

基于 FPGA 的多运动目标实时检测系统设计

薛浪¹ 曾庆立¹ 钟戴元¹ 周佳凯¹ 唐瑞东¹

XUE Lang ZENG Qingli ZHONG Daiyuan ZHOU Jiakai TANG Ruidong

摘要

针对运动目标检测速度慢、检测不准确并且不适于多数应用场景的问题,利用 FPGA 的低功耗、实时性、高速率等优势,文章设计了一种基于 FPGA 的多运动目标实时检测系统。系统主要由图像处理模块,数据缓存模块和 LCD 显示模块 3 个核心部分构成。其中,使用 SDRAM 作为数据流的实时缓存,确保了数据流的高效处理;在图像处理环节,系统采用帧差法提取运动目标轮廓,并对提取轮廓中的每个像素进行形态学处理,以消除噪点和填补目标物的孔洞;然后基于距离阈值识别并区分多个目标,并在原始图像上绘制出包围盒,从而实现精确的目标分割与跟踪;在算法设计完后,通过 LCD 显示屏实时呈现检测结果。实验结果表明,系统能快速检测出视频中的多个运动目标,且资源消耗较低。此外,系统在分辨率为 800 px×480 px 的情况下,能够有效地实时检测出多个运动目标,显示出了良好的实时性和准确性,为计算机视觉领域中快速检测运动物体提供了一定的参考和应用价值。

关键词

运动目标;FPGA;阈值;帧差法;包围盒

doi: 10.3969/j.issn.1672-9528.2025.01.006

0 引言

运动目标检测是计算机视觉领域的一个重要部分,在视频监控^[1-2]、自动驾驶^[3]、机器人导航^[4]等领域有着广泛的应用。其基本原理是在视频序列中检测出运动物体,并对其进行跟踪和分析。

运动目标检测常用的方法包括背景差分法^[5]、光流法^[6]和帧差法^[7]等。背景差分法是通过统计得到背景模型,与当前帧于背景模型进行比较,将差异较大的像素点标记为运动区域^[8]。该方法能够适应场景的变化,但可能在背景模型更新不及时或场景变化较大时出现误检;光流法通过估计像素点的运动矢量来检测运动物体,这种方法能够检测出较为精确的运动区域,可计算出运动目标的速度^[9],但容易受噪声和光照变化的影响,且计算量复杂,对硬件资源的需求较大,不适于硬件实现;帧差法通过比较相邻两帧或三帧之间的图像差异来检测运动物体,该方法简单易行、稳定性好、计算量小^[10-11]。

此外,运动目标检测在实时性要求高的应用场景中尤为重要,要求算法准确快速。因此,研究者们正致力于开发更高效、更准确的目标检测算法,以满足不断增长的行业需求。田洪金等人^[12]考虑到背景差分法易受背景干扰、执行速度慢等问题,对该算法进行了改进,结果检测效果良好,但未充

分考虑到目标检测的实时性。冯天骄等人^[13]采用帧差法进行检测,检测结果快速准确、实时性较好。马明哲^[14]针对环境干扰对检测效果的影响进行系统设计,结果表明该系统能够有效检测场景中的运动目标时。但这些只是针对单个运动目标的实时检测,然而,在多数情况之下,实际场景中的运动目标并非只有一个,而是多个目标的同时出现。因此,不适用于大多数场景的需求。刑凯^[15]在背景变化的情况下采用相应算法,在 FPGA 上实现了多运动目标的检测,但是算法较为单一,存在一定的局限性。

针对上述问题,本文提出了一种基于 FPGA 的多运动目标检测系统,充分利用 FPFA 在低延迟、高速率、实时处理方面的优势。该系统通过精心设计的硬件架构和优化的算法实现,能够快速准确地从视频流中检测出多个运动目标。这一成果为实时监控、智慧交通等领域的应用提供了坚实的技术基础。

1 系统总体结构设计

图 1 展示了 FPGA 实现多运动目标检测的总体系统架构设计,该架构由以下关键模块构成:首先,原始视频输入系统后,通过“RGB 转 YCbCr”模块进行色彩空间的转换,为后续处理打下基础;接着,系统将存储转换后的视频帧,利用“帧差运算”模块对连续帧进行比较,以识别运动目标;然后,利用“形态学图像处理模块”对检测到的运动区域进

