

一种高效目标搜索方法及其在图象分析中的应用

黄可为 赵雁南 孙发杰 王家钦

(清华大学智能技术与系统国家重点实验室, 清华大学计算机系, 北京 100084)

摘要 针对自然环境中(自然光照和复杂背景条件下)的特定目标模式搜索, 利用视觉神经生理学和认知心理学实验结果, 归纳出一种仿生目标搜索策略——BOES(Bionic Object Exploring Strategy)策略。该策略应用在车牌识别系统中, 设计出简捷实用的高效模式搜索算法, 解决了快速准确定位车牌的难题。

关键词 模式搜索 BOES策略 PPC特征 车牌定位

0 引言

自然环境中的特定目标搜索是目标探测、自动识别等领域的一项关键技术。在自然环境中, 目标和背景的成像条件一般是不可控制的, 随机变化的因素(尤其是光照条件)和复杂的背景信息给目标搜索带来巨大困难。不同光照下, 目标的颜色、亮度、明暗对比度都有很大变化; 背景信息往往比目标信息更加复杂, 某些背景域又可能与目标有不小的差异; 再加上观察距离、角度的不同, 要从种种干扰中区别出目标是十分困难的。现有的很多目标搜索算法, 由于其特征检测器只考虑了边缘等共性特征, 没有考虑在特定搜索环境中的空间特性, 所以效果并不理想。阈值选择困难, 难以限定空间尺度范围。实际应用中, 常常又有实时性要求; 特别是当目标在整幅图象中所占的比例比较小的情况下, 从整幅图象中寻找小目标必然要在大量的背景信息中搜索, 如果没有高效率的搜索方法, 就需要耗费很多计算时间和存储空间。

由于模式识别起源于模仿生物知觉的愿望, 图象中的目标搜索更以人类高效率的视觉系统为模仿目标, 所以在本文的研究中, 结合生物视觉的一些特性, 利用神经生理学和认知心理学实验结果, 研究了自然环境中目标模式的搜索方法, 归纳为BOES策略。现在, 成功的视觉系统或图象分析系统, 大都遵循了这个原则。这里作者明确提出来, 并有意识地作为指导性原则应用于工程实践, 取得了良好的效果,

证明了这个策略是行之有效的。该策略在一个车牌识别系统中得到验证。本文将自然环境下车牌的自动定位为应用背景, 介绍这种模式搜索策略及其实现过程。

1 某些生物视觉特性的启发

(1) 选择性和主动视觉。生物视觉的最显著的特点之一是有选择性, 即定向主动性, 观察者的注意力总是有目的地指向最感兴趣的东西。一个成功的视觉系统往往不能缺少视觉目的的引导。

(2) 恒常性模式特征。图象中各点的灰度是多种因素综合作用的结果。这些因素的任何变化都会改变图象的灰度分布。但是人眼仍然能够从图象中分辨出各个物体。可见, 人类视觉具有某种恒常性。Marr理论^[1]认为, 这种恒常性的产生, 是因为大脑不但把点状的传感信息聚集成整体, 而且经过了一个因素分解过程; 然而, 工程实现因素分解并不是那么容易的。我们认为, 在一定条件下, 恒常性不必经过因素分解就可以获得。人类视觉恒常性的产生, 是物体可变的外表后面不变的恒常性模式特征——PPC(Pattern Permanent Characters)特征标识作用的结果。PPC特征是对随机因素不敏感的特征。选用目标模式的PPC特征, 是最关键的问题。

(3) 特征检测器。Hubel和Wissel发现了动物大脑皮层视区中的神经元感受野特性: 可以检测在特定位置、特定方位(最适方位)的直线或边缘, 可以敏感直线和边缘的方向和运动, 并且作为刺激的直

线的长度和边缘的宽度都有最适值,即在动物的视觉系统中存在某种形式的特征检测器。

(4) 组织归纳图象数据。格式塔心理学家发现,人具有对图象数据进行组织归纳的能力,可以在多个层次上发现图象数据的规则性、一致性和连续性等空域和频域整体特性。人类视觉的这些特性,可以看作是空域滤波和频域滤波作用的结果。

