

# 基于不变性变换的运动目标检测的研究

袁 杰 都思丹 高敦堂

(南京大学电子科学与工程系, 南京 210093)

**摘 要** 通常采用基于三阶自相关变换的算法来对运动目标进行检测, 是因为这种算法对目标的平移、旋转和伸缩具有不变性的特点, 特别适于目标有水平位移、前后移动和旋转等情况下的检测, 但是直接对二维目标进行三阶自相关变换的运算量很大, 为了提高效率, 提出了一种基于三阶自相关变换的快速算法。应用该算法和理论, 对不同姿态的静态目标进行了检测实验, 并对含有运动车辆的高速公路视频图像进行了车辆检测, 结果表明, 该算法在兼顾效率的同时还可获得较高的目标检测准确率。

**关键词** 运动目标识别 运动目标检测 不变性变换

**中图分类号:** TP391.41 **文献标识码:** A **文章编号:** 1006-8961(2004)11-1284-06

## The Research of Moving Objects Detection Using HOS

YUAN Jie, DU Si-dan, GAO Dun-tang

(Department of Electronic Science and Engineering of Nanjing University, Nanjing 210093)

**Abstract** In the first part of the article, three orders correlation transformation method is applied to detect the moving targets, in series video frames. The obvious advantages of this approach is that it can assure the varginer targets to sustain the same eigen-values when it is displaced, rotated or resized. Therefore, it is especially suitable for the situations when the moving target have displacement, rotation or sizing in video series. The second part of the article provides an effective fast algorithm research on three orders correlation transformation. After applying the fast algorithm, the calculation time can be reduced to less than 6% of that with normal algorithm. In the third parts of the paper, experiments are made to test the algorithms and theories mentioned in the previous part of the article. Through the experiments, different stance of the same target in different images is detected successfully. Moving vehicles are detected in express way video series followed by the comparison of using the normal frames' differential method and the one in this article, from which it shows that the latter method is more accurate and more non-confused than the former one.

**Keywords** moving object recognition, moving object detection, invariant transformation

## 1 引 言

众所周知, 图像信号的许多变换方法均具有不变性, 如图像的直方图就是对图像的平移、旋转具有不变性的二次信号, 而图像信号的能量谱、圆函数展开、Fourier-Mellin 变换等都是基于原始图像灰度值而得到的二次信号, 也具有一定的不变性, 但由于这些二次信号在取得不变性的同时, 也不同程度的丢失了原始图像的几何特征, 从而在应用上受到限制。

图像处理和图像识别的研究已有近 50 年的历史, 其中序列图像的处理是图像处理和识别的一个重要分支, 它是对活动目标进行采样而得到的图像序列进行处理, 以获得从单一图像所不能获得的信息, 用以跟踪运动物体和研究活动目标的运动规律以及进行变化场景的分析、行动的理解等方面的研究, 而且序列图像处理与人工智能、计算机视觉的研究有着密切的关系。序列图像中运动目标的检测与跟踪是计算机视觉和图像编码研究的主要内容, 并且这种运动目标的检测与跟踪也是应用视觉研究领域的一个重要课题。miura 提出了运动观察者对目

标进行识别的计算模型及其在视频检测中的应用<sup>[1]</sup>, Mesh 提出了抗噪声动态摄像机检测目标的方法<sup>[2]</sup>。其目标跟踪是随着数字视频技术的发展而产生的一个新的研究课题。Cédras 则对当前基于运动目标识别的技术和方法作了一个概述<sup>[3]</sup>。由于目标跟踪的结果蕴含了场景中每一个运动个体的大量时空信息(实际反映了该视频流中主要的内容信息),因此该研究课题在基于内容的视频传输、视频检索以及数字监控等各个领域有着重要的意义。Barceli 给出了一种从视频序列中提取背景、前景和模板的框架<sup>[4]</sup>。现有的目标跟踪算法,从目标跟踪的依据来看,可以分为基于先验知识和非基于先验知识两大类;从结果的侧重点来看,又可以分为侧重于运动目标区域和侧重于运动目标运行轨迹两大类。这些算法各有特色,但都面临着实时性局限和背景噪声干扰两个问题。要解决这两个问题必须先对目标跟踪算法的基本思想进行分析。由于所有目标跟踪算法的基本思想都围绕着特征提取和特征匹配两个问题,因此从排除噪声干扰的角度来说,多种特征的提取以及复杂的特征匹配规则显然是有利的,但由于这样同时增加了算法的复杂度,也影响了算法实现的实时性,所以,如何快速而又能强抗噪声地从视频流中提取出代表性的特征就成为解决问题的关键。

