点特征, 质心算法用来确定目标位置。根据数值研究结果的分析 and 讨论, 可以得出结论, 提出的方法具有较高的目标跟踪精确度,

总之, 对于嵌入式开发, 特别是对于军事目标跟踪具有比较好的现实意义, 而且算法已经在开发板上运行, 但从纯粹的算法角度看, 它仅适用于跟踪小目标, 未来将逐渐弥补这些设计缺陷。