

# 基于多层次联视觉显著性模型的肇事车辆锁定方法

柴桢亮 臧 笛

(同济大学计算机科学与技术系 上海 201804)

(同济大学嵌入式系统与服务计算教育部重点实验室 上海 200092)

**摘 要** 肇事车辆的锁定是智能交通系统中一个十分重要的问题,因此针对肇事车辆的锁定,提出了一种基于多层次联视觉注意模型的肇事车辆匹配方法。在模型的每一层中,基于传统视觉注意模型的思想,通过生成显著图的方式提取车辆的一个显著性特征,如颜色、车标,并将其与肇事车辆进行匹配,过滤掉特征不相似的车辆,经过多次显著性特征提取和匹配,最终获得唯一的肇事车辆。实验结果表明,该模型可以准确地从车辆数据库中锁定肇事车辆,且对光照变化和噪声有较强的鲁棒性。

**关键词** 肇事车辆,车辆匹配,计算机视觉,视觉注意模型,车牌识别

**中图分类号** TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.4.059

## Localization of Causing-traffic-trouble Vehicle with Multi-level Cascaded Visual Attention Model

CHAI Zhen-liang ZANG Di

(Department of Computer Science and Technology, Tongji University, Shanghai 201804, China)

(The Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 200092, China)

**Abstract** The localization of causing-traffic-trouble vehicle is one of the most key problems for intelligent transportation system (ITS). This paper proposed a multi-level cascaded visual attention model to localize the causing-traffic-trouble vehicle. In each level of the proposed model, one significant feature of the vehicle such as color or vehicle logo is extracted and compared to the vehicle which has caused an accident. Then vehicles that have no similar features can be filtered. By performing feature extraction and feature comparison for several times, only the causing-traffic-trouble vehicle will be left behind. The experimental results demonstrate that the proposed approach is able to locate the causing-traffic-trouble vehicle accurately and is robust to luminance and noise.

**Keywords** Causing-traffic-trouble vehicle, Vehicle matching, Computer vision, Visual attention model, Plate recognition

## 1 引言

随着社会的不断发展,交通给人们的生活带来了巨大的便利,但与此同时,机动车数量的急速增长,也使得交通拥堵、事故频发、违章逃逸等问题愈发严重。基于视频的智能交通系统对于提高城市交通管理水平、降低事故率、节省人力资源、保障安全出行等都具有重要意义。在实时监控或已存储的视频库中准确并快速地锁定肇事车辆,是智能交通系统的重要功能,也是交通管理的基本保证。

车辆匹配是图像匹配在智能交通领域上的具体应用。早期的方法是利用灰度相关的模板匹配方法<sup>[1]</sup>来进行车辆匹配,该算法具有原理简单、实现简便等优点,但其由于直接以灰度图像作为输入进行计算,因此对于噪声、光照变化等极为敏感。其后,图像匹配主要利用图像的全局或局部特征(如轮廓特征、边缘特征、统计特征、纹理特征等)进行匹配。文献[2]通过颜色直方图来实现匹配,文献[3]提出了一种基于最

小均方误差的图像匹配方法。这些方法具有一定的抗噪性,但通常准确性不高。随后,Lowé 等人提出了具有尺度、旋转不变性的 SIFT 算子<sup>[4]</sup>,在此基础上出现了很多改进的基于 SIFT 算子的匹配方法<sup>[5-8]</sup>,但是这类算法计算量较大,匹配时间通常较长。

通过对人类视觉系统的研究发现,人类视觉系统在复杂的场景中会将注意力集中在少数显著的视觉对象上,这就是视觉选择注意,计算机视觉注意模型受此启迪而诞生。视觉注意模型分为自底向上(Bottom-Up)<sup>[9]</sup>的视觉注意模型和自顶向下(Top-Down)<sup>[10]</sup>的视觉注意模型。目前,视觉注意模型的研究主要集中于自底向上的视觉注意模型,其中 Itti-Koch 模型<sup>[11]</sup>是当前最为流行的一种模型,其核心思想是将多个图像特征融合成一张显著图,进而检测出注意力焦点的位置。在模型的建立过程中,颜色、亮度和边缘方向这几个底层的特征被并行地提取出来并进行多尺度描述和中心-边缘操作,形成特征图;再将所有的特征图整合到一张单独的显著

到稿日期:2014-05-28 返修日期:2014-08-26 本文受国家自然科学基金(61103071),教育部博士学科点新教师基金(20110072120065),留学回国人员科研启动基金,2012年科技部国际合作专项(2012DFG11580)资助。

