基于修正相关系数的运动目标检测*

董安国，李　聪，王丹丹
（长安大学 理学院，西安 710064）

摘　要：通过计算相邻两帧之间对应行和列的相关系数，利用运动目标区域相关系数相对较小的特点，分割出运动目标区域，进而利用逐点匹配算法准确地检测出运动目标。考虑到计算相关系数的运算量大，为了提高检测效率，对相关系数进行了修正。该算法可以使用于复杂背景下多运动目标的检测，具有对噪声不敏感等特点。实验表明，该算法可以有效地检测出所有运动目标，稳定性好，适合于对运动目标的实时检测。

关键词：运动目标检测；修正相关系数；图像配准
中图分类号：TP391　**文献标志码**：A　**文章编号**：1001-3695(2016)07-2197-04
doi:10.3969/j.issn.1001-3695.2016.07.060

Moving target detection algorithm based on modified correlation coefficient

Dong Anguo[1], Li Cong[2], Wang Dandan[3]
(*School of Science, Chang' an University, Xi' an* 710064, *China*)

Abstract: By computing the the correlation coefficient of corresponding rows and ranks between two adjacent frames, it segmented the moving object region by using the trait of target area which correlation coefficient was relative small, then tested moving target accurately through point by point matching algorithm. Because of the number of the operations for the correlation coefficient was tremendous, this paper modified the correlation coefficient to improve the efficiency of the detection. This arithmetic had the character of little calculating amount and not sensitive to the noise, it could be used to the multiple moving target detection under complication back-ground. Experiments show that the algorithm can detect all the moving target accurately, and it is reliable, so it fit for real time detection of moving targets.

Key words: moving target detection; modified correlation coefficient; image registration

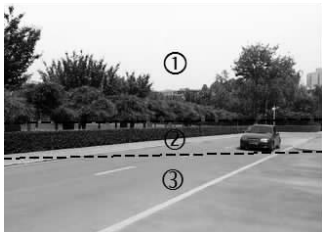
运动目标检测是从图像序列中将运动目标从背景图像中检测出来，运动目标的有效分割对目标的跟踪、分离等后期处理起着至关重要的作用，但背景的不规则变化给目标提取带来了很大的挑战。目前，在摄像机固定的前提下，运动目标的主要检测方法有帧间差分法[1,2]、背景差分法[2~5]、光流法[6,7]、混合高斯法[8,9]等。帧间差分法是通过对视频序列中相邻两帧图像进行差分运算来获得运动物体轮廓，但由于相邻两帧图像的运动目标有重叠部分，所以检测出的物体是前后两帧相对变化的部分，并没有检测到重叠部分，出现“空洞”现象，而且提取的目标比实际大，出现“双影”现象。背景差分法实现简单，运算速度快，但要求前景像素的灰度值和背景像素的灰度值存在一定的差别，且检测结果易受噪声的影响，使准确度下降。光流法采用了光流场信息，能够较好地从背景中检测出前景目标，从而实现摄像机运动过程中独立运动目标的检测。然而，多数光流法要遍历视频序列中所有帧的所有像素点，运算量大，算法复杂度高，实用性不足。高斯混合模型是利用多个高斯分布建立背景模型，通过背景更新不断调整背景模型中的高斯分布的组成，因而具有一定的场景适应性，但是计算速度慢、复杂度高，且容易丢失部分的目标信息，导致目标检测不完整。本文采用计算相邻帧灰度图像之间相关系数的方法[10]来分割运动目标，该方法对于较小的目标也可以进行准确的提取，对于复杂背景下多运动目标的检测，采用分层、分块的思想进行分割，最后用逐点匹配的方法[11~13]进行精细化修正。

