

利用国产 FPGA 的高帧率目标跟踪方法

王向军^{1,2}, 朱 慧^{1,2}

(1. 天津大学精密测试技术及仪器国家重点实验室, 天津 300072;
2. 天津大学微光机电系统技术教育部重点实验室, 天津 300072)

摘要: 由于复杂度较高的跟踪算法多数采用国外芯片实现, 基于国产 FPGA 的跟踪方法得到了研究。针对易于 FPGA 设计的模板匹配算法鲁棒性不足的问题, 基于绝对差和 (SAD) 相似性度量方法, 提出了求窗口内最小绝对差和 (SMAD) 的方法。利用最大值和最小值滤波预先处理窗口内图像数据, 再求最小绝对差, 将 SMAD 算法的加减法器资源消耗降低为原来的 31.8%。提出的类金字塔的方法在节省逻辑资源的同时, 提高了对目标尺度变化的适应性。OTB 数据集的实验表明: 提出的方法与鲁棒性较高的 DDIS 算法相比, 跟踪成功率和平均重叠率分别提高了 1.18%、0.13%, 又有易于流水线设计的优势。利用国产 FPGA 实现了帧率可达 100 fps 的实验系统并验证其跟踪性能, 为在国产 FPGA 上实现高帧率目标跟踪系统提供了解决方案。

关键词: 国产 FPGA; 目标跟踪; 模板匹配; 高帧率

中图分类号: TP391.41 **文献标志码:** A **DOI:** 10.3788/IRLA20220905

0 引言

现场可编程逻辑门阵列 (Field Programmable Gate Array, FPGA), 性能好、灵活性高, 在军工、医疗等领域有着重要的作用。但目前 FPGA 几乎被美国垄断, 中国的国产化需求迫切。国产 FPGA 已取得突破性进展^[1], 国产的中低端芯片已达到基本应用需求, 但其仍需在应用中发展。在开发时可能会遇到逻辑混乱、时序违例、走线困难、扇出能力差等诸多难点^[2], 但只有不断暴露问题, 才能逐渐提升国产 FPGA 的可靠性。相比基于 DSP 的跟踪系统^[3], FPGA 有并行运算的优势, 研究其自主可控是必然趋势^[2, 4]。局限于高性能跟踪算法的复杂性, 目标跟踪应用多数都采用了国外的高性能芯片实现^[5-6], 至今很少有关于跟踪系统国产化的文献发表。所以研究一个基于当前国产 FPGA 的跟踪算法具有重要的研究价值。

目标跟踪算法主要分为生成式和判别式。生成式算法基于目标的灰度、几何等特征对物体进行跟踪^[7], 如经典的模板匹配算法。而判别式包括相关滤波^[8]和深度学习^[9]等算法, 一般有更高的跟踪准确度, 但很难在保证高帧率的前提下实现。若用国产芯

片实现目标跟踪, 除了考虑逻辑资源的局限性, 还需考虑实现方式的劣势。利用进口 FPGA 实现复杂算法的较大优势是可直接用 C++、python 语言进行开发, 极大地减小了开发难度。国产芯片多数需用 Verilog 开发, 即使是利用高级综合工具实现了跟踪功能的人员也不能轻易将其国产化。而且目前国内各厂家 FPGA 的基础架构不同, 考虑该研究在其他芯片中应用的可行性, 文中目标是研究一个易于 Verilog 设计、有通用性的跟踪算法, 并提高实时性和鲁棒性。所以文中选择适合用流水线设计的模板匹配作为基础算法, 其因简单性和准确性得到广泛应用, 分为基于灰度和特征相关两类。用灰度值进行匹配的方法简单, 但对于尺度变化和部分遮挡的情况效果不佳^[10]。基于特征的算法可提高鲁棒性^[11], 但复杂的特征提取也有硬件需求较高的问题, 会受到芯片资源的限制。基于灰度的模板匹配一直被研究与改进, 基本方法有绝对差异和 (Sum of Absolute Difference, SAD)、差分平方和 (Sum of Squared Difference, SSD) 和归一化互相关 (Normalized Cross Correlation, NCC)。这三种方法衡量相似性时会严格一对一地对比模板和候选窗口

收稿日期: 2022-12-29; 修订日期: 2023-02-20

作者简介: 王向军, 男, 教授, 博士, 主要从事光电传感与测试、计算机视觉与图像分析、微小型光机电系统及 MEMS 方面的研究。

图像中的所有像素点,抗干扰能力较差,当目标变化复杂时,跟踪性能严重下降。而 BBS^[12]方法是在两个点集之间进行相似性测量,当存在较大背景噪声、遮挡和非刚性变形时,依然有较精准的跟踪效果。后又有 DDIS^[13]方法在此基础上,进一步提高了模板匹配的抗干扰能力。但这两种方法需求取两个 patch 集合的两两匹配分数,若用中端 FPGA 实现将会面临资源不足的问题。后来 QATM^[14]等方法则是结合深度学习的匹配算法,目前不适合用国产 FPGA 实现。

文中针对 SAD 模板匹配算法缺乏合适更新模板机制以及鲁棒性不足等问题,提出了求窗口内最小绝对差和 (Sum of Minimum Absolute Difference, SMAD) 的方法,解决了 SAD 衡量相似性过于严格的问题。同时针对基本逻辑单元有限的问题,文中进一步改进了 SMAD 方法,大大降低了其对加法器的消耗。此外,文中还提出一种易于 FPGA 硬件实现的类金字塔更新策略,从而更好地适应目标的尺度变化。验证文中算法的跟踪性能后,将其用紫光 FPGA 实现并搭建出了一个实时目标跟踪系统,进一步说明了目标跟踪系统国产化设计的可行性。

1 SMAD 模板匹配算法

1.1 SAD 模板匹配算法原理

模板匹配是在一帧图像中提取和模板同样大小的候选窗口图像,通过衡量两图像相似度的方法搜索与目标最相似的位置。其中 SAD 的方法由于没有乘除法运算,用其在 FPGA 中实现目标跟踪,可以节省大量逻辑资源。假设模板 T 尺寸为 mn , 视频图像 I 尺寸为 MN , 则 SAD 值可用如下公式求出:

$$S(x,y) = \sum_{c=1}^m \sum_{r=1}^n |T(c,r) - I(x+c-1,y+r-1)| \quad (1)$$

式中: $1 \leq x \leq M - m + 1, 1 \leq y \leq N - n + 1$; $S(x,y)$ 为以 (x,y) 为左上角的候选子图的 SAD 值。

