文章编号:1001-9081(2011)01-0266-04

doi:10.3724/SP. J. 1087.2011.00266

# 交通标识牌字符提取算法

## 王 振,魏志强

(中国海洋大学 计算机科学与技术系,山东 青岛 266100) (wangoucqd@gmail.com)

摘要:为解决智能交通系统中道路标识牌字符提取问题,提出了一种快速的基于颜色与笔画的新算法。首先采用主元分析(PCA)方法提取标识牌颜色特征并进行定位;然后对确认后的标识牌区域进行仿射处理,获得容易进行文字提取的图像;最后根据形态学的 top hat、skeleton 算子以及区域生长等算法得出道路标识牌字符清晰的二值化图像,送光学符号识别(OCR)软件识别。实验结果显示该算法具有很强的准确性和鲁棒性。

关键词:主元分析;仿射变换;笔画算子;区域生长

中图分类号: TP391.41 文献标志码: A

## Text extraction algorithm for traffic signs

WANG Zhen, WEI Zhi-qiang

(Department of Computer Science and Technology, Ocean University of China, Qingdao Shandong 266100, China)

Abstract: A fast and robust approach for the text extraction of traffic signs based on color and stroke was proposed. First, a new color model derived from Karhunen-Loeve (KL) transform was applied to find all possible traffic sign candidates. Then, affine transformation was performed to restore traffic signs to let every road sign seem to be vertical to the camera optical axis which can improve the accuracy in detecting texts embedded in traffic signs. Finally, mathematical morphology and region growing algorithms were used to obtain a clearer binary picture which was sent to Optical Character Recognition (OCR) software. The experimental results demonstrate great robustness and efficiency of the proposed algorithm.

Key words: Principal Component Analysis (PCA); affine transformation; stroke-based algorithm; region growing

### 0 引言

近年来,随着交通运输业的飞速发展,交通事故已成为当前世界各国所面临的严重问题。车辆导航和车辆辅助驾驶系统目前正成为基于视觉的现代智能交通研究热点之一。该研究主要是利用传感器技术、计算机技术、通信技术等先进手段,掌握驾驶员、道路以及周围车辆的状态信息,并做出综合的分析和判断,以便在存在安全隐患的情况下给司机发出劝告和报警,提醒司机注意规避危险。

本文将道路环境中蓝底白字标识牌文字的定位与提取作为研究重点。道路标识牌字符可以告诉驾驶员当前车辆大致位置信息,确定正确行使道路等,而道路环境信息提取是车辆导航和车辆辅助驾驶的先决条件。

字符定位与分割算法广泛地应用于车牌识别和基于内容的图像检索中,现有的方法大致分为三类:基于连通域的方法、基于纹理的方法和基于边缘的方法<sup>[1]</sup>。这些定位方法在基于静态图片和视频帧的文字提取方面取得较好的效果,但是在道路环境中噪声较多的情况下,错误率较高。为了适应在道路环境中定位道路标识牌字符,本文提出了一种快速鲁棒的交通标识牌字符提取算法。

#### 1 交通标识牌检测定位算法

每一个交通标识牌通常都有一种特定的颜色,而且与背景具有较大的对比度。这种颜色信息能够有效地用于减小标识牌的定位区域。

图像处理中最常用的是三基色 RGB 色彩空间,即以红(R)、绿(G)、蓝(B)三种颜色作为三基色,再通过三基色的加权混合形成各种颜色。RGB 模型是构成颜色表示的基础,其他颜色表示方法可以通过对 RGB 模型的变换得到。虽然交通标识的颜色是确定的,但图像中的目标亮度却随着照射光源的变化差异很大。在 RGB 空间中,由于对于颜色的亮度和色度的度量不是独立的,所以采用 RGB 空间做颜色搜索是不合适的。为了将亮度与色彩彻底分离,研究者们更多采用HSV(Hue 色调或色彩,Saturation 色饱和度,Value 明度值)空间或者用 intensity, brightness, lightness 等代替 value,得到HIS、HSB、HSL 等空间模型。但上述颜色空间对于标识牌的准确定位效果并不理想。本文利用了基于主元分析的颜色模型方法,有效地进行了标识牌的初始定位。