1. 吉首大学通信与电子工程学院 湖南吉首 416000

行腐蚀和膨胀操作，以优化目标的轮廓；紧接着，系统执行“多目标检测”过程，精确识别出视频中的所有运动目标；最终，“绘制包围盒”模块在原始视频帧上为每个检测到的目标绘制包围盒，实现目标的精确定位和标记。此外，“SDRAM 控制模块”负责数据的存储和读取，确保整个系统的高效运行。处理后的视频流通过 LCD 显示模块以 60 Hz 的帧率呈现，整个系统通过各模块的紧密协作，确保多运动目标检测的准确性和实时性。

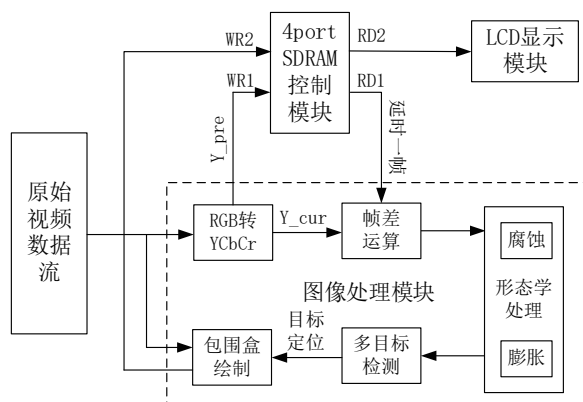


图1 系统设计框架图

2 算法原理及硬件实现

2.1 灰度化处理

灰度化处理即将彩色图像转换为灰度图像，一般将RGB565转换为YCbCr，由于FPGA不善于处理浮点数，便将浮点数转为定点数，式(1)是各项系数扩大256倍的结果，式(2)是向右移8 bit的公式。

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.144B \\ Cb &= 0.568(B - Y) + 128 = -0.172R - 0.339G + 0.511B + 128 \\ Cr &= 0.713(R - Y) + 128 = 0.511R - 0.428G - 0.083B + 128 \end{aligned} \quad (1)$$

$$\begin{aligned} Y &= (77R + 150G + 29B) \gg 8 \\ Cb &= (-43R - 85G + 128B) \gg 8 + 128 \\ Cr &= (128R - 107G - 21B) \gg 8 + 128 \end{aligned} \quad (2)$$

为了防止运算过程出现负数，对式(2)进行进一步变换，得到公式为：

$$\begin{aligned} Y &= (77R + 150G + 29B) \gg 8 \\ Cb &= (-43R - 85G + 128B + 32768) \gg 8 \\ Cr &= (128R - 107G - 21B + 32768) \gg 8 \end{aligned} \quad (3)$$

2.2 帧差法原理及实现

基于摄像头采集视频所呈现出的连续性特点，如当场景内无运动物体，连续的帧变化微弱；反之，如场景内存在运动物体，连续的帧之间的运动区域出现明显变化^[16]。因此，帧差法即对时间上连续两帧或三帧图像像素点进行差分运算，并对相减后的图像进行阈值化处理来提取图像中的运动

区域。式(4)是差分图像公式，其中 f_n 和 f_{n-1} 分别为视频序列中第 n 帧和第 $n-1$ 帧图像， $f_n(x, y)$ 和 $f_{n-1}(x, y)$ 分别是连续两帧对应像素点的灰度值，将两帧图像灰度值相减并取绝对值便得到差分图像 $D_n(x, y)$ ， T_1 是图像二值化阈值，差分图像大于该阈值便认为是运动目标区域，赋值为1；否则是静止区域，赋值为0。

$$D_n(x, y) = \begin{cases} 1, & |f_n(x, y) - f_{n-1}(x, y)| > T_1 \\ 0, & \text{others} \end{cases} \quad (4)$$

帧差法的核心是对相邻两帧灰度图像进行运算，图2展示了FPGA实现帧差法的具体过程，主要包括帧缓存模块和帧差模块，帧缓存模块需要使用SDRAM的乒乓操作实现，在缓存完一帧后，新一帧开始时，同步从缓存中读出已存储的一帧数据，同时继续写入到帧缓存中。该模块需要合理设计一个双端口SDRAM控制器。一侧读写端口用作帧缓存，另一侧端口用来缓存图像处理后的视频流。具体为将原始视频流经灰度处理的数据分为两路，一路写入SDRAM的WR1端口侧，作为帧图像数据缓存；另一路通过场同步信号控制延时从SDRAM的RD1端口读出。对于帧差模块，需要根据视频图像设定合理的阈值 T_1 ，并将当前帧图像与上一帧图像数据进行差分，若差分结果大于设定的阈值便认为是目标运动区域，并将值设置为255，否则将值置为0。