## 2 三阶自相关变换及其特性

设  $h(t)$  为二维确定性连续图像信号,  $t$  为表示图像像素位置的二维矢量,  $t = [x \ y]^T$ 。则它的三阶自相关变换<sup>[5]</sup>定义为

$$h_3(t_1, t_2) = \iint_{t \in R^2} h(t)h(t+t_1)h(t+t_2)dt \quad (1)$$

对于离散图像, 设  $h(t)$  为能量有限的离散二维图像信号, 图像的尺寸为  $N \times N$ 。  $t \in [0, 1, \dots, N-1] \times [0, 1, \dots, N-1]$ , 则三阶自相关变换可定义为

$$h_3(t_1, t_2) = \frac{1}{N^2} \sum_t h(t)h(t+t_1)h(t+t_2) \quad (2)$$

其中,  $t_1, t_2$  为二维矢量,  $(t_1, t_2) \in [0, 1, \dots, N-1] \times [0, 1, \dots, N-1]$ 。

三阶自相关变换具有如下主要性质:

### (1) 唯一性

图像的一、二阶自相关函数是关于图像信息的不完全描述, 存在信息丢失问题, 且不同图像可具有相同的一、二阶自相关函数, 而三阶自相关变换则是

图像信息的全描述, 只要  $h_3(t_1, t_2)$  非全零, 则原始信号与它的三阶自相关变换就存在一一对应关系<sup>[6]</sup>, 由于从二维图像域到三阶自相关域的变换不丢失图像信息, 因此对原始图像的检测可以转换为对其三阶自相关变换的检测。

### (2) 平移不变性

设两个二维图像信号  $g(t) = h(t+t_0)$ , 其中  $t_0$  为对  $t$  进行平移后的二维矢量, 则其三阶自相关变换存在以下关系:  $g_3(t_1, t_2) = h_3(t_1, t_2)$ , 这说明图像在三阶自相关变换不受图像信号在平面平移的影响。

### (3) 对称性

三阶自相关变换的对称性表明存在冗余信息, 根据这个特性可以考虑采用快速算法来减少其计算复杂度。本文充分利用了这一特性进行优化计算。

### (4) 旋转和伸缩不变性

设二维图像信号  $g(t) = h(Mt)$ ,  $M$  为图像信号的幅度和旋转变换矩阵, 即

$$M = k \begin{bmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{bmatrix} \quad (3)$$

$k$  为尺度变化因子, 则其三阶相关函数存在以下关系:  $g_3(t_1, t_2) = h_3(Mt_1, Mt_2)$ , 这说明对于信号的幅度和旋转变换, 两个偏置信号矢量  $t_1, t_2$  在三阶相关域的变换是相同的, 也说明图像的三阶自相关变换不受图像信号在平面旋转和伸缩的影响。

### (5) 对噪声的不敏感性

三阶相关量对加性零均值高斯噪声和线性对称分布噪声具有较强的抑制性, 且对于其他类型的噪声还可采用以下两种方法来提高信噪比: ①在可以得到多个样本时, 通过样本平均的方法可很好地消除对称分布噪声; ②当只能得到模式的单个样本时, 在一定范围内的三阶相关量的平均可以很好地抑制噪声。

### (6) 傅里叶变换特性

设二维图像信号  $h(t)$  的三阶自相关变换为  $h_3(t_1, t_2)$ , 其傅里叶变换  $H_3(f_1, f_2)$  与图像信号频谱的关系为

$$H_3(f_1, f_2) = H(f_1)H(f_2)H(-f_1 - f_2) \quad (4)$$

其中,  $H_3(f_1, f_2)$  为如下二维图像信号  $h(t)$  三阶自相关变换  $h_3(t_1, t_2)$  的傅里叶变换,  $f_1 = [u_1, v_1]$ ,  $f_2 = [u_2, v_2]$  为二维频率矢量

$$H_3(f_1, f_2) = \iint h_3(t_1, t_2) e^{-jU_1 t_1} e^{-jV_2 t_2} dt_1 dt_2 \quad (5)$$

