

自动多目标识别和跟踪系统

邬敏鸣

(中国科学院上海技术物理研究所, 上海, 200083)

摘要: 研制的自动多目标识别和跟踪系统采用了三级分类识别、动态景物分析等技术, 不仅改善了微弱运动目标的捕获能力、实现了运动多目标识别, 还具有预计目标遮掩及实现目标再捕获的能力, 使系统具有高的目标识别率, 当目标数为10个时, 系统识别速率为10帧/s.

关键词: 目标识别, 目标跟踪, 动态景物分析.

引言

自动目标识别系统(简称ATR系统)的基本功能是对目标进行探测、分类和识别, 它能有效地代替人进行目标检测和目标识别, 因而具有广泛的应用前景. 随着红外成像探测器技术的迅速发展, 人们对目标识别系统性能的要求日益提高, 这些要求主要是: (1) 能够实现自动目标检测(捕获)、识别、优选、预计; (2) 能够对运动多目标同时跟踪, 尤其能在较复杂的背景及目标遮掩、较低对比度等情况下.

智能目标识别系统的研究始于本世纪80年代初, 它采用先进、完善的算法及大规模集成电路技术, 以最少的硬件增加, 实现多目标跟踪, 并达到接近零的目标丢失率. 本文采用较先进的算法, 把目标识别、目标跟踪结合于一体, 通过帧间分析、动态景物分析等先进技术, 采用高速双DSP微处理器, 能直接接收、处理标准RS-170电视视频信号, 研制成能自动实时完成多目标识别和跟踪的实验系统.

1 基本原理及系统实现

自动目标识别、跟踪系统的整个处理过程一般可以分为信号预处理、目标检测、目标提取、特征量计算、目标分类、目标跟踪等, 本系统的各处理功能单元如图1所示, 其中帧间分析等则由软件完成.

系统能实现目标的实时自动检测、识别和跟踪. 为了使系统具有相当的适用性, 采用了自适应门限, 硬件实时二维滤波, 系统能实时估计背景, 突出弱目标信号, 从而保证系统具有高的检测概率. 目标提取采用硬件流水线结构, 比较一般目标提取的空域处理法,

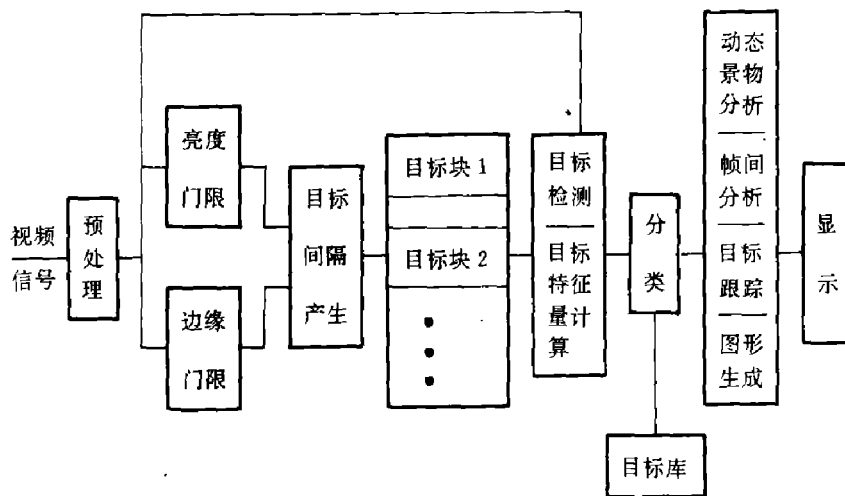


图1 系统处理框图

Fig. 1 The basic processing approach

大大减少了处理时间，从而实现实时处理。

由于目标图像的复杂性和随机性，本系统采用三级分类识别技术使系统在尽量少的计算量下，保证系统具有高的目标识别率及低的目标虚警率。其中第一级是去除干扰、假目标；第二级是把可能的潜在目标经计算特征量分类为10类目标之一；第三级将识别分类的结果进一步进行帧间分析，从而使识别方法从通常的帧内横向分析进一步到图像序列的纵向分析，保证系统具有高的目标识别率和低的虚警率。