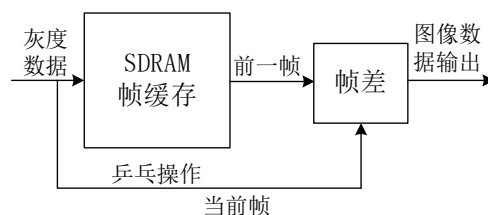


图2 帧差模块设计

2.3 形态学处理

帧差运算得到的二值图像通常包含较多噪声和不完整的目标边缘，原因是帧差法对微小光照变化和图像噪声敏感，容易导致运动目标的轮廓出现断裂或伪影。为提高检测精度，形态学操作被用于有效去除这些噪声^[17]。具体来说，该操作（先腐蚀后膨胀）能去除小的边界和孤立点，并填充目标内部的空洞，从而清晰界定运动目标轮廓。这不仅减少了图像噪声，还增强了目标的连通性，为后续目标分析和跟踪提供更精确的图像信息。

对于1 bit的 3×3 阵列，可通过腐蚀和膨胀操作进行处理。腐蚀操作在 3×3 窗口中，只要有一个像素不为1，则将其腐蚀为0；而膨胀操作只需一个像素为1，即可将其膨胀为9个有效像素。然而，这种方法可能导致过度腐蚀或膨胀，结果

不够准确。为此，可以在 3×3 窗口中引入阈值，当窗口内有 n 个 1 时，结果为 1，否则为 0。获得 3×3 窗口后，可以开始腐蚀 / 膨胀运算，整个过程可分为 3 个步骤。

(1) 计算窗口中 3 个像素的累加和。

$$\begin{aligned} \text{Sum1} &= P11 + P12 + P13 \\ \text{Sum2} &= P21 + P22 + P23 \\ \text{Sum3} &= P31 + P32 + P33 \end{aligned} \quad (5)$$

(2) 计算窗口中 9 个像素的累加和。

$$\text{Sum} = \text{Sum1} + \text{Sum2} + \text{Sum3} \quad (6)$$

(3) 将累加和 (Sum) 与阈值 (Thresh) 进行比较，若 Sum 大于或等于 Thresh，则输出 1 作为腐蚀 / 膨胀结果，否则输出 0。此外，还需判断 3×3 窗口的中心像素是否位于图像边界，若位于边界，则输出 0 作为腐蚀 / 膨胀结果。

根据上述的计算步骤，图 3 是采用 FPGA 的流水线方式设计出腐蚀 / 膨胀算法的计算框图。通过这样的计算，只消耗掉 2 个时钟周期便能实现运算的结果，大大提升了系统的计算效率。

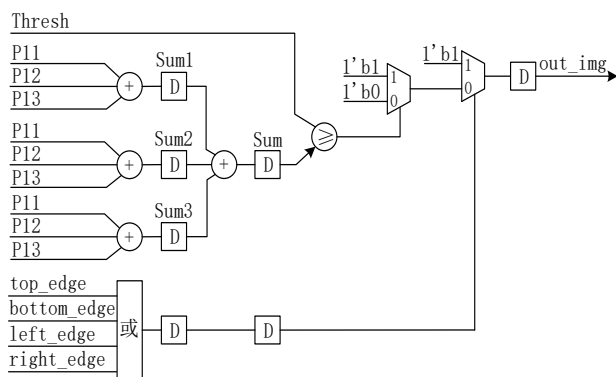


图 3 腐蚀 / 膨胀电路设计模块

2.4 多目标检测与包围盒绘制

2.4.1 距离阈值法的多目标定位

设计多运动目标检测时，关键在于准确区分和定位多个运动目标。经过二值化处理的图像中，每个目标表现为一个连通域，但同一目标可能会被分割成多个较小的连通域。因此，首先需要根据定义设定一个距离阈值 T_2 ，用以衡量各连通域之间的接近程度。通过计算连通域之间的距离，可以确定它们是否属于同一目标。如果两个连通域之间的距离小于设定的阈值，则它们将被归类为同一目标；反之，如果距离大于 T_2 ，则它们被认为是不同的目标。阈值的选取需要根据具体情况进行细致调整，以确保目标检测的准确性和可靠性。如果阈值设定过高，可能会导致本不属于同一目标的连通域被错误地识别为同一目标；而阈值过低，则可能导致同一目标的不同连通域被误判为不同的目标。