### 3 基于三阶自相关变换的图像特征提取

根据三阶自相关变换的性质可知,由于运动目标在二维平面内有移动和旋转时,经过三阶自相关变换后的数据保持不变,而且在抑制加性零均值高斯噪声和线性对称分布噪声方面有着很好的特性,因此可以根据这些变换后的数据进行目标检测。由于变换后的数据维数和变换前是一样的,因此为了有效地进行检测,必须要对变换后的数据进行特征提取。本文按照如下步骤进行特征提取:

设存在二维信号  $g(t) = h(Tt + \hat{t})$ , 其中  $\hat{t}$  为  $t$  平移变换后的二维矢量,  $T$  为幅度和旋转变换矩阵,  $l$  为幅度变换因子,  $\varphi$  为旋转角度, 则

$$\hat{t} = [\hat{x}, \hat{y}]^T \text{ (上标 T 为矩阵转置)}$$

$$T = l \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \quad (6)$$

根据三阶自相关变换的性质有

$$g_3(t_1, t_2) = h_3(T(t_1 + \hat{t}), T(t_2 + \hat{t})) = h_3(Tt_1, Tt_2) \quad (7)$$

若对上式进行傅里叶变换, 则根据三阶自相关变换的傅里叶特性和傅里叶变换的尺度变换特性可以得到

$$G_3(f_1, f_2) = G(f_1)G(f_2)G(-f_1 - f_2) = l^{-3}e^{-3j\varphi}H(f_1e^{j\varphi})H(f_2e^{j\varphi})H(-lf_1e^{j\varphi} - lf_2e^{j\varphi}) = l^{-3}e^{-3j\varphi}H_3(lf_1, lf_2) \quad (8)$$

其中,  $G_3(f_1, f_2)$  为  $g_3(t_1, t_2)$  的傅里叶变换,  $G(f)$  为  $g(t)$  的傅里叶变换。再在两边取幅度, 即得到

$$|G_3(f_1, f_2)| = l^{-3}|H_3(lf_1, lf_2)| \quad (9)$$

如果将  $|G_3(f_1, f_2)|$ ,  $|H_3(lf_1, lf_2)|$  在二维区间  $\begin{bmatrix} 0 & 2\pi \\ 0 & 2\pi \end{bmatrix}$  上作幅度归一化, 则两者就是完全一致的结果, 可以作为特征量进行目标检测。

### 4 三阶自相关变换的快速算法

由于特征提取的过程中要计算二维图像的三阶自相关变换, 其计算非常费时, 因此为了实现实时处理, 需要进行快速算法的研究。

对  $m \times n$  图像区域可按照如下公式进行三阶自相关变换:

$$h_3^{(x)}(x, y) = f(x, y)f((x+x_1)_m, y)f((x+x_2)_m, y)$$

$$h_3^{(x,y)}(x, y) = h_3^{(x)}(x, y)h_3^{(x)}(x, (y+y_1)_n)h_3^{(x)}(x, (y+y_2)_n)$$

$$h_3(t_1, t_2) = \frac{1}{mn} \sum_{y=0}^{n-1} \sum_{x=0}^{m-1} h_3^{(x,y)}(x, y)$$

$$t_1 = [x_1 \quad y_1]^T, t_2 = [x_2 \quad y_2]^T$$

其中,  $(x+x_1)_m, (x+x_2)_m, (y+y_1)_n, (y+y_2)_n$  为求模运算。在进行试验的时候, 明显可以看到, 计算一个  $n \times n$  的 24 位彩色图像的三阶自相关变换需要进行的乘法运算为  $12n^4$ , 需要进行的加法运算次数为  $3n^4$ , 如何将如此大的运算量进行优化是算法从理论到实用的关键步骤。本文所做实验仅针对灰度图像进行处理, 此时需要进行的乘法运算次数为  $4n^4$ , 需要进行的加法运算次数为  $n^4$ 。如果先将目标区域划分为  $2 \times 2$  的子块, 再对每个子块求灰度平均值, 然后对这些灰度平均值组成的新区域进行三阶自相关运算, 则可以进一步降低运算量。通常在保证检测结果的前提下, 可以进一步划分成  $3 \times 3, 4 \times 4$  子块。表 1 是划分成不同子块的运算量比较表。

