

基于混合高斯模型的背景差分法的 FPGA 实现

韩剑辉, 崔猛强, 袁耀辉

(哈尔滨理工大学 计算机科学与技术学院, 黑龙江 哈尔滨 150080)

摘要: 针对视频中运动目标检测实时性问题, 提出了一种基于混合高斯模型的背景差分法的 FPGA 实现方法. 背景差分算法是运动检测算法中最常用的方法, 该算法首先对实时得到的图像序列进行建模, 得到运动场景的背景模型. 然后在运动目标分割中及时的根据场景变化更新背景. 考虑到实时性要求的高帧率、高分辨率视频分割任务, 利用 FPGA 并行处理和硬件实现的优点, 对混合高斯模型进行修改来对它进行定制, 以适合于硬件实现. 最后, 通过在 FPGA 开发平台进行了仿真实验, 结果验证了算法在硬件实现的正确性, 以及在实时性能上的提高.

关键词: 运动检测; 背景差分; 混合高斯模型; FPGA

中图分类号: TP399 **文献标志码:** A **文章编号:** 1007-2683(2013)04-0089-05

FPGA Implementation for Background Subtraction Based on Gaussian Mixture Models

HAN Jian-hui, CUI Meng-qiang, YUAN Yao-hui

(School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China)

Abstract: Aiming at the real-time problem of moving object detection in video, this paper proposed a FPGA realization method of background subtraction algorithm based on Gaussian mixture model. Firstly, the algorithm model is built for image sequence, and the background model of moving scene is obtained. Then the background model is updated timely according to scene change in moving object segmentation. Considering the real-time requirements of high frame rate and high resolution, the paper modified the Gaussian mixture model for FPGA hardware implementation. The test results show that the algorithm is correct in the hardware implementation, and the real-time performance is improved.

Key words: motion detection; background subtraction; Gaussian mixture models; FPGA

0 引言

在现实生活中, 运动图像包含有大量有意义的

视觉信息. 运动目标检测技术作为计算机视觉、运动图像分析、人机交互和智能视频监控技术的基础已广泛应用于军事、计算机辅助设计、航空航天、智能机器人等领域, 其检测结果的好坏直接影响到目标

收稿日期: 2012-11-16

基金项目: 2012 年国家大学生创新创业训练计划项目(201210214019)

作者简介: 韩剑辉(1962—), 男, 研究员, 硕士生导师, E-mail: hanjianhui@hrbust.edu.cn;

崔猛强(1987—), 男, 硕士研究生;

袁耀辉(1990—), 男, 学士.

定位、跟踪以及行为理解等后续处理^[1,9]。运动目标检测的图像差分法一般可以分为两种:帧间差法与背景差法。其中背景差法能够快速、准确地分割出运动物体,但容易受环境和光照的变化影响,而且现实生活中的场景具有复杂性和不可控的变化,主要表现在如下几种情况:①背景中静止物体移动。当背景中原先静止的物体开始运动,或者场景中的静止物体被移到另一位置保持静止,都会产生虚影(ghost)造成误检;②运动物体停在场景中;③场景的光照变化,包括自然光照的渐渐变化,天气阴晴突然变化以及室内的开关灯等;④混乱场景中有摆动的树枝树叶、闪烁的显示器、波动的水面及喷泉等。为了解决上述问题,提出了很多建立背景模型的方法^[2]。本文针对混合高斯模型进行背景差分,为了提高实时性和处理速度,对高斯模型进行修改使其适合FPGA硬件实现。

1 背景差分法

1.1 背景差分法基本原理

背景差法(background subtraction)是目前运动检测中的主流方法,它是利用含有运动目标的当前图像与背景图像的差分来检测运动区域的一种技术。

背景减法基本思想是:首先,用事先存储或者实时得到的图像序列对每个像素进行统计建模,得到背景模型 $D_k(x, y)$;其次,将当前每一帧图像 $f_k(x, y)$ 和背景模型 $B_k(x, y)$ 相减,其中 $D_k(x, y)$ 为作差后的结果;计算出在一定阈值 T 限制下当前图像中出现的偏离背景模型值较大的那些像素,则判定此像素为运动目标上的像素点,相减的阈值操作后得到的结果直接给出了目标的位置、大小、形状等,从而得到较完整的目标信息。

$$D_k(x, y) = |f_k(x, y) - B_k(x, y)| \quad (1)$$

$$D_k(x, y) = \begin{cases} 1 & D_k(x, y) > T \\ 0 & \text{else} \end{cases} \quad (2)$$