动态景物分析是一种先进的技术，对运动平台的情况特别有利，不仅能实现运动多目标识别，而且使目标识别系统能胜任动态环境。

本系统采用两个高速微处理器 TMS320C25，CPU1 主要用于控制采集图像信息，由目标间隔逻辑产生目标块，计算目标特征量，确定目标位置、大小，实现目标分类，以及与 CPU2 的通讯等。CPU2 系统主要是根据目标数据进行动态景物分析，以及与第一微机系统通讯等等。

1.1 目标提取

目标提取的功能是从图像中提取出潜在目标。本文的方法是先从每一条扫描线中提取出“目标间隔”，然后将所有的“目标间隔”分别整理、连接成各自的潜在目标块，整个处理过程都是随串行视频信号实时流水线完成。

1.1.1 自适应门限

自适应门限是根据图像信号的变化，其门槛高度能自动随局部图像对比度的变化而变化，从而能适应复杂背景的情况，抑制了属慢变化的背景和干扰，突出了弱目标信号，使系统具有高的检测概率，能适应复杂的实际情况。自适应门限由两个二维处理器（边缘滤波器和背景滤波器）实现，视频图像经过二维边缘处理，获得“边缘”信号，经过亮度门限，获得“亮度”信号。

二值边缘信号主要由边缘滤波器实现，边缘滤波器其实是一个二维的微分处理器，图像经过它的处理，目标边缘被突出，再经过边缘门限，产生边缘逻辑信号。二值亮度信号

主要由背景滤波器实现, 背景滤波器是一个自适应二维回归滤波器, 它在图像的每一点上产生二维局部背景信号的平均值. 视频信号减去背景后, 再经过亮度门限, 产生亮度逻辑信号.

1.1.2 目标间隔的产生

当某一条扫描线上由自适应门限产生的二值亮度逻辑信号和边缘逻辑信号相吻合时, 产生目标间隔. 间隔起始于亮度信号和边缘信号的一致, 并结束于两个信号的吻合. 因此, 目标间隔产生的关键是如何合理选取间隔起始和结束的判据. 一个有效目标间隔的产生应该是基于过去、现在和将来的边缘, 及亮度信号, 并且具有双向间隔提取功能. 所谓双向间隔提取法, 即当该算法在对某一条扫描线进行运算时, 不论是从左到右操作, 还是从右到左操作, 其结果是一致的. 在流水线单向结构的实时硬件秘理电路中, 要实时进行双向操作是比较困难的. 系统采用的算法具有双向功能, 即当单向操作时, 能产生合适的目标间隔.

1.1.3 目标块的产生

由于硬件实现目标间隔提取, 故需要处理的数据量大大减少, 对软件处理很有利. 从自适应门限出来的目标间隔数据, 在图像扫描回扫时, 读入数据存储器, 输入数据格式以行序排列. 这部分算法的目的是将按扫描线排列的一维目标间隔, 组合、匹配成二维的目标块, 组成各可能的目标, 它是采用中点的对称匹配算法完成的.

1.2 去除假目标分类器

由于图像的复杂性和随机性, 经目标分割提取的目标块可能很多, 其中包括许多假目标, 因此必须有一些计算简单的特征量来区分目标和假目标, 先从中去除绝大多数假目标. 这些特征量一般选择目标的统计特征和结构特征 (强度量和几何量), 这里我们选取了如下几个作为特征量: 平均目标强度为 8bit, 平均目标对比度 8bit, 总面积 10bit, 周长 7bit. 由于处理量比较大, 分类器必须简单、有效. 一个线性分层分类器被用来区分、去除假目标. 这个分类器由一系列接连的门限域值组成, 每一个目标块仅要求经四次比较, 就可通过分类器, 这从计算量来说是很有利的.