各连通域之间的距离关系可分为 3 种情况。图 4 中黑点代表目标位置，方框（包围盒）代表目标的邻域。第一种情况是两个目标的领域不相交，说明是两个独立的运动目标；第二种情况是邻域相交，但都没落到互相的邻域之内，这种情况也视为两个目标；第三种情况便认为是同一运动目标，因为两个目标的邻域重合且都落到了互相的邻域之内。

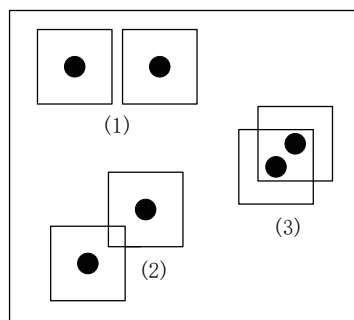


图 4 连通域距离关系图

多个目标区分开后，需要定位多目标所在的位置即确定各个包围盒的坐标。该包围盒是由左上角点 ($X_{\min}, Y_{\min}$) 和右下角点 ($X_{\max}, Y_{\max}$) 的坐标所构成的方框，因此确定这两个坐标点便可定位出各个目标所在的位置。定义 n 组寄存器存储包围盒的坐标，每个包围盒的坐标以 ($X_{\min}[i], Y_{\min}[i]$)、($X_{\max}[i], Y_{\max}[i]$) 表示。由左至右，由上至下遍历每个像素点，当像素点有效时，判断该像素点与某个包围盒的距离是否小于 T_2 ，如果是，那么认为该像素点属于此包围盒，然后再判断该像素点的横坐标是否小于包围盒的最小横坐标，或者大于其横坐标最大值，是的话将包围盒的横坐标更新为当前像素点的横坐标，同理可更新包围盒的纵坐标。包围盒的判断顺序是从小到大的，一旦某个像素点属于某个包围盒，便不会再进行后续包围盒的判断。若该像素点不属于任何一个现有包围盒，则生成一个新的包围盒，并根据像素点的坐标更新包围盒的边界。

2.4.2 多运动目标检测设计

图 5 是多目标检测的框架图，首先将经过形态学处理得到的二值化图像作为输入。随后，该图像被送入 n 个单目标定位模块，每个模块独立识别并记录该目标的包围盒坐标信息（包括 $X_{\min}$ 、 $Y_{\min}$ 、 $X_{\max}$ 、 $Y_{\max}$ ）以及包围盒的有效性。这些单目标定位模块是串联的，确保每个模块的输出成为下一个模块的输入。同时，使用计数器对有效包围盒进行计数，便于多目标定位模块整合出所有有效包围盒的信息，并根据这些坐标信息，使系统在原图像中绘制出每个目标的包围盒。最终，当所有包围盒绘以流水线方式制完成后，系统输出包含目标包围盒的视频流，而每一个包围盒的绘制会消耗一个时钟周期，因此，该模块会延迟 n 个时钟周期。