1.2 背景建模

背景建模是背景差分最关键的步骤,背景模型的选择决定后面的步骤采用的方法和具体实现,所以,很多背景差分的分类方法就是根据背景模型来区分。不仅如此,背景模型的选择也决定背景差分算法的好坏。背景模型最简单的就是使用和每帧图像同样大小图像,在这种情况下,后面的背景差分步骤就和上面的帧间差方法一样,把当前真与背景模

型图像差分即可。总之,背景模型选择的原则是要能较好的描述和表示背景的特点,并且此特点能用来区分背景和一前景,而且背景模型越简单越好,因为背景模型在整个过程算法运行过程中需要维护,简单的模型有利于保证算法的效率和实时性。但是在现实的环境中,背景的场景不可能一成不变,背景会随着光照变化、运动物体停在背景中或者背景中固定物体移除等特殊情况的变化而变化。考虑到这些问题,一个自适应的背景建立成为研究背景差分法的关键。因此,近些年,提出许多背景建模方法,大致可分为:基本背景模型(basic background modeling)、基于统计学的背景模型(statistical background modeling)、模糊背景模型(fuzzy background modeling)、背景估计模型(background estimation)等^[3]。

2 混合高斯模型分析

Stauffer等提出的经典的自适应混合高斯背景提取模型,这种方法是基于像素的背景建模方法,它是根据视频中的每个像素在时间域上的分布情况来构建各个像素的颜色分布模型,以此达到对背景进行建模的目的。混合高斯背景模型是有限个高斯函数的加权和,它能够描述像素的多峰状态,适合于对光照渐变、树木摇摆等复杂背景进行准确建模^[4]。

2.1 混合高斯模型建模

假设图像中的每一点的像素观测值和其他像素点的观测值相互独立,在某时刻 t ,设像素点的观测值为 X_t ,则时刻 t 观测值为 X_t 的概率可以用具有 K 个高斯分布的混合高斯模型建模^[5]:

$$P(X_t) = \sum_{i=1}^k \omega_i \varphi(X_t | \mu_i, \Sigma_i) \quad (3)$$

$$\varphi(X_t | \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{n/2} |\Sigma_i|^{1/2}} \times \exp\left\{-\frac{1}{2}(X_t - \mu_i)^T \Sigma_i^{-1} (X_t - \mu_i)\right\} \quad (4)$$

其中: K 是高斯分布的个数,一般取3至5个; ω_i 是第 i 个高斯分布的权重,也就是用该模型将下一帧出现的先验概率,并且 $\sum_{i=1}^k \omega_i = 1$; μ_i 是时刻 t 第 i 个高斯分布的均值向量; Σ_i 是其相应的协方差矩阵; φ 是高斯分布的概率密度函数。这里假设 X_t 互不相关,所以协方差就是一个对焦矩阵,可以用一个 n

维的方差 $\sigma_{i,t}^2$ 表示. 假设图像中某像素点, 在 $1 \sim t$ 时刻的状态为 $\{X_1, X_2, \dots, X_t\}$, 权重 $\omega_{i,t}$ 反应了像素点从 $1 \sim t$ 时间段内满足第 i 个高斯分布的后验概率 $\{X_1, X_2, \dots, X_t\}$ 中属于第 i 个高斯分布的观测样本越多, 权重 $\omega_{i,t}$ 就越大, 方差 $\sigma_{i,t}^2$ 越小.

