

## 从移动背景红外序列图像中检测运动目标

曹炬<sup>\*\*\*</sup> 谭毅华<sup>\*\*</sup> 马杰<sup>\*\*</sup> 田金文<sup>\*\*</sup>

<sup>\*</sup>(航天科技集团一院十二所 北京 100854)

<sup>\*\*</sup>(华中科技大学图像所, 图像信息处理与智能控制教育部重点实验室 武汉 430074)

**摘要:** 该文提出了一种从背景移动红外图像中自动检测运动目标的算法。该算法首先采用图像灰度互相关度量的匹配算法对连续的6帧序列进行配准, 用第1帧和第4帧配准, 第2帧和第5帧配准, 第3帧和第6帧配准, 然后用配准后的图像对分别做差分运算, 再将3个差分图像按像素相乘, 在运动目标处得到了非常尖锐的相关峰。这为进一步自动跟踪目标提供了一个跟踪窗口的中心点。实验结果验证了该方法的有效性。

**关键词:** 运动目标检测, 红外图像, 配准, 相关度量, 图像差分

中图分类号: TP391.4 文献标识码: A 文章编号: 1009-5896(2005)01-0043-04

## Detecting Moving Objects from Infrared Image Sequence on Displacing Background

Cao Ju<sup>\*\*\*</sup> Tan Yi-hua<sup>\*\*</sup> Ma Jie<sup>\*\*</sup> Tian Jin-wen<sup>\*\*</sup>

<sup>\*</sup>(Beijing Aerospace Autocontrol Institute, Beijing 100854, China)

<sup>\*\*</sup>(State Key Laboratory of Education Ministry for Image Processing and Intelligent Control, Inst. for Pattern Recognition and Artificial Intelligence, HUST, Wuhan 430074, China)

**Abstract** An approach to detecting moving objects from infrared image sequence on displacing background is proposed in this paper. First the correlation matching technique is used for registering five pictures with the sixth one. Then the registered images are processed to calculate the image differences between the first and fourth, the second and fifth, the third and sixth. After these difference images are multiplied, a very high correlation peak is obtained. Using this method a central point for tracking window can be provided. The experimental results illustrate that this approach is available to this situation.

**Key words** Detecting moving objects, Infrared image, Registration, Correlation measurement, Image differences

### 1 引言

由运动目标所形成的图像序列可分为两种情况: 静止背景和变化背景。前一种情况通常发生在摄像机相对静止的状态, 如摄像头固定的监控设备; 后一种情况发生在摄像机也相对运动的状态, 如车载摄像机、扫描的摄像机云台、运动目标跟踪等。从处理方法上看, 一般是采用突出目标或消除背景的思想<sup>[1]</sup>。对静止背景的情况可以采用消除背景的方法, 即对序列图像进行帧间简单差分或自适应背景对消的方法。对变化背景的情况, 若采用消除背景的方法, 则需要先进行帧间稳像和配准, 若采用突出目标的方法, 则需要先在配准的前提下进行多帧能量积累和噪声抑制。此外, 可以从运动速度上区分不同的目标和背景, 即采用光流法<sup>[2]</sup>确定序列图像中运动物体的速度矢量, 并在基本光流场方程的基础上提出各种各样的约束条件和计算方法, 如马尔可夫随机场方法等<sup>[3]</sup>。但是由于红外成像设备一般采用能量积累的方法, 即

使是静止的物体, 其图像的帧间差别也较大, 若直接采用光流法或帧间差分方法, 都会产生很多虚假的目标。即便是可见光 CCD 成像, 自然景物环境也会因风吹树叶, 太阳位置改变导致阴影的变化而永远不会绝对静止, 这就意味着必须考虑背景的变化。光流法由于采用迭代近似求解或用最小二乘法解超定方程, 其时间开销很大, 对于从红外图像中计算光流存在较大误差, 因而实用性较差。差图像法可分为两类: 一类是用序列中的每一帧与一个固定的静止的参考帧做图像差。另一类是用序列图像中的两帧进行差分, 然后二值化该差分图像来提取运动信息。然而这种方法常常会引起两种不良的后果: 一种是两帧间运动物体重叠部分没有检测出来, 只检测出物体的一部分; 另一种是检测出物体在两帧中的目标信息比真实的物体大很多。Sayrol 等人<sup>[4]</sup>提出序列图像信息是不断累积的, 其冗余度可以在高阶统计过程中用来减少附加噪声的影响。文献[5]中提供了一种通过傅里叶变换对运动估计的方法, 它将每一列像素灰度相加形成一维阵列,