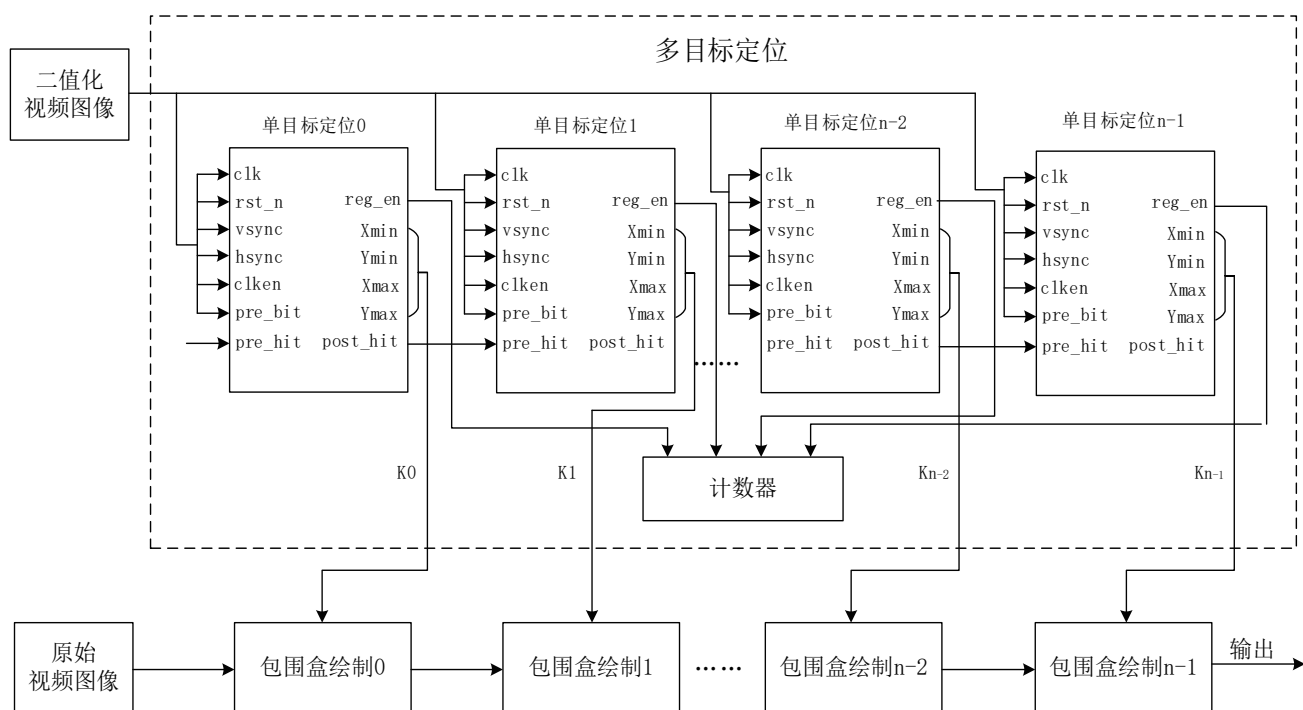


图 5 多运动目标检测框图

3 实验结果验证与分析

3.1 实验平台搭建

为了验证多运动目标检测的效果，整个检测过程采用 Verilog 硬件描述语言进行编程。在 Quartus ii 13.1 软件对算法进行编译、分配管脚以及综合，确保代码的逻辑正确和资源优化。该系统采用型号为 EP4CE10F17C8 的 FPGA 芯片来实现多目标检测的全部算法，其主控芯片时钟频率为 100 MHz，图像分辨率定为 800 px×480 px，刷新率为 60 Hz。

3.2 结果分析

本系统需要设置两个阈值，帧差二值化阈值和多运动目标识别的距离阈值，这两个阈值会影响检测结果，因此需要根据具体情况设置合理的阈值，使结果更准确。实验将图像二值化阈值 T_1 设置为 50，距离阈值 T_2 设置为 80。图 6 和图 7 分别是多目标检测的结果，图中分别检测了 2 个和 5 个运动物体，识别效果都达到 100%，准确性到进一步提升。根据测试，本实验分别增加对运动目标检测的个数，最多可识别 16 个运动目标，当运动物体个数增多、运动速度加快时，需要根据场景合理选取距离阈值，因而若个数超过 7 个时，识别效果会相对有所下降。为了体现系统的实时性，将该算法在 Matlab 上进行对比，算法所花费的时间远高于在 FPGA 上实现的时间。

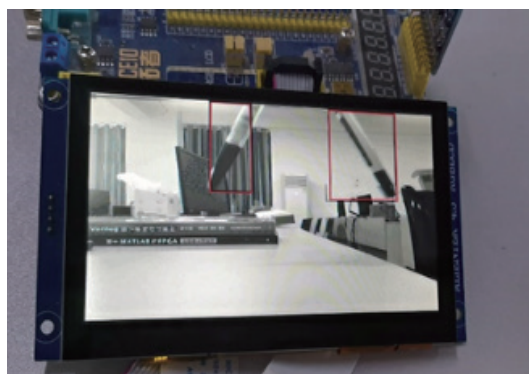


图 6 两个运动目标检测结果



图 7 五个运动目标检测结果

3.3 功耗分析

该芯片共有 10 320 个 LE（logic element）、11 172 个寄存器、180 个可用引脚、423 936 个 memory Bit、2 个 PLL。当检测的目标数增加时，会消耗更多的逻辑资源，但是整个工程所