柴桢亮(1990-),男,硕士生,主要研究领域为图像处理与应用;臧笛(1975-),女,副教授,主要研究领域为计算机视觉、模式识别,E-mail: zangdi@tongji.edu.cn(通信作者)。

图中,最后通过最大值检测和 Winner-Take-All 的方法从显著图中找到最显著点的位置。该模型具有很高的计算效率和较好的适用性,但是该模型不能应对一幅图像需要提取多种不同显著特征的情况。除 Itti 模型外,Stentiford 模型<sup>[12]</sup>、光谱剩余假说模型<sup>[13]</sup>以及 Hu-Rajan-Chia 模型<sup>[14]</sup>也是比较经典的视觉注意模型,但是这几种模型主要适用于显著物体细小的情况。

在综合考虑 Itti 视觉注意模型优缺点的基础上,本文提出了一种基于多层级联视觉显著性模型的肇事车辆匹配方法,从每一层分别提取车辆的一种显著性特征,并与肇事车辆进行匹配,通过匹配过滤掉该特征不同于肇事车辆的图像,并将余下的车辆图像输入模型的下一层进行再次匹配。在模型的最后一层,通过提取车辆的唯一标识特征——车牌,并利用卷积神经网络对车牌字符进行识别,从而实现车辆的唯一锁定。相比于单纯依靠车牌来进行匹配,本模型由于经过前几个特征的匹配,可以过滤掉大部分的车辆图像,从而减少了车牌识别错误带来的影响,提升了车辆锁定的准确率。实验表明,本模型具有较好的准确性和鲁棒性。

## 2 视觉注意模型

由于 Itti 模型是本模型的基础,因此本节首先介绍 Itti 视觉注意模型及相关概念。

Itti 视觉显著模型的算法流程如下:

(1)首先对输入图像进行特征提取,通过计算彩色图像的 R、G、B 3 个通道的平均值得到图像的亮度特征;通过色彩的竞争机制得到图像的颜色特征(红-绿特征图和蓝-黄特征图);最后通过 Gabor 方向金字塔得到 0°、45°、90°、135°这 4 个方向的特征图。

(2)分别对 3 种特征图进行高斯金字塔多尺度描述和中心-边缘算法操作,使得每个特征得到 6 张不同尺度下的特征图。

(3)将所有的特征图归一化后进行线性融合,得到最终的物体显著图。

为方便后文的说明,先介绍上面提及的 3 个概念:图像高斯金字塔、中心-边缘操作和特征图融合方法。

图像高斯金字塔是一种图像处理技术,它的本质是图像的多尺度表示方法。在一个典型的高斯金字塔中,每一层图像是上一层尺度的一半。金字塔的底部是图像的高分辨率表示,而顶部是低分辨率表示,当金字塔向上移动时,图像尺寸和分辨率就降低。在层与层之间,使用高斯滤波器对图像进行滤波,再对滤波后的图像进行隔行采样得到金字塔下一层的图像。高斯金字塔可以用公式递归地表示为:

$$\begin{cases} G_0(x, y) = G(x, y) \\ G_{k+1}(x, y) = \sum_{i=-2}^2 \sum_{j=-2}^2 w(i, j) G_k(2x+i, 2y+j) \end{cases} \quad (1)$$

其中,  $k=0, 1, \dots, 8$ ,  $G_k(x, y)$  是金字塔第  $k$  层的图像,  $G(x, y)$  是输入图像,  $w(i, j)$  是高斯卷积核。

在视觉注意模型中,“中心”定义为图像金字塔中高分辨率的图像,而“边缘”定义为内插到高分辨率的低分辨率图像,那么中心-边缘操作其实就是将低分辨率图像内插至高分辨率图像后与原高分辨率图像进行相减,得到两者图像的差异

结果。中心-边缘操作可以用公式表示为:

$$G_c \ominus G_s = G_c - Inter_{\leftarrow, \rightarrow}(G_s) \quad (2)$$

其中,  $G_c$  为中心高分辨率图像,  $G_s$  为边缘低分辨率图像,  $\ominus$  为中心-边缘操作符,  $Inter$  为内插操作。

一般取  $c \in \{2, 3, 4\}$ ,  $s \in \{3, 4\}$ ,  $s = c + \sigma$ , 也就是在 2-5, 2-6, 3-6, 3-7, 4-7, 4-8 层之间进行中心-边缘操作。

最后将中心-边缘操作后的 6 幅图像进行融合,得到特征显著图的公式如下:

$$P = N\left(\bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} N(G(c, s))\right) \quad (3)$$

