

直方图反向投影多目标检测优化算法^①

张辰^{1,2}, 赵红颖², 钱旭¹

¹(中国矿业大学(北京)机电与信息工程学院, 北京 100083)

²(北京大学遥感与地理信息系统研究所, 北京 100871)

摘要: 针对目标检测中图像背景信息复杂以及反向投影方法目标颜色类型单一的问题, 提出了一种改进的直方图反向投影目标检测优化算法. 在建立多个目标模型弥补一般反向投影方法不能同时检测含有不同颜色信息目标这一不足的同时, 通过对目标概率图像进行一系列优化处理以减小复杂背景对目标检测的干扰. 实验结果表明, 该算法在满足对视频图像实时处理要求的同时可准确检测目标位置, 且检测效果及计算效率优于传统反向投影算法及其他方法.

关键词: 目标检测; 反向投影; 目标概率; 颜色直方图

Optimized Multi-Object Detection Method Based on Histogram Back-Projection

ZHANG Chen^{1,2}, ZHAO Hong-Ying², QIAN Xu¹

¹(School of Mechanical Electronic & Information Engineering, China University of Mining and Technology, Beijing 100083, China)

²(Institute of Remote Sensing and Geographical Information System, Peking University, Beijing 100871, China)

Abstract: Aiming at the complexity of background image in object detection and the singleness of object's color type in common back-projection method, an optimized object detection method based on histogram back-projection was presented. In this algorithm, multiple object models were established to make up the defect that multiple objects with different colors cannot be detected by using common back-projection method. Besides, in order to reduce the interference of complex background, a series of optimizing process steps for the probability image was added. Experimental results show that the proposed algorithm not only can accurately detect the position of objects in real-time, but also with better detection effect and higher operation efficiency than the common back-projection algorithm and other methods.

Key words: object detection; back-projection; object probability; color histogram

计算机视觉跟踪中的目标检测是对目标进行跟踪、分类、行为理解等高级应用的基础, 主要目的是从视频图像中提取出目标并获得目标的颜色、形状、轮廓等特征信息. 较为常见的目标检测方法主要有背景差法^[1]、相邻帧差法^[2]、光流法^[3]等. 其中背景差法利用当前图像与背景图像的差来检测运动区域, 其原理和算法设计简单, 但受光线、天气等外界条件变化的影响较大. 相邻帧差法通过比较图像序列中相邻帧中所有对应位置的像素点, 判断相邻帧之间的像素差是否大于某一阈值以确定目标, 这种方法稳定性较好,

能够适应各种动态环境, 但这种方法只能提取对象边界, 不能对目标的完整区域进行提取. 1980年, B.K.P. Horn 和 B.G.Schunck 提出了光流法的概念^[4], 其原理是根据各个像素点的速度矢量特性对图像进行动态分析, 通过检测运动目标和邻域背景二者速度矢量的差异判断目标位置, 光流法具有广泛的适用性和良好的准确性, 但计算复杂度较高, 运行效率低, 不利于视频图像序列的实时处理. 这些方法有一个共同特点, 即分析的是图像中的动态目标, 而对例如建筑、道路、植被、河流等静态目标的检测往往无法获得理想效果,

^① 收稿时间:2013-07-23;收到修改稿时间:2013-08-27

为此 M. Swain 在 1991 年提出了一种在图像中利用颜色特征检测目标位置的方法, 命名为直方图反向投影法 (Histogram Back-Projection)^[5].

在处理视频图像序列时, 图像背景的色彩信息往往非常复杂, 经常会有目标与背景颜色重叠的现象, 直接利用直方图反向投影检测目标的效果并不理想. 针对这些问题, 本文提出的算法主要在如下几方面做出优化: (1) 通过对目标概率图像进行优化处理以减小复杂背景对目标的干扰; (2) 建立图像金字塔提高运算效率以达到对视频图像的实时处理要求; (3) 通过建立多个目标颜色直方图模型, 弥补一般反向投影方法不能同时检测多个含有不同颜色信息目标的不足.