1　运动目标特征描述及检测指标确定

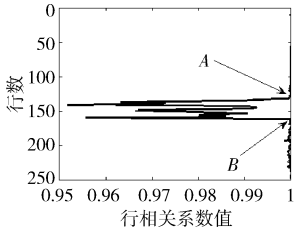
设 P、Q 分别表示视频序列中相邻两帧图像($m\times n$)，若 P 和 Q 的第 i 行均没有运动目标，则 P 和 Q 的第 i 行 $P_{i,\cdot}$ 和 $Q_{i,\cdot}$ 的取值对应于背景中相同的物理对象，尽管存在灰尘、光线等随机因素，但 $P_{i,\cdot}$ 和 $Q_{i,\cdot}$ 的相似性很高；若 P 或 Q 的第 i 行有运动目标，则 P 和 Q 的第 i 行 $P_{i,\cdot}$ 和 $Q_{i,\cdot}$ 的取值对应于不同的物理对象，其相似性较低。基于上述特征的分析，引入行相关系数 r_i($P_{i,\cdot}$ 和 $Q_{i,\cdot}$ 的相关系数)作为两个对象相似性指标。这样，可以通过计算 $r_i(i=1,2,\cdots,m)$，来检测出目标在纵向所处的位置，如图 1(b)所示。对于第 i 列 $P_{\cdot,i}$ 和 $Q_{\cdot,i}$ 的分析情况完全类同，引入列相关系数 c_i($P_{\cdot,i}$ 和 $Q_{\cdot,i}$ 的相关系数)，通过计算 $c_i(i=1,2,\cdots,n)$，来检测出目标在横向所处的位置，如图 1(c)所示。

收稿日期：2015-04-19　**修回日期**：2015-06-24　**基金项目**：国家自然科学基金资助项目(11171043)；国家自然科学基金资助项目(11201038)；中央高校基本科研业务费专项资金项目(CHD2012TD015)

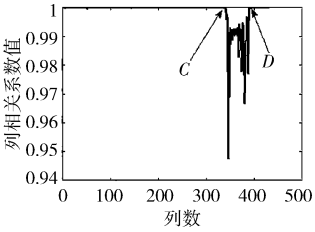
作者简介：董安国(1964-)，男，浙江象山人，教授，硕士研究生，主要研究方向为数值代数、数字图像处理、复杂交通系统建模及优化(donganguo@chd.edu.cn)；李聪(1989-)，女，陕西西安人，硕士研究生，主要研究方向为数字图像处理；王丹丹(1989-)，女，河南新乡人，硕士研究生，主要研究方向为数字图像处理.



(a)视频序列其中一帧



(b)行相关系数



(c)列相关系数

图 1　视频序列中的一帧及行列相关系数

以行相关系数为例,如图 1(b)所示,A 点之前的行为图 1(a)中标号为①的部分,由于这部分几乎没有噪声,其相关系数均等于 1;而 A,B 之间对应的行为(a)中标号为②的部分,有运动目标存在,其相关系数明显小于 1;B 点之后对应的行为图 1(a)中标号为③的部分,由于路面有灰尘(噪声),其相关系数小于且接近于 1。基于上述分析,可以通过对一维信号$\{r_i\}$和$\{c_i\}$的分析,分别确定运动目标在纵向及横向的位置(即图 1(b)中 A、B 及图 1(c)中 C、D 的坐标),从而检测出运动目标。

2　相关系数的修正

假设有两个变量 X、Y,那么两变量间的相关系数可以通过以下公式进行计算:

$$\rho_{X,Y}=\frac{\mathrm{cov}(X,Y)}{\sigma_X\sigma_Y} \tag{1}$$

在计算相关系数时需计算两个变量的标准差,其大概上反映了数据相对于其均值的平均偏离量。由于标准差的计算量大,为了提高运算效率,用离均差和的平均值来代替式(1)中的标准差,即用 $q_X=\frac{1}{n}\sum_{i=1}^{n}|x_i-\bar{x}|$、$q_Y=\frac{1}{n}\sum_{i=1}^{n}|y_i-\bar{y}|$ 分别代替 σ_X、σ_Y 实现对相关系数的修正。图 2 给出了图像每一行的标准差 σ_x 和修正的 q_x 的关系:如图 2 所示,红线代表 σ_x 的值,蓝线代表 q_x 的值(请见电子版),其值基本相同,所以可以用离均差和的平均值 q_x 来代替标准差 σ_x 的计算,以减少运算量。下文提及的相关系数均指修正后的相关系数。

3　基于修正相关系数的运动目标检测

由视频序列特征可知,对于相邻两帧图像,属于背景的行(列)的相关系数均等于或接近于 1,有运动目标的行(列)的相关系数较小,为此将利用相关系数直方图来确定运动目标的位置。以图 1 行相关系数$\{r_i\}$为例,其直方图如图 3 所示。