其中,  $G(c, s) = G_c \ominus G_s$ ,  $N(\cdot)$  为归一化操作,  $P$  是最终得到的显著图。

## 3 基于多层级联视觉注意模型的肇事车辆匹配

### 3.1 系统结构

车辆的一些视觉显著性特征,比如车辆颜色、车标、车牌等,都可以用来区分车辆,因此对于肇事车辆的锁定,本模型采用一种基于视觉显著性模型的多层级联式的匹配方法。模型共 4 层,每一层针对每种特征的特点将不同的底层特征输入视觉注意模型并生成相应的显著图,通过显著图提取车辆的显著特征,然后利用该显著特征进行匹配,筛选过滤掉该特征明显不同于肇事车辆的车辆图像,并将余下的车辆输入下一层,继续进行匹配,直到通过最后一层后唯一地锁定肇事车辆图像。

本模型中提取的车辆的视觉显著性特征分别是车辆的颜色分布、车标、年检标志以及车牌。整个匹配的流程如图 1 所示。其中虚线框内的部分是本模型方法的 4 个层级。

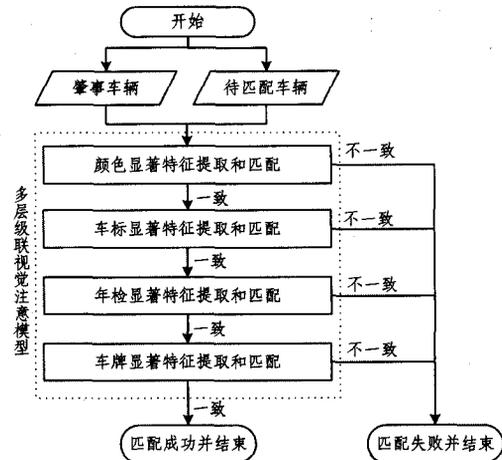


图 1 系统流程

下面 4 小节分别介绍每种显著性特征的提取方法以及匹配方法。

### 3.2 颜色显著性特征

车辆颜色是一种十分直观的视觉特征,每辆车的颜色以及色彩分布不尽相同,因而可以将车辆的颜色作为一个视觉特征进行匹配。

为了得到车辆的颜色显著图,进行以下操作:

首先为了去除光照变化带来的影响,对原始输入图像进行亮度调整,将原图像中的  $r, g, b$  3 通道分别减去它们的平均值,公式如下:

$$\begin{cases} r' = r - r_A \\ g' = g - g_A \\ b' = b - b_A \end{cases} \quad (4)$$

$r, g, b$  是输入图像的 3 个通道的值,  $r_A, g_A, b_A$  是输入图像 3 个通道的平均值。

然后再根据“颜色对”模型, 将  $r', g', b'$  转化为 4 原色: 红色、绿色、蓝色和黄色, 分别表示为  $R, G, B, Y$ 。两者之间的对应转化关系为:

$$\begin{cases} R = r' - \frac{g' + b'}{2} \\ G = g' - \frac{r' + b'}{2} \\ B = b' - \frac{r' + g'}{2} \\ Y = \frac{r' + g' - |r' - g'|}{2} - b' \end{cases} \quad (5)$$

我们将上述 4 原色合成两个“颜色对”, 作为  $R-G$  颜色子通道和  $B-Y$  颜色子通道, 分别为:

$$\begin{cases} RG = R - G \\ BY = B - Y \end{cases} \quad (6)$$

对上述两通道分别构建高斯金字塔, 并进行中心-边缘操作后相融合得到车辆的车身图, 融合公式为:

$$C = N\left\{\bigoplus_{c=2}^3 \bigoplus_{s=c+3}^{c+4} [N(RG(c, s)) + N(BY(c, s))]\right\} \quad (7)$$

其中,  $c, s$  代表了高斯金字塔的不同层的层号,  $RG(c, s)$  表示对  $R-G$  通道的第  $c$  层和第  $s$  层进行中心-边缘差分得到的差分图,  $N(*)$  表示归一化操作。

图 2 中, 左图是车辆原图, 中间的图像是经过上述视觉注意模型后得到的车身图, 右图是将中间的车身图经过二值化处理得到的图像。将二值化的车身图作为掩膜作用到原图上去, 可以得到车辆的彩色车身图, 这样在提取颜色时可以提取到纯粹的车辆的颜色, 从而去除背景颜色带来的影响。