1.3 特征量的计算

二维的 Moment 不变量作为目标的特征量, 已经被用于各种目标的识别. 它相对于目标位置、大小、及方向不变, 因此具有相当的实用价值. 设 $I(X, Y)$ 是目标的强度分布, 则 (p, q) 阶 Moment 距可表示为

$$M_{pq} = \sum_{x=1}^m \sum_{y=1}^n I(X, Y) X^p Y^q, \quad (1)$$

这里 $N = n \times m$, 是总的像元数. 根据式 (1), 目标的质心可表示为

$$\begin{cases} \bar{X} = \frac{M_{10}}{M_{00}}, \\ \bar{Y} = \frac{M_{01}}{M_{00}}. \end{cases} \quad (2)$$

由式 (2) 可进一步得

$$\mu_{pq} = \sum_{x=1}^m \sum_{y=1}^n I(X, Y)(X - \bar{X})^p (Y - \bar{Y})^q. \quad (3)$$

再经归一化得

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{(p+q)/2+1}}. \quad (4)$$

由此可以证明 η_{pq} 是关于目标位置、大小不变的量. 在识别目标时, 一般取到三阶, 共有 η_{00} 、 η_{11} 、 η_{20} 、 η_{02} 、 η_{21} 、 η_{12} 、 η_{03} 、 η_{30} 等 (其中 $\eta_{10} = \eta_{01} = 0$). 再考虑目标旋转的影响, 进一步有 7 个实用的不变量如下:

$$\begin{aligned} \varphi(1) &= \eta_{20} + \eta_{02}, \\ \varphi(2) &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \\ \varphi(3) &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2, \\ \varphi(4) &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2, \\ \varphi(5) &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2], \\ \varphi(6) &= (\eta_{20} - \eta_{02})[(\eta_{30}\eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}), \\ \varphi(7) &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]. \end{aligned} \quad (5)$$

有的不变量之间有较强的相关性, 经过分析, 实际上只取四个左右就可, 如取 $\varphi(1)$, $\varphi(2)$, $\varphi(4)$, $\varphi(6)$, 这样可大地减少特征量的计算量.

实际应用时, 一般有三种不同形式的不变量, 即: 强度不变量, 这里用 $I(X, Y)$ 代表目标的强度分布; 单位不变量, 这里假设目标强度均为单位值, 即 $I(X, Y) = 1$; 边界不变量, 这里假设目标边界点强度为单位值, 而其它点的强度为零.

可以看到, 虽然强度不变量携带的目标信息量较多, 但它的计算量比较大; 边界不变量的计算量最少, 但表征目标的信息太少; 相对来说, 单位不变量比较适中. 实践表明, 用强度不变量或单位不变量进行目标识别时, 其识别概率和虚警概率相差无几.

1.4 目标的分类

经前面第一次筛选后, 目标已较少. 第二次分类采用较严格的距离不变量作为目标识别的特征量, 分类结果可靠. 为了实现目标分类, 必须预先设置各类目标足够丰富的目标数据库, 以此作为目标分类依据. 计算每个被识别目标特征量与已知目标库中各类目标的距离, 然后根据最小距离, 确定被识别目标属于哪一类. 其中距离可表达成

$$D(X, Y) = \sum_{i=1}^n |X_i - Y_i|, \quad (6)$$

其中 $X = (X_1, X_2, \dots, X_n)^T$, $Y = (Y_1, Y_2, \dots, Y_n)^T$. 用常见的 $K-NN$ 分类器, 实现目标的分类. 选取距离最小的五个库目标, 其中五个库目标中占大多数的目标类型就是被识别目标的目标类型. 目标分类完成后, 输出一个帧内各个目标的类型、目标位置、大小等数据.

1.5 帧间分析、目标预计

许多原型系统没有达到满意的系统性能,最主要的是没有利用前后关系信息,如目标分类过程只依靠描述物体的特征量,没有利用物体周围环境或在前面时刻的状态等信息,本文帧间分析和动态景物分析技术通过利用前后关系信息,进一步改善了系统性能.帧间分析是要实现的主要功能之一.它是整个系统目标分类过程的第三步,通过积累多帧图像的分类结果,计算目标概率,从而作出更可靠、更准确的目标识别.

由于系统记忆目标数据,包括运动参数等,故对再入目标和目标重叠等情况,系统能根据先前数据,作出分析、判断,从而也提高了系统的目标识别率降低了虚警率.