#### 1.1 基于主元分析的标识牌特征提取及定位

主元分析(Principle Component Analysis, PCA)是一种基于目标统计特征的最佳正交变换, 称其为最佳变换是由于它具有重要的优良性质:

- 1) 使变换后产生的新的分量正交或不相关;
- 2) 以部分新的分量表示原矢量,并使其均方误差最小;
- 3)使变换矢量更趋于确定、能量更集中等,这使它在特征选取、数据压缩等方面都有着极为重要的应用。

首先,本文在不同的光照条件和气候条件下,收集 298 张的道路牌和非道路牌影像,选择 RGB 颜色空间,某图片中第 *i* 个像素可以表示为:

$$x_i = \left[ x_{i,r} x_{i,g} x_{i,b} \right] \tag{1}$$

收稿日期:2010-07-07;修回日期:2010-08-18。 基金项目:山东省中青年科学基金资助项目(2005BS011001)。

作者简介:王振(1971 - ),男,山东青岛人,副教授,博士,主要研究方向:计算机图像处理、模式识别; 魏志强(1968 - ),男,山东青岛人,教授,博士生导师,主要研究方向:计算机图像处理、普适计算。

通过统计 298 张训练样本的 R、G、B 通道的颜色分布,进而得到相对应的协方差矩阵  $\Sigma$ 。利用 Karhunen-Loève 变换,首先得到协方差矩阵  $\Sigma$  的特征值  $\lambda_i$  和特征向量  $e_i(i=1,2,3)$ 。其中  $\lambda_1 \ge \lambda_2 \ge \lambda_3$ ,然后利用特征向量对 R、G、B 分量进行线性组合形成三个新的颜色特征:

在实验中利用 Karhunen-Loève 变换得到三个特征向量  $e_i(i=1,2,3)$  分别为: (0.3396,0.3392,0.3212), (0.4896,0.0923,-0.4181), (0.2898,-0.4823,0.2279)。

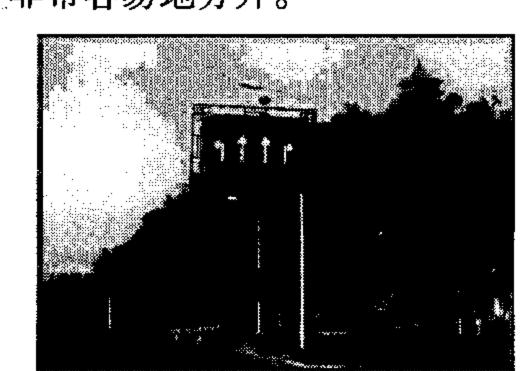
与最大特征值相对应的颜色特征  $C_1$  用于颜色灰度的变换,约为:

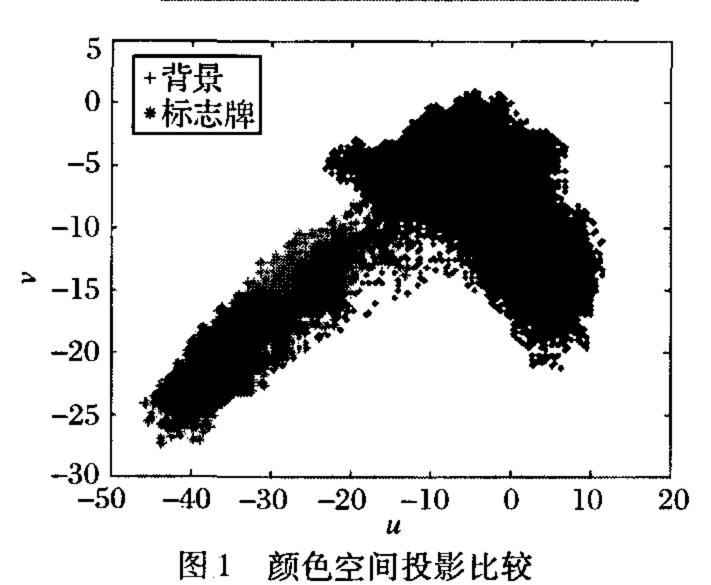
$$C_1 = \frac{1}{3}R + \frac{1}{3}G + \frac{1}{3}B \tag{3}$$