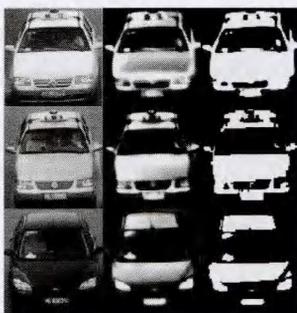


图 2

对上面得到的彩色车身图再次应用式(5)便可得到车身的 4 张颜色显著图, 即颜色层的显著图。

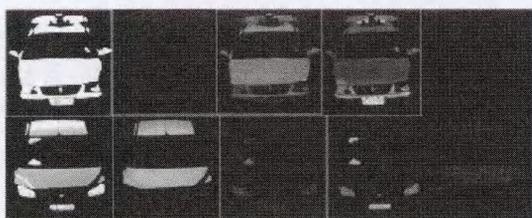


图 3

图 3 中最左边一列是车辆的车身彩图, 右边 4 列图像分别是红、绿、蓝、黄 4 颜色分量的显著图。

颜色匹配采用以下方法: 分别计算肇事车辆和待匹配车辆的 4 原色直方图, 使用巴氏距离计算两者之间的相似度。巴氏距离计算公式如下:

$$D_B(p, q) = -\ln(BC(p, q)) \quad (8)$$

$$BC(p, q) = \sum_{x \in X} \sqrt{p(x)q(x)}$$

其中,  $p, q$  是直方图分布函数。

利用巴氏距离公式可以计算得到 4 个相似度  $D_B(R_1, R_2), D_B(G_1, G_2), D_B(B_1, B_2), D_B(Y_1, Y_2)$ 。当且仅当这 4 个相似度均小于某个阈值时, 我们认为两辆车在颜色上是一致的。

### 3.3 车标志显著性特征

车标也是车辆的一个十分明显的视觉特征, 具有不同车标的两辆车不可能是同一辆车, 因而在第二层我们提取车标作为一个视觉显著性特征。

车标定位一般采用基于车牌定位的方法来实现, 这种方法的优势在于能够准确定位。本模型中车牌的定位方法会在第 3.5 节作详细介绍。根据车牌定位可以获取车牌矩形的 4 个边界的坐标, 并得到车牌的宽度  $d$  与高度  $h$ , 接着在车牌上方搜索以  $0.8d$  为宽、 $4h$  为高的矩形区域, 实现车标的粗定位。

将粗定位的图像灰度化后分别进行垂直和水平方向上的 sobel 边缘检测, 然后分别对垂直边缘图和水平边缘图进行高斯金字塔多尺度描述和中心-边缘差分操作及特征图融合, 之后得到垂直显著图和水平显著图, 并用最大类间方差 (Otsu) 算法将显著图进行二值化。

图 4 示出经过粗定位后的车标图像以及经过 sobel 边缘检测后的边缘图像, 图 5 示出车标的水平和垂直方向的显著图以及它们的二值图像。



图 4 从左至右为: 粗定位车标图像, sobel 垂直边缘图, sobel 水平边缘图

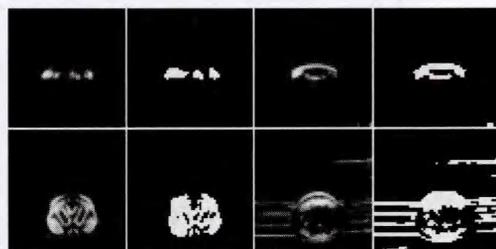


图 5 从左至右为: 垂直显著图, 二值化后的垂直显著图, 水平显著图, 二值化后的水平显著图

通过图 5 可以发现,显著图中会保留汽车的进气栅,这会影 响对于车标的精确定位,因此需要将进气栅从显著图中去 除。考虑到车辆的进气栅一般只有两种形态:水平或者垂直, 它只在水平显著图和垂直显著图中的一张中出现,所以二 值化后的两幅显著图中所含前景像素较少的那一张可认为是 不包含进气栅的,我们选择不包含进气栅的显著图来进行车 标的精确定位。

对于选择的垂直显著图或水平显著图,包含其中所有前 景点的最小外接矩形即可认为是车标所在位置。

车标的匹配采用以下方法:计算当前车辆与肇事车辆车 标的长宽比,如果比值超出一定范围,则认为车标不一致,式 (9)给出了具体的公式和阈值。

$$r = \frac{w1/h1}{w2/h2}; \text{if } r > 1.25 \text{ or } r < 0.8 \text{ then fail} \quad (9)$$