(5) 分层结构。人类视觉系统的一个重要特点是其信息加工是分层分级进行。从一幅图象中认出物体,就需要低层首先分辨出表示物体的基本元素,如明暗、颜色、纹理、轮廓等;高层再根据以往的知识 and 经验,经过分析推理识别物体的类别。

受上述生物视觉特点的启发,要求目标模式的搜索算法:(1) 要模仿生物视觉的定向主动性,具有强烈的目的性和特征选择性;(2) 抓住模式的 PPC 特征,设计 PPC 检测器对图象滤波,屏蔽背景干扰信息;(3) 使用空域滤波和频域滤波技术,利用目标



图 1 自然环境下拍摄的车辆图象

2.1 车牌 PPC 特征检测器与图象背景信息屏蔽

BOES 策略的原则(1)和(2)指出,要针对目标的特点,抓住目标模式的 PPC 特征,设计 PPC 特征检测器对图象滤波,屏蔽背景干扰信息。由于图象的边缘处包涵了丰富的信息,也是最基本的特征,所以我们猜测,人眼在图象中寻找车牌时可能运用了类似边缘的模式特征。然而,各种通用的边缘特征检测器,由于只考虑了边缘的共性,没有考虑在特定环境中边缘所处空间的特征,所以在检测车牌区域的边缘时,效果并不理想。图 2(b)、(c)、(d)分别显示了 Sobel 算子、Laplacian 算子和 Marr-Hildreth 算子等经典边缘检测算子的处理结果。

图 2 示例表明,通用特征检测器不适用于特定目标的搜索。对于特定的车牌模式搜索任务,人眼不是寻找普通意义的边缘,而是具有 PPC 特征的边缘,即具有车牌区域空间分布特点的边缘。对于车牌检测,需要针对车牌的局部模式特性,设计 PPC 特征检测器。在下面的讨论中,我们称之为 PPC 滤波器(Pattern Permanent Characters filter)。

如图 1 所示,虽然车速、车辆种类、车牌位置有

模式的整体特性进一步组织归纳图象数据;(4) 应当将任务分成若干层次,不同层次完成不同功能,由粗到精地逐步排除干扰,实现目标模式的快速准确定位。这 4 个原则,有机地构成了一种搜索策略,我们称之为仿生目标搜索策略——BOES 策略。下面以车牌识别系统中的车牌定位算法为例,说明该策略的实现过程。

2 采用 BOES 策略的车牌定位算法

要在自然环境下摄取的车辆图象中识别车牌,车牌定位是难点。不同的自然光照条件、车牌与摄像机之间的距离和角度时常变化、车体和周围环境构成的较复杂背景信息等因素(见图 1),给车牌定位的精度和速度都带来很大影响。基于 BOES 策略的车牌定位算法,可以在自然环境中快速准确地定位车牌。

所不同,车牌与摄像头的距离和角度有一定变化,但是车牌搜索属于大视野中小目标的搜索,在物距变化不大的情况下(1m~2m),车牌大小相差很少,特别是车牌字符的笔画宽度变化很小,而且车牌区域局部背景较单纯,车牌区域字符具有中央亮周边暗的象素分布,这些字符的笔画,就是由具有这样特点的象素聚集而成。这种特征实际上就是车牌区域字符点的一种 PPC 特征。将具有这种局域特征的点作为候选点从图象中抽取出来,可以滤掉大量的背景干扰信息,实现背景信息屏蔽,为车牌准确定位奠定基础。另外,前面分析已经表明,不限定特征尺度,边缘抽取就会将各个尺度的边缘都抽取出来,搀杂进大量干扰信息。在图象目标搜索时,特征尺度表现为分辨率设定和特征模板尺寸的选择。车牌的 PPC 滤波器就是针对其特征尺度设计的。在实验设备设定参数的控制下,图象中车牌字符的笔画宽度基本上是 3—5 个象素。鉴于此,特征尺寸控制为 9×9 象素。

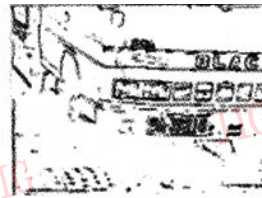
车牌的 PPC 滤波器由空间特征算子组 DC (Dimensional Characteristic Detector Combination)及其对应的条件算子组 COC (Conditional Operators Combination)构成。