颜色平面(u,v)垂直于轴(1/3,1/3,1/3),由其他两个特征向量张成,定义如下:

$$u = \frac{2(R-B)}{5}, v = \frac{R+B-2G}{5} \tag{4}$$

图1显示了交通标识牌与非交通标识牌像素的投影结果,\*号区域表示对蓝色交通标识牌的二次投影结果,+号区域区域为背景像素投影结果。本文对树、天空的区域与蓝色标识牌进行投影比较,可以发现在(u,v)空间上通过合适的分类器能将二者非常容易地分开。





本文使用简单的径向基函数(Radial Basis Function, RBF)神经网络进行分类工作。图 2 是部分的训练样本图片。

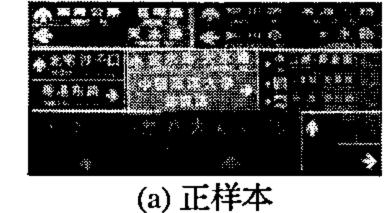




图 2 部分训练样本

#### 1.2 候选区域的筛选

由上节所得到的道路牌候选区并不一定全是需要的,如图 3(b)所示,少量的树叶或是车灯等均会造成框选到不正确的候选区。所以利用简单的几何判断式作为过滤。首先,先假设候选区的长和宽分别为 L、W,而 E 为候选区内所含有的边缘信息的总数。依据此假设,对于影像大小 640 × 480,不满

足下列条件的区域者则会将其滤除:

1)长宽比例不合适:

$$L \times 5 \leq W, W \times 5 \leq L$$
 (5)

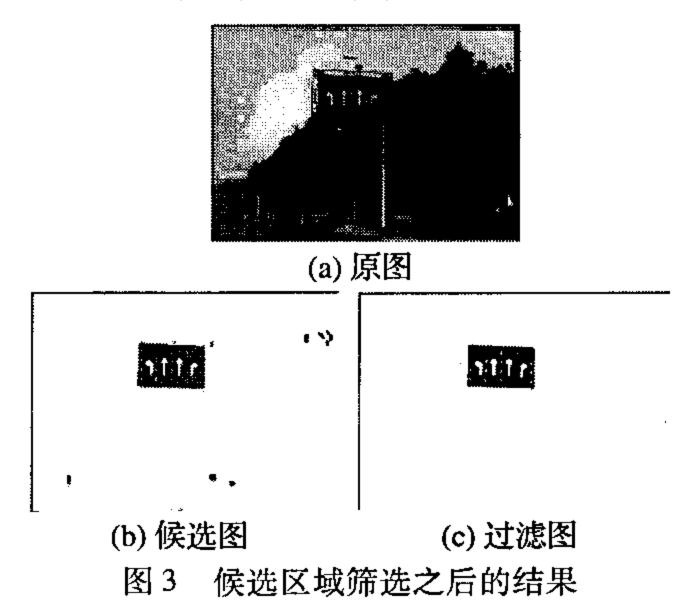
2)区域过小:

$$L < 10, W < 10$$
 (6)

3)边缘像素比例过少:

$$E/(L \times W) < 0.02 \tag{7}$$

条件1)是因为每一种道路牌的长宽比例是固定的,在考虑倾斜的情况下,根据所能取得的实验资料做分析,所以订定长宽比限制在5.0以内,若超过此比例,则将其滤除。条件2)是为了滤除掉树叶或是车灯等所造成的少量噪声。条件3)是由我们所要求的道路牌内文字来取得信息,文字能产生边缘信息的特性,而道路牌内也含有大量文字,所以若候选区内所含有的边缘信息不足,则表示此候选区并不含文字或是因为候选区距离相机太远的机会很大,所以我们也将其滤除。候选区经过筛选的结果如图3(c)所示。



上述三个条件均满足时,则进行区域的跟踪及仿射处理。 标识牌的跟踪采用比较简单的覆盖率的方法。只要两个区域 的重合面积覆盖率大于 0.6 即可。当跟踪对象到达图像上边 缘或右边缘时,根据车速的高低,在高速路上取倒数第 10 帧 中对象进行仿射变换处理以及文字定位;在市区车速低于 60 km时,取倒数第 6 帧中对象进行仿射变换处理以及文字定 位。