表 1 三阶自相关变换运算量比较表

| 子块  | 乘法运算次数                                 | 加法运算次数                             | 减少的运算量 (%) |
|-----|--|------------------------------------|------------|
| 1×1 | 4n <sup>4</sup>                        | n <sup>4</sup>                     | 0          |
| 2×2 | n <sup>4</sup> +0.25n <sup>2</sup>     | 0.25n <sup>4</sup> +n <sup>2</sup> | 75         |
| 3×3 | 0.44n <sup>4</sup> +0.11n <sup>2</sup> | 0.11n <sup>4</sup> +n <sup>2</sup> | 89         |
| 4×4 | 0.25n <sup>4</sup> +0.06n <sup>2</sup> | 0.06n <sup>4</sup> +n <sup>2</sup> | 94         |

为了进一步减少运算量, 进行实验的时候还利用了三阶自相关变换的对称特性来进行优化, 优化时, 只需要计算在  $x \in [0, n/2], y \in [0, x]$  区域内的  $h_3(t_1, t_2)$  数值就可以了。下面对这个优化的正确性进行论证:

根据三阶自相关变换的对称性, 有

$$h_3(t_1, t_2) = h_3(t_1 - t_2, -t_2) = h_3(t_2 - t_1, -t_1) = h_3(t_2, t_1) = h_3(-t_1, t_2 - t_1) = h_3(-t_2, t_1 - t_2)$$

对于二维图像的三阶自相关变换, 当变量  $t_1, t_2$  的维数超出图像的维数  $m, n$  的时候, 则采取循环求余的方法, 即

$$t = (x, y)^T = \begin{cases} (x, y)^T & x < m, y < n \\ (x, y_n)^T & x < m, y \geq n \\ (x_m, y)^T & x \geq m, y < n \\ (x_m, y_n)^T & x \geq m, y \geq n \end{cases}$$

根据上式, 将  $h_3(t_1, t_2)$  循环平移  $(\frac{m}{2}, \frac{n}{2})$  单位

后,有如图 1 所示的对称关系。

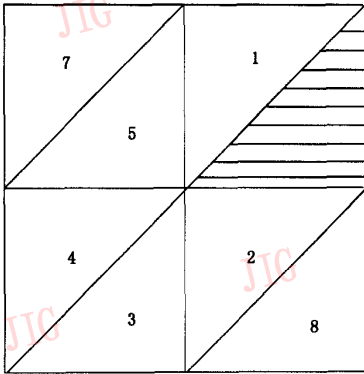


图 1 三阶自相关变换对称性示意图

图 1 中阴影部分的三阶自相关变换数值和其余的三角形部分的三阶自相关变换数值存在一一对应关系。

阴影部分为循环平移 $(n/2, n/2)$ 单位前的 $\{h_3(t_1, t_2) | x \in [0, n/2], y \in [0, x]\}$ 区域,也就是现在图 1 中的区域 3。不妨设 $m=n$ ,图 1 中各区域关系分析如下:

(1)图 1 中区域 1 为循环平移 $(n/2, n/2)$ 单位前的 $\{h_3(t_1, t_2) | x \in [0, y], y \in [0, n/2]\}$ 区域,也就是现在图 1 中的区域 4。根据 $h_3(t_1, t_2) = h_3(t_2, t_1)$ 得知,阴影部分和区域 1 存在一一对应关系。

(2)图 1 中区域 2 为循环平移 $(n/2, n/2)$ 单位前的 $\{h_3(t_1, t_2) | x \in [0, n/2], y \in [n-x, n]\}$ 区域,也就是现在图 1 中的区域 7,根据 $h_3(t_1, t_2) = h_3(t_1 - t_2, -t_2)$ 得知,阴影部分和区域 2 存在一一对应关系。