再进行傅里叶变换得到两个波峰,波峰处的频率与物体的运动速度有关。Dubuisson 提出利用 3 帧图像计算出两个差图像<sup>[6]</sup>,再令它们对应像素相乘,但这些方法对于变化背景的红外图像处理依然存在较多的伪运动信息。本文提出利用 6 帧序列图像得到 3 幅差分图像,然后令它们对应像素相乘,得到非常高的运动目标相关峰,以此相关峰作为跟踪窗口的中心点,可以较好地检测出复杂变化环境下红外图像的运动目标。

## 2 变化背景红外图像的预处理

红外图像最大的弱点是成像噪声很大,因此需要先采用中值滤波方法去除那些孤立噪声点。

我们对变化背景的图像采用差分方法必须先将差分的两幅图像配准。图像匹配算法大致分为两大类,第一类是统计匹配算法,主要方法有平均绝对差算法 MAD、平均平方差算法 MSD 以及基于灰度的去均值归一化积相关算法 NPROD-DMEAN,统计匹配算法在同种图像匹配时由于匹配精度高而成为实际应用中常用的方法,但计算相关函数时因数据量大而非常耗时,而在异种图像匹配时需要转换成相同特征的图像,因此使用范围受到限制。第二类是基于特征的匹配算法,即从图像中提取边界、纹理、熵、能量等特征,利用这些特征进行匹配,以达到快速的目的。图像特征匹配可以利用图像的边缘、物体的形殊点、线特征、拐点特征等,从而可将图像的匹配转化成少量特征的匹配,以达到提高速度、抗几何失真及灰度变化的目的。特征模板匹配方法解决了异种图像的匹配问题,它同时具有特征匹配法和统计匹配法的特点。多子区域相关算法用于克服几何失真的影响,它将图像分为若干小的子区域,计算每个子区域与实测图像的匹配度量,求出它们的局部峰值点,然后利用这些峰值点估计几何失真变换的参数。不变矩作为描述图像的一个特征,也可以用于图像匹配。但从序列红外图像中提取特征难以做到一致,且图像信息损失很大,经过实验比较,我们采用基于灰度互相关度量的匹配算法计算最佳匹配位置。

图像统计匹配算法中最常用的方法是子图像模板匹配,它将模板在图上移动遍历图像并计算相关函数,相关值最大处即为匹配最好处。假定尺寸为  $m \times n$  的小图像为模板图像  $w(x,y)$ ,尺寸为  $M \times N$  的大图像为待匹配图像  $f(x,y)$  ( $m < M$ ,  $n < N$ ),如果要寻找  $w(x,y)$  在  $f(x,y)$  中的位置,可以用  $w(x,y)$  在  $f(x,y)$  上移动,每移动一个位置,  $w(x,y)$  在所覆盖位置就有一个子图像  $g(x,y)$ ,最简单的情况  $w(x,y)$  和  $g(x,y)$  的相关函数可以写为

$$c(s,t) = \sum_{x=1}^m \sum_{y=1}^n w(x,y) \times g(x,y) \quad (1)$$

其中  $g(x,y) = f(x+s,y+t)$ ,坐标  $(s,t)$  为大图像中模板所在位置的起始点坐标(一般取左上角为起始点)。

考虑到式(1)对于图像的幅度值变化比较敏感,我们采用归一化的相关函数:

$$\gamma(s,t) = \frac{\sum_{x=1}^m \sum_{y=1}^n [w(x,y) - \bar{w}][g(x,y) - \bar{g}]}{\sqrt{\sum_{x=1}^m \sum_{y=1}^n [w(x,y) - \bar{w}]^2 \times \sum_{x=1}^m \sum_{y=1}^n [g(x,y) - \bar{g}]^2}} \quad (2)$$

其中  $s=1,2,\dots,M-m$ ;  $t=1,2,\dots,N-n$ ,  $\bar{w}$  是  $w(x,y)$  的均值,  $\bar{g}$  是  $g(x,y)$  的均值,相关函数  $\gamma(s,t)$  已尺度变换到区间  $[-1, 1]$ ,所以与图像的幅度无关。模板在大图像上遍历以后,可以得到  $(M-m) \times (N-n)$  个相关系数值,找出最大值时的  $(s,t)$  即为最佳匹配位置。