## 2 仿射处理

使用摄像机或数码相机拍摄的图像会因为相机主轴和文字所在的平面不垂直的原因而发生仿射形变。这种形变会极大的影响到文字行的识别效果,因此需要进行形变的恢复。

为了将发生了仿射形变的文字从图像平面恢复到文字的真实平面,以前的研究需要利用摄像机参数和匹配点信息。本文尝试了使用平面间的 Homography 操作进行平面仿射变换,避免了在文字恢复时依赖于相机的参数,从而会给实际应用带来更大方便。对于平面间的 Homography 运算,需要 4 个以上的匹配点(任意三点不共线)就可以将一个平面中的点投影到另一个平面。

假设相机是小孔相机,图像坐标系是 o-xyz,世界坐标系为 O-XYZ,则图像坐标和平面的三维实际坐标之间存在如下关系:

$$\begin{bmatrix} x' \\ y' \\ h \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \\ H_{41} & H_{42} & H_{43} & H_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$
(8)

(12)

其中:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{h} \begin{bmatrix} x' \\ y' \end{bmatrix} \tag{9}$$

其中的平面转换、旋转等投影变换完全有一个矩阵 H 实现, 即 Homography 矩阵。根据式(8)可以得到:

$$XH_{11} + YH_{12} + ZH_{13} + H_{14} - XxH_{31} - YxH_{32} - ZxH_{33} - xH_{34} = 0$$

$$XC_{21} + YH_{22} + ZH_{23} + H_{24} - XyH_{31} - YyH_{32} - ZyH_{33} - yH_{34} = 0$$

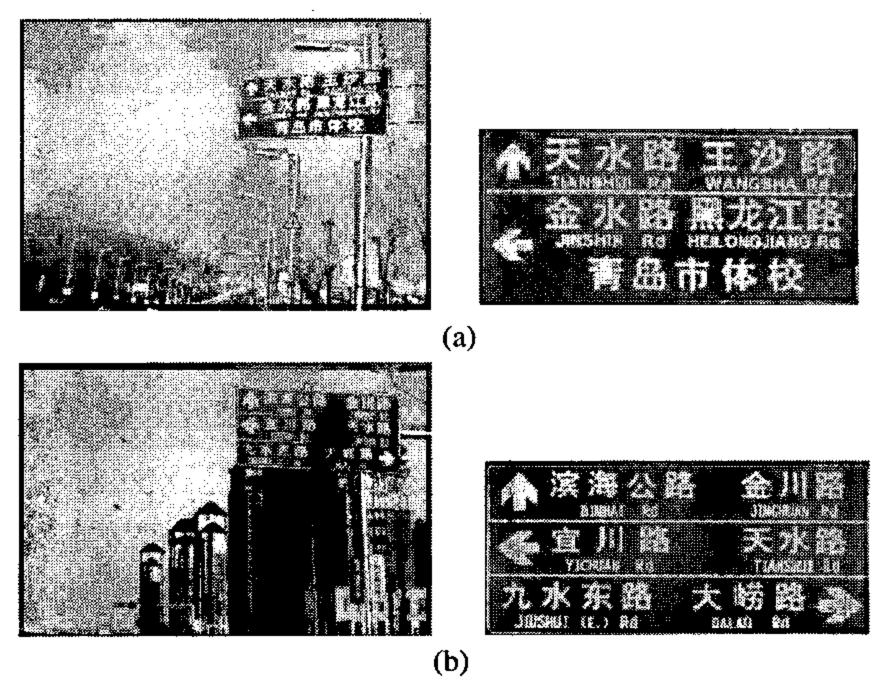
$$(10)$$

因为文字位于平面上, 式中 Z 坐标的数值在世界坐标系 中始终为零。在假设系数的情况下,得到:

$$\begin{bmatrix} X^{1} & Y^{1} & 1 & 0 & 0 & 0 & -X^{1}x^{1} & -Y^{1}x^{1} \\ 0 & 0 & 0 & X^{1} & Y^{1} & 1 & -X^{1}x^{1} & -Y^{1}x^{1} \\ X^{2} & Y^{2} & 1 & 0 & 0 & 0 & -X^{2}x^{2} & -Y^{2}x^{2} \\ 0 & 0 & 0 & X^{2} & Y^{2} & 1 & -X^{2}x^{2} & -Y^{2}x^{2} \\ X^{3} & Y^{3} & 1 & 0 & 0 & 0 & -X^{3}x^{3} & -Y^{3}x^{3} \\ 0 & 0 & 0 & X^{3} & Y^{3} & 1 & -X^{3}x^{3} & -Y^{3}x^{3} \\ X^{4} & Y^{4} & 1 & 0 & 0 & 0 & -X^{4}x^{4} & -Y^{4}x^{4} \\ 0 & 0 & 0 & X^{4} & Y^{4} & 1 & -X^{4}x^{4} & -Y^{4}x^{4} \end{bmatrix} \begin{bmatrix} H_{11} \\ H_{12} \\ H_{14} \\ H_{21} \\ H_{22} \\ H_{24} \\ H_{31} \\ H_{32} \end{bmatrix} = \begin{bmatrix} x^{1} \\ y^{1} \\ x^{2} \\ y^{2} \\ x^{3} \\ y^{3} \\ x^{4} \\ y^{4} \end{bmatrix}$$

通过给定的 4 个不共线的对应点 $\{(X^1,Y^1) - (x^1,y^1)\}$ ,  $\{(X^2,Y^2)-(x^2,y^2)\},\{(X^3,Y^3)-(x^3,y^3)\},\{(X^4,Y^4)-(x^3,y^3)\}\}$  $(x^4,y^4)$  就能求出式(12) 中的未知系数。如果有多于4个的 对应点,可以式(12) 置为超定方程进行求解。在求得所有系 数以后,则可以将图像平面 o-xyz 内的文字平面的像素点投影 到其原来的真实平面。

仿射变换后的部分结果如图 4 所示。



仿射变换结果

### 标识牌文字定位与提取

在经过标识牌定位以及仿射变换处理之后,由于标识牌 绝大多数有两色构成:前景文字是白色;背景为蓝色的。所以 可以通过灰值化处理,再经过二值化处理,大部分就可以送 OCR 软件识别,如图 5(b)所示。

但是我们发现笔画较密集的文字由于分辨率和对比度的 原因容易造成笔画粘连,如"黑"、"路"、"青"等,比较常用的 几种二值化方法以及基于颜色的 K 均值聚类方法,结果均不 理想(结果见实验与分析部分)。

由于字符笔画与背景的区别不仅体现在灰度级、颜色上,

还体现在局部的几何结构上。因此,本文提出了一种基于笔 画算子的种子区域增长方法来进行文字的分割。其实在此之 前已有许多学者都试图通过建立笔画模型来进行字符图像的 分割,几种比较好的笔画检测方法有: (Logical Level Technique,LLT)<sup>[2]</sup>、filiformity<sup>[3]</sup>和双边缘技术<sup>[4]</sup>。这类算法 在背景复杂时不能很好地解决字符图像分割的问题,但是当 背景简单时却能有效地削弱图像中大面积的背景,高效地实 现文字的分割问题。

第31卷

本文利用形态学的结合笔画宽度信息的 TopHat 变换,有 效地检测出图像中笔画的双边缘,而且相对于上述的笔画检 测算法,其检测效果更好,实现更简单。根据边缘像素的比例 到达2%时,标识牌文字笔画的宽度一般不超过5个像素,所 以设置 W = 5 来提取标识区文字。

通过上述笔画提取,如图 5(c)所示,可以清楚地发现大 量的背景像素得到抑制,而大于宽度为5个像素的笔画也被 抑制,通过大津率阈值二值化可以得到笔画特征二值化图,再 利用形态学的骨化算子来获取安全的区域增长种子点,如图 5(d)所示,在基于颜色距离的区域增长算法里,先将 RGB 颜 色空间转换为前文介绍的经过 KL 变换的 u-v 颜色空间,颜色 距离域值可以通过实验比较容易地设置,为避免笔画粘连,应 将域值设置较低。