该算法步骤为: 1) 从初始帧中提取模板; 2) 将模板在下一帧图像中逐像素点移动, 当前模板覆盖区域为候选子图。通过求二者像素点的 SAD 来得到该子图的差异值, 直到求出整幅图的所有候选子图差异值为止; 3) 以相似度最大 (即差异值最小) 的位置, 作为本帧的目标位置; 4) 对每一帧图像执行步骤 2)~3)。

由于该方法测量相似性时会把模板和候选窗口中的所有像素逐一比对, 在跟踪目标变化小, 图像背景噪声少的环境下可以实现高效跟踪。但当目标自身变化、存在遮挡、光照改变时, 跟踪成功率会大大降低, 所以 SAD 的相似度衡量方法有待改善。

1.2 SMAD 方法的提出

针对传统模板匹配受复杂场景、遮挡、形变影响严重的缺点, BBS^[9]和 DDIS 算法^[10]将模板和候选窗口分成包含多个 patch 的集合, 然后求每个 patch 的最佳匹配, 依据双向匹配数量确定目标位置。但是这两种方法用硬件实现的复杂度较高, 如果要保证不丢帧地实时处理图像数据流, SAD 需并行求出 $m \times n$ 对像素点的相似性, 但这两种方法需求出 $(m/p \times n/p)^2$ 对 patch 的相似性 (p 一般取 3), 并行计算复杂度远高于 SAD。因此, 文中也从避免模板与待匹配子图的逐像素点严格求差值的角度考虑, 提出了一种求窗口内最小绝对差之和 (SMAD) 的方法。基本思想是当评价候选子图与模板的相似性时, 不直接求模板某像素值与待匹配子图对应位置的像素值之差, 而是求其与待匹配子图对应位置周围 9 个像素点的最小差值。如图 1 所示, 当模板滑动到视频图像的 (x,y) 位置时, 在求模板图像 (c_0,r_0) 位置像素点 $T(c_0,r_0)$ 与待匹配子图的相似性时, 需考虑图像中红框内的 9 个位置像素点与 $T(c_0,r_0)$ 的差值, 以其中的最小绝对差作为 (c_0,r_0) 的差异度。而在求模板的边缘位置的窗口内最小绝对差时, 不考虑周围像素点, 只求对应像素点的绝对差即可, 如图 1 中的蓝色像素对所示。

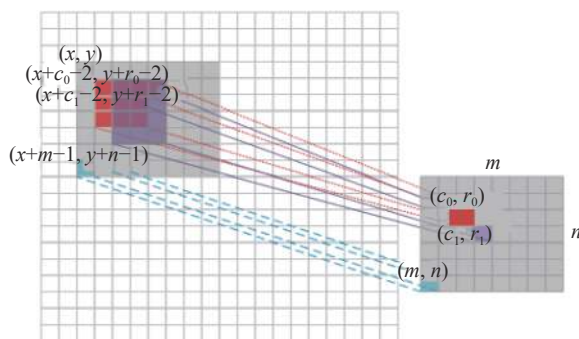


图 1 SMAD 方法示意图

Fig.1 Schematic diagram of SMAD method

如果用 $D(x,y)$ 表示视频图像中 (x,y) 位置待匹配图的 SMAD 值, 则可用公式 (2) 表示:

$$D(x,y) = \sum_{c=1}^m \sum_{r=1}^n \min_{\substack{-1 \leq p \leq 1 \\ -1 \leq q \leq 1}} (|T(c,r) - I(x+c'-1, y+r'-1)|) \\ c' = \min(\max(c+p, 1), m) \\ r' = \min(\max(r+q, 1), n) \quad (2)$$

SMAD 方法通过窗口改进 SAD 方法对每个像素异常值都敏感的劣势,从而提高了模板匹配算法对噪声、形变、局部遮挡情况的鲁棒性。当视频图像中某个像素点存在较大噪声时,可以通过周围像素点是否与对应模板像素值接近来滤除这些异常值的影响。而目标形变导致的部分特征发生偏移,局部遮挡导致的部分区域出现异常值,也可通过窗口内求最小绝对差的思想来提高匹配的成功率。

1.3 SMAD 方法的资源优化

如果采用的模板大小为 32×32 , 计算模板与同大小的候选子图时,用 SAD 方法需要 1024 次减法求绝对差,然后采用如图 2 所示的多级流水加法器求和,总共需要综合出 $32 \times 32 \times (1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \dots + \frac{1}{1024}) = 2047$ 个加减法器。如果采用 SMAD 方法实现模板匹配,理论上会有更多的减法器和比较器开销,待匹配子图的边缘区域需要 31×4 个减法器,其余区域的窗口内绝对差总共需要 $30 \times 30 \times 9$ 个减法器,加上多级流水求和的 1023 个加法器,需要 9247 个加法或减法器,约是传统 SAD 方法的 4.5 倍。

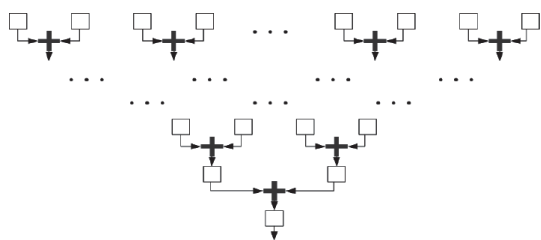
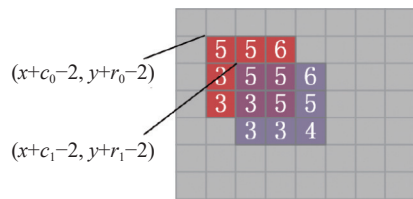


图 2 多级流水线加法器

Fig.2 Multistage pipelined adder

紫光同创 FPGA 的查找表资源有限、块 RAM 充足,基于文中研究出了一种基于最大值滤波和最小值滤波的优化方法,如图 3 所示。

图中表格的“/”前后分别为 $T(c_0, r_0)$ 和 $T(c_1, r_1)$ 的最小绝对差。先将图像数据进行最大值滤波与最小值滤波,得到原像素值周围的 3×3 窗口内的最大与最小值,之后在计算 SMAD 时,不比较窗口内的 9 个值与



Find the minimum absolute difference

$T(c_0, r_0)/T(c_1, r_1)$	Before optimization	After optimization
2	1/1	1/1
4	1/0	0/0
5	0/0	0/0
7	1/1	1/1