1 直方图反向投影计算原理

1.1 反向投影方法概述

反向投影是一种通过给定目标的像素分布直方图模型计算图像中目标概率的方式, 即在图像中利用目标颜色直方图模型寻找目标. 为了获得反向投影图像, 首先要计算目标颜色直方图 M 和待检测图像颜色直方图 I 的比例 ξ , 即图像中像素与目标模型的相似度, 通过这一结果得到目标概率图像, 图像中明亮部分即为检测出的目标位置. 利用反向投影方法对原图像计算得到的目标概率图像图像如图 1 所示.

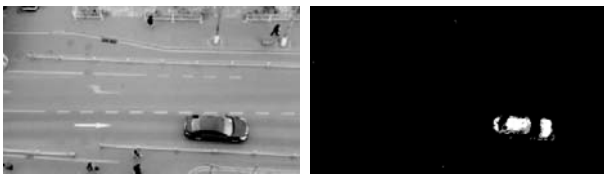


图 1 目标概率图像

1.2 目标模型建立

建立目标模型是计算反向投影图像的核心步骤, 也是整个算法的基础. 在理想条件下, 反向投影方法中目标与背景的颜色频带不能重叠, 因此为了提高目标检测的准确性, 应选择目标中面积较大且与背景颜色不重叠的颜色信息作为模型. 图 2 为目标颜色的 H-S 直方图模型 (H 代表色度, S 代表饱和度).

计算目标直方图模型首先要将目标图像由 RGB 图像转化为 HSV 图像^[6]. HSV 色彩模型可描述为由色调 (Hue)、饱和度 (Saturation)、亮度 (Value) 这三个参数所建立的圆柱坐标系 (如图 3 所示), 这个坐标系中的任

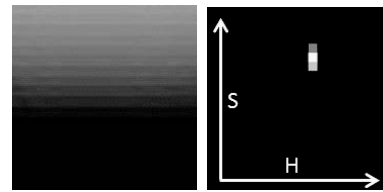


图 2 直方图反向投影目标模型



图 3 HSV 色彩模型

意坐标即为某一颜色的值. 已知某颜色在 RGB 色彩模型中的坐标为 (r, g, b) , 可由以下公式计算该颜色在 HSV 模型中相应的坐标 (h, s, v) :

$$v = \max(r, g, b) \quad (1)$$

$$s = \begin{cases} 0 & v = 0 \\ 1 - \frac{\min(r, g, b)}{v} & v \neq 0 \end{cases} \quad (2)$$

$$h = \begin{cases} \text{undefined} & s = 0 \\ 60^\circ \times \frac{g - b}{v - \min(r, g, b)} & v = r \\ 60^\circ \times \frac{b - r}{v - \min(r, g, b)} + 120^\circ & v = g \\ 60^\circ \times \frac{r - b}{v - \min(r, g, b)} + 240^\circ & v = b \end{cases} \quad (3)$$

1.3 目标概率的计算

建立目标直方图模型后, 要利用该模型计算视频图像序列中每一帧图像的目标概率, 计算过程中同样需要先将原始的 RGB 图像转化为 HSV 图像, 并分割为单通道并建立单通道图像数组, 再利用已有的目标模型判断当前帧中的每一个像素是否为已知目标的一部分. 计算后得到的目标概率图像删除了目标模型以外的颜色信息, 只保留了原图像直方图与目标模型直方图重叠部分的颜色信息, 即原图像中和目标颜色一致的部分.