假设每一个像素点出现时背景的概率高于前景 K 个状态按照 $\omega_{i,t}/\sigma_{i,t}$ 的比例值降序排列. 这个比例值是相应状态的稳定性的指示. 当 $\omega_{i,t}/\sigma_{i,t}$ 变大, 表示一个状态收敛或者变得稳定, 说明这些点在 $1 \sim t$ 时间段内出现的概率较大, 因此这些分布可以作为背景的描述. 前 B 个状态出现的先验概率的和大于一个门限 T 时 ($T \in (0, 1)$) 为

$$B = \arg \min_b \left(\sum_{k=1}^b \omega_{i,t} > T \right) \quad (5)$$

将被标一记为背景, 其余的标记为前景. 因此, 在大部分时间前景都是在场景中占较少的部分的应用中, GMM 模型将表现的更好.

2.2 模型更新

当有新的观测点 X_{t+1} 来的像素值分别与 K 斯分布的均值 $\omega_{i,t}$ 相比, 同时计算观测点落入相应高斯分布的概率, 并按某一判断法则选择匹配的高斯分布^[6], 法则为

$$\rho = \frac{|X_{t+1} - u_{i,t}|}{\sigma} \leq I = 2.5 \quad (6)$$

据此, 就可以选择出符合判断法则的高斯分布. 当存在匹配的高斯分布时, 则需要根据当前像素 X_{t+1} , 对这些高斯分布的权值、均值和方差参数进行更新处理.

$$\omega_{i,t} = (1 - \alpha) \omega_{i,t} + \alpha M(t) \quad (7)$$

$$M(t) = \begin{cases} 1 & \text{匹配} \\ 0 & \text{不匹配} \end{cases} \quad (8)$$

$$u_{i,t+1} = (1 - \beta) u_{i,t} + \beta X_{t+1} \quad (9)$$

$$\sigma_{i,t+1}^2 = (1 - \beta) \sigma_{i,t}^2 + \beta (X_{t+1} - u_{i,t+1})^T (X_{t+1} - u_{i,t+1}) \quad (10)$$

$$\beta = \alpha \times \frac{1}{\sqrt{2\pi\sigma_{i,t}^2}} \times e^{-\frac{(X_t - u_{i,t})^2}{2\sigma_{i,t}^2}} \quad (11)$$

其中 α 是学习速率, 可以由用户指定, β 是分布参数的更新速率. 在这些匹配的高斯分布中, 还需要按照各自的 $\omega_{i,t}/\sigma_{i,t}$ 比值进行排序, 然后从中选择最能代表背景的高斯分布, 并以此来最终确定表征背景的高斯分布.

2.3 混合高斯模型优点

混合高斯模型算法与其他目标检测分割算法相比, 混合高斯模型有一下几方面的优点: 首先, 混合

高斯模型能够适应快速的光照变化. 根据学习速率 α , 可以调整对光照变化的响应速率; 与固定背景模型相比, 混合模型的优越之处在于其可以及时更新背景, 从而大幅减小了错误检测的概率. 其次, 混合高斯模型算法通过用多个高斯分布对背景进行建模, 典型的混合高斯模型中高斯分布的数量在 $3 \sim 6$ 个之间. 假设有 3 个高斯分布用于高斯模型建模, 其中 2 个用于背景建模, 1 个用于前景建模, 这意味着背景建模中有一个高斯分布可以描述背景中反复运动的背景物体(如树叶摇摆)为背景而不被误判为前景运动物体. 最后, 混合高斯模型算法可以有效的减少由于摄像头抖动而造成的误判.

3 混合高斯模型算法的 FPGA 实现

FPGA 是一种基于查找表结构的可编程逻辑器件, 具有大量的灵活的逻辑单元, 很好的在线可编程性和高速并行的结构特点. 它可以方便地实现组合逻辑和时序逻辑功能^[8]. FPGA 较小的基本逻辑单元(LC), 可以通过级联实现较长的数据通路, 适合实现流水线结构. 本文中, 选择 Altera 公司的 Cyclone II 系列 FPGA 作为处理平台. 在硬件上实现算法在提高实时性能的同时对算法的复杂性也会有所要求, 相对软件平台实现的复杂算法在硬件平台需要消耗大量硬件资源, 有些甚至无法实现, 下面对目前文献中提到的各种算法进行对比^[7], 如下表所示:

表1 比较不同的分割算法

| | FD | Median | LPF | GMM | KDE |
|------|-----|------------|--------|----------------|------------------|
| 算法性能 | 快 | 快 | 中 | 中 | 低 |
| 内存需求 | 1 帧 | 50 ~ 300 帧 | 1 帧的平均 | 1 帧的 K 个高斯参数 | N 帧中 K 个高斯参数 |
| 分割质量 | 差 | 低 | 可接受 | 好 | 很好 |
| 复杂度 | 非常低 | 中 | 低 | 低 | 高 |

由表1可以看出, 综合考虑各种算法, GMM 最符合硬件实现, 但是直接在 FPGA 上实现混合高斯模型算法的话, 需要很长的设计周期和消耗大量的硬件资源, 使其综合性能大大削减. 因此, 在这一部分将对此算法进行修改降低其复杂性, 适宜于硬件实现, 权衡了算法在实际运用中的复杂性和精确度.

3.1 混合高斯模型算法的修改

该算法在考虑合理的设计周期和硬件资源消耗

的前提下在硬件实现,对原始的算法进行修改时必不可缺的步骤.但是,修改后的算法应该几乎和原始算法具有相同的性能的同时还必须适宜硬件实现.本文中对算法的修改主要有两点如下:

1) 对分布的更新参数(在式(11)中定义的 β)设置为常数,减少了硬件计算复杂度从而减少硬件资源消耗.

2) 为了保证算法的精确度,对于原始算法其他的计算部分保持不变.但是,为了提高硬件实现的效率,需要对计算的操作进行修改.比如:在式(6)中,把除法运算改为乘法和加法运算来实现,从而节省了设计周期和硬件资源.

对于本文中混合高斯模型算法的修改,在适合硬件实现的基础上更加接近原始算法的精确度.

3.2 算法的硬件实现

此算法模块基于FPGA硬件实现总共分为控制匹配模块、参数更新和排序模块、参数存储模块和前/背景分割模块等几个模块.总体系统结构框图如图1所示.从图中可以看出,在整个算法结构中,参数更新模块是计算量最大的模块,在典型的软件实现混合高斯模型的算法中,由于计算量大很难符合实时性要求.此处利用FPGA的并行处理和吞吐量大的特点,完全可以达到实时性要求.本文中,采用3个高斯模型,主要考虑到树叶、亮度等情况,参数更新和排序模块的部分verilog代码如下所示:

```
always@(posedge clk) begin
    if( match) begin
        if( r == 1) w_out[1] = w[1] + B * (1 - w[1]);
        else w_out[1] = w[1] - A * w[1];
        if( r == 2) w_out[2] = w[2] + B * (1 - w[2]);
        else w_out[2] = w[2] - A * w[2];
        if( r == 3) w_out[3] = w[3] + B * (1 - w[3]);
        else w_out[1] = w[3] - A * w[3];
    end
    else begin
        w_out[1] = w[1];
        w_out[2] = w[2];
        w_out[3] = w[3];
    end
end
```

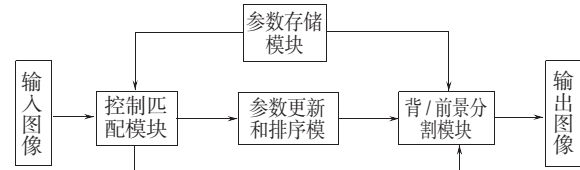


图1 系统框图

通过 Synplify Pro 工具综合后,参数更新和排序模块的RTL视图实现原理如图2所示.