如果长宽比相近,把一个车标缩放到另一个车标的大小, 利用模板法作差进行比较。如果差分图像近似于空白,则认 为车标一致,否则判定车标不一致。

### 3.4 年检显著性特征

除了车辆的颜色分布、车标以外,车窗上贴着的年检标志 也可以作为车辆的一个视觉显著性特征。

首先对车辆图像进行二值化,由于车窗部分在二值图像 中是背景像素,我们认为从  $0.2 \times$  车辆图像高度处向下找到 的第一个具有一定宽度的连续背景区域就是车窗,而且大多 数年检标志都是贴于车窗的左上角的,因此,只取左上角  $1/4$  车窗图像,得到它的亮度图。亮度图计算公式如下:

$$I = \frac{r+g+b}{3} \quad (10)$$

对于亮度图,用高斯金字塔对它进行多尺度表示,并应用 中心-边缘算法,可以得到年检标志的显著图。

图 6 左图是车辆原始图像,中图是定位得到的包含年检 标志的车窗图像,右图是年检显著图。



图 6

对于年检的匹配,首先我们计算年检标志占整个车窗的 比值,如果当前车辆以及肇事车辆的比值比较接近,那么继续 下一步的匹配,否则认为两车的年检标志不匹配;接着我们将 包含年检的图像均分成  $3 \times 2$  小块,分别计算每一小块中年检 标志所占的百分比,这样可以得到一个六维向量;然后用巴氏 距离计算当前车辆及肇事车辆向量间的相似距离,如果距离 小于预先设定的阈值则判定两车具有相似的年检标志,年检 匹配成功。

### 3.5 车牌显著特征

由于每块车牌独一无二的特点,因此在最后一层,我们通

过车牌来完成肇事车辆的唯一锁定。本模型使用视觉显著模 型进行车牌的定位,使用卷积神经网络来进行车牌字符的识别。

卷积神经网络是深度学习的一种,它通常由 2 个卷积层、 2 个下采样层以及 1 个全连接层构成。卷积层用以提取输入 图像的局部特征,如边缘信息和角点特征等;下采样层可以防 止参数爆炸;全连接层用以进行特征分类。卷积神经网络权 值共享的特点降低了网络模型的复杂度,从而使图像可以直 接作为网络的输入,省去了繁琐的特征提取过程。另外由于 它的底层卷积层会提取图像最基础的局部特征,因此网络对 目标物体或图像的形变、旋转具有较好的抵抗性。与传统的 BP 神经网络相比,卷积神经网络具有训练参数少、可直接输 入二维图像等优点。

#### 3.5.1 车牌定位

以车牌的颜色对抗色、亮度和车牌边框方向作为输入图 片的视觉特征,分别构建高斯金字塔。通过对高斯金字塔进 行中心-边缘差分并归一化,获取各个特征的多张特征图,对 特征图进行跨尺度融合及归一化后,形成对应于车牌颜色、亮 度和边框方向的特征显著图,最后将这 3 张特征显著图组合 成视觉显著图,并根据车辆颜色对车牌显著图进行优化处理, 突出其中的车牌区域。

##### (1) 特征提取

颜色特征:由于中国的车牌以蓝底白字和黄底黑字居多, 因此仅选取蓝色-黄色颜色对作为颜色特征,公式如下:

$$\begin{cases} B = b - \frac{r+g}{2} \\ Y = \frac{r+g}{2} - \frac{|r-g|}{2} - b \end{cases} \quad (11)$$

$$BY = B - Y$$

亮度特征:亮度特征图  $I$  由式(10)求得。

方向特征:对亮度图  $I$  进行 Gabor 滤波,考虑到车牌边框 正常情况下在  $0^\circ$  和  $90^\circ$  方向信息最明显,只进行  $0^\circ$  和  $90^\circ$  方 向上的滤波,得到方向特征图  $O(\theta)$ ,  $\theta \in \{0^\circ, 90^\circ\}$ 。

##### (2) 计算特征显著图

在上一步中已经获取了颜色、亮度、方向的特征图,接着 对它们分别构建高斯金字塔,并按照式(12)获得其进行中心- 边缘操作后的特征显著图。

$$\begin{aligned} C &= N \left[ \bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} N(BY(c,s)) \right] \\ I &= N \left[ \bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} N(I(c,s)) \right] \end{aligned} \quad (12)$$

$$O = N \left\{ \sum_{\theta \in \{0^\circ, 90^\circ\}} N \left[ \bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} N(O(c,s,\theta)) \right] \right\}$$

##### (3) 显著图融合

将 3 个特征显著图线性组合成视觉显著图,用  $S$  表示。 但由于 3 个特征在凸显车牌区域的显著性时,各自的力度是 不同的,因此三者线性组合时的系数也不相同。车牌颜色与 周围颜色的不连续性使得颜色特征最为重要,而车牌基本上 都有边框,因此边框方向特征居次位,最后才是亮度特征。由 于实验颜色、方向、亮度的系数分别取 0.65, 0.2, 0.15 时显著 图效果较佳,因此车牌显著图  $S$  可以表示为:

$$S=0.65C+0.2O+0.15I \quad (13)$$

(4)凸显车牌区域

通过图3不难发现利用式(5)得到的4原色图像可以大致地勾勒出汽车车身的轮廓,因此可以用4原色图像将汽车车身从车牌显著图S中去除,得到更加精确的车牌区域。

如图7所示,最左边为输入的车辆图像,第二张图为使用式(13)计算后得到的车牌显著图,第三张图为使用4原色图像去除车身后的车牌显著图,最后一张图是最终定位到的车牌。