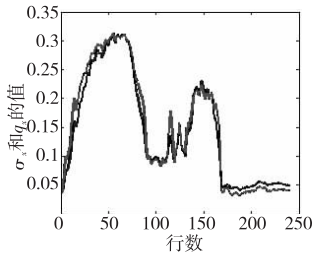


图 2　行的 σ_x 和 q_x 值

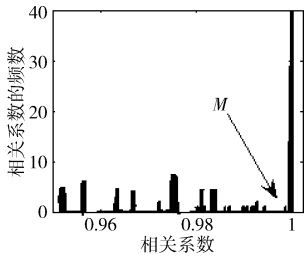


图 3　行相关系数直方图

从图 3 可以看出,相关系数等于或接近于 1 的行(即属于背景的行)的频数较多,而相关系数较小的行(即有运动目标存在的行)频数相对较少,这样,可以以右边第一个极小值(右谷点)作为背景与运动目标的相关系数分界值。为了提高算法的稳定性,避免对锯齿状数据求极小值,在求右谷点前先对数据作平滑处理。算法过程如下:

a)对相关系数 $r_i(i=1,2,\cdots,m)$ 以 $\Delta r(\Delta r=1/K)$ 为步长(K 为给定整数),统计出相关的频数 $f_t(t=1,2,\cdots,K)$;

b)对 $f_t(t=1,2,\cdots,K)$ 作 k-均值滤波:记 $N=\{1,2,\cdots,n\}$,$\Omega_i=\{i-k,\cdots i-1,i,i+1,\cdots,i+k\}\cap N$,$g_i=\frac{1}{|\Omega_i|}\sum_{j\in\Omega_i}f_i(i=1,2,\cdots,K)$,$g_i$ 即为 f_i 的滤波结果($i=1,2,\cdots,K$);

c)找出 $g_i(i=1,2,\cdots,K)$ 右边第一个极大值点 i_0,从 i_0 开始依次向左搜索,找到第一个局部最小值点 i^*,i^* 对应的灰度值即为背景和运动目标的相关系数阈值 r^*。

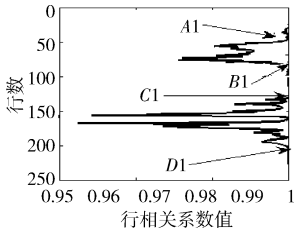
3.1　基于修正相关系数的单运动目标检测

以图 1 为例(取 $K=300$,取 $k=2$),其阈值为 0.9985、0.9974,求得车辆横向位置是 335 ~ 342 列,纵向位置是 124 ~ 152 行,检测结果如图 4 所示。

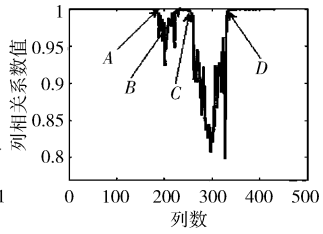
对于单运动目标,直接运用整幅图像的行、列相关系数就可以较好的检测到运动目标区域,然而在多运动目标场景中,以图 5(a)为例,计算其行、列相关系数,如图 5(b)、(c)所示。

图 4　单运动目标检测结果

(a)视频序列其中一帧



(b)行相关系数



(c)列相关系数

图 5　多运动目标场景检测

对于图 5(a),利用上述方法计算整幅图像的行相关系数,检测到 $A1$、$B1$、$C1$、$D1$ 的位置,计算其列相关系数,检测到 A、B、C、D 的位置,检测结果如图 6 所示。从图 6 可以看出,计算整幅图像的行(列)相关系数,会检测出四个部分,而其中只有两个运动目标,其中一个目标所在行和另一个目标所在列的交叉处的背景部分也被误检出来,于是对多运动目标采用分层的思想进行提取,以下介绍对多运动目标的检测过程。

3.2　基于相关系数的多运动目标检测

对于多运动目标场景,以图 5(a)为例进行分析,计算其列相关系数$\{c_i\}$,得到图 7。从图 7 可以看出,列相关系数明显呈现双峰状态,A、B 之间对应的列和 C、D 之间对应的列分别为两个运动目标所在的列。因此,利用其列相关系数直方图确定出 A、B、C、D 的位置,取 B、C 中点对应的列为分割线(如图 7 竖线处所示),将图像分成两部分,那么每一部份就只包含一个运动目标,再利用与检测单运动目标相同的做法,分别计算