## 3 图像差分相乘原理

如果一个阶跃边缘信号从第 1 帧  $f_1$  中一个位置依次移到第 2 帧  $f_2$  中某一位置和第 3 帧  $f_3$  中的某一位置和第 4 帧  $f_4$  中的另一位置,则第 1 帧和第 3 帧的差如图 1 中  $|f_1 - f_3|$  的信号所示,第 2 帧与第 4 帧的差如  $|f_2 - f_4|$  信号所示。如果让这两个差值信号相加,则这两个差值信号的中间部分会出现较大的值,而其它有信号差的位置也会累加,如图中  $|f_1 - f_3| + |f_2 - f_4|$  信号的黑粗线所示。如果让这两个差值信号相乘,则只有这两个差值信号的中间部分会出现较大的值,而其它位置还将保持为零,如图中  $|f_1 - f_3| \times |f_2 - f_4|$  信号的黑粗线所示。

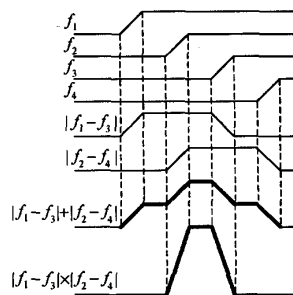


图 1 图像差分相乘原理

二维图像的每一个像素都可以看成是图 1 中的差分相乘点,对于图像中某一点,如果没有运动或者没有强的边缘通过,则在该点处的灰度差在差分相乘图像中是不能产生高的相关峰的。对于二维图像来说,如果图像中没有运动物体,则两幅图像的灰度差为零,只有在运动重叠的地方才会在差分相乘图像中产生高的相关峰。理论上,在差分相乘图像中,只有运动目标对应的像素点位置为非零,但实际上红外图像由于光照变化及各种噪声的影响,使得差分相乘图像中在运动目标点以外的很多像素点位置上值都不为零,总有很多虚假的运动信号。为确保有效抑制虚假运动目标信号就必须使差分相乘图像中的相关峰更加尖锐。经过实验,我们采用 6 帧图像做差分相乘的方法可以达到满意的效果,产生了高的

相关峰。差分相乘方法按照下面公式进行计算：

$$d(x,y) = |f_1(x,y) - f_4(x,y)| \times |f_2(x,y) - f_5(x,y)| \times |f_3(x,y) - f_6(x,y)| \quad (3)$$

其中  $f_1(x,y), f_2(x,y), \dots, f_6(x,y)$  分别为配准后的序列图像像素,  $d(x,y)$  为差分相乘图像像素值。

有了这个尖锐的相关峰, 我们可以轻而易举地确定跟踪窗口的中心点。

### 4 目标检测算法实现

由于目标在图像中占据较小的成分, 而背景的移动代表了图像的主要运动特征, 因此采用灰度互相关度量的匹配算法能够确定图像的最佳匹配位置。

我们采用第 1 帧图像与第 4 帧图像配准, 第 2 帧与第 5 帧配准, 第 3 帧与第 6 帧配准的方法。这样做可以使移动背景的位置发生变化, 在差值图像相乘时对结果影响较小。假定待匹配图的中心像素位置为  $(x_0, y_0)$ , 该点的 8 邻域为

$$\begin{matrix} (x_0-1, y_0-1) & (x_0-1, y_0) & (x_0-1, y_0+1) \\ (x_0, y_0-1) & (x_0, y_0) & (x_0, y_0+1) \\ (x_0+1, y_0-1) & (x_0+1, y_0) & (x_0+1, y_0+1) \end{matrix}$$

目标检测算法的主要步骤如下：

步骤 1 从第 4 帧图像中截取如图 2 虚线所示的图像作为匹配参考图, 每边保留 10 个像素的宽度以保证最终搜索到最佳匹配位置。

步骤 2 在第 1 幅图像中, 以  $(x_0, y_0)$  为中心位置在待匹配图中截取与参考图同样大小的匹配图, 按照式(2)计算两幅图像的相关系数。

步骤 3 分别以 8 邻域的位置作为中心点移动裁剪框的位置, 得到 8 幅不同位置的匹配图, 按照式(2)分别计算与参考图的相关系数。

步骤 4 比较这 9 个相关系数, 取其中最大相关值的位置作为最佳匹配位置, 假定如图 3 所示  $(x_0-1, y_0+1)$  位置为最佳匹配窗口的中心点位置。