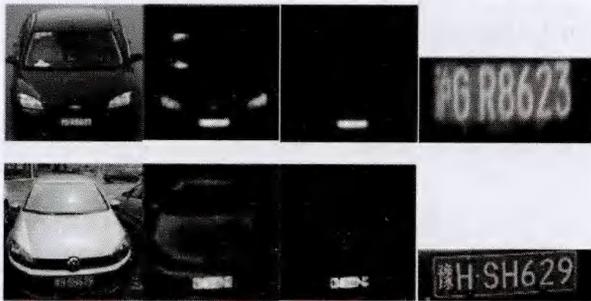


图7 从左至右为:原图,车牌显著图,去除车身后的车牌显著图,定位得到的车牌

### 3.5.2 车牌匹配

得到车牌后,需要进行车牌字符分割以及车牌字符识别才能完成车牌的匹配。

车牌字符分割使用投影法,将字符二值化后向X轴方向进行投影,投影的波谷即可认为是两个字符之间的间隔区域。

车牌字符识别通过卷积神经网络来实现。

首先利用训练数据训练得到汉字网络以及字母数字网络,之后将某个字符图像输入其中一个网络,进行一次前向传播,从输出层的结果就可以判定这个字符的类别,即完成了识别,流程如图8所示。

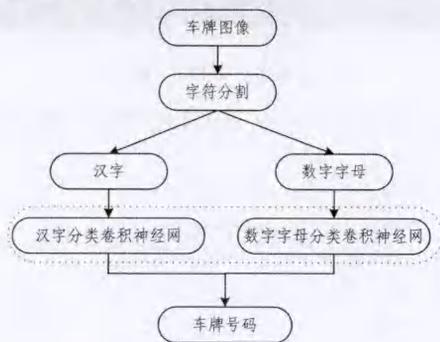


图8 车牌字符识别流程

将肇事车辆与当前匹配车辆的字符图像分别输入两个卷积神经网络,可以得到两个车牌号码,再判断是否一致,如果一致,则表明当前车辆就是肇事车辆;反之则不是肇事车辆。

## 4 实验与分析

本文实验用的车辆图像总共500张,主要来源于上海市某路口交通视频。通过对视频图像进行高斯背景建模得到场景的背景图像,并利用视频图像和背景图像作差得到分割后的车辆图像。图9展示了部分实验图像。

本文实验在Intel Core 2 2.80GHz CPU、2G内存的Windows平台下进行。



图9 实验车辆图像

### 4.1 实验一

从车辆图像集中选取了8张不同颜色并且出现过两次以上的车辆图像作为肇事车辆,将图像集中的其余499张图像分别与它们进行匹配,寻找出相同的车辆。

表1 肇事车辆锁定实验结果

肇事车辆	颜色	车标	年检	车牌	锁定车辆
	20	4	2	1	
	82	14	6	1	
	41	26	18	1	
	170	66	12	1	
	14	8	5	1	
	45	17	9	1	
	153	41	22	1	
	9	5	2	1	

匹配共分4个阶段,每一阶段利用第2节中所述的方法提取车辆的一种显著特征,将其与肇事车辆的该特征进行匹配,经过匹配将特征相似的车辆图像保留,将其余图像剔除。下一阶段只匹配上一阶段匹配后保留下来的图像,减少匹配次数。表1列出了8张肇事车辆图像经过每一次特征匹配后剩余的車輛数目以及最终锁定的車輛。

由表1可知,利用本文提出的分层视觉注意模型确实可以从車輛数据集中找出与肇事車輛图像相同的車輛图像,并且每一层提取的显著特征可以有效地缩小检索和匹配范围。

#### 4.2 实验二

將車輛图像集中的每一张車輛图片与其他所有图片进行匹配,匹配时依次匹配提取的4个显著特征,如果某个特征匹配失败則車輛匹配立即终止,判定两車不同;否則继续直到4个特征均匹配完成,然后输出两車匹配的结果。