(a) 原始图象



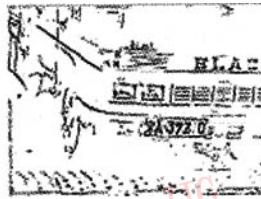
(b) Sobel 算子处理结果



(c) Laplacian 算子+局部灰度方差统计



(d) Marr-Hildreth 算子+局部灰度方差统计



(e) PPC 滤波器处理结果

图2 PPC 滤波器和传统边缘检测算子的处理结果

(1) 1型空间特征算子组 DC_1 及其条件算子组

COC_1

$$DC_1 = \{DC_1^{\alpha_i} | \alpha_i = 45^\circ \times i, i = 0, 1, 2, 3\}$$

$$DC_1^{0^\circ} = \frac{3}{4} \times \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad DC_1^{45^\circ} = \frac{3}{4} \times \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$DC_1^{90^\circ} = \frac{3}{4} \times \begin{bmatrix} 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad DC_1^{135^\circ} = \frac{3}{4} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

由于人眼视觉对空间某位置的感觉来自其局部范围的总光量,而不是每一点的平均光强,所以上面的各个空间特征算子是求加权和而不是求加权均值。

记1型算子目标候选像素点的集合为 S_1 。像素点 $P(x, y)$ 处象素的灰度值为 $f(x, y)$ 。记 $P(x, y)$ 周围灰度分布矩阵为:

$$D(x, y) = \begin{bmatrix} f(x-1, y-1) & f(x, y-1) & f(x+1, y-1) \\ f(x-1, y) & f(x, y) & f(x+1, y) \\ f(x-1, y+1) & f(x, y+1) & f(x+1, y+1) \end{bmatrix}$$

$COC_1 = \{COC_1^{\alpha_i} | \alpha_i = 45^\circ \times i, i = 0, 1, 2, 3\}$, 其中, $COC_1^{\alpha_i}(P(x, y)) \in \{\text{TRUE}, \text{FALSE}\}$

$$COC_1^{0^\circ}(P(x, y)) = (DC_1^{0^\circ} \otimes D(x, y) - DC_1^{0^\circ} \otimes D(x-3, y) > \text{Threshold}) \wedge (DC_1^{0^\circ} \otimes D(x, y) - DC_1^{0^\circ} \otimes D(x+3, y) > \text{Threshold})$$

$$COC_1^{45^\circ}(P(x, y)) = (DC_1^{45^\circ} \otimes D(x, y) - DC_1^{45^\circ} \otimes D(x-3, y+3) > \text{Threshold}) \wedge (DC_1^{45^\circ} \otimes D(x, y) - DC_1^{45^\circ} \otimes D(x+3, y-3) > \text{Threshold})$$

$$COC_1^{90^\circ}(P(x, y)) = (DC_1^{90^\circ} \otimes D(x, y) - DC_1^{90^\circ} \otimes$$

$$D(x, y-3) > \text{Threshold}) \wedge (DC_1^{90^\circ} \otimes D(x, y) - DC_1^{90^\circ} \otimes D(x, y+3) > \text{Threshold})$$

$$COC_1^{135^\circ}(P(x, y)) = (DC_1^{135^\circ} \otimes D(x, y) - DC_1^{135^\circ} \otimes D(x-3, y-3) > \text{Threshold}) \wedge (DC_1^{135^\circ} \otimes D(x, y) - DC_1^{135^\circ} \otimes D(x+3, y+3) > \text{Threshold})$$

$$\text{像素点 } P(x, y) \in S_i \iff \bigvee_{\alpha_i} COC_1^{\alpha_i}(P(x, y)) =$$

TRUE

(2) 2型空间特征算子组 DC_2 及其条件算子组 COC_2 与1型类似,只是空间特征算子组改为:

$$DC_2^{0^\circ} = DC_2^{45^\circ} = DC_2^{90^\circ} = DC_2^{135^\circ} = \frac{9}{16} \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

$$\text{总目标候选象素点集合 } S = \bigcup_i S_i.$$

其中,每个类型的算子组中 Threshold 是相同的。这里的阈值是图象字符最低对比度阈值,低于此阈值的像素点算法将不予考虑。这样做是模仿了人眼的分辨特性:在明暗对比度低到一定程度时即分辨不清。

