

视觉注意力机制在货运列车车号定位中的应用

葛 莉, 赵歆波

(西北工业大学 计算机学院, 陕西 西安 710072)

摘 要: 为了定位货运列车车厢上的车号信息, 提出并实现一种基于视觉注意力机制的定位新方法. 该方法利用眼动跟踪技术获取真实眼动数据, 结合提炼的车厢图像显著特征, 建立车号注意力机制模型, 并利用该模型预测车号显著区域, 最后采用图像处理技术完成车号区域的分析定位. 实验表明, 相比传统基于灰度的处理方法, 该算法更能适应不同光照条件下各类车厢的车号定位, 具有更好的鲁棒性和普适性.

关键词: 眼动跟踪; 车号定位; 视觉注意力; 显著性

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-7180(2014)08-0034-06

Application of Vision Attention Mechanism in Freight Car License Location

GE Li, ZHAO Xin-bo

(School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: To locate license number of every freight car, a novel localization method is presented and implemented based on the visual attention mechanism. The method utilizes eye-tracking technology to collect real eye movement data, combines eye-tracking data with salient features extracted from car images to establish a visual attention model of car license. Then, the model is used to predict salient license areas and. Finally, image processing technology is applied to analyze and locate the license area. As experimental results show, this algorithm is more insensitive to different types of car and illumination variance, and performs better in robustness and universality than traditional methods based on intensity.

Key words: eye-tracking, license location, visual attention, salient model

1 引言

目前, 铁路货车车号定位主要采用无线射频技术^[1], 此类方法硬件设备复杂, 不利于维护, 而且由于电子标签丢失和污损等原因, 车号的识别难以获得 100% 的准确率. 为此, 研究者开始探索利用图像处理技术^[2-4], 通过对拍摄的车厢图像进行处理分析, 实现车号的定位. 这类方法通常先将原图二值化或者将提取的边缘图像二值化, 然后利用投影法^[2]、扫描线法^[3]或者将二值图进行形态学处理^[4]后利用连通域分析等方法, 实现车号的定位. 尽管基于视频处理的方法具有所需设备简单, 安装费用低, 操作方

便等优点, 但是, 由于存在实场环境下光照条件复杂, 而且车厢上非车号字符区域干扰等客观因素的影响, 使得此类方法存在难以克服的图像二值化难题, 对不同车型定位的鲁棒性不好.

视觉注意力机制^[5]是人类视觉信息加工过程中重要的心理调节机制, 利用视觉感知的选择能力, 可将计算资源优先分配给那些在特定目标下容易引起观察者注意的显著性区域. 由于车号定位具有明确的目标选择性, 将该机制引入车号定位, 可以优化图像分析所需的计算资源.

为此, 本文拟将视觉注意力机制应用于列车车号的定位过程, 利用眼动跟踪技术获取真实眼动数

收稿日期: 2013-09-12; 修回日期: 2013-11-08

基金项目: 国家自然科学基金(61171156); 西北工业大学研究生创业种子基金(Z2013138)

据,结合提取的车厢图像显著特征,建立车号注意力机制模型,并利用该模型预测车号显著区域.最后结合图像处理技术完成车号区域的分析定位.

2 基于视觉注意力机制的车号定位

视觉注意力首先由视觉内容中与众不同的部分,即与比周围其他部分显著的部分引起,然后根据个体的兴趣,去选择注意焦点.在提取车厢图像中车号信息这一任务的驱动下,观察者将注意力有选择地集中在车号区域,这就是视觉注意力机制在车号定位过程中的体现.本文依据视觉注意力机制建立针对车号定位的视觉注意力模型,预测图像中可能包含车号的感兴趣区域,并用灰度图表示,即车号显

著图,接下来利用显著图完成车号的定位.

2.1 引入视觉注意力机制的车号定位流程

基于视觉注意力的车号定位首先完成车号显著性模型训练,然后利用该模型进行车号定位,其流程如图1所示.图中左部分为车号显著性模型建立的过程,包括建立车厢图像库、获取真实眼动数据、提取显著特征和模型训练模块;右部分是车号定位过程,包括图像预处理、提取显著特征、计算显著图 and 提取车号区域模块.

2.2 车厢图像库建立

使用图像采集装置,全天候采集10 000幅列车车厢图像(包含车号区域),图像分辨率为 780×390 ,从而建立起车厢图像库.