在进行区域生长算法之前,首先在骨化特征图中进行连 通域的分析,将连通域中长度大于标识牌宽度 1/4 的删除,高 度大于标识牌高度的 1/4 也被删除。通过以上步骤,可以有 效滤除掉标识牌中的长线。然后根据文献[5]介绍的基于边 缘像素的区域增长方法,只是本文改为基于骨架像素的区域 增长方法对文本进行预定位,根据笔画算子的特点,只有笔画 宽度小于制定阈值的才被保留,所以,标识牌当中的大箭头在 笔画特征二值化图中只有箭头的少量部分保留,如图 5(c)所 示,在骨化图中密度相对于文字部分密度较少,因此,通过该 步骤,可以将标识牌中的箭头部分滤除,不进行文字的识别。 当然,箭头对于方向的指示具有识别的意义,但在本文中不做 进一步的研究,这也是我们以后需要进一步改进的地方。区 域增长的结果如图 5(e) 所示。



(a) 灰值化图

JANSHUI Rd HEILONGJIANG Rd

(b) 大津律阈值二值化结果



(c) 笔画算子提取结果

(d) 骨化图

青岛市体校

(e) 区域生长结果图 文字定位与提取步骤

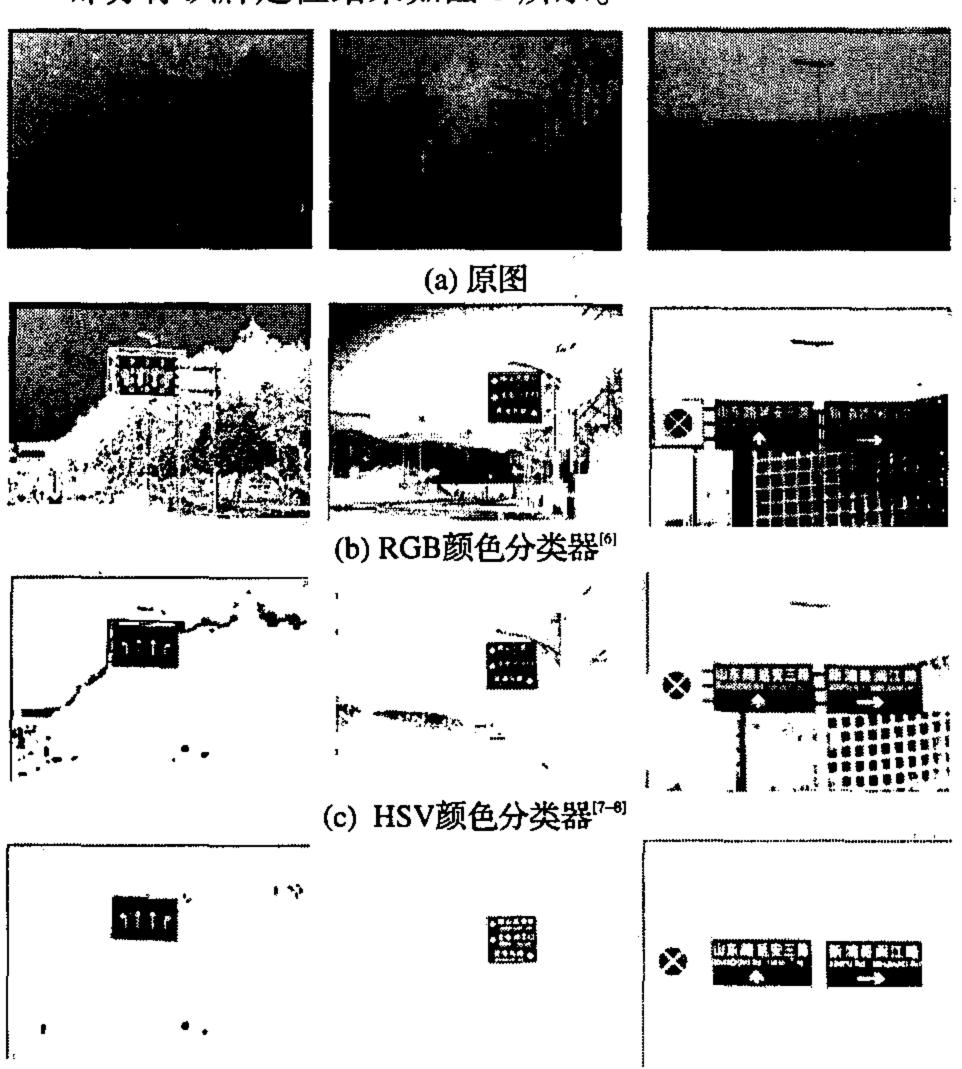
基于颜色的区域增长方法的最终文字区域将略大于基于 骨架像素的预定位文字区域,可以通过简单的连通区域分析 方法,在预定位区域的基础上,取外延笔画像素的最大外接矩 形即可。定位后的二值图像送 OCR 软件进行文字识别。