上述算子组的几何意义是很明显的:1型算子能较好地抽取车牌模式边缘,但是滤噪能力不强,所以其阈值 Threshold 取得较高,这样做的结果是一些目标点未被抽出;2型算子滤噪能力很强,但是对笔画较为细密的车牌字符点(例如字母、汉字)抽取性能不大好。1型和2型配合起来效果就相当好。从图2(e)示例可以看出,使用PPC滤波器可以将车牌模式特征边缘点基本检出,屏蔽大量背景信息,处理效果明显优于其他边缘检测器。

我们将 PPC 各个算子中的阈值设置为所能承受的最低对比度下的阈值,其负面影响是其它区域的候选点增多,这种影响可以通过后续的滤波步骤加以消除。

下面以图 3 为例演示车牌定位的几个连续处理步骤。PPC 滤波处理结果见图 3(b),其中,黑色点为

本步骤选出的候选点。由图 3(b)可见,具有车牌字符模式特征的象素点基本上被抽取出来,而大量的背景信息被屏蔽掉了。PPC 滤波器得到的是二值滤波图象,大量背景信息的屏蔽,节省了后续处理的计算时间和存储空间,适应了实时性要求。

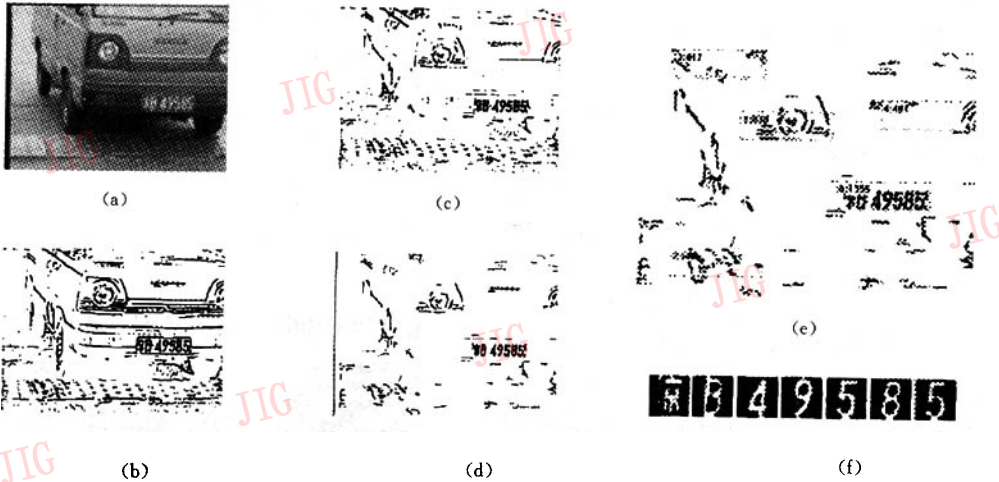


图 3 车牌定位各个步骤处理结果

2.2 空域滤波和频域滤波

BOES 搜索策略原则(3)指出,可以使用空域滤波和频域滤波技术,利用目标模式的整体特性进一步组织归纳图象数据。模仿视觉的空域滤波特性,我们设计了不同空间尺度的滤波器:长程滤波与颗粒滤波,模仿视觉的频域滤波特性,设计了频域滤波器。

3.2.1 空域滤波:长程滤波与颗粒滤波

(1) 长程滤波

从图 3(b)中看到,很多候选点连缀成长程带状曲线。而车牌区域不会有这种情形。如果不作相应处理,就会给车牌定位带来干扰。因此,在定位前应当先作长程滤波。

$\forall P(x_i, y_i) \in S$, 初始化长线集合 $L(x_i, y_i) = \{P(x_i, y_i)\}$, 记集合

$$B(x, y) = \left\{ \begin{array}{l} P(x-1, y-1), P(x, y-1), P(x+1, y-1), \\ P(x-1, y), P(x, y), P(x+1, y), \\ P(x-1, y+1), P(x, y+1), P(x+1, y+1) \end{array} \right\}$$

$$L(x_i, y_i) = \{P(x, y) | (P(x, y) \in S) \wedge (B(x, y) \cap L(x_i, y_i) \neq \emptyset)\}$$