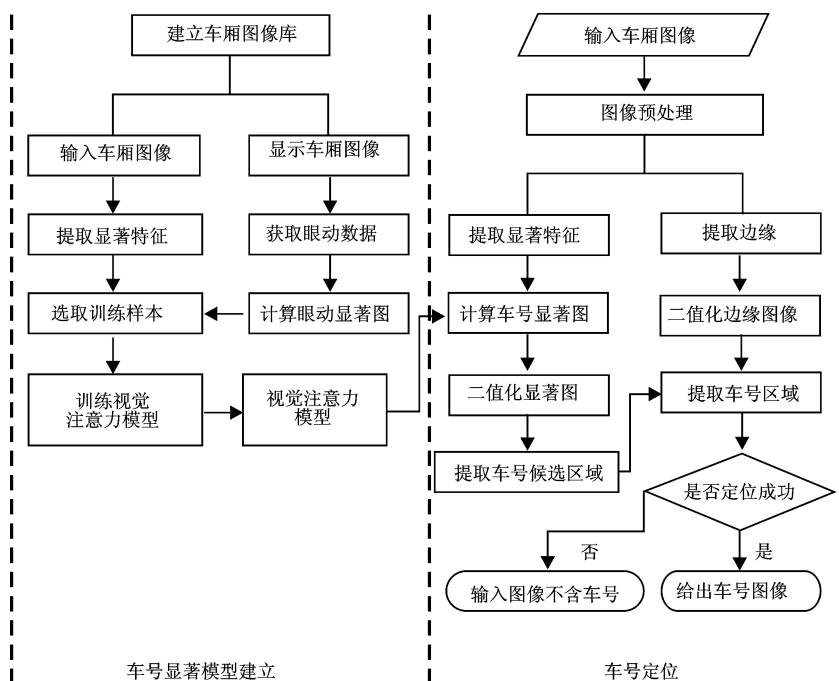


图1 基于视觉注意力机制的车号定位流程

2.3 建立视觉注意力模型

2.3.1 获取眼动数据

使用桌面式眼动仪 TobiiX120 采集眼动数据,该眼动仪的精度为 0.5° ,采样频率为120 Hz.显示器分辨率为 $1\,920 \times 1\,280$,观察者距离显示屏600 mm,车厢图像居中显示在屏幕上,显示时间为1 s,图像之间显示0.5 s的黑屏作为分隔.要求观察者注视车厢图像中的车号区域,使用TobiiX120记录观察过程中的眼动数据.图2(a)为注视图,显示了注视点的位置,可以看到,注视点集中在车厢号区域.由于眼动注视点是一些离散的点,将注视图用高斯核进行卷积,得到连续的注视图,即眼动显著图(如图2(b)所示).

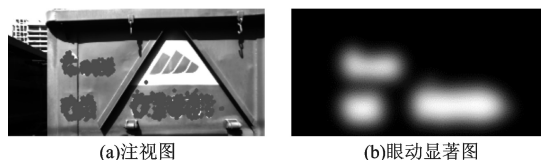


图2 真实眼动数据

2.3.2 显著特征提取

(1) 亮度特征

列车车厢上印刷的车号字符和车厢皮通常具有一定的灰度差异.亮度特征代表了像素点的亮暗信息,有利于增强暗底亮字或亮底暗字的车号区域的显著性.设 R 、 G 、 B 分别表示输入图像的红色、绿色和蓝色通道,则亮度特征 I 通过计算 R 、 G 、 B 平均值得到,见式(1).

$$I = \frac{R+G+B}{3} \quad (1)$$

(2) 全局对比度特征

列车车厢的车号字符一般具有相近的灰度,而且和车厢其他区域具有一定的对比度.全局对比度特征通过计算一个区域和整个图像的对比度得到,能够将车号和周围环境分离开.

本文采用文献[6]中提出的基于全局灰度对比度的显著性计算方法,具体做法是:首先,计算亮度特征 I 的直方图以及灰度级 a_n 出现的频率 b_n ($n=0,1,\dots,255$);其次,计算图像中任一像素点 I_k 与每一个像素点 I_i 的加权灰度距离之和,作为像素点 I_k 的特征值 S_k (式(2)),式中, I_k 和 I_i 的灰度值分别为 a_m 和 a_n , 权值取为灰度级 a_n 出现的频率 b_n ;最后,依照单点特征值的计算方法,计算出图像中所有像素点的特征值,作为该图的全局对比度特征.

$$S_k = \sum_{\forall I_i \in I} b_n |I_k - I_i| = \sum_{n=0}^{255} b_n |a_m - a_n| \quad (2)$$

(3) 方向特征

车号字符通常具有各个方向上的特征,而列车车厢上存在的水平、竖直或者倾斜的凹凸槽,具有某一方向上的明显特征,在其他方向上没有分布,可见,车号字符和车厢其他区域的方向分布存在差异.因而,提取车厢图像 I 在多个方向上的特征,可以增加车号字符的显著性.