帧间分析采用平滑算法,它主要取决于景物的统计性(主要是目标的先验概率);目标的优选、重要性;分类器的偏差,目标识别概率.其中前两项依赖于环境,可以由操作者定义,后一项是由整个处理系统的性能决定的.

1.6 动态景物分析

动态景物分析技术对运动平台的情况特别有利,它采用符号目标匹配的算法,将每一帧里的目标(包括目标形状,强度,和位置信息)用符号来描述,然后将当前帧与前几帧进行符号目标匹配,实现图像对准.一旦相关目标被匹配,景物模式,包括运动平台及目标的运动特性,进一步被新的景物模式所代替.由于所有目标相对于景物的运动均被检测,就实现了多目标的同时跟踪,而且可以预计目标遮掩及将来的目标情况.由于记忆了过去的景物历史,景物模式可能被扩展到比现有的视场大,从而具有目标再捕获、再跟踪等功能,能弥补由于跟踪平台的运动使目标多次进出视场的情况.

先进的动态分析技术使系统在不增加硬件仅增加较少计算量的情况下,消除平台运动的影响,实现多目标的同时跟踪,并具有目标再捕获、再跟踪,预计目标遮掩等功能,同时还能改善、提高微弱运动目标的捕获能力.动态景物分析的关键技术是:包括检测和识别功能的有效的景物提取;能描述目标和平台运动、目标、背景信号、目标遮掩等情况的景物模式;当前帧与前几帧的精确匹配.

2 实验结果

由于缺乏实际的视频图像资料,而且,多目标运动的目标模拟实验装置难以实施,因此实验中采用的是事先制作的模拟图像.目标是卫星模型,通过摄录像将目标图像读入计算机,由计算机产生的模拟图像在显示终端以 NTSC 电视制式的视频信号输出送入系统,就如同直接从电视摄像机来的信号一样.这样,一旦需要就可以直接将摄像关接入目标系统.在已建立的模拟图像的基础上,通过大量的实验,分析并优化选取目标的特征量分类器的结构、目标数据库的大小等,最终使整个系统具有良好的性能,以下是部分实验结果.

2.1 目标特征量实验

为了检验系统性能而采用 6 个模拟目标,它们都是 256 灰度的图像.为证明所用目标特征量的性能,该目标被取样的 6 个不同旋转角度和 6 个不同大小比例的特征量数据如表 1 所示.从表中可以看到,我们所取的目标特征量 $\varphi(1)$ 、 $\varphi(2)$ 、 $\varphi(3)$ 、 $\varphi(4)$ 、 $\varphi(5)$ 、 $\varphi(6)$ 、 $\varphi(7)$ 等是相对目标平移、大小变化和旋转不变的量,因此,作为识别目标的依据是可靠的.

在研究选取特征量时, 我们用下式计算目标特征量之间的互相关值:

$$R = \frac{\sum_{i=1}^7 \varphi(i) \varphi'(1)}{\sqrt{\sum \varphi^2(i) \sum \varphi'^2(i)}} \quad (7)$$

实践表明: (1) 同一目标不同取样的特征量之间有很大的相关性, 一般其互相关值达到 0.99996 以上, 而不同目标的特征量之间的相关性就小得多, 而且随目标不同而变化, 要提高目标的识别能力, 实际上就是要提高这两个互相关的比值; (2) 由于特征量之间有一定的相关性, 选取适当减少的目标特征量, 其效果反而比采用所有特征量好, 选取的依据是上述两个互相关的比值. 因此, 我们实际上只采用 $\varphi(1)$ 、 $\varphi(2)$ 、 $\varphi(4)$ 、 $\varphi(6)$ 作为目标识别的特征量.

