

视频图像中运动目标检测算法的提高^①

毛育佼, 陈苒君, 唐莉萍

(东华大学 信息科学与技术学院, 上海 201620)

摘要: 基于士兵对抗射击系统, 研究复杂环境下运动目标检测算法的提高, 从提高算法处理效率和降低外界干扰对算法准确性影响两方面进行讨论. 对士兵对抗射击系统中大尺寸的视频流图像, 先在下采样图像中判断运动目标的位置, 再根据运动目标的位置, 到原图中进行精确分析, 从而提高算法的处理速度. 通过背景融合方法更新图像背景, 用不断更新的图像背景结合背景差分法获取运动目标, 可以有效消除场景光线变化和相机抖动等因素对检测结果的影响, 提高算法的准确性.

关键词: 运动目标检测; 下采样; 背景更新; 背景差分

Optimization of the Moving Target Detection Algorithm in Video Images

MAO Yu-Jiao, CHEN Ran-Jun, TANG Li-Ping

(Communication and Information System, Donghua University, Shanghai 201620, China)

Abstract: The optimization of the moving target detection in complex environment which based on the Soldiers Shooting Confrontation System is proposed in this article. The algorithm is discussed from two aspects, the efficiency and the accuracy effected by the interference. To improve the efficiency of the algorithm, the large images from the video stream are first down-scaled, so that the location of the moving targets can be quickly determined. After that, the targets will be accurately analyzed in the original images according to their positions. In addition, the effect of the camera shaking and light changing can be eliminated by blending the current image with the background. Combining with background subtraction, background update can effectively improve the accuracy of the algorithm.

Key words: moving target detection; down-scaling; background update; background subtraction

运动目标检测技术是视频流图像处理领域中的重要研究课题. 从复杂背景中快速、准确地分离出运动目标是后续图像处理的基础, 在智能交通、自动导航、智能楼宇等系统中有着广泛的应用. 因此, 在视频图像中检测运动目标有重要的研究意义和应用前景.

背景静止的运动目标检测方法主要包括背景差分法^[1]、相邻帧差分法^[2]和光流法^[3]. 用这些方法提取运动目标时通常需要解决两个问题, 一是图像尺寸过大造成运算速度的降低, 二是视频拍摄过程中光照和抖动等外界因素的干扰对检测准确性的影响.

为了提高对大尺寸图像的处理速度, 通常以降低图像分辨率为代价, 即对图像进行下采样处理. 常用

的下采样方法包括最近邻域法(KNN)^[4]、双线性差值法(Bilinear interpolation)^[5]和三线性插值法(Bicubic interpolation)^[6]. 最近邻域法能快速降低图像分辨率, 但较易产生失真; 双线性法通常用于纹理平滑, 该方法能有效地克服图像失真, 但在一定范围外会丢失大量像素, 算法较为复杂; 三线性插值法能较好地保留细节质量, 但运算量较大, 处理速度较慢.

针对光线变化和相机抖动的干扰, 国内外学者也提出了多种有效的抗干扰方法. 背景建模^{[7][8]}以及动态偏移场建模^[9]等方法可以用来消除相机抖动造成的图像模糊. 通过高斯模型卡尔曼滤波法^[10]、多帧平均法^[11]以及 surendra 算法^[12]对背景进行实时更新, 在更

^① 收稿时间:2013-08-28;收到修改稿时间:2013-09-26

新过程中消除光线变化噪声的干扰。

本文研究视频对抗系统中运动士兵的目标检测,为了准确地检测到场景中的运动目标,系统使用高分辨率(4096*1080)的工业摄像机以15帧/秒的速度对场景进行拍摄,每一帧图像约500万个像素点。拍摄过程中,相机抖动和场景光照变化等外界因素不可避免。本文在对已有方法研究的基础上,从算法效率和准确性两个方面对算法进行优化,使之适合于视频对抗系统中运动士兵的目标检测。

1 传统的运动目标检测方法及其问题

士兵视频对抗的场景建立在室内的环境下,分为左右两个场景,场景中均包含掩体(石块)。两组士兵(A和B)在各自的场景中移动并进行射击,过程中可躲藏于掩体后。两个相机分别放置于两个场景的正前方,拍摄士兵移动和射击的过程。两个场景的正前方分别有两个大屏幕,实时播放相机在对方场景获取的视频,即A组士兵可在屏幕中看到B组士兵的视频,获知对方的运动状态,反之亦然。两组士兵分别向屏幕里运动中的对方士兵射击,进行对抗训练,根据士兵被击中的位置和暴露面积判断胜负。