图 3 SMAD 方法的优化前后对比

Fig.3 Comparison before and after optimization of SMAD method

模板的差异,而是只比较与最大值、最小值的差异。以 $T(c_0, r_0)$ 和 $T(c_1, r_1)$ 所对应的两个小窗口为例,最大值滤波后的值都为 6,最小值滤波后的值都为 3。如果模板中 $T(c, r)$ 的值小于等于 3 或者大于等于 6,则窗口内最小绝对差值一定为模板像素与 3 或 6 的差距。当 $T(c, r)$ 的值在最大值滤波与最小值滤波结果范围内时,考虑大多数像素值变化的平滑性,默认窗口内存在与模板相同的像素值,即令最小绝对差值为 0,如图 3 表格中 $T(c, r)$ 为 5 的情况所示。当出现如表格中 $T(c, r)$ 为 4 的例外情况,最小绝对差理论值与通过最大值与最小值滤波优化后的结果存在差异,但这种差异较小,对最终通过多级加法求和后的 SMAD 值影响不大,可忽略这种情况带来的偏差。通过此方法处理数据,需要的加减法器仅为 2947 个,只用了理论上 31.8% 的加减法器,大大降低了资源消耗。

1.4 类金字塔的模板更新策略

当形变和遮挡现象较为严重,目标尺度变化较大时,仅用 SMAD 策略不足以支持目标的持续跟踪,需要一种合适的模板更新机制来进一步提升跟踪效果。目前更新模板主要从时间和形变两个角度考虑,即设置更新模板的时间间隔或者当最佳匹配的差值超过阈值后更新模板。这两种更新策略都没有考虑到目标尺度变化较大时,模板大小是否还合适的问题。当目标变小后更新模板会导致模板包含较多背景信息,当目标变大后不能提取到重要特征,进而导致跟踪精度下降。文中在阈值更新策略的基础上,提出了一种类金字塔的更新策略,通过在更新模板时提供的目标尺度信息,决定新模板的提取尺度。

提出的类金字塔模板更新策略和经典的高斯金字塔不同的是,文中主要从节省资源角度出发,并没有构建多个图层,而是只利用一幅模板图和候选子图实现了多尺度的滑窗匹配。如图 4 所示,上面的 3 个模板图 T 为从视频流中提取的 $m \times n$ 的模板,下面的 3 个图为 $M \times N$ 的某帧图像 I 。为保证模板以某个像素位置为中心,模板的行和列数都设为奇数。当求模板 T 与图像 I 中某 $m \times n$ 候选子图的相似度时,应如中间图所示,按像素坐标一一对应。为了能够同时测量现目标尺度相对于模板目标尺度变小或变大后的相似度,分别采用了图 4 中左图和右图的对对应关系。如果当前目标大小几乎变为原来的一半,则模板降采样后和含有目标的待匹配子图的中心区域有较高的相似度;同理,如果当前目标大小几乎变为原来的两倍,则含有目标的待匹配子图降采样后和模板的中心区域有较高的相似度。文中直接通过间隔取像素点的方式来达到降采样的目的,如图 4 中左上角和右下角图中标黄的点为降采样后有效的像素值,再分别与待匹配子图或模板中心的 $(m-1)/2 \times (n-1)/2$ 子图按照蓝色线对应。然后依据文中提出的 SMAD 方法来计算候选窗口与模板的三种差异值 $D1 \sim D3$ 。其中, $D1$ 最小时表明当前目标变小,更新模板时提取行列数各缩小为之前的 $1/2$; $D3$ 最小时则表明当前目标变大,更新模板时将提取行列数各增加为之前的两倍。

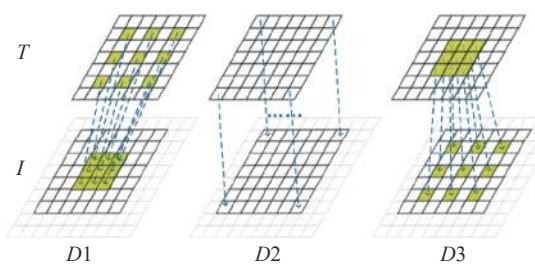


图 4 类金字塔策略

Fig.4 Pyramid-like strategy

2 目标跟踪系统的设计

2.1 国产 FPGA 分析与选用

在功能和性能上,国内 FPGA 都与进口芯片存在一定的差距。如表 1 所示,在产品种类、工艺水平、门数、IP 核、市场、技术支持等方面,国外 FPGA 都存在较大优势。目前中国 FPGA 以 40、28 nm 工艺为

主,但早在 2011 年国外就发布了采用 28 nm 工艺的 FPGA 工程样片。在 EDA 工具方面,国内水平相对落后,存在布通率较低等问题。但使用国内芯片具有可控性高、自主性强的优势,用其研发各个领域的技术产品已经成为主流趋势,况且国产 FPGA 性能方面能够满足中、低端的市场需求。不过国产芯片开发方式单一,在门电路延迟、扇出能力、布局布线方面存在劣势,需要通过优化逻辑、添加时序约束、物理约束改善。为了研究能达到较高帧率、适用于国产 FPGA 的跟踪算法,文中选择紫光同创 Logos2 系列的 PG2 L100 H-6 IFBG676 芯片将其实现并验证。

表 1 国内外 FPGA 对比

Tab.1 Comparison of domestic and foreign FPGAs

Metrics	Foreign FPGA	Domestic FPGA
Types	More	Fewer
Process	7 nm	28 nm
Gates	100 million	Most: 10 million
IP cores	More	Fewer
Market	High/mid/low end	Mid/low end
Technology	Mature	Immature
Autonomy	Poor	Better

文中基于该芯片构建了一个国产化的实时目标跟踪系统,并进行算法设计与验证,整体框图如图 5 所示。首先通过 IIC 配置 OV5640 相机分辨率为 1280×720 ,然后将分两个时钟输出的 8 bit RGB565 数据流转换成 16 bit 像素数据。之后通过帧读写控制模块依据摄像头的场同步信号切换 DDR3 写入某帧数据的起始地址。当读取下一帧时,读地址设置为当前写地址的前一帧起始地址,保证读和写的地址不冲突。此外该模块还需例化 2 个异步 FIFO 来解决跨时钟域问题。从 DDR3 读出图像数据后,将数据按照 VGA 标准转换成 1280×720 视频流。视频流一路转换成灰度图传到目标跟踪模块求取当前帧目标位置,另一路传到 HDMI 视频编码芯片,将视频实时输出显示。此外,文中在将图像数据转成 VGA 标准时,依据跟踪位置实时画出目标框。同时,还设计了 UART 通信模块用于实现上位机点选和跟踪目标位置的回传。通过提高视频数据读出时钟并添加时序约束,该系统的帧率可达 100 fps,有较好的实时性。