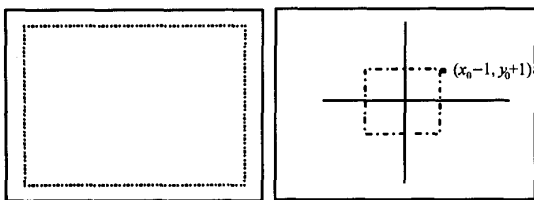


图 2 匹配参考图制备

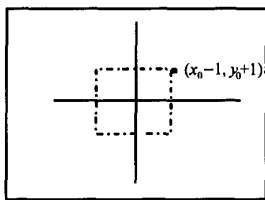


图 3 待匹配图的中心位置及其 8 邻域

步骤 5 再以  $(x_0-1, y_0+1)$  和它的 8 邻域作为裁剪图像的中心点计算相关系数, 找到新的最佳匹配位置。

步骤 6 重复 4, 5 步骤的过程, 直至最佳匹配位置为 8 邻域的中心点位置 (例如第一次匹配的中心点为  $(x_0, y_0)$ )。

步骤 7 对第 2 幅图像与第 5 幅做最佳匹配。

步骤 8 对第 3 幅图像与第 6 幅做最佳匹配。

步骤 9 将匹配好的图像按照式(3)做差分相乘运算, 得到尖锐的相关峰。

### 5 实验结果及讨论

图 4 所示为一序列红外图像, 由于摄像机的移动, 使得背景发生了位移。如果单靠一两帧图像, 我们很难发现路上移动的行人; 由于图像左边灰度值很高, 若直接分割图像, 也很难分割出目标。如果使用播放工具连续播放序列图像, 由于行人的移动方向与大块背景的移动方向有明显差异, 人眼存在视觉暂留, 所以人们可以在连续播放几帧之后很快发现移动的行人。



图 4 序列红外图像

如果采用光流法寻找运动背景和不同的光流场, 由于红外图像的成像特点, 将使计算得到的光流速度矢量非常杂乱, 难以确定真实的运动目标。如果采用特征匹配的方法, 序列图像分割出来的特征差别很大, 难以实现配准。而灰度互相关度量的匹配算法能够利用图像的最大信息确定图像的最佳匹配位置。由于采用整块图像匹配且匹配次数不多, 所以匹配运算的速度远大于计算光流场的迭代速度。

图 5(a) 是配准后第 1 帧和第 4 帧的差值图像  $d(x,y) = |f_1(x,y) - f_4(x,y)|$ ; 图 5(b) 是配准后第 2 帧和第 5 帧的差值图像  $d(x,y) = |f_2(x,y) - f_5(x,y)|$ ; 图 5(c) 是配准后第 3 帧和第 6 帧的差值图像  $d(x,y) = |f_3(x,y) - f_6(x,y)|$ 。从图中可以看出: 虽然图像已经配准, 但道路的轮廓和很多没有运动的地方在差值图像中都被保留下来, 这为进一步的分割带来困难。图 6 是三幅差值图像累加的结果, 从图中可以看出: 虽然目标得到增强, 但道路的边缘和图中没有运动的背景也得到增强, 也会给进一步运动检测带来不利影响。

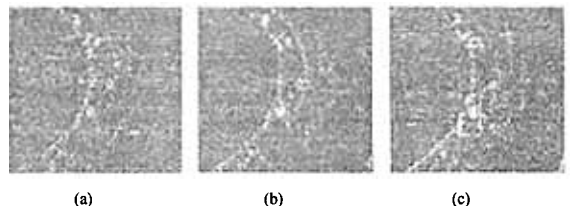


图 5 配准后的差值图像



图 6 3 幅差值图像累加 图 7 3 幅差分图像相乘

图7是3幅差分图像相乘得到的图像,从图中可以看出除了运动目标的尖锐相关峰之外,变化的背景跟相关峰相比变的很弱,几乎无法看到,从而提高了运动目标检测的信噪比。图8是3幅差分图像相乘得到很尖锐的相关峰,从图中可以看出:每个移动目标处都有一个尖锐的相关峰,这为进一步跟踪提供了便利。图9是3幅差分图像累加产生的相关峰,从图中可以看出:除了运动目标以外,很多地方都得到增强,导致进一步的检测可能产生错误。