(3)图 1 中区域 3 为循环平移 $(n/2, n/2)$ 单位前的 $\{h_3(t_1, t_2) | x \in [n/2, n], y \in [n/2, x]\}$ 区域,也就是现在图 1 中的阴影区域,根据 $h_3(t_1, t_2) = h_3(t_2 - t_1, -t_1)$ 得知,阴影部分和区域 3 存在一一对应关系。

(4)图 1 中区域 4 为循环平移 $(n/2, n/2)$ 单位前

的 $\{h_3(t_1, t_2) | x \in [n/2, n], y \in [x, n]\}$ 区域,也就是现在图 1 中的区域 1,根据 $h_3(t_1, t_2) = h_3(-t_1, t_2 - t_1)$ 得知,阴影部分和区域 4 存在一一对应关系。

(5)图 1 中区域 5 为循环平移 $(n/2, n/2)$ 单位前的 $\{h_3(t_1, t_2) | x \in [n/2, n], y \in [0, x - n/2]\}$ 区域,也就是现在图 1 中的区域 8,根据 $h_3(t_1, t_2) = h_3(-t_2, t_1 - t_2)$ 得知,阴影部分和区域 5 存在一一对应关系。

图 1 中的区域 7 和区域 8 为图像的高频区域,因其对于特征值提取没有很大的作用,所以可以不用计算。

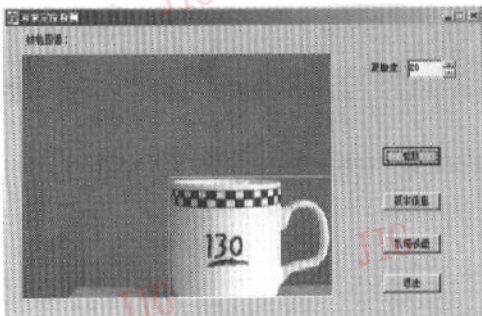
根据这样的对称关系,在对图像进行三阶自相关变换的时候,由于只需要计算原先的 1/8 数据就可以了,后面进行的实验就采用了这样的计算方法,因而极大程度地缩短了运算时间。

### 5 实验和结论

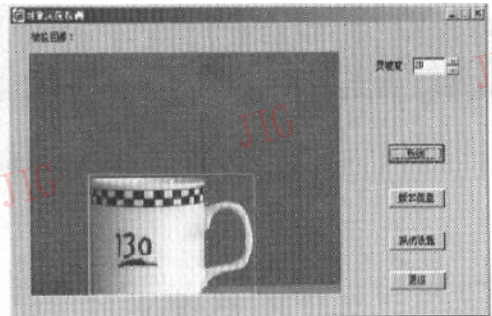
第 1 个实验是一个不同位置和形态杯子的检测实验。通过实验结果(如图 2 所示)可以看出,在目标平移、旋转和倾斜等不同姿态下,采用本文提出的算法都能对目标进行很好的识别。

图 3 是采用帧差分算法和本文提出的算法对高速公路上的一段视频图像进行目标识别的实验结果。从图 3 中可以看出,本文算法在定位精度和准确度方面都优于帧差分算法。

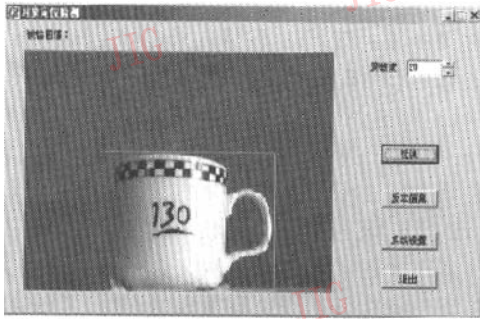
由于三阶自相关变换对目标的平移、旋转和伸缩具有不变性的特点,因此本文在这个变换的基础上进行了创新,即在保留三阶自相关变换的优点的前提下,精简了运算量,使之能够快速有效地检测运动目标。在对高速公路上的运动车辆进行的检测和识别实验取得了满意的结果。