视频对抗系统中,士兵在掩体间移动,通过视频向对方士兵射击进行对抗训练。计算机需要快速并准确地从摄像机获取的视频流中检测到运动目标,确定其位置。对于背景静止的运动目标检测系统,通常采用背景差分法获取当前帧与背景帧的差值信息,通过对差值图像的分析,检测出运动目标。

传统的算法检测目标的步骤如下:

(1) 读取第1帧图像作为背景 Bkg 。

(2) 读取第 i 帧图像 Img , 与第一张背景图片相减获得差值图像 d , 即:

$$d(x, y, i) = |Img(x, y, i) - Bkg(x, y, 1)| \quad (1)$$

式中 Img 表示当前帧图像, Bkg 表示背景帧图像, i 表示第 i 帧图像, x, y 为图像中像素点的坐标, $Bkg(x, y, 1)$ 表示第1帧背景图像, 即原始背景。

(3) 对差值图像以 T 为阈值进行二值化, 即:

$$ID(x, y, i) = \begin{cases} 255 & d(x, y, i) \geq T \\ 0 & d(x, y, i) < T \end{cases} \quad (2)$$

式中 ID 为二值图像, i 表示第 i 帧, x, y 为图像中像素点坐标, d 为当前第 i 帧与背景帧的差值, T 为阈值。

(4) 扫描二值化后的差值图像, 判断图中是否出

现异常, 即非零点。

(5) 对获取的非零点进行跟踪并合理合并, 同时排除噪声点。

士兵对抗系统的目标是快速而准确地从视频中检出运动士兵并对其进行跟踪。需要对上述算法解决以下两个问题:

(1) 工业摄像机获得的视频图像尺寸较大(4096*1080), 对每张图片的各个点都进行扫描判断, 影响算法的运行速度。

(2) 在差值图像中存在相机抖动或光照变化形成的边界线和大量噪声, 对运动目标的检测有一定的干扰。

2 提高算法速度的方法

士兵对抗系统使用高分辨率的工业摄像机以15帧/秒的速度拍摄现场, 每一帧的图像尺寸约为五百万个像素点。用 $i5$ 处理器、500G 硬盘、2G 内存的计算机, 在 win7 操作系统下运用传统算法对拍摄到的60秒视频进行处理, 处理完整个视频共需36分12秒, 每张图片处理速度为2.54秒, 算法效率较低。

为了提高算法的运算速度, 本文考虑从定性和定量两个层面对算法进行优化。先定性判断图像中是否存在运动目标, 当判断出图像中出现运动目标后, 再对该目标进行定量分析, 获得其面积与形状。这样, 检测的主要时间花费在定性分析上, 定量分析所消耗的时间可以大大减少。

为了减少定性分析的时间, 采用下采样的方法降低图像分辨率, 先在下采样获得的小尺寸图像中进行定性分析, 判断有无运动目标的出现, 当出现运动目标后再根据运动目标的位置, 到原图中进行定量分析。

与一般下采样的目的有所不同, 本文下采样并不需要考虑对原图的保真, 而是希望通过对图像的缩小, 快速地找出图中是否有运动目标出现, 即找出当前帧与背景的差异。因此对下采样得到的图像要求具有如下的特点:

(1) 要使运动目标出现带来的信息最大化, 需要保留下采样产生的失真, 该失真在与背景相减后能产生差异, 使得算法的触发更为灵敏。

(2) 对士兵对抗环境中间隔较大的掩体和士兵边缘, 不易产生交叉失真。

采用最近邻域法对图像进行下采样, 将每一个 3×3 像素区域用一个像素代替, 在行和列上分别取第

一个像素,使各图像的尺寸在行和列两个方向上均缩小三分之一,整体图像缩小为原图的九分之一.对视频流中某一帧图像下采样结果如图 2(c)所示.采样后的图像中灰度值变化较小,图像中的扰动信息不明显.

为了更好地获得图像扰动信息,达到图像扰动最大的效果,利用背景图像与后续图像相比,背景图像较暗,灰度值较小,当有运动目标出现时,灰度值增大的特征,本文采用改进的最近邻域算法,使图像通过背景差分后获得的结果最为明显,保留当前图像中灰度值最大的部分.

改进算法如图 1 所示,图左侧表示读取一个 3*3 的子块,即 9 个像素点,实线表示选取一个像素点,每个像素点包括 R、G、B 三个值,分别取 9 个像素点中 R、G、B 三个值的最大值组成下采样后的像素,即图右侧像素,将每个 3*3 的子块用一个像素代替.