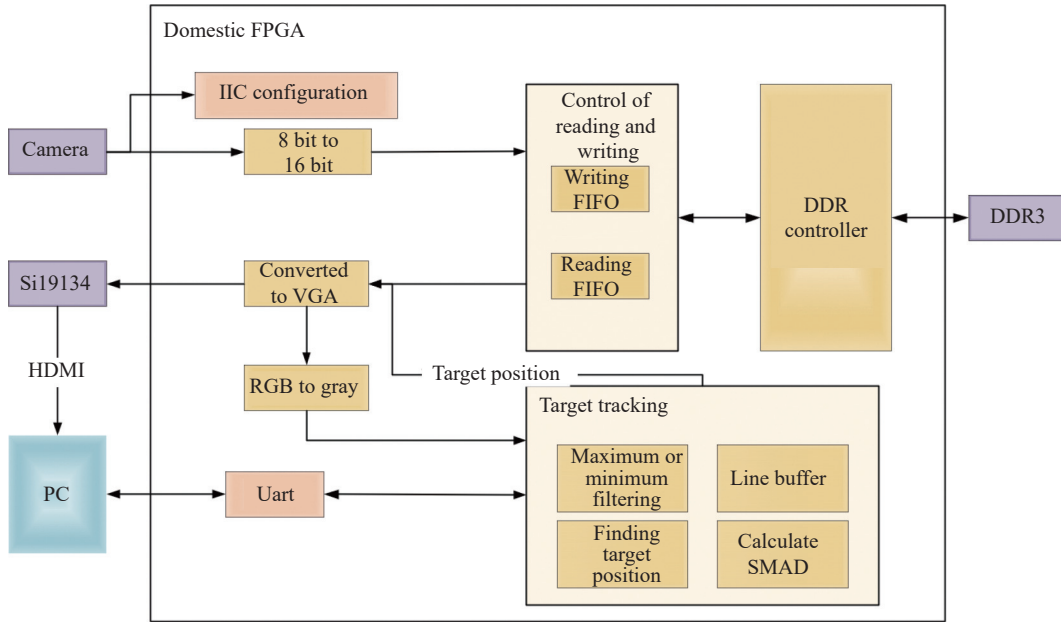


图 5 目标跟踪平台结构框图

Fig.5 Structural diagram of the target tracking platform

2.2 最大值与最小值滤波模块

最大值和最小值滤波模块为文中优化 SMAD 方法的核心模块,其硬件架构如图 6 所示。用 FPGA 实现滤波算法,需要利用多个 FIFO 实现多行缓存。FIFO 深度设为图像一行的大小, FIFO 数量取决于模

板大小,构成最终的 Line Buffer 模块。先通过两个深度为一行的 FIFO 得到 3 行并行数据,之后通过触发器得到 3×3 窗口。文中实现了能在 3 个数据中找最大值和最小值的模块,通过例化 4 个该模块并用两级流水线即可求出 9 个数据中的最大和最小值。此外,

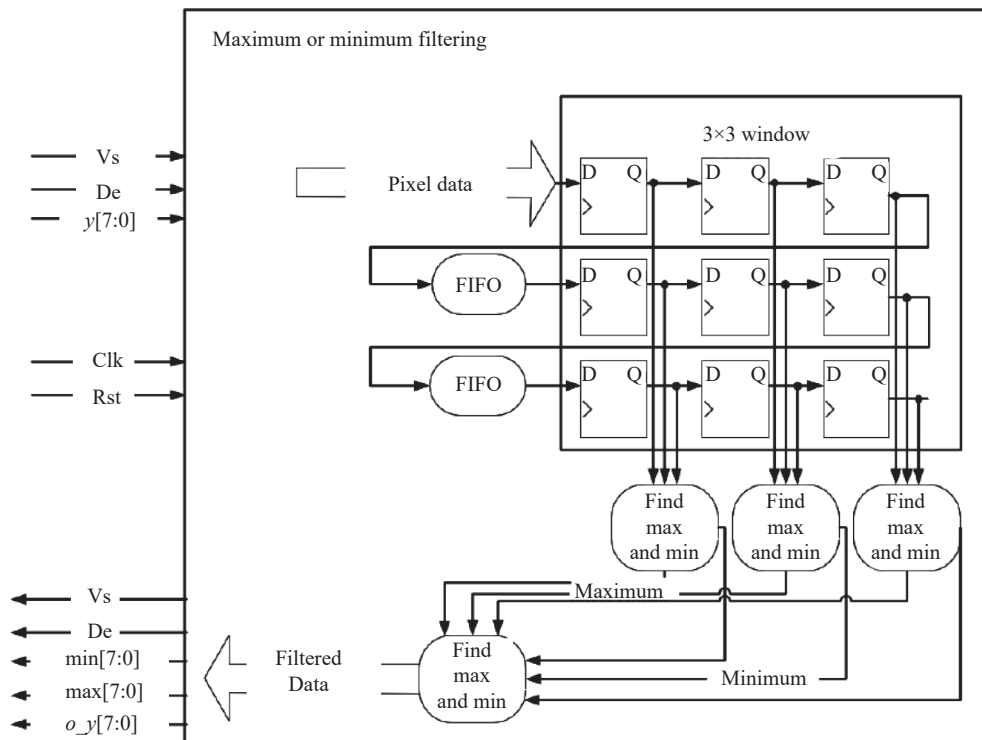


图 6 最大值和最小值滤波的硬件架构

Fig.6 Hardware architecture for maximum and minimum filtering

该模块还需输出和最大值、最小值时序对齐的原像素数据、场信号和数据有效信号,之后输入到目标跟踪模块。

2.3 类金字塔的 SMAD 计算模块

类金字塔的 SMAD 计算模块需要 Line Buffer 模块来形成符合模板尺寸的窗口数据,模板的寄存也在该模块中实现。依据上位机传来的初始目标位置提取,为了方便求取模板中心,文中模板大小设为 31×31 。跟踪过程中,如果尺度发生改变则会通过降采样的方式增加模板的提取范围,相当于将模板尺度分为 31×31 、 62×62 、 124×124 三级。该模块的输入包括: 1) 31 行并行的最大值和最小值滤波数据及其有效信号; 2) 像素数据位置,依据场同步和有效信号利用计数器产生,为模板提取时提供数据流的坐标; 3) 模板更新请求,该信号有效时,需在下一帧提取新的模板; 4) 模板位置,初值为点选目标位置,有更新模板请求时,变为当前的目标跟踪位置。而输出信号包括: 1) SMAD 数据,滤波后的图像数据流和模板求绝对差后通过多级流水求得; 2) SMAD 有效信号; 3) 尺度信号,利用类金字塔思想产生的尺度标志。