將匹配结果与实际情况相比较,可以得到4种情况,如表2所列。

表2 4种比较结果

		匹配结果	
		相同車輛	不同車輛
实际情况	相同車輛	R1	R2
	不同車輛	R3	R4

4种情况中,R1和R4是正确结果,R2和R3表示匹配过程出现了错误。可以用3个指标来衡量匹配算法的性能:正确率、查准率、召回率。正确率用来衡量匹配成功或者失败的正确性,查准率用来衡量匹配成功的图像中真实相同的比例,召回率用来衡量相同車輛匹配成功的比例。三者的计算公式如下:

$$\text{正确率} = \frac{R1+R4}{R1+R2+R3+R4} \quad (14)$$

$$\text{查准率} = \frac{R1}{R1+R3} \quad (15)$$

$$\text{召回率} = \frac{R1}{R1+R2} \quad (16)$$

本文方法的实验结果如表3所列。

表3 本文方法结果

		匹配结果	
		相同車輛	不同車輛
实际情况	相同車輛	35	2
	不同車輛	0	124713

由表3可知,本文方法的正确率为99.9%,查准率为100%,召回率为94.6%。

为了验证本文提出的算法,另外使用一种融合颜色直方图特征和SIFT特征的匹配方法进行本组車輛匹配实验,实验结果如表4所列。

表4 颜色+SIFT特征方法实验结果

		匹配结果	
		相同車輛	不同車輛
实际情况	相同車輛	34	3
	不同車輛	3324	121389

由表3可知,该方法的正确率为98.1%,召回率为91.9%。

通过上面两表的比较可以发现,本文提出的肇事車輛匹配方法具有较高的正确率、查准率以及召回率。这是由于本

模型使用了车牌作为肇事車輛匹配最后的显著特征;另外,卷积神经网络较高的识别准确率保证了不同車輛不易被检测为相同車輛,从而使得表2中的R4较高而R3较低,因而利用本模型进行实验得到的查准率通常接近于100%,而召回率会略低于准确率。SIFT方法由于SIFT特征具有旋转和尺度不变性,并且融合了颜色特征,因此能够较准确地完成車輛匹配,但是这种方法对于颜色及外形结构均相似的車輛无法予以区分,因而会有部分不同車輛被误判为相同車輛。

实验二表明利用本文提出的分层视觉注意模型进行車輛匹配具有较高的准确性。

#### 4.3 实验三

为验证光照和噪声环境对匹配准确率的影响,分别调整車量图像的亮度以及添加噪声进行测试。

##### (1) 亮度

对所有500张車輛图像调整亮度,得到500张亮度调整后的車輛图像。然后将每张車輛图像分别与相对应的亮度调整后的車輛图像及其他一张不同車輛进行匹配,总共进行1000次匹配。

图10为亮度调整后的車輛图像,分别为Gamma值为0.3、0.5、0.7、1、1.5、2时的图像,Gamma值愈小,亮度越高;Gamma值愈大,亮度越低。Gamma值为1表示亮度不变。

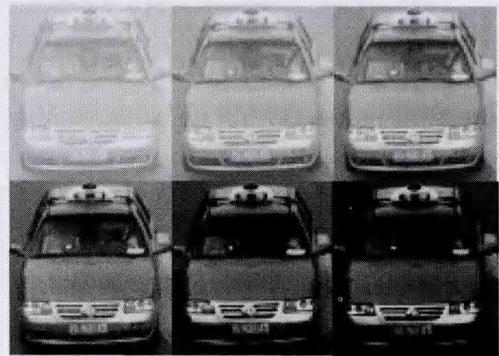


图10 调整了亮度的車輛图像

图11为随亮度Gamma值的变化,准确率和召回率的变化曲线。

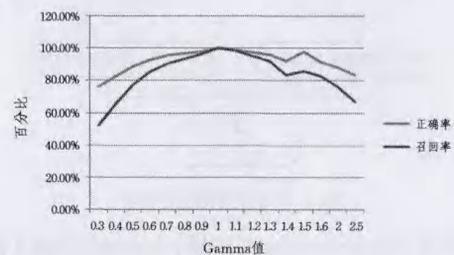


图11 車輛图像亮度与匹配正确率和召回率

图11表明,本模型对于亮度变化不敏感,正确率始终可以保持在80%以上。这是因为在进行匹配前,对亮度进行了调整,从而在一定程度上去除了亮度变化带来的影响。

##### (2) 噪声

类似于亮度变化,对所有500张車輛图像添加不同方差的高斯噪声,得到500张添加噪声的車輛图像,然后将每张車輛图像分别与相对应的带噪声的車輛图像及其他一张不同車

辆进行匹配,总共进行 1000 次匹配。

图 12 为添加了高斯噪声后的车辆图像,从左至右分别为添加了方差为 0(原图)、0.01 和 0.05 高斯噪声的车辆图像。



图 12 添加了噪声的车辆图像

图 13 为高斯噪声对本模型正确率及召回率的影响曲线。在一般的噪声(方差低于 0.02)污染下,本模型的准确率可以保持在 90%以上,并且对于十分严重的噪声污染(方差 0.05 左右),匹配准确率仍能保持在 70%以上,这是由于使用了高斯金字塔对图像进行多尺度描述,因此整个模型具有较强的抗噪能力。