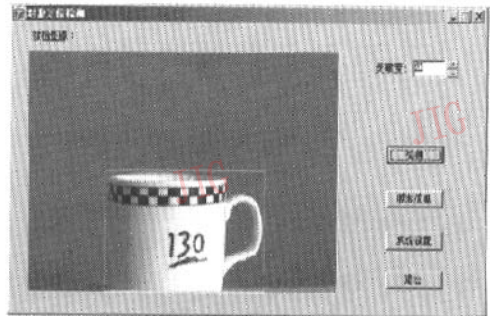
(a) 被检测到的目标



(b) 目标水平移动后的检测结果



(c) 目标倾斜运动后的检测结果



(d) 目标旋转运动后的检测结果

图 2 对不同姿态的目标进行检测的实验结果



(a) 帧差分算法单目标检测结果



(b) 帧差分算法多目标检测结果



(c) 本文算法对单个目标的检测结果



(d) 本文算法对多个目标的检测结果

图 3 本文算法目标检测结果及与帧差分算法的检测结果对比

参考文献

- Miura Ken-ichiro. A computational model for the detection of object motion by moving observer using self-motion signals[J]. Information Science, 2000, 123:55~73.
- Mesh Roland. A noise robust method for 2D shape estimation of moving objects in video sequences considering a moving camera [J]. Signal Process, 1998, 66:203~217.
- Cédras Claudette. Motion-based recognition [J]. Image and Vision Computing, 2001, 13(2):129~155.
- Barceló Lluís. Bayesian video mosaicing with moving objects[J]. International Journal of Pattern Recognition and Artificial

Intelligence, 2002, 16(3):341~348.

- 5 Lohman A W. Triple correlation[J]. Proceedings of the IEEE, 1984, 72:889~901.
- 6 Alshebeili S A. Cumulant-based identification approaches for non-minimum phase FIR system[J]. IEEE Transactions. on Signal Process, 1993, 41(4): 1576~1588.



**都思丹** 1962年生,1998年获得南京大学电子科学与工程系博士学位,现任南京大学电子科学与工程系博士生导师。研究方向为电路与系统。



**袁 杰** 1975年生,2000年获得南京大学电子科学与工程系信号与信息处理专业工学硕士学位,2003年获得南京大学电子科学与工程系无线电物理专业理学博士学位,目前留校任教。研究方向为数字图像处理 and 电路与系统。

E-mail:yuanjie@nju.edu.cn



**高教堂** 1940年生,现任南京大学电子科学与工程系博士生导师。研究方向为数字图像处理。

## 《中国图象图形学报》

### 关于公式及正文中符号表示的一些约定

根据有关规定,对外文符号作如下约定:

1、凡是量符号均用斜体表示,其中标量用白斜体,矢量(向量)、矩阵、集合用黑斜体(整数、有理数、自然数集用正黑体)。

2、凡代表某一名称(含意)的符号用白正体,如:最大(max)、最小(min)、左(left)、右(right)、边缘(edge)有时用一个字母  $e$  表示,若用斜体则成了一个量符号,所以正斜体一定要区分。一般名称符号用于限定量符号,放量符号的上下角,如:  $X_{\max}$ ,  $A_{\text{left}}$ ,  $A_{\text{edge}}$ 。

3、一篇文章中,同一个量用同一符号,不同量用不同符号(包括希文字母、英文字母、大小写、黑体、白体),如,一个字母的黑、白体即为两个符号,因此不要一个量有时用黑体,有时用白体,或有时用大写,有时用小写(易误解为两个量)。同一个量正文中与公式中应一致。

4、一个符号表示一个量,最好不用多个字母代表一个量,主量符号居中,其余限定(说明)符号放上、下角。如  $A$  代表面积,海洋的面积则为  $A_{\text{sea}}$ ,第  $i$  块海洋面积则为  $A_i^{\text{sea}}$ 。一般限定符号均放下角,但有时有几个限定符号,既有正体,又有斜体,不好放在一起,正体的符号可放左上角(因不易与量的幂相混)。若量符号放左上角,又不代表量的幂,则应加括号。如  $A_i^{(n)}$ ,指迭代  $n$  次形成的第  $i$  块输出。

5、单位符号(如 s、m、dB),微分符号  $d$ ,三角函数符号  $\sin$ 、 $\cos$ ,自然对数底  $e$ ,圆周率  $\pi$  等用正体。