此模块以上位机点选标志作为复位信号,当串口模块输入有效点选位置才开始真正工作。此时提取模板模块会依据当前数据流的位置信息决定是否在

本帧提取模板,如果数据流已经超过了模板框的位置,则等到下一帧再提取模板,否则直接在本帧提取。同时通过寄存器缓存的待匹配子图的像素值、最大值和最小值滤波数据,将这 3 个数据按照图 4 所示的类金字塔组合方式输入到求绝对差值的模块,最终会得到 3 个尺度的绝对差数据。模板数据如果小于最小值,则输出最小值与其的绝对差,如果大于最大值,则输出最大值与其的绝对差,否则输出为 0。然后,将 3 个绝对差输入到并行的多级流水加法器模块,得到 3 个 SMAD 值。以 D_2 为相似度的衡量,作为分数信号输出,而尺度信号表明了 D_1 、 D_2 与 D_3 的大小关系。由于 D_2 是 $m \times n$ 个差值的求和,加数是 D_1 和 D_3 的 4 倍,所以,将 D_1 、 D_3 与 D_2 右移两位的值进行比较。当产生更新模板请求后,存储模板图时会依据该尺度信息在三级模板尺度中进行切换。

2.4 目标搜寻模块

目标搜寻模块是上述模块的后续流水设计,该模块主要用来寻找 SMAD 数据流的最小值。其记录最小值所对应的像面位置,并以此为最终的跟踪目标位置输出到 VGA 转换模块(依据跟踪结果画出跟踪框)和串口通信模块(将位置信息传给上位机)。该模块的实现逻辑如图 7 所示。其输入主要为类金字塔

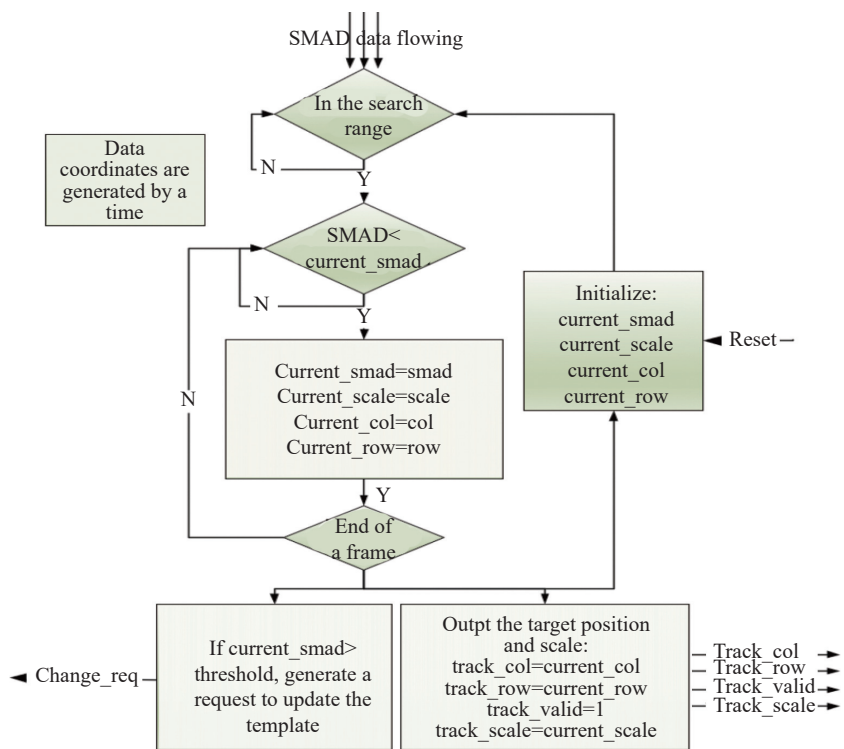


图 7 目标搜索模块流程图

Fig.7 Flowchart of the target search module

的 SMAD 计算模块输出的 SMAD 数据流。依据输入数据有效和场同步信号,用计时器产生了数据流的坐标位置 col 和 row,之后依据该坐标判断数据流是否在搜索范围内,不在则不对数据流采取任何操作。如果是在搜索范围内,则需判断该分数数据流 SMAD 值是否为最佳。文中采用 current_smad、current_scale、current_col、current_row 这 4 个寄存器变量来存储某一帧中目前最佳匹配对应的差异值、尺度、列坐标和行坐标,在每一帧开始之前,将 current_smad 赋值为其位宽范围内的最大值。如果输入分数在搜索范围内且小于 current_smad,则更新上述四个寄存器变量直到场同步信号拉低。一帧数据处理完成后,依据最佳匹配差异值 current_smad 与阈值的关系判断是否产生模板更新请求。同时输出目标跟踪位置 track_col 和 track_row,跟踪有效信号拉高并输出跟踪目标的尺度信息 track_scale。

3 实验

3.1 算法验证

文中利用 Matlab 实现了提出的跟踪算法,并用 OTB 数据集进行测试。目标跟踪算法性能的评价指标主要从目标中心像素误差和矩形框重叠率两个角度出发。中心像素误差指算法跟踪结果的中心点与真实目标中心点位置的欧氏距离,其缺乏尺度的考量。重叠率是指算法跟踪结果的波门与真实波门的重合度。

为了验证 SMAD 相较于 SAD 方法的提升效果,选取有旋转、尺度变换、遮挡、快速运动属性的图像序列,测试二者跟踪每一帧图像的重叠率,如图 8 所示。跟踪前期二者的指标差距不大,但随着目标发生变化且被遮挡后,SAD 算法重叠率变为 0,即跟踪失败,而 SMAD 方法依然保持 0.28 的重叠率。为了验证类金字塔策略在更新模板时的正确性,文中利用如图 9 所示的飞机模型进行测试,当切换模板时显示当前模板在图中的位置。图中红框内图像表示模板,当目标尺度变化较小时,维持当前模板尺度,随着模型尺度增大,模板将会切换到更大尺度。