上式给出了每一个线集合 $L(x_i, y_i)$ 的膨胀算法。长程滤波算子对于线集合 $L(x_i, y_i)$ 的空间尺度

的容忍度由下面两个式子定义:

$$Dimx(x_i, y_i) = \max\{d_x | \forall P(x, y) \in L(x_i, y_i), d_x = |x - x_i|\}$$

$$Dimy(x_i, y_i) = \max\{d_y | \forall P(x, y) \in L(x_i, y_i), d_y = |y - y_i|\}$$

定义长程滤波算子 $Ero_{long}(S)$:

$$Ero_{long}(S) = S - L(x_i, y_i), \forall P(x_i, y_i) \in S,$$

$$Dimx(x_i, y_i) > Thresh_x \vee$$

$$Dimy(x_i, y_i) > Thresh_y$$

长程滤波算法基于这样的思想:如果许多点连缀成一个长程曲线 L (如上面图中显示的那样),那么长程滤波器考察 L 在 x 方向和 y 方向的空间延展程度 $Dimx$ 和 $Dimy$ 。如果超过阈值,即将所有属于 L 的候选点从候选点集合 S 中滤掉。处理结果见图 3(c)。

(2) 颗粒滤波

长程滤波之后,图象中仍有许多候选点集结成小型的彼此分离的颗粒状,而车牌区的候选点很少有这种情况。人眼在寻找车牌模式时,对这种颗粒状的小尺度象素点聚集应该有过滤作用。模拟这种特性,进行了颗粒滤波处理。

定义集合 $B_{m \times n}(x, y) =$

$$\left\{ P\left(x - \frac{n-1}{2}, y - \frac{m-1}{2}\right), \dots, P\left(x + \frac{n-1}{2}, y - \frac{m-1}{2}\right) \right\}$$

$$\left\{ P\left(x - \frac{n-1}{2}, y + \frac{m-1}{2}\right), \dots, P\left(x + \frac{n-1}{2}, y + \frac{m-1}{2}\right) \right\}$$

$$B_{granule}(x, y) = B_{m \times n}(x, y) - B_{(n-1) \times (m-1)}(x, y)$$

定义颗粒滤波算子 $Ero_{granule}(S)$:

$$Ero_{granule}(S) = S - B_{(n-1) \times (m-1)}(x, y),$$

$$\forall P(x, y) \in S, B_{granule}(x, y) \cap S = \phi$$

颗粒滤波基于这样的想法:对于某一个候选点颗粒状集合,在它的周围,一定存在一个小矩形区(矩形区的大小与所考察的颗粒粒度有关),这个小矩形区的四周,包含非常少的候选点(理想情况是没有候选点,如 $Ero_{granule}(S)$ 的定义式中描述的那样)。那么,对全图进行小矩形区扫描,检查矩形区周边上的候选点个数是否小于阈值。如果小于,则认为该矩形区内的所有候选点,形成了孤立的颗粒状集合,抹去这些点,就达到了颗粒滤波的目的。由于颗粒大小不一,可以进行若干次不同粒度的颗粒消抹。处理结果见图 3(d)。

2.2.2 频域滤波:候选区选择

具有车牌模式特性的像素在图象中的分布频度不同,在车牌区域特别集中,这种聚集性表现了一种频域特性。模仿视觉频域滤波特性,我们设计了频域滤波器。

首先定义有限集合 B 的“度” $Deg(B)$ 为 B 中元素的个数。

定义频域滤波算子 $Cym(S)$:

$$Cym(S) = \{ B_{m \times n}(x, y) | Deg(B_{m \times n}(x, y) \cap S) > ThreshCondition \}$$

$Cym(S)$ 得到一组集合,每一个集合代表了图象中候选点最为集中的一个区域,作为车牌的一个候选区域。定义式中的 $ThreshCondition$ 不是一个单纯的阈值,而是集合 $B_{m \times n}(x, y)$ 满足的一组位置和度的比选条件。其中,度比选条件起到了频域滤波作用;位置比选条件则避免了几个候选区聚集在非常临近的地方,防止造成错误的引导。