### 4 实验与分析

为验证本方法的有效性,本文通过在高速公路、市区街道上所拍得的多段视频来对整个系统做出评估,其中包括不同的道路情况(直的、弯的)和不同的天气条件(晴天、多云)。

视频拍摄时,将摄像机置于坐车的右前座,镜头与道路呈30°仰角。摄像机为家用 JVC GZ-HD300AC,元件像素305 万,有效像素 1/4.1 in(1 in = 2.54 cm)。实验中采集了 50 段视频,包含有 186 个蓝底白字的道路标识牌(指示牌、告示牌和限速牌),设 TP (True Positive)表示被正确定位的标识牌数,FP (False Positive)表示被错误标识为标识牌的虚警数,FN (False Negative)表示被漏检的标识牌数。本文采用 Precision (精确性)和 Recall(召回率)两个指标进行了定量比较,并与基于 RGB 颜色阈值方法<sup>[6]</sup>、HSV 颜色空间<sup>[7-8]</sup>的标识牌定位方法进行了比较。

部分标识牌定位结果如图 6 所示。



(d) 本算法分类器 图 6 标识牌定位算法比较

从图 6 可以明显看出,本文提出的标识牌定位方法效果显著高于其他方法,定位准确,虚警少,漏检率也较低。

标识牌经过定位、确认以及仿射变换处理后,就要进行文字的识别,本文算法通过与基于颜色的聚类方法、原图经灰值化一二值化处理的方法进行了比较实验,部分结果如图 7 所示。在 30 段视频的 56 个标识牌中,人工统计共 952 汉字,632 英文单词,其他符号认为是虚警。

统计结果如表1所示,其中:

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

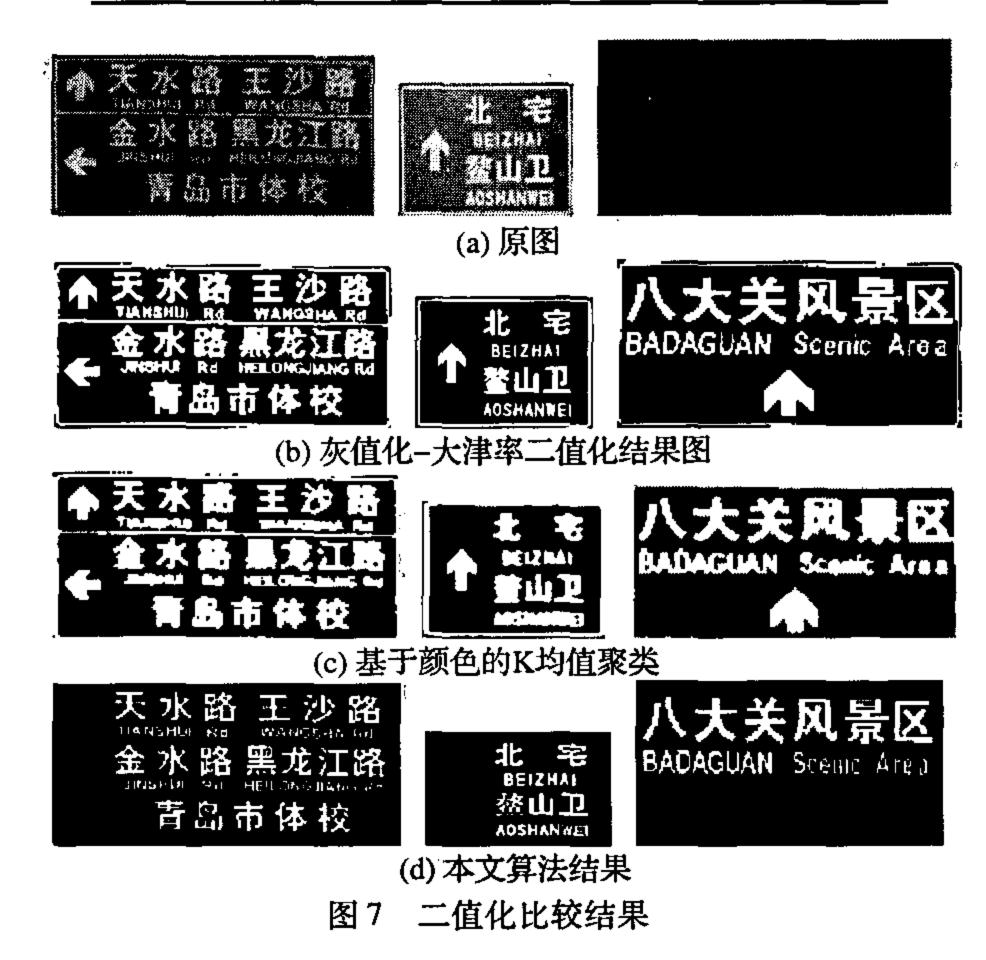
从实验结果看,本文提出的文本定位与提取算法在两个性能指标中均是最高的,说明了本文算法能够适应复杂道路环境和光照阴影变化,具有较好鲁棒性。同时本文算法可以与其他需要进行定位的标识牌颜色一起作为正样本训练神经

网络,从而进行更广泛的定位应用。

表 1 几种方法的统计结果

%

方法	指标 -	对象	
		汉字	英文单词
大津率	Precision	85.70	67.20
二值化	Recall	86.30	69.10
K均值	Precision	81.50	47. 20
聚类	Recall	84.50	50.10
本文	Precision	96.20	87.30
算法	Recall	97.10	88.70



#### 参考文献:

- [1] LYU M R, SONG J Q, CAI M. A comprehensive method for multi-lingual video text detection, localization, and extraction [J]. IEEE Transactions on Circuit and System on Video Technology, 2005, 15 (2): 243 255.
- [2] KAMEL M, ZHAO A. Extraction of binary character/ graphics images from grayscale document images[J]. CVGIP: Graphical Models and Image Processing, 1993, 55(3): 203 217.
- [3] DJEZIRI S, NOUBOUD F, PLAMONDON R. Extraction of signatures from check bank ground based on a filiformity criterion [J]. IEEE Transactions on Image Processing, 1998, 7(10): 1425 1438.
- [4] YE XIANGYUN, MCHERIET M, SUEN C Y. Stroke-model-based character extraction from gray-level document images [J]. IEEE Transactions on Image Processing, 2001, 10(8):1152-1161.
- [5] YE QIXIANG, HUANG QINGMING, GAO WEN. A robust text detection algorithm in images and video frames[J]. Image and Vision Computing, 2005, 23(6): 565-576.
- [6] de ESCALERA A L, MORENO L E, SALICHS M A. Road traffic sign detection and classification [J]. IEEE Transactions on Industrial Electronics, 1997, 44(6), 848 859.
- 7] VITABILE S, POLLACCIA G, PILATO G, et al. Road signs recognition using a dynamic pixel aggregation technique in the HSV color space [C]// Proceedings of IEEE International Conference on Image Analysis and Processing. Washington, DC: IEEE Computer Society, 2001: 572 577.
- [8] VITABILE S, GENTILE A, SINISCALCHI S M, et al. Efficient rapid prototyping of image and video processing algorithms [C]// Euromicro Symposium on Digital System Design. Washington, DC: IEEE Computer Society, 2004:452 458.