模板匹配的核心就是模板和视频图像中子窗口的相似度衡量,为了进一步验证文中研究算法的性能,文中除了选用 SAD、SSD、NCC 这三种传统的度

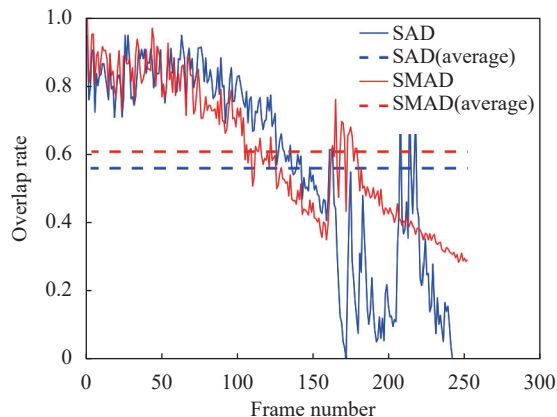


图 8 SMAD 和 SAD 跟踪指标对比

Fig.8 Comparison of tracking indicators between SMAD and SAD

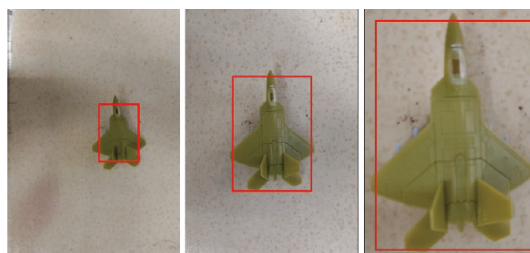


图 9 模板的尺度切换

Fig.9 Scale switching of the template

量方法,还选用了具有高鲁棒性的无参数 BBS 和 DDIS 算法进行对比。由于国产 FPGA 性能的局限性,涉及深度特征的算法文中不予考虑,且 BBS 和 DDIS 算法是基于灰度和位置信息进行跟踪的,如果采用其他复杂特征更会增加 FPGA 实现的难度。为了直观地对比各个算法的跟踪性能,文中截取了汽车序列(如图 10 所示)的其中四帧,图中左上角数字即为图像序列帧号。当汽车尺度变大并逐渐被树干遮挡时,基于 SAD、SSD 和 NCC 三种相似度衡量的跟踪算法都会跟踪失败。而 BBS 和 DDIS 算法具有较高的抗遮挡能力,能在遮挡后依然跟踪成功,文中提出的 SMAD 算法也能成功跟上汽车目标,证明了文中算法的抗遮挡和尺度适应能力。

为了测试跟踪算法在各个场景下的总体性能,可通过统计多次跟踪结果的平均指标进行评价,文中采用平均重叠率(Average Overlap Rate, AOR)和成功率(AOR 大于一定阈值的占比)两个指标来进行算法之间的比较,得到表 2。其中 SMAD-0 表示未添加类金字塔策略的 SMAD 方法,和经典的 SAD 方法相比,各

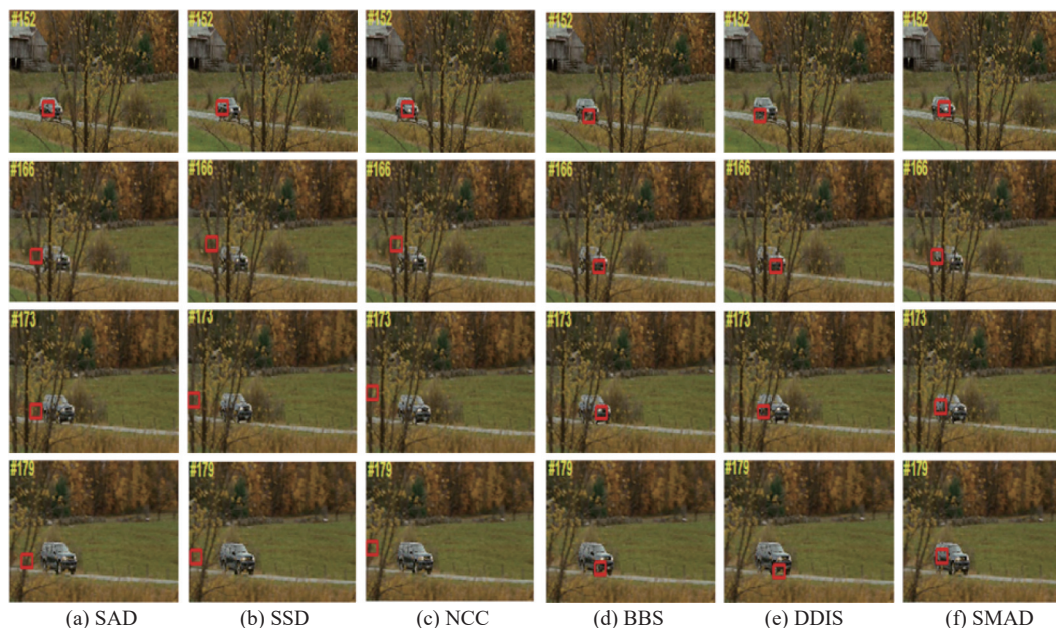


图 10 各模板匹配算法跟踪汽车序列的性能比较

Fig.10 Performance comparison of each template matching algorithm tracking the car sequence

表 2 各跟踪算法的指标对比

Tab.2 Comparison of metrics for each tracking algorithm

	Rotational deformation		Scale change		Partial occlusion		Various scenarios	
	Success rate	AOR	Success rate	AOR	Success rate	AOR	Success rate	AOR
SAD	0.7604	0.39548	0.8012	0.5364	0.7412	0.33176	0.7680	0.3818
SSD	0.7563	0.36219	0.7786	0.4997	0.7454	0.33879	0.7511	0.3695
NCC	0.7588	0.39248	0.7567	0.4404	0.7362	0.30951	0.7467	0.4228
BSS	0.7642	0.45265	0.8385	0.6817	0.7847	0.50696	0.7851	0.4974
DDIS	0.7683	0.45625	0.8273	0.6188	0.7802	0.50951	0.8197	0.5399
SMAD-0	0.7640	0.44162	0.8261	0.5583	0.7775	0.48105	0.7842	0.4974
SMAD	0.7634	0.44186	0.8594	0.6821	0.7791	0.48060	0.8294	0.5406

个场景下的跟踪指标都有所提升,验证了该相似性度量方法的抗干扰能力。对比 SMAD-0 和 SMAD 在尺度变化和综合场景下的跟踪指标,成功率和 AOR 都有所提高,验证了类金字塔更新策略的有效性。

对比表中 SAD 和 SMAD 的跟踪指标可以验证文中优化策略的整体提升效果,其中成功率提高了 8.0%,AOR 提高了 41.5%。文中算法和目前有高鲁棒性的 DDIS 匹配方法相比,在综合场景下成功率和 AOR 分别提高了 1.18% 和 0.13%。从三种特定场景的跟踪指标进行分析,DDIS 算法在旋转形变和部分遮挡场景下的跟踪性能更优,但当目标发生尺度变化时文中算法有更好的跟踪效果。从 FPGA 流水线设