各部分行(列)的相关系数,用以检测出目标区域。于是对图 5(a)进行检测得到图 8 的结果。

图 6　对图 5(a)的检测结果

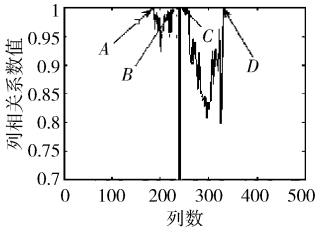


图 7　图 5(a)列相关系数

图 8　对图 5(a)的检测结果

对于较简单的多运动目标场景,分部分计算行(列)相关系数,可有效避免多检测出车辆所在行、列的交叉部分,检测效果比较好。但对于更为密集的多运动目标场景(如图 9(a)所示),对其运用上述方法进行检测,可得到图 9(b)的结果:

(a)视频序列其中一帧　　(b)对(a)的检测结果

图 9　视频序列中的一帧及检测结果

对于图 9(a),运用上述方法可将其分为三列,对两边只包含一个运动目标的部分可以直接用上述方法检测出目标位置,但对于中间两辆车的部分因为其面积比较小,样本点太少,直接用行(列)相关系数进行检测,不能很好地将背景去除,于是采用块相关系数提取运动目标。

3.3　块相关系数去除背景

块相关系数的思想是将视频序列相邻两帧图像分割成大小相同、相互独立的块,计算对应块之间的相关系数,利用运动目标区域块相关系数相对较小的特点,分割出运动目标区域。此处,以图 9(b)两辆车的部分为例,计算其块相关系数,画相关系数直方图如图 10 所示。

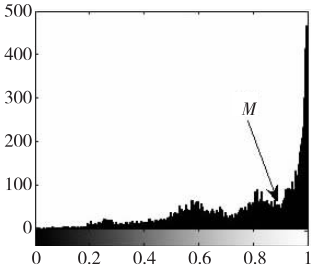


图 10　块相关系数直方图

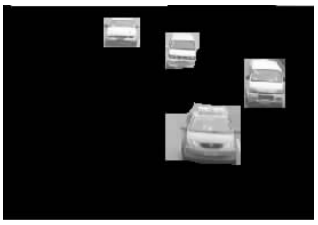


图 11　块相关系数检测的结果

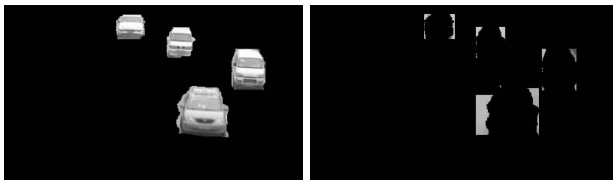
从图 10 可以看出,相关系数等于或接近于 1 的块(即属于背景的块)的频数较多,而相关系数较小的块(即有运动目标存在的块)频数相对较少,基于此特点,用与第一部分相同的方法计算块相关系数阈值,利用块相关系数直方图来确定运动目标的位置,对于只含有一个目标的部分,由于背景比较少,相关系数等于或接近于 1 的块比较少,此处用图 10 计算的阈值即可,对图 9(b)进行检测,得到图 11 的结果。

综上所述,基于相关系数的运动目标检测,可以很好地提取出目标所在行和列的位置,找到运动目标区域。对于多运动目标场景,结合分层、分块的思想,用计算相关系数的方法能更准确地检测出运动目标区域,如图 11 所示;但在图 11 中可以看到目标周围仍有很多的背景像素,为了进一步精确地提取运动目标,用逐点匹配的方法从目标的周围进行精细化的背景去除,这种方法运算量小,稳定性好,准确度高。

4　精细化背景去除

点精细化背景去除的思想是对处理过的图像 f 从目标四周逐点选取,以选取点为中心,取一个 5×5 的块,记作 Ω,然后在它的前一帧图像 g 中相同位置处也选取同样大小的像素块,计算它们的相似度指标 SI:$SI=\sum_{(i,j)\in\Omega}|f(i,j)-g(i,j)|$。

如果相似度指标 SI 小于一个阈值,则说明该点为背景点,将该点赋值为 0;若 SI 大于阈值,说明该点为前景点,像素值保持不变;然后更新图像 f,重复上面的步骤继续做下去,直到图像的像素值不改变为止。此处,阈值的选择方法同上。以图 11 为例对目标周围进行点精细化的背景去除,结果如图 12 所示。