在待检测图像颜色分布的频带范围大于目标颜色这一条件成立的情况下, 文献[7]中将目标概率用以下

公式表示:

$$\zeta(c) = \frac{M(c)}{I(c)} \tag{4}$$

$$BPI(c) = 255 \times \zeta(c) \tag{5}$$

其中, $M(c)$ 为颜色 c 在目标直方图中的值, $I(c)$ 为颜色 c 在待检测图像直方图中的值, $\zeta(c)$ 即为待检测图像中颜色 c 目标概率. $BPI(c)$ 为颜色 c 在反向投影图像中的灰度值, $BPI(c)$ 越大颜色 c 在反向投影中相应位置出现的部分越明亮, 目标出现的概率也就越高.

2 直方图反向投影优化算法

2.1 算法流程

在目标颜色与背景颜色频带图不发生重叠的理想状态下, 通过反向投影方法得到目标概率图像后即完成了对目标的检测. 但在实际情况中, 人们所处理的视频图像序列中通常包含许多较为复杂的色彩信息, 因此会出现目标与背景颜色重叠、目标尺寸变化及噪声干扰等现象, 这些影响因素导致了直接利用基本反向投影方法检测目标的效果往往并不理想.

本文所设计的基于反向投影的目标检测算法流程如图 4 所示. 首先建立一个或多个目标的 H-S 直方图模型, 随后输入待检测的视频图像序列. 为了提高运算效率, 本算法通过对原图像建立高斯图像金字塔, 在保留图像有效信息的情况下大幅减小了计算量, 随后利用目标模型计算出反映当前帧中目标出现概率的反向投影图像, 并对图像进行一系列优化处理以达到消除噪声同时突出目标的目的, 最后输出目标检测结果并载入下一帧图像.

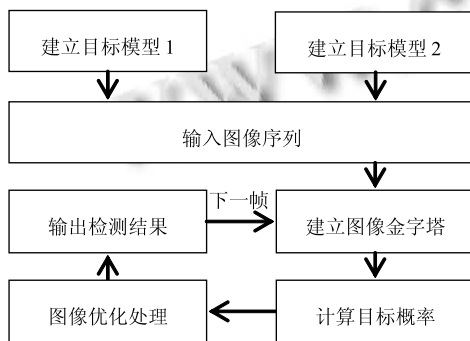


图 4 基于反向投影的目标检测算法流程图

2.2 建立图像金字塔

随着近年来高清技术的普及, 视频图像的主流分

辨率已由过去的 QVGA(320*240)和 VGA(640*480)级别提升到 720P(1280*720)或 1080P(1920*1080)高清级别, 分辨率的提升无疑加大了图像处理的计算量. 在对高分辨率图像数据进行处理时, 通过对原始图像建立高斯金字塔^[8], 可大幅减少后期处理的计算量, 从而有效提高运算效率.

图像金字塔主要用于机器视觉和图像压缩, 一幅图像的金字塔是一系列以金字塔形状排列的分辨率逐步降低的图像集合, 金字塔的底部是待处理图像的高分辨率表示, 顶部是低分辨率的近似. 当向金字塔上层移动时, 图像尺寸和分辨率就降低. 如图 5 所示, 若从金字塔第 i 层(G_i)生成第 $i+1$ 层(G_{i+1}), 需用高斯核对 G_i 进行卷积, 删除偶数行和偶数列, 得到的 G_{i+1} 面积是 G_i 的四分之一, 进行循环操作即可产生整个金字塔. 对大小为 $M*N$ 的 G_0 建立高斯金字塔, 计算过程可由公式(6)表示:

$$G_{i+1}(x+y) = \sum_{m=-p}^p \sum_{n=-p}^p w(m,n)G_i(2x+m,2y+n) \tag{6}$$

其中 $0 \leq x < M/2t, 0 \leq y < N/2t, 0 \leq i \leq t$ (t 为金字塔层数), m 和 n 的取值范围是 $[-p, p]$, 高斯核 w 的大小为 $(2p+1)*(2p+1)$.

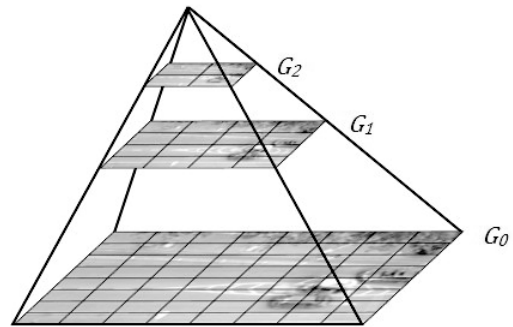


图 5 图像金字塔的建立

2.3 反向投影图像优化处理

在实际应用中, 由于背景噪声的干扰以及目标本身颜色信息的复杂性, 根据计算后得到的目标概率直接检测目标位置效果往往并不理想, 因此需通过对反向投影图像进行一系列形态学优化处理以获得更好的检测效果^[9,10]. 同时, 对同一图像建立多个目标模型, 分别计算不同颜色目标的目标概率并进行叠加, 可弥补传统单一模型反向投影方法无法检测不同颜色目标

的不足. 文中算法对反向投影图像优化处理的步骤及效果如图 6 所示.

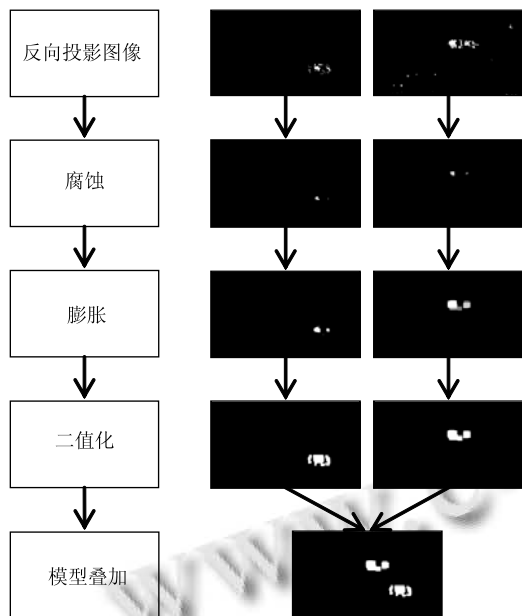


图 6 反向投影图像优化处理步骤

对图像进行腐蚀-膨胀操作, 可在有效去除图像背景中的噪声干扰的同时突出目标. 文献[11]中将腐蚀操作定义为公式(7):

$$A \ominus B = \{z | (B)_z \subseteq A\} \quad (7)$$

其中 A 和 B 是二维整数空间 Z^2 的集合, 核 B (图 7(a)) 通常是一个中间带有参考点的实心正方形或圆盘^[12]. 通过这一公式可将 B 对 A 的腐蚀理解为 A 中所有点平移 B 后得到的点 z 集合, 如图 7(b)所示. 同理, A 被 B 膨胀由公式(8)定义:

$$A \oplus B = \{z | [(\hat{B})_z \cap A] \subseteq A\} \quad (8)$$

A 中所有点平移 \hat{B} 后得到的点 z 的集合, 其中 \hat{B} 为 B 的反射, 如图 7(c)所示. 随后对腐蚀-膨胀后的反向投影图像进行二值化分割, 得到最终的目标区域.

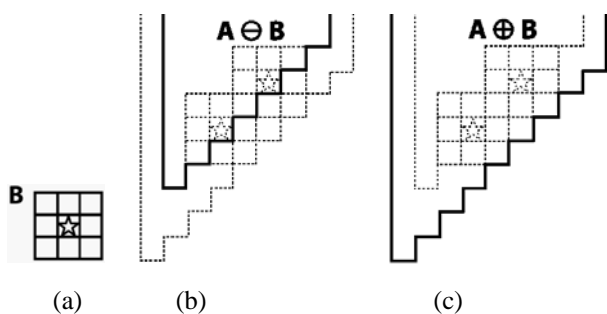


图 7 形态腐蚀与膨胀示意图

3 实验结果与分析

以上算法基于 Visual Studio 2010+OpenCV 2.4.4 平台实现, 数据为一段每秒 30 帧分辨率为 1280*720 的 MP4 格式视频, 对视频中含有两种不同颜色信息的车辆目标进行检测, 同时与背景差法、帧间差法、单一颜色模型的传统反向投影算法的检测效果及运算效率进行了比较.

实验中的背景差法和相邻帧差法在文献[2]中有较为详细的论述. 背景差法的原理是利用当前图像与背景图像的差来检测运动区域, 可由下式表示:

$$d = |I_L(x, y, i) - B_L(x, y)| \quad (9)$$

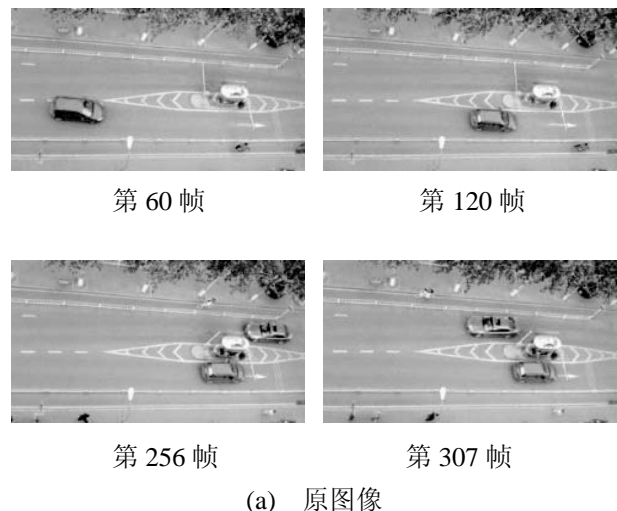
$$ID_L(x, y, i) = \begin{cases} d & d \geq T \\ 0 & d < T \end{cases} \quad (10)$$

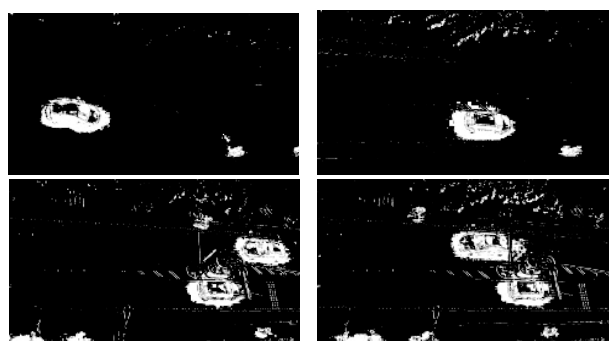
其中, ID_L 是背景帧差图, B_L 是背景的高度分量, i 表示帧数($i=1, 2, \dots, N$), N 为图像序列总帧数, T 为阈值. 帧间差的原理是计算相邻帧对应像素点的差值, 如果大于一定阈值则认为该点对应位置有运动目标存在, 由下式表示:

$$d = |I_L(x, y, i) - I_L(x, y, i-1)| \quad (11)$$

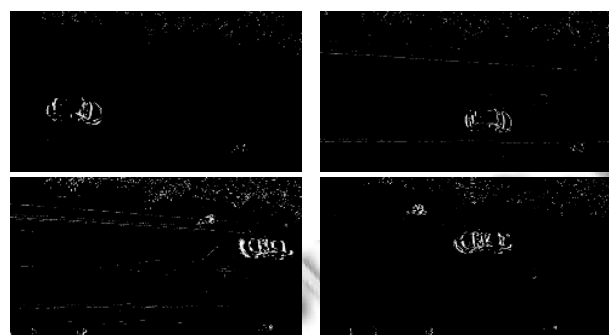
$$ID_L(x, y, i) = \begin{cases} d & d \geq T \\ 0 & d < T \end{cases} \quad (12)$$

其中, ID_L 是相邻帧差图像, $I_L(x, y, i)$ 和 $I_L(x, y, i-1)$ 分别是第 i 帧和第 $i-1$ 帧的高度分量, i 表示帧数($i=1, 2, \dots, N$), N 为图像序列总帧数, T 为阈值. 对于传统直方图反向投影算法, 仅对单一颜色模型的目标概率图像进行了二值化处理. 图 8 中显示了不同检测方法的在视频中各帧的检测效果, 表 1 中对比了不同方法处理单帧图像所花费的运算时长及处理整段视频后单帧图像所花费的平均时长.

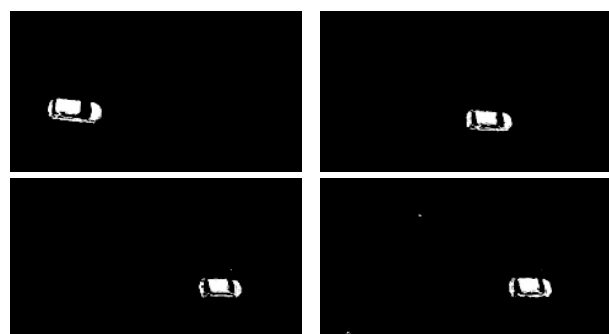




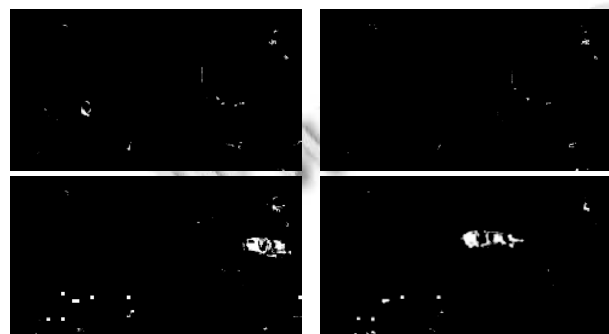
(b) 背景差法



(c) 帧间差法



(d) 反向投影(目标 1)



(e) 反向投影(目标 2)



(f) 本文方法

图 8 不同目标检测方法效果对比

表 1 不同目标检测方法运算效率对比(单位: ms)

方法	第 60 帧	第 120 帧	第 256 帧	第 307 帧	平均
背景差	21.98	28.86	20.95	29.86	25.23
帧间差	15.68	14.94	15.99	15.20	15.75
反向投影(目标 1)	31.98	30.99	30.92	32.10	31.25
反向投影(目标 2)	31.02	30.99	31.01	31.00	31.29
本文方法	14.98	15.97	15.95	15.00	15.60

由实验结果可以看出,相对于背景差法、帧间差法以及传统的直方图反向投影算法,本文算法可在准确检测目标的同时有效去除复杂背景对目标的干扰,具有更好的检测效果.在运算效率方面,改进后的反向投影优化算法通过建立图像金字塔大幅减少了计算量,即使叠加了多个目标模型,运算速度仍优于传统的反向投影算法及其他方法.

4 结语

本文针对视频图像背景信息复杂以及传统反向投影方法目标颜色类型单一的问题,提出一种基于直方图反向投影的目标检测优化算法.实验结果表明,当图像背景信息较为复杂时,对反向投影图像进行腐蚀-膨胀-二值化操作可达到消除背景噪声同时突出目标的目的;若图像尺寸较大,可通过建立图像金字塔有效提高运算速度.此外,对视频图像中多个目标进行实时检测时,应根据视频中的目标特点对目标模型进行优化,避免建立过多的目标模型,以满足实时处理的要求.

参考文献

- 1 刘亚,艾海舟,徐光佑等.一种基于背景模型的运动目标检测与跟踪算法.信息与控制,2002,31(4):315-319,328.
- 2 李刚,邱尚斌,林凌等.基于背景差法和帧间差法的运动目标检测方法.仪器仪表学报,2006,27(8):961-964.
- 3 严强,黄增喜,曹丽萍等.光流法在机车安全行驶中的应用.计算机应用研究,2013,30(4):1219-1222,1233.

(下转第 218 页)

调试技术的实时性要求;

③ 简单实用, 主机端能够在 Matlab、Excel 中使用 COM 接口对接受到的 DSP 数据进行分析处理。

4 结语

本文针对设置断点不能有效地对实时嵌入式系统软件进行代码调试的问题, 从分析 RTDX 技术的原理入手, 提出了利用 RTDX 开展了弹载软件代码调试的技术途径。实验结果表明, 利用 RTDX 技术能够正确地将调试所需要的数据从 DSP 传输至主机进行分析, 其传输时间和传输速率也能够满足当前导引头信息处理软件的实时性要求, 为开展实时嵌入式软件代码调试提供了新的技术途径。

此外, RTDX 技术支持主机客户端发送数据至 DSP 软件, 这为下一阶段利用 RTDX 开展弹载 DSP 软件故障注入、建立闭环的弹载 DSP 仿真验证环境等工作提供了研究方向。

参考文献

1 Texas Instruments Incorporated. 彭启琮, 张诗雅, 常冉等译。

(上接第 182 页)

- 4 Horn BKP, Schunck BG. Determining optical flow. *Artificial Intelligence*, 1981, 17(1): 185-203.
- 5 Swain M, Ballard D. Color indexing. *International Journal of Computer Vision*, 1991, 7(1): 11-32.
- 6 杨康叶, 邬春学. 基于 RGB 模型颜色相似性的彩色图像分割. *计算机系统应用*, 2013, 22(3): 128-131.
- 7 Lee J, Lee W, Jeong W. Object tracking method using back-projection of multiple color histogram models. *Proc. of ISCAS' 03*, 2003, 2(2): 668-671.
- 8 Adelson EH, Anderson CH, Bergen JR, et al. Pyramid methods in image processing. *RCA engineer*, 1984,

TI DSP 集成化开发环境使用手册. 北京: 清华大学出版, 2005: 184-185.

- 2 彭启琮, 管庆等. DSP 集成开发环境. 北京: 电子工业出版社, 2004: 163-164.
- 3 刘红刚, 尹岗. RTDX 在运动数字铭牌识别中的应用. *自动化应用*, 2011, (2): 3-4.
- 4 顾长怡, 黄佩伟. 利用 RTDX 技术实现计算机实时语音分析. *计算机工程*, 2001, (2): 155-156.
- 5 胡敏露, 徐刚锋, 李飏. RTDX 在图像处理中的应用. *国外电子元器件*, 2003, (5): 21-24.
- 6 倪骏, 徐向东. 运用 RTDX 技术的 TMS320C6201 的软件验证方法. *空军雷达学院学报*, 2002, (4): 40-42.
- 7 许海丽, 张茂青, 严震宇, 沈芳. 基于 MATLAB 的 DSP2812 RTDX 实时绘制电机转速曲线. *苏州大学学报(工科版)*, 2013, (5): 47-50.
- 8 张传清, 安钢, 樊新海. RTDX 与 Matlab 实现基于 DSP 的 FIR 滤波器设计. *装甲兵工程学院学报*, 2001, (3): 64-69.

29(6): 33-41.

- 9 熊哲源, 樊晓平, 黎燕等. 基于数学形态学边缘检测的车牌字符分割算法. *计算机系统应用*, 2010, 19(9): 155-158.
- 10 毛若羽, 陈相宁. 改进的 HSI 空间形态学有噪彩色图像边缘检测. *计算机应用研究*, 2013, 30(2): 635-637.
- 11 Gonzalez RC, Woods RE. *Digital image processing (International ed.)*. 2008.
- 12 Bradski GR, Kaehler A. *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, Incorporated, 2008: 115-118.