对同一帧图像通过改进的下采样方法,得到的灰度图像如图 2(d)所示.

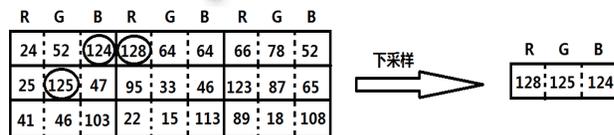


图 1 改进的下采样算法

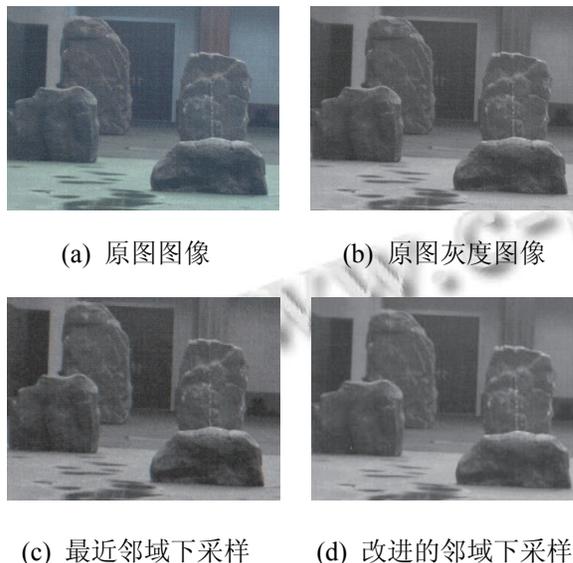


图 2 最近邻域下采样及其改进算法效果图

由图 2 可见,改进后的最近邻域法对差值图像进行下采样的结果,相较于一般最近邻域法获得的下采样结果中非零点更多,即对后续算法的触发更为敏感.

通过 matlab 分别计算图 2(c)和图 2(d)灰度值的方差和熵值,得到图 2(c)灰度值的方差为 38.7115,熵值为 7.1223,图 2(d)灰度值的方差为 39.3338,熵值为 7.1583.图像灰度值的方差和熵值越大说明图像的信息量越大.改进后的最近邻域下采样法获得的下采样图像可以最大限度地保留图像变化信息,使扰动最大,因此本文选用改进的最近邻域下采样方法对图像尺寸进行缩小.

3 外界干扰对算法准确性的影响的消除

视频图像处理过程中,相机抖动、光照、遮挡物阴影等因素都会对场景产生噪声影响,这些外界条件的变化带来的干扰对算法的准确性有着很大的影响.

由于相机抖动,当前帧与背景帧相减后,会出现明显的边界线,计算机检测过程中可能将其错判为运动目标.场景中的光照、阴影等外部条件的变化,会使后续帧的背景发生改变.如果不对背景图像进行更新,通过背景差分法获得的差值图像中很可能会出现“伪目标”.

因此,在视频流图像的处理过程中,需要考虑相机抖动、光照等外界因素变化对系统的影响,优化算法,提高算法准确性.

3.1 相机抖动的消除

针对相机抖动产生的边界线噪声引起的运动目标误判,通过以下的处理步骤进行改进.

(1) 读取第一帧背景图像 Bkg,使用 Canny 算法对背景帧进行边缘检测,获得背景帧中的边缘 A (如图 3(a)所示).

(2) 对边缘 A 进行膨胀获得区域 B (如图 3(b)所示白色部分),膨胀后的区域 B 即为相机抖动后可能产生边界线的区域.为避免目标检出延迟,选择膨胀 2 次.

(3) 从背景 Bkg 中除去区域 B 后获得新的检测区域 C (如图 3(b)所示黑色部分),该检测区域中不包含相机抖动产生的边界线.

$$C(\text{检测区}) = Bkg(\text{整个图像}) - B(\text{相机抖动区域})$$

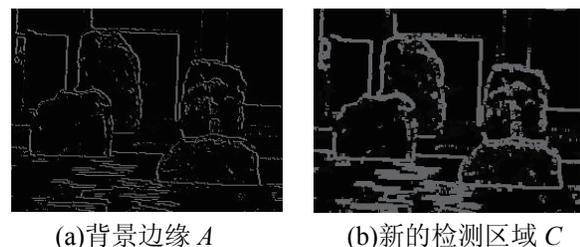


图 3 相机抖动的消除

将相机抖动的区域从检测范围内去除,抖动产生的边界线就不会被检测到,也就不会被误判为捕捉目标.虽然这种处理会导致运动目标的出现延迟几帧,但是由于本文处理的是视频流中的图像,并不影响最终检测结果.