(a)精细化去除背景的结果　　(b)与图 10 相比去掉的背景部分

图 12　对图 11 去除背景的结果

点精细化的背景去除能够完整地提取出运动目标,具有很好的检测性,对于有多运动目标的检测仍有很好的检测效果。

综上所述,基于相关系数的运动目标检测算法具体实现步骤如下:

a)计算视频序列图像帧行(列)的相关系数;

b)利用相关系数直方图确定相关系数阈值,将相关系数大于阈值的行(列)判断为背景,使其像素值赋值为 0,相关系数小于阈值的行(列)判断为运动目标所在的行 (列),使其像素值不变;

c)将未分开的含有运动目标的区域进行分块,计算块相关系数,将相关系数大于阈值的块判断为背景块,使其像素值赋值为 0,相关系数小于阈值的块判断为运动目标所在的块,使其像素值不变;

d)点精细化去除背景,计算相似度指标 SI,如果相似度指标 SI 小于一个阈值,则说 明该点为背景点,将该点赋值为 0;若 SI 大 于阈值,说明该点为前景点,像素值保持不 变;然后更新图像 f,重复上面的步骤继续 做下去,直到图像的像素值不改变为止。

5　数值实验

下面分别用混合高斯法、帧差法、背景差法和本文提出的方法对不同视频进行相邻两帧的运动目标提取,第三列是给第一列视频加入噪声的实验,图 13 就是对目标提取的结果。

实验对比可知:帧间差分法不能完全提取所有相关的特征像素点,在运动目标内部易产生空洞现象,而且提取的目标比实际大,出现“双影”现象,不利于进一步的目标分析与识别。混合高斯模型容易丢失目标信息,导致运动目标有空洞,而且

混合高斯法对背景变化的敏感度低,运算速度慢;背景差分法对提取背景的精度要求很高,且检测结果易受噪声的影响,准确度不高。对于加入噪声的视频,用帧间差分法、混合高斯法、背景差分法提取的效果都不尽理想,由于它们对噪声的影响比较敏感,背景部分去除不彻底,而本文提出的基于修正相关系数运动目标检测的方法,具有对噪声不敏感的特点,算法鲁棒性好,图 13 中第三列是对加噪声图像的检测,可以看出本算法效果明显优于其他算法。

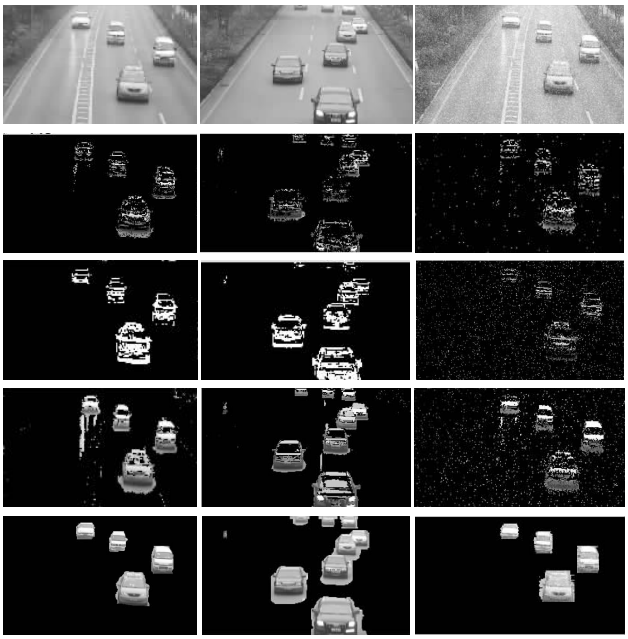


图 13 目标撮结果对比

为了进一步说明本文算法的检测有效性和运算速率,此处将第三列加入噪声的视频,运用 ROC 曲线对几种运动目标检测算法进行性能评估。目标检测从机器学习理论来讲本质就是两类分类问题:判断像素点是目标或者不是目标的问题。ROC 分析的基础即是两类分类问题,曲线上的点代表同一检测算法在不同阈值下的虚警率和检测概率,因此可以有效地描述算法的检测性能。将 ROC 技术运用于运动目标检测,设 N 为"理想"检测参考图中背景点的个数,P 为"理想"检测参考图中运动目标点的个数。利用已知分类器可得到以下四种结果,TP 代表运动目标被分类为运动目标的个数,FN 代表运动目标被分类为背景的个数,FP 代表将背景点分类为运动目标点的个数,TN 代表将背景点分类为背景点的个数,则可计算正确检测率 R_{TP} 与虚警率 R_{FP} 两个分类性能评估指标:

$$R_{FP}=\frac{FP}{FP+TN}=\frac{FP}{N},\ R_{TP}=\frac{TP}{TP+FN}=\frac{TP}{P}$$

评估方法的主要工作流程如下:

a)将待评估的检测算法,分别按一定的间隔设置 8 个不同的阈值(范围覆盖常用的取值区域)对输入影像进行检测,分别得到 32 个运动目标检测结果。

b)确定"理想"检测参考图。将检测结果图集进行像素统计和 ROC 分类,得到候选目标检测图,计算每一等级的候选目标检测图与目标检测结果图集直接的卡方检验参数,取最大值作为最优目标等级,其对应的候选目标检测图即为"理想"检测参考图。

c)将各检测算法在不同阈值下的检测图与检测参考图进行分析,得到多个工作点的正确检测率 R_{TP} 与虚警率 R_{FP},并连接成为 ROC 曲线图。

由 ROC 曲线图性质,越靠近左上角,算法的准确性就越高。因此,图 14 说明本文算法较前几种算法有更好的检测性和更高的准确性。运算速率比较如表 1 所示。

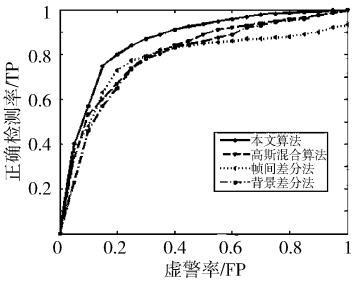


图 14 ROC 曲线图

表 1 运算速率比较

算法	速率/s
混合高斯法	10.29
帧差法	6.45
背景差分法	5.21
本文方法	4.07

由表 1 可以看出,本文算法较前几种算法运行速度快,适合于对运动目标的实时检测。

6 结束语

本文运用修正相关系数对运动目标进行检测,分割出运动目标区域,得到完整的目标信息,再运用逐点匹配的方法对得到的目标进行精细化修正,检测出更为精确的目标轮廓,为实现运动目标的识别、跟踪、分离奠定更为坚实的基础。由于摄像镜头与道路存在确定的夹角,当车辆密度增大时,左右车道中的目标往往做不到垂直(或水平)可分,使行(列)相关系数直方图的峰值增多,对此可以根据摄像镜头的角度对图像进行旋转,最后用本文算法进行检测,可望能提高检测效果,对于这些问题,后面将作进一步研究。

参考文献:

[1] 甘明刚,陈杰,刘劲. 一种基于三帧差分和边缘信息的运动目标检测方法[J]. 电子与信息学报,2010,32(4):894-897.

[2] 薛丽霞,罗艳丽,王佐成. 基于帧间差分的自适应运动目标检测方法[J]. 计算机应用研究,2011,28(04):1551-1554.

[3] Liu Xiaoming,Zhang Zhonghui,Li Guanyu. Research on technology of traffic video incidents detection under highway condition[J]. The Journal of China Universities of Posts and Telecommunications,2010,23(17):79-83.

[4] Lee S,Lee J,Hayes M H. Adaptive background generation for automatic detection of initial object region in multiple color-filter aperture camera-based surveillance system[J]. IEEE Trans on Consumer Electronics,2012,58(1):104-110.

[5] 李宏,杨廷梧,任朴舟,等. 基于光流场技术的复杂背景下运动目标跟踪[J]. 光电工程,2006,33(10):13-17.

[6] Zhao Long,He Xinhua. Adaptive gaussian mixture learning for moving object detection[C]//Proc of the 3rd IEEE International Conference on Broadband Network and Multimedia Technology. 2010:112-117.

[7] Gauglitz S,Hollerer T. Evaluation of interest point detectors and feature descriptors for vidual tracking[J]. International Journal of Computer Vision,2011,94(3):335-360.

[8] 李明之,马志强,单勇,等. 复杂条件下高斯混合模型的自适应背景更新[J]. 计算机应用,2011,31(7):1831-1834.

[9] 李明,赵勋杰. 改进的基于高斯混合模型的运动目标检测算法[J]. 计算机工程与应用,2011,47(8):204-206.

[10] 原凯敏,韦春桃,邹瑄. 一种具有旋转不变性的相关系数影像校正方法[J]. 计算机应用研究,2014,31(7):2220-2223.

[11] 张阿飞,沈海斌. 基于差分的块匹配运动估计算法[J]. 计算机工程与应用,2013,49(6):127-130.

[12] 薛丽霞,王祥,王佐成. 基于空间关系的改进菱形块匹配策略[J]. 计算机应用与软件, 2014,31(2):164-166.

[13] 马灵飞,孟昭鹏. 图像配准在中药贴剂图像处理中的应用[J]. 计算机工程与应用, 2011,47(20):155-158.