计角度考量,文中算法有较大的优势。

3.2 设计模块的功能仿真

文中将如图 5 所示的各个模块用 QuestaSim 软件进行了功能验证。为了测试跟踪结果的正确性,文中将初始帧目标的中心坐标设为(40,20),每一帧目标都向右下角平移一个像素,即在 x 和 y 方向位置都加 1。得到仿真结果如图 11 所示,其中点选位置 target_col 和 target_row 设为初始帧中目标的中心,跟踪模块输出的跟踪结果为 track_col 和 track_row,二者每一帧都加 1,与理论值相符。同时仿真分析各模块的延迟时间,将图像数据转灰度并进行跟踪处理的整个流水线的延时时间为 16 个行同步周期+37 个时钟周期。

其中,行延迟取决于模板大小和滤波模块的窗口大小,时钟延时是由于多级加法器(10个时钟周期)、窗口数据寄存(15个时钟周期)和其他时序优化导致。之后使用紫光FPGA进行硬件开发,整个跟踪系统的查找表(LUTs)、寄存器(Registers)、专用RAM(DRM)和输入输出接口(I/O ports)消耗如表3所示。其中每个 DRM 有 36 kBits 的存储单元,可独立配置为两个 18 kBits 单元。LUTs 和 Registers 作为基本逻辑资源,

主要消耗在目标跟踪算法的多级流水加法器部分,目标跟踪模块使用了 67% 的 LUTs、40.5% 的 Registers,其余模块的消耗量相对较少。利用的 DRM 中有 30 个用于实现目标跟踪模块的 Line Buffer,其余 6.5 个用于实现 DDR 读写的 FIFO 缓存、最大值与最小值滤波、摄像头参数存储。I/O ports 用于与摄像头、DDR、视频编码芯片和串口相连。

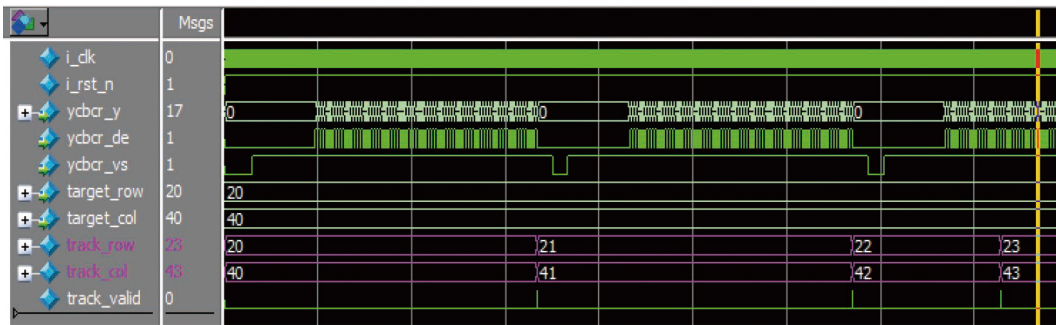


图 11 目标跟踪模块功能仿真波形图

Fig.11 Functional simulation waveform of the target tracking module

表 3 FPGA 资源消耗

Tab.3 Resource consumption of the FPGA

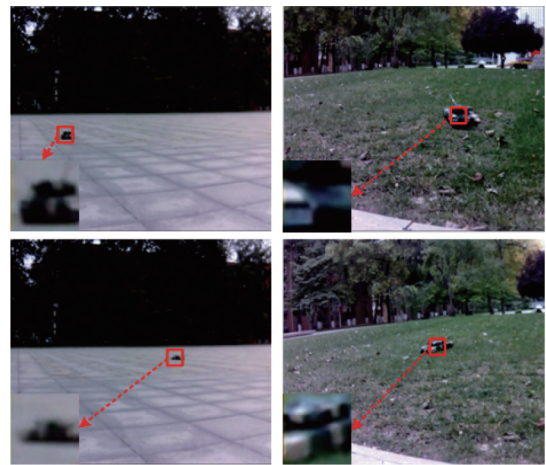
Resources	Consumption	Consumption percentage
LUTs	48457	72.76%
Registers	58587	43.98%
DRM	36.5	23.55%
I/O ports	102	34.00%

3.3 跟踪实验

文中使用 Qt Creator 开发了与基于国产 FPGA 的目标跟踪平台适配的上位机应用程序。其主要实现以下功能: 1) 读取 FPGA 传来的视频数据,并实时显示; 2) 能实现目标点选功能; 3) 通过串口接收 FPGA 回传的目标跟踪位置。为了测试文中搭建的目标跟踪系统,利用上位机程序和该系统进行联合实验。基于豪威科技的 OV5640 摄像头,把坦克缩比模型作为跟踪目标对平台进行了测试。

文中选取不同背景,在目标发生尺度变化、运动方向和速度变化时,测试该平台的跟踪情况。先选取比较空旷的广场作为背景,测试跟踪算法的基本性能。在上位机点选目标后,令坦克模型进行无规则运

动,若跟踪框中心为目标像素点,则判定跟踪有效。测试结果如图 12(a) 所示,验证了文中搭建系统的实用性。然后,文中选取林地草坪为背景,控制坦克模型沿任意方向运动。由于草地与缩比模型颜色相近且存在模型底部遮挡问题,更能测试跟踪平台的鲁棒性。如图 12(b) 所示的实验表明,在存在干扰的背景



(a) 广场测试 (a) Testing on the lawn (b) 草坪测试 (b) Testing on the square

图 12 目标跟踪系统的跟踪实验

Fig.12 Tracking experiment for a target tracking system

下,该目标跟踪平台仍然可以成功跟踪到目标。

此外,文中也进行了跟踪平台的抗遮挡能力测试实验,如图 13 所示。当模型被草丛挡住一部分时,文中的跟踪平台仍能成功跟踪到目标,进一步验证了文中提出方法的可行性。



图 13 目标跟踪系统的抗遮挡测试

Fig.13 Anti-occlusion test of target tracking system

4 结 论

为打破技术垄断,提高目标跟踪应用产品的自主性,文中从国产 FPGA 性能和跟踪帧率两个角度进行考察,改进了传统的 SAD 模板匹配方法,提出了一种 SMAD 方法。并通过最大值滤波与最小值滤波策略降低了其流水线设计的逻辑资源消耗,总共需要的加减法器资源减少为原来的 31.8%。针对跟踪目标尺度变化较大的情况,文中研究出了一种类金字塔的模板更新策略来提升算法的跟踪性能。基于 OTB 数据集的实验结果表明,文中提出的算法与 SAD 方法相比,跟踪成功率和 AOR 分别提高了 8.0%、41.5%。相比于目前具有较强抗干扰能力的 DDIS 算法,跟踪成功率提高了 1.18%,AOR 提高了 0.13%,且有着更适合利用 FPGA 进行流水线设计的优势。然后,文中在国产 FPGA 上构建了实时的目标跟踪系统,跟踪帧率可以达到 100 帧/s。其次,文中对其进行了跟踪性能测试,结果表明该系统的效果较好,为高帧率目标跟踪系统的国产化提供了参考方案。