用的引脚、memory Bit 和 PLL 不会因为目标数的增加而增加。系统硬件资源占用情况如表 1 所示。

表 1 逻辑资源占用情况

检测个数	LEs/%	Total registers
2	37	2644
3	40	2761
4	43	2878
5	46	2995
6	49	3112
7	52	3229

4 结语

本文提出的实时多运动目标检测系统，充分利用了 FPGA 的多项优势。该系统能够在各种复杂环境中，如视频监控、交通管理、工业自动化等领域，快速准确地识别和跟踪多个运动目标。此外，系统还具备高度的可扩展性和灵活性，能够根据实际应用需求进行定制和优化，以满足不同场景下的性能要求。

参考文献：

[1] 刘洋. 全景拼接视角下室内视频监控中的人群目标检测技术研究 [D]. 北京：北方工业大学, 2024.

[2] 方勇, 张建军. 面向智能视频监控中的运动目标检测方法研究 [J]. 计算技术与自动化, 2022, 41(4): 79-83.

[3] 周燕, 许业文, 蒲磊, 等. 自动驾驶场景下的图像三维目标检测研究进展 [J]. 计算机科学, 2024, 51(11): 133-147.

[4] 计大威, 邱洋. 基于单目视觉定位与运动目标检测算法的机器人在竞赛系统中的研究 [J]. 自动化与仪器仪表, 2024(4): 201-205.

[5] 刘思杨, 程方. 一种融合背景差分 and 帧间差分的运动目标检测方法 [J]. 舰船电子工程, 2024, 44(2): 45-48.

[6] SZOLGAY D, BENOIS-PINEAU J, MEGRET R, et al. Detection of moving foreground objects in videos with strong camera motion [J]. Pattern analysis and applications, 2011, 14(5): 311-328.

[7] 黄名政, 李彬华. 动态帧差法及其 FPGA 设计与实现 [J]. 光学技术, 2023, 49(2): 231-237.

[8] 谢有浩, 墨士玉, 姜阔胜. 集成 ROI 的改进背景差分法运动工件定位跟踪 [J]. 滁州学院学报, 2022, 24(5): 49-53.

[9] DITRIA E, JINKS E, CONNOLLY R M. Automating the

analysis of fish grazing behaviour from videos using image classification and optical flow [J]. Animal behaviour, 2021, 177(7): 31-37.

[10] 屈晶晶, 辛云宏. 连续帧间差分与背景差分相融合的运动目标检测方法 [J]. 光子学报, 2014, 43(7): 213-220.

[11] 杨庆江, 冯新宇, 李钊枢, 等. 基于改进帧间差分的动态目标检测算法 [J]. 黑龙江科技大学学报, 2022, 32(6): 779-783.

[12] 田洪金, 殷玲, 刘云. 一种基于背景更新率的运动目标检测算法 [J]. 科技与创新, 2024(2): 171-174.

[13] 冯天桥, 毛征宇, 彭向前. 基于 FPGA 的运动目标实时检测系统设计 [J]. 电子技术应用, 2023, 49(6): 99-103.

[14] 马明哲. 基于 FPGA 的运动目标检测和跟踪算法的研究与实现 [D]. 天津：天津科技大学, 2023.

[15] 邢凯. 基于 FPGA 的多运动目标检测技术研究 [D]. 昆明：昆明理工大学, 2020.

[16] 韩帅. 基于改进帧间差分法与背景差分法车辆检测 [J]. 农业装备与车辆工程, 2021, 59(6): 88-92.

[17] 钱成铎. 基于 ZYNQ 的图像采集与表面缺陷检测系统 [D]. 淮北：淮北师范大学, 2022.

【作者简介】

薛浪 (1999—), 女, 贵州盘州人, 硕士研究生, 研究方向: 图像处理、目标检测, email: 453234593@qq.com.

曾庆立 (1975—), 通信作者 (email: 2878627@qq.com), 男, 湖南永州人, 博士在读, 高级实验师, 研究方向: 机器视觉、嵌入式系统.

钟戴元 (2000—), 男, 湖南怀化人, 硕士研究生, 研究方向: 图像处理, email: 1594299348@qq.com.

周佳凯 (2000—), 男, 湖南永州人, 硕士研究生, 研究方向: 基于 FPGA 的图像处理, email: 2928715787@qq.com.

唐瑞东 (2000—), 男, 湖南娄底人, 硕士研究生, 研究方向: 图像处理, email: 1515871076@qq.com.

(收稿日期: 2024-10-15)