实际调研发现车厢凹凸槽通常与水平线呈 0° 、 60° 、 90° 和 120° 的夹角.本文采用二维 Gabor 滤波器 (见式(3)) 提取图像在 0° 、 60° 、 90° 和 120° 这四个方向上的特征.

$$h(x, y) = \frac{1}{2\pi\delta_u\delta_v} \exp\left\{-\frac{1}{2}\left[\frac{u^2}{\delta_u^2} + \frac{v^2}{\delta_v^2}\right]\right\} \cos(\omega u)$$

$$u = x\cos\theta + y\sin\theta, v = x\sin\theta - y\cos\theta \quad (3)$$

式中, θ 是滤波器的方向, δ_u 、 δ_v 是高斯包络在 u 轴和 v 轴上的标准差, ω 表示复正弦函数的频率.

(4) 边缘特征

车号区域往往具有一定的边缘,因而可以提取图像的边缘作为显著特征.

本文采用 Sobel 算子提取原始图像的边缘,得到其边缘特征. Sobel 算子是一阶导数边缘检测算子,由一对 3×3 的卷积核组成,将图像 I 与 Sobel 算子做卷积 (见式(4)),再根据式(5)即可得到边缘特征.

$$g_x = S_x * I(i, j), g_y = S_y * I(i, j), \quad (4)$$

式中, S_x 和 S_y 分别是水平方向和竖直方向的 Sobel 算子, $*$ 表示卷积运算.

$$M(x, y) = \sqrt{g_x^2 + g_y^2} \quad (5)$$

2.3.3 训练视觉注意力模型

将车厢图像缩放至 400×200 , 提取其显著特征. 把车厢图像的注视图同样缩放至 400×200 , 从注视图中随机选取 10 个显著度前 20% 的像素点标记为“1”, 其对应的特征向量作为正样本; 随机选取 10 个显著度后 60% 的像素点标记为“-1”, 其对应的特征向量作为负样本, 因此, 对于车厢图像库共得到 200 000 个训练样本.

模型的建立需要解决标准支持向量机的分类问题, 本文采用双线性支持向量机模型训练正负样本, 从而得到适合车号定位的视觉注意力模型.

2.4 定位车号区域

对车厢图像进行预处理以改善图像的质量, 然后提取显著特征, 并利用建立的车号显著性模型预测其车号候选区域, 再分析定位图像的车号区域.

2.4.1 图像预处理

实际采集到的车厢图像受到不同光照条件的影响, 而且在采集和传输过程中, 不可避免的引进噪声数据. 本文首先采用同态滤波减轻不同光照条件的影响, 然后采用中值滤波去除噪声. 为了减少运算量, 将图像大小缩放至 400×200 像素.

2.4.2 提取显著特征

提取输入图像的显著特征, 即上文提到的亮度、全局对比度、方向和边缘特征, 共得到 7 个特征图 (如图 3 所示).



图 3 特征显著图

2.4.3 计算显著图

将输入图像每个像素点对应的显著特征向量作为输入样本,利用建立的视觉注意力模型预测每个像素点的显著度,得到初步的显著图,然后用高斯核对该显著图进行卷积,得到连续的显著图,即车号显著图。

2.4.4 提取车号区域

定位过程需要依据车号喷涂的先验知识,即:车号由上面车型号和下面车厢号两部分组成,这两部分的左端基本是对齐的,车厢号是7个数字,车厢号一般分两部分印刷,这两部分水平位置基本对齐,左边印刷2、3或4个,相应的,右边印刷5、4或3个,总之,车号由三个矩形子块组成,三个子块的长宽比范围是:0.5~5,而且三者之间的相对位置基本保持

不变。

利用已得到的车号显著图提取车号区域的流程如图4所示。首先采用Otsu算法对车号显著图进行二值化,然后将二值图扩展到 780×390 像素。然后,获取连通域的最小边框的高度及长宽比信息,舍弃人眼都无法看清的高度小于10像素的区域,保留长宽比在0.5~5范围内且满足车号区域相对位置的候选区域。接下来,采用Canny算法提取候选区域对应的边缘图像,按行扫描统计跳变点个数、首尾跳变点所在的列坐标以及跳变点之间的间距和(大于子图像宽度的 $1/10$)。最后,保留连续的扫描行行数大于20的图像,即为车号图像。按照以上过程完成一次车号区域的定位,如图5所示。