参考文献:

- [1] Khan S M, Mann A, Peterson D. The semiconductor supply chain: Assessing national competitiveness[M]. US: Center for Security and Emerging Technology, 2021.
- [2] Yang Dezhen, Li Kaifeng, Cai Jiayi, et al. Realization of infrared image processing based on domestic FPGA [J]. *Laser & Infrared*, 2022, 52(8): 1223-1229. (in Chinese)
- [3] Wang Xiangjun, Guo Zhiyi. Long-term target tracking algorithms in force-constrained platform [J]. *Infrared and Laser Engineering*, 2020, 49(1): 0126003. (in Chinese)
- [4] Wu Lei, Wang Yongjia, Jiang Xianyang. Efficient migration of LwIP without operating system based on Ziguang FPGA [J]. *Journal of Computer Applications*, 2021, 41(S2): 183-186. (in Chinese)
- [5] Przewlocka-Rus D, Kryjak T. Towards real-time and energy efficient siamese tracking—a hardware-software approach[C]// Design and Architecture for Signal and Image Processing: 15th International Workshop, 2022: 162-173.
- [6] Cui Z, An J. Heterogeneous siamese tracking system based on PYNQ framework[C]//2020 6th International Conference on Control, Automation and Robotics (ICCAR), IEEE, 2020: 16-20.
- [7] Chen Fangfang, Song Daiping. Adaptive scale target tracking algorithm based on dynamic template [J]. *Laser & Optoelectronics Progress*, 2023, 60(4): 0410018. (in Chinese)
- [8] Huang Yueping, Li XiaoFeng, Lu Ruitao, et al. Research on infrared object tracking algorithm via adaptive label and sparse-learning correlation filter [J]. *Chinese Journal of Scientific Instrument*, 2022, 43(12): 199-210. (in Chinese)
- [9] Zhang Jinpu, Wang Yuehuan. A survey of siamese networks tracking algorithm integrating detection technology [J]. *Infrared and Laser Engineering*, 2022, 51(10): 20220042. (in Chinese)
- [10] Wu Xiaojun, Zou Guanghua. High performance template matching algorithm based on edge geometric features [J]. *Chinese Journal of Scientific Instrument*, 2013, 34(7): 23-30. (in Chinese)
- [11] Kim J R, Jeon J W. Real-time scale and rotation invariant multiple template matching[C]//2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM), IEEE, 2022: 1-5.
- [12] Dekel T, Oron S, Rubinstein M, et al. Best-buddies similarity for robust template matching[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 2021-2029.
- [13] Talmi I, Mechrez R, Zelnik-Manor L. Template matching with deformable diversity similarity[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 175-183.
- [14] Cheng J, Wu Y, AbdAlmageed W, et al. QATM: Quality-aware template matching for deep learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 11553-11562.

High frame rate target tracking method using domestic FPGA

Wang Xiangjun^{1,2}, Zhu Hui^{1,2}

(1. State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China;

2. MOEMS Education Ministry Key Laboratory, Tianjin University, Tianjin 300072, China)

Abstract:

Objective Target tracking plays an important role in the military, medical and other fields, and Field Programmable Gate Array (FPGA) is widely used in the direction due to its good performance and high flexibility. However, at present, limited to the complexity of high-precision tracking algorithms, most of the target tracking systems are implemented by foreign high-performance chips, which leads to weak autonomy and controllability. If domestic chips are used to achieve target tracking, it will face the problem that there are few IP cores and most modules need to be designed in Verilog. In addition, the feasibility of the research algorithm in other domestic FPGAs needs to be considered. Therefore, the objective is to study a tracking algorithm that is easy to design in Verilog, has generalization, and improves real-time and robustness.

Methods Template matching is easy to design with pipelines and is selected as the basic algorithm, which is widely used due to its simplicity and accuracy. Among them, the template matching algorithm based on Sum of Absolute Difference (SAD) has no multiplication and division operation, which is suitable for FPGA implementation with limited resources. This tracking algorithm has too strict constraints, which leads to the problem of insufficient robustness. Based on the Sum of Absolute Difference (SAD) similarity measurement method, a method for finding the Sum of Minimum Absolute Difference (SMAD) in the window is proposed. In order to reduce resource usage, the maximum and minimum filtering (Fig.3) is used to preprocess the image data in the window and then the minimum absolute difference is obtained, which reduces the resource consumption of the addition and subtraction of the SMAD method to 31.8%. Moreover, a pyramid-like update strategy (Fig.4) that is easy to implement by FPGA hardware is proposed to better adapt to the scale change of the target. In order to verify the tracking performance of proposed algorithm, Unigroup FPGAs are used to implement it and build a real-time target tracking system (Fig.5).

Results and Discussions Based on the two indicators of Average Overlap Rate (AOR) and success rate, the algorithm comparison experiment was carried out with the OTB dataset. It verified that the proposed algorithm has certain anti-occlusion and scale adaptability (Fig.10). Compared with the SAD method, the tracking metrics in each scenario are improved. In the scale change and synthesis scenario, after the SMAD method is added to the pyramid-like strategy, its success rate and AOR are improved, which verifies the effectiveness of the pyramid-like update strategy (Tab.2). Compared with the robust DDIS algorithm, the proposed method improves the average success rate and overlap rate by 1.18% and 0.13%, respectively, and is easier to design with FPGA. Then the target tracking system is implemented by domestic FPGAs. The delay time is 16 line synchronization cycles plus 37 clock cycles and tracking frame rate can reach 100 frames per second. Different outdoor backgrounds were selected to test the tracking of the system when the target changed in scale, direction of motion and speed (Fig.12). The anti-occlusion test experiment of the tracking platform shows that (Fig.13) when the target is partially occluded, the tracking system can still successfully track the target, which further verifies the feasibility of the proposed method.

Conclusions In order to break the technology monopoly and improve the autonomy of target tracking application products, the traditional SAD template matching method is improved considering the limitations of domestic FPGA and the performance of tracking algorithms. The SMAD method is proposed and its resource consumption is optimized. Combined with a pyramid-like template update strategy, its tracking performance is improved. The experiments of OTB dataset and domestic tracking system verify its tracking effect. It provides a reference scheme for the localization of high frame rate target tracking system.

Key words: domestic FPGA; target tracking; template matching; high frame rate