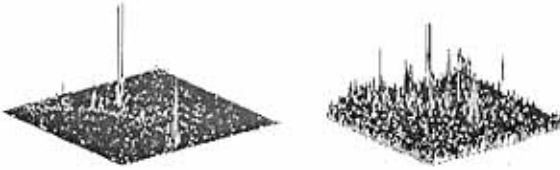


图8 3幅差分图像相乘相关峰 图9 3幅差分图像累加相关峰

图10所示为一序列红外图像,由于摄像机的移动,使得背景发生了位移。如果单靠一两帧图像,我们很难发现移动的坦克(图中方框所示);如果使用播放工具连续播放序列图像,由于坦克的移动方向与大块背景的移动方向有明显差异,人眼在连续几帧之后很快就会发现移动的坦克。

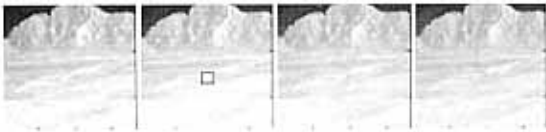


图10 包含移动坦克的红外序列图像

图11是2幅差分图像相乘得到的图像  $d(x,y) = |f_1(x,y) - f_4(x,y)| \times |f_2(x,y) - f_5(x,y)|$ ,从中可以看出还有一些山的边沿被保留下来。图12是3幅差分图像相乘得到的图像,从图中可以看出除了运动目标的尖锐相关峰之外,变化的背景跟相关峰相比变的很弱,几乎无法看到,从而提高了运动目标检测的信噪比。图13是3幅差分图像相乘得到很尖锐的相关峰。

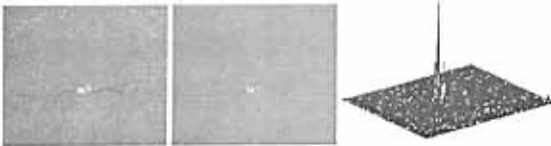


图11 2幅差分图像相乘 图12 3幅差分图像相乘 图13 3幅差分图像相乘的相关峰

本文所述方法适合于移动背景中运动目标不很大且目标慢速移动的情况,即序列图像两帧之间背景和只移动一两个像素,这样才能利用占主要运动特征的背景进行配

准。如果运动目标在背景中占据较多的成分(例如:达到20%),则可能影响到匹配精度,进而影响到检测结果。如果是高速移动的背景或运动目标(如飞机机载成像设备),则需要对配准的方法开展研究,采用更适合的配准方法,针对这些情况要对本方法的适用性做进一步的验证。

### 6 结论

本文模拟人眼的生物学特征,对6幅红外序列图像先采用两两灰度互相关度量的匹配算法计算最大相关值,将此时的匹配位置作为最佳配准位置,然后用配准后的图像做差分运算,用第1帧减去第4帧,第2帧减去第5帧,第3帧减去第6帧,再将3个差分图像按像素相乘,在运动目标处得到了非常尖锐的相关峰。以此相关峰作为跟踪窗口的中心点,就可以用跟踪窗口自动锁定运动目标,从而实现从移动背景中自动检测并跟踪运动目标的功能。实验结果验证了该方法的有效性,该方法可以试用于红外探测器在复杂背景中自动发现并锁定运动目标的情况。

### 参考文献

- [1] Milan Sonka, Vaclav Hlavac, Roger Boyle. Image Processing, Analysis, and Machine Vision[M]. A Division of Thomson Asia Pte Led, USA, 1999: 679 - 721.
- [2] Horn B K P, Schunck B G. Determining optical flow[J]. *Artificial Intelligence*, 1981, 17 (1-3): 185 - 204.
- [3] Heiz F, Bouthemy P. Multimodal estimation of discontinuous optical flow using Markov random fields [J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1993, 15(12): 1217 - 1232.
- [4] Sayrol E, Gasull A. Motion estimation using higher order statistics [J]. *IEEE Trans. on Image Processing*, 1996, 5(6): 1077 - 1083.
- [5] Gonzalez R C, Woods R E 著, 阮秋琦, 阮宇智译. 数字图像处理(第二版)[M]. 北京: 电子工业出版社, 2003: 507 - 514.
- [6] Dubuisson M P, Jain A K. Contour extraction of moving objects in complex outdoor scenes [J]. *International Journal of Computer Vision*, 1995, 14(1): 83 - 105.

曹 炬: 男, 1966年生, 高级工程师, 博士生, 主要从事飞行器导航自动控制、图像处理与目标识别方面的研究工作。  
田金文: 男, 1962年生, 教授, 博士生导师, 研究方向为图像处理、模式识别与人工智能。