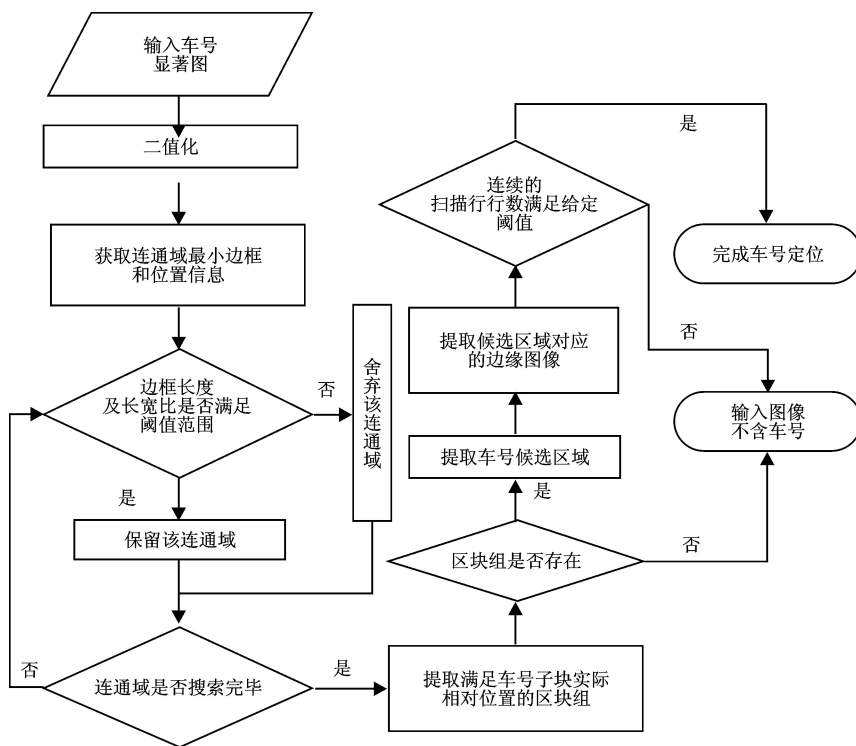


图4 提取车号流程图

3 实验及讨论

为了验证本文方法,在铁路旁安装图像采集设备,全天候采集主要的且具有代表性的C64K、C70A和C80这三种车型的列车图像各1000幅(白天和晚上各500张),建立相应的测试图像库。摄像机与车厢的距离在一定的范围内,且摄像机与车号区域基本保持正投影,因而图像中的车号大小范围一定,这对于设置定位参数和提高定位准确率有很大的帮助。本文实现了文献[2]和文献[3]的定位算

法,以和本文算法进行对比试验。不同算法的定位准确率和平均定位时间的对比情况如表1所示。

由表1的实验结果可以看出,对于三种车型图像,一方面,文献[2]和文献[3]比本文算法的定位准确率低,而且文献[2]对于C80车厢图像的定位准确率为0%,较另外两种车厢图像低40%以上,文献[3]对于C64K的定位准确率较其他两种车厢图像低10%以上,本文算法对于三种车厢图像的定位准确率差异在2%以内,这说明本文算法对于不同车型图像有较好地的鲁棒性。另一方面,文献[2]、文献



图5 车号定位过程

[3]和本文算法对三种车厢图像的定位平均时间基本相同,本文较文献[2]和文献[3]定位时间长,但能够满足实时性的需要。

表1 对比试验结果

测试集	定位准确率/%			定位平均时间/ms		
	文献[2]	文献[3]	本文	文献[2]	文献[3]	本文
C64K	44.8	32.7	88.5	258	162	432
C70A	45.6	56.3	90.4	267	154	420
C80	0	57.7	90.6	264	152	426
全部	30.1	48.9	89.8	263	156	426

基于实验结果,分析三种算法定位失败的原因。文献[2]采用的迭代二值化方法对于强光、弱光和不均匀光照图像的二值化效果不好,后续的投影法定位抗干扰能力差,而且该算法默认车号字符比车厢皮的灰度值大,即只适用于车号较亮车厢皮较暗的车厢图像,而实验采集的C80车厢图像中车号较暗车厢皮较亮,故该算法不能定位C80车厢图像中车号。文献[3]提取图像的Sobel边缘图像并将其二值化,对于车号字符与车厢背景对比度低的情况,譬如:车号字符本身喷涂、粉尘遮挡、雨水冲刷以