从运算效率上来看,将检测区域由整张图缩小为区域 C 后,在防止检测出边界线的同时也可以减少不必要的运算量,提高系统的处理速度.

3.2 背景更新对干扰的消除

背景差分法实现简单,算法复杂度较低,能提取出较为完整的目标图像,但该方法受到光照、阴影等外界条件影响较大,故本文通过背景更新的方法来消除这些外界的干扰,提高算法准确性.

光照、遮挡物阴影等外界条件变化会影响图像的背景强度,若始终用当前帧与原始背景进行差分,则会在差分图像中出现大量由于图像背景强度变化产生的噪声.本文采用当前帧与背景帧融合的方法,实时更新背景.确保当前背景与当前帧的环境变化相一致.具体步骤如下:

(1) 获取当前帧与原始背景帧中各个像素 R 、 G 、 B 的值

(2) 将当前帧 i 与原始背景中 R 、 G 、 B 的值分别按一定比例进行融合,即:

$$Bkg'(x,y,i) = a \times Im g(x,y,i) + b \times Bkg(x,y) \quad (3)$$

式中 Bkg' 为更新后的新背景, i 表示第 i 帧, x 、 y 为图像中像素点的坐标, a 、 b 为背景融合系数, $a+b=1$.

(3) 将当前帧减去新的背景帧获得差值图像 d' , 即: $d'(x,y,i) = |Im g(x,y,i) - Bkg'(x,y,i)|$ (4)

式中 Img 表示当前帧图像, Bkg' 表示更新后的背景帧图像, d' 表示当前第 i 帧与更新后新背景图像的差值, i 表示第 i 帧, x 、 y 为图像中像素点的坐标.

将当前帧与更新后的背景进行差分,能较好地降低外界干扰对系统准确性的影响.

4 实验结果分析

为了验证本文算法的可行性,从运算速度、抗干扰及背景融合比例三个方面进行了比较实验.

(1) 比较算法优化前后的运算时间,验证优化后算法的实时性.

为了提高算法准确性,本文运用背景更新再次优化了算法.经测算,优化后的算法处理完上述视频流-

需 28 分 03 秒,平均每帧图像处理时间为 1.97 秒.表 1 为算法优化前后运行时间对比.

表 1 算法优化前后运算时间对比

| | 传统算法 | 下采样 | 下采样结合背景更新 |
|---------|-----------|-----------|-----------|
| 处理整个视频流 | 36 分 12 秒 | 26 分 40 秒 | 28 分 03 秒 |
| 平均每帧用时 | 2.54 秒 | 1.87 秒 | 1.97 秒 |
| 速度提高 | / | 26.38% | 22.44% |

由表 1 可见,通过下采样结合背景更新对传统算法进行优化,优化后的算法处理速度较传统算法提高了 22.44%,该算法优化方法提高了算法准确性的同时保证了算法的实时性.

(2) 通过实验分析背景融合前和背景融合后,光照等环境因素改变产生的干扰对实验结果的影响.实验结果如图 4 所示.

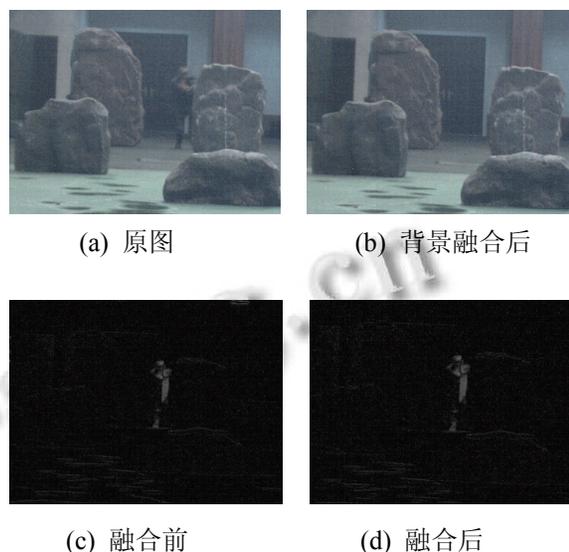


图 4 背景融合前后效果图对比

图 4(b)为图 4(a)与背景帧按一定比例融合后获得的新背景,图 4(a)中可见的运动士兵在融合后消失,获取了当前帧的背景,同理,当前帧产生的细小噪声也能被去除.图 4(c)为当前帧与原始背景图 4(a)相减后获得的差值图像,图 4(d)为当前帧与融合后的新背景图 4(b)相减后获得的差值图像.可见图 4(d)获得的运动士兵更为清晰明亮.