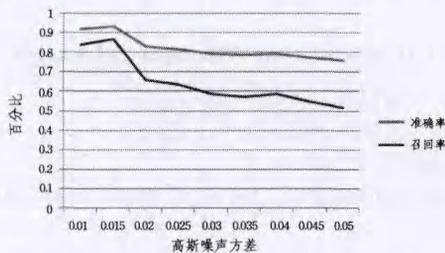


图 13 高斯噪声方差与匹配正确率和召回率

**结束语** 本文提出了一种基于多层级联视觉显著性模型的车辆锁定方法。即借鉴 Itti 模型的思想,在提取不同显著特征时选用了不同的底层特征作为输入,并对得出的显著图做了不同的处理,从而能够更加有效地区分显著特征和背景。实验表明,本模型具有较好的正确率,并且对光照变化与噪声污染具有较强的鲁棒性。

### 参 考 文 献

- [1] 罗钟铤,刘成明. 灰度图像匹配的快速算法[J]. 计算机辅助设计与图形学学报,2005,17(5):966-970
- [2] Piccardi M, Cheng E D. Multi-frame moving object track matching based on an incremental major color spectrum histogram matching algorithm[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, 2005. IEEE,2005:19
- [3] Liu S, Sun J, Dang J. A Linear Resection-Intersection Bundle Adjustment Method[J]. Information Technology Journal,2008,7(1):220-223
- [4] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision, 2004, 60(2):91-110
- [5] Hu X, Tang Y, Zhang Z. Video object matching based on SIFT algorithm[C]//2008 International Conference on Neural Networks and Signal Processing. IEEE,2008:412-415
- [6] Ke Y, Sukthankar R. PCA-SIFT: A more distinctive representation for local image descriptors[C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004(CVPR 2004). IEEE,2004,2:506-513
- [7] Alharin F, Wang C, Ristic-Durrant D, et al. Improved SIFT-Features Matching for Object Recognition[C]//BCS International Academic Conference, 2008:178-190
- [8] Shan Y, Sawhney H S, Kumar R. Unsupervised learning of discriminative edge measures for vehicle matching between non-overlapping cameras[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2008,30(4):700-711
- [9] 曾志宏,李建洋,郑汉垣. 融合深度信息的视觉注意计算模型[J]. 计算机工程,2010,36(20):200-202
- [10] Baluch F, Itti L. Training top-down attention improves performance on a triple-conjunction search task[J]. PloS one,2010,5(2):e9127
- [11] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on pattern analysis and machine intelligence,1998,20(11):1254-1259
- [12] Stentiford F W M. Attention-based image similarity measure with application to content-based information retrieval[C]//Electronic Imaging 2003 International Society for Optics and Photonics,2003:221-232
- [13] Hou X, Zhang L. Saliency detection: A spectral residual approach[C]//IEEE Conference on Computer Vision and Pattern Recognition,2007(CVPR'07). IEEE,2007:1-8
- [14] Hu Y, Rajan D, Chia L T. Adaptive local context suppression of multiple cues for salient visual attention detection[C]//IEEE International Conference on Multimedia and Expo, 2005(ICME 2005). IEEE,2005:4
- [15] 楼汉琦,蔡晓东,李长俊. 一种实现快速车标定位的方法[J]. 国外电子测量技术,2013(6):72-74
- [16] 毛颂安. 汽车车标检测方法的研究与实现[D]. 成都:电子科技大学,2012
- [17] 周小龙,张小洪,冯欣. 基于视觉显著图的车牌定位算法[J]. 光电工程,2009,36(11):145-150
- [18] Siddiqui F U, Isa N A M. Enhanced moving K-means (EMKM) algorithm for image segmentation[J]. IEEE Transactions on Consumer Electronics,2011,57(2):833-841
- [19] Guo Y, Hsu S, Sawhney H S, et al. Robust object matching for persistent tracking with heterogeneous features [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007,29(5):824-839
- [20] Zhu L J, Hwang J N, Cheng H Y. Tracking of multiple objects across multiple cameras with overlapping and non-overlapping views[C]//IEEE International Symposium on Circuits and Systems,2009(ISCAS 2009). IEEE,2009:1056-1060
- [21] 李星,郭晓松,郭君斌. 基于 HOG 特征和 SVM 的前向车辆识别方法[J]. 计算机科学,2013,40(11A):329-332