图 3(e) 中,5 个虚线方框代表 5 个选定的候选区。车牌区域通过对几个候选区的进一步比较得出。

候选区的大小取决于车牌区在图象中所占的比例和算法性能要求。我们将候选区尺度设定为车牌尺度的 1.5—2 倍,这样能够较快地确定候选区;再根据

推理规则,对几个候选区进行筛选,选出车牌所在的区间,实现大尺度的车牌粗定位;而后在候选区内部进行车牌的精确定位。这样,按照 BOES 策略原则(4)的要求,有针对性地突出车牌,弱化背景干扰,将不可能区域逐步排除,由粗到精地将车牌位置勾勒出来。图 3(f) 为车牌精确定位后字符切分的结果。考虑到论文篇幅限制,字符切分算法这里不讨论了。

2.3 实验结果

我们在阴天、晴天、早、午、晚和不同角度随机拍摄了 200 幅样本(分辨率为 480×360 像素),包括背光、反光等各种光照条件。在 Pentium II 300(64M) 上实现该算法,0.5 秒内完成定位任务,测试结果为定位准确率 96%。如图 4 所示。



图 4 一些处理实例

错误定位主要是由特殊的光照条件引起的,车牌背光、反光都会影响定位精度。对于这种情况,我们认为需要通过改善光学系统性能来补偿。

3 结 论

本文针对自然环境中特定目标模式搜索,利用视觉神经生理学和认知心理学实验结果,归纳出仿生目标搜索策略——BOES 策略。该策略应用在车牌识别系统中,设计出简捷实用的模式搜索算法,解决了快速准确定位车牌的难题。本文还提出了恒常性模式特征——PPC 特征的概念;在车牌模式定位问题中,通过 PPC 滤波器、空域和频域滤波等技术,具体实现了 BOES 策略。本文讨论的思想方法,很有实用价值。

参 考 文 献

1 Marr D. Vision. W. H. Freeman and Company, 1982.

- 2 荆其诚等. 人类的视觉. 北京:科学出版社,1987.
- 3 崔 屹. 数字图象处理技术与应用. 北京:电子工业出版社, 1997.
- 4 王润生. 图象理解. 长沙:国防科技大学出版社,1995.
- 5 刘立峰. 基于物理模型的颜色恒常性研究及应用[学位论文]. 清华大学,1996.
- 6 Yu Zhong, Kalle Karu, Anil K J. Locating text in complex color images. *Pattern Recognition*, 1995,28(10):1523~1535.



黄可为 清华大学智能技术与系统国家重点实验室硕士研究生。目前研究方向为图象分析,模式识别等。

赵雁南 清华大学智能技术与系统国家重点实验室副教授。主要研究方向为智能机器人,模式识别,虚拟现实等。

孙发杰 清华大学智能技术与系统国家重点实验室助教。目前研究方向为智能机器人控制器,虚拟现实等。

王家钦 清华大学智能技术与系统国家重点实验室教授,博士生导师。主要研究方向为人工智能,虚拟现实,智能机器人等。

An Efficient Object Exploring Strategy and Its Application in the Image Analysis

Huang Kewei, Zhao Yannan, Sun Fajie and Wang Jiaqin

(State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

Abstract Pattern recognition is derived from the will to imitate the sense of the creatures, object searching in an image aims at imitating the human's visual ability. In this paper, utilizing the results of the vision neurophysiology and the psychology of cognizance, and aiming at the specific pattern location in the natural environment, we put forward a Bionic Object Exploring Strategy—BOES. In a license plate recognition system, an algorithm of succinctness and practicality on this strategy has developed to solve the difficulty of locating license plates rapidly and precisely.

In our algorithm, to deal with the complex, uncontrollable background of the natural environment and the variability of the color, brightness and size, we extract the license plate pattern's most permanent characters, design a set of filters to mask the background, organize the image data in a series of steps by spatial and frequencial filtering, and divide the task into several layers. This system can locate the license plate in an image of complex environment within 0.5 second on a Pentium II 300Hz machine, and the rate of success is over 95%.

Keywords Pattern location, BOES strategy, PPC character, License plate location