(3) 对式(3)中背景融合系数 a 进行调整,运用不同融合系数对图像进行融合,对结果进行分析,判断

出最佳融合比例。

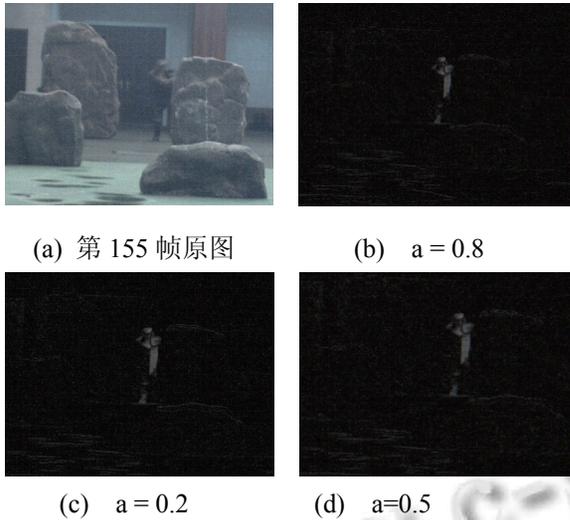


图 5 a 取不同值进行背景融合后的效果

分别以融合系数 a 的值为 0.2、0.5 和 0.8 的比例通过式(3)对当前帧进行背景融合并更新背景,再用当前帧减去融合后的背景获得的差值图像如图 5(b)(c)(d)所示,当 $a=0.8$ 时,检测出的运动士兵较为模糊, $a=0.5$ 时,运动士兵较为明显,但士兵周围离散噪声较大。经试验证明,当原图和背景以 0.2: 0.8 的比例融合,即 $a=0.2$ 时,融合效果最佳,较好地保证了算法的准确性。

5 结论

本文针对士兵对抗系统中可能产生的问题,从算法运算效率和准确性两个因素对算法进行优化。提出改进的最近邻域下采样方法缩小图像尺寸,从定性和定量两个层面对图像进行处理,减小算法运算量从而提高算法效率。通过对检测区域的缩小,消除了相机抖动对运动目标检测带来的噪声影响,结合背景融合的方法对背景进行实时更新,克服外界变化带来的干扰对算法准确性的影响。实验证明本文对算法的优化能提高复杂环境中运动目标的检测的实时性和准确性。

参考文献

1. Piccardi M. Background subtraction techniques: a

Review. Systems, Man and Cybernetics, IEEE International Conference on. 2004, 4. 3099–3104.

2 Foresti GL. Object recognition and tracking for remote video surveillance. Circuits and Systems for Video Technology, IEEE Trans., 1999, 9(7): 1045–1062.

3 Sun D, Roth S, Black MJ. Secrets of optical flow estimation and their principles. Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. 2010. 2432–2439.

4 Weinberger KQ, Saul LK. Distance metric learning for large margin nearest neighbor classification. The Journal of Machine Learning Research, 2009, 10: 207–244.

5 Wang S, Yang K. An image scaling algorithm based on bilinear interpolation with VC++. Techniques of Automation and Applications, 2008, 27(7): 44–45.

6 Peng X, Tatsuki T. Performance comparison of bilinear interpolation, bicubic interpolation, and B-spline interpolation in parallel phase-shifting digital holography. Optical Review, 2013, 20(2): 193–197.

7 Fergus R, Singh B. Removing camera shake from a single photograph. ACM Transactions on Graphics, 2006, 25(3): 787–794.

8 Afonso MV, Bioucas-Dias JM. Fast imagerecovery using variable splitting and constrained optimization. Image Processing, IEEE Trans. on, 2010, 19(9): 2345–2356.

9 Zou MY, Liu Y. Dynamic displacement field model: A new model for describing image sequence distortion. Journal of Electronics & Information Technology, 2008, 30(9): 2143–2147.

10 刘鑫,刘辉等.混合高斯模型和帧间差分相融合的自适应背景模型.中国图象图形学报,2008, 13(4):729-734.

11 Fan Y. A real-time algorithm of dynamic background extraction in image sequence. Machine Learning and Cybernetics, Proc. 2005 International Conference on. 2005, 8. 4997–5000.

12 徐方明,卢官明.基于改进 surendra 背景更新算法的运动目标检测算法.山西电子技术,2009,5:39–40.