表 1 目标多次取样的特征量数值

Table 1 The characteristic values of target features samples

特征量	平均值 \pm 均方差	均方差 / 平均值
$\varphi(1)$	$(1.6093 \pm 0.00374) \times 10^{-1}$	0.00232
$\varphi(2)$	$(7.3662 \pm 0.01354) \times 10^{-4}$	0.00184
$\varphi(3)$	$(1.6001 \pm 0.01002) \times 10^{-4}$	0.00626
$\varphi(4)$	$(3.1500 \pm 0.02015) \times 10^{-4}$	0.00640
$\varphi(5)$	$(-1.0487 \pm 0.03197) \times 10^{-9}$	0.03049
$\varphi(6)$	$(-1.5680 \pm 0.01632) \times 10^{-8}$	0.01041
$\varphi(7)$	$(-2.1986 \pm 0.03128) \times 10^{-9}$	0.01422

2.2 目标识别分析实验

表 2 是图像经假目标排除、目标分类后, 其中几帧的分类结果. 其中“T”表示识别为目标, “U”表示识别为无目标, “F”表示为假目标. 这个序列说明了帧间分析、结果平滑的概念, 表中第 6 帧与第 8 帧的分类结果“U”根据前后概率分析显然应为“T”, 另外, 注意到表中假目标不是每一帧都在同一位置出现, 因此也是一个识别假目标的重要线索.

表 2 图像序列的目标分类结果

Table 2 Object classification for image sequence

		帧 数							
		1	2	3	4	5	6	7	8
目 标	a	T	T	T			T	T	T
	b	T	T	T		T	T	T	T
	c	T	T	U	T	T	T	T	U
	d	T	T	T	T	T	U	T	U
	e	F				F			F
	f	F	F	F	F	F	F	F	F

2.3 系统目标识别实验

表 3 是系统识别目标的结果, 实验中所有目标以及假目标各输入 50 个次.

表 3 目标分类结果
Table 3 Results of object classification

	a	b	c	d	e	f	N	总计
a	50	0	0	0	0	0	0	50
b	0	48	0	0	0	0	2	50
c	0	0	49	0	1	0	0	50
d	0	0	0	49	0	0	1	50
e	0	0	1	0	49	0	0	50
f	0	0	0	0	0	48	2	50
N	1	1	0	0	0	0	48	50
总计	51	49	50	49	50	48	53	350

从表 3 可知: 目标分类精度 = $339/350 = 96.9\%$, 目标识别分类虚警率 = $2/350 = 0.6\%$, 目标丢失率 = $5/350 = 1.4\%$. 通过实验, 我们证实了系统的目标识别率是高的, 并且注意到目标库还不够完善, 造成有时出现识别差错. 当然, 由于目标模拟、背景模拟的局限, 系统有待进一步的测试实验.

3 结论

本文提出的自动目标识别系统, 将先进算法和高速双微处理器有机结合, 使系统具有较好的功能. 实践证明该实验系统具有良好的多个运动目标识别、跟踪性能, 识别目标数目前为 10 个, 识别速率为 10 帧/s, 目标识别率 95% 以上, 目标虚警率很小, 达到设计要求.

致谢: 感谢匡定波、周起勃先生和其它同志的热情支持、精心指导.

参考文献

- 1 A L Gilbert et al. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1980,2:47
- 2 Bir Bhanu. *IEEE Trans. Aerosp. Electron. Syst.*, 1986,AES-22(4):364
- 3 Hu Ming Kuei. *IRE Transactions on Information Theory*, 1962,IT-8:179
- 4 DE Soland et al. *SPIE*, 1979,178:175
- 5 EM Rounds et al. *SPIE*, 1980,238:126
- 6 Yun-Kung Lin J. *SPIE*, 1981,302:117
- 7 模式识别与图像处理, 戚飞虎等译, 上海: 上海交大出版社, 1989
- 8 WK 普拉特著. 数字图像处理学, 高荣坤等译, 北京: 科学出版社, 1984
- 9 J斯克关斯基等著. 模式分类器和可训练机器, 阎平凡等译, 北京: 科学出版社, 1987

MULTI-TARGET RECOGNITION AND TRACKING SYSTEM

Wu Minming

(Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China)

Abstract: Based on the techniques of three levels classification and dynamic scene analysis, the prototype multi-target recognition and tracking system can not only segment faint moving targets from slowly varying background, but also track moving multi-target, predict impending occlusion, reacquire and track targets which wander in and out of the field of view. The system has high percentage of correct identification, and performs real-time processing. The processing rate of target classification is 10 frame/second.

Key words: target recognition, target tracking, dynamic scene analysis.