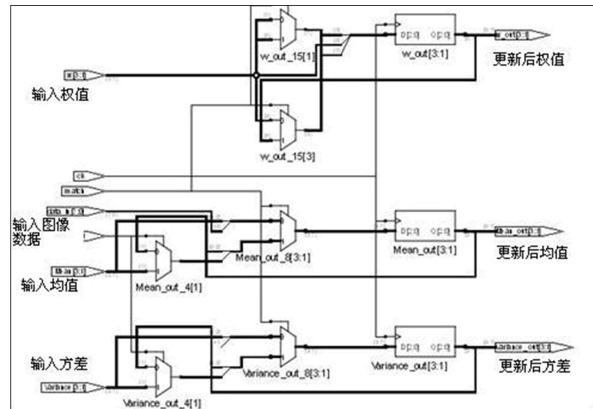


图2 参数更新和排序模块综合RTL视图

3.3 基于FPGA的仿真结果

本文中,对视频序列的第36帧和112帧的图像进行处理,采用Quartus II和Modelsim 6.5b联合Matlab进行仿真.首先采用Matlab将视频帧原始图片转换mif文件作为Quartus II的输入,经过Quartus II仿真后得到仿真网表,将该网表作为Modelsim 6.5b的输入,最终通过将Modelsim的仿真输出结果输出到txt文件,然后再通过Matlab读取文件如图3所示.

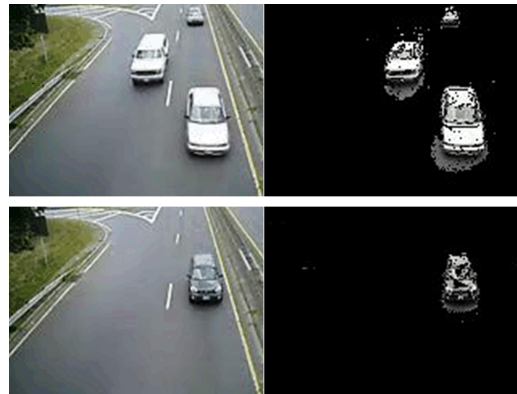


图3 仿真结果

从仿真结果中可以看出,在精度方面,基于FPGA实现的混合高斯模型算法很好的分割了运动物体与背景图像,有效的解决了树叶摆动和光照变化

对运动目标的干扰,保证了原算法的精确度.在速度方面,时钟频率为 25 MHz 时可达到 30fps,完全满足了实时性要求,适合高速运动物体的检测.

4 结 语

运动目标检测是目标跟踪和识别等后期处理的基础,也是研究成败的关键.本文将典型的基于混合高斯模型的背景差分方法利用 FPGA 实现,对于运动目标检测效果显著,而且有效提高了算法的速度,增强了在实时性方面的性能.同时,本文研究过程中也存在一定缺点,比如对阴影的影响^[10],但是在固定场景或者光照弱的情况下足以满足要求,在后续的研究中将在这方面进行改善.

参 考 文 献:

- [1] 张鹤. 运动目标检测算法研究[D]. 武汉: 武汉科技大学, 2011.
- [2] 李海霞, 范红. 基于背景差法的几种背景建模方法的研究[J]. 工业控制计算机, 2012, 25(7): 62-64.
- [3] BOUWMANS T, ELBAF. Background Modeling Using Mixture of Gaussians for Foreground Detection-A Survey[J]. Recent Patents on Computer Science, 2008: 219-237.
- [4] 吴晓军, 宋展. 基于高斯混合模型的视频背景建模技术初探[J]. 先进技术研究通报, 2010, 4(7): 75-79.
- [5] CHRIS Stauffer, GRIMSON W E L. Adaptive Background Mixture Models for Real-time Tracking[J]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1999, 2: 246-252.
- [6] 朱碧婷, 郑世宝. 基于高斯混合模型的空间域背景分离法及阴影消除法[J]. 中国图象图形学报, 2008, 13(10): 1906-1909.
- [7] JIANG H, ARDO H, OWALL V. A Hardware Architecture for Real-time Video Segmentation Utilizing Memory Reduction Techniques[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009: 226-236.
- [8] 潘松, 黄继业. EDA 技术实用教程[M]. 北京: 科学出版社, 2005.
- [9] 王洋, 曾雪琴, 范剑英. 汽车牌照字符识别系统设计[J]. 哈尔滨理工大学学报, 2012, 7(1): 90-95.
- [10] 刘雪, 常发亮, 王华杰. 运动目标检测中的阴影去除方法[J]. 微处理机, 2008(5): 116-117.
- [11] 徐柱. 运动目标检测与阴影消除的研究[D]. 镇江: 江苏大学, 2010.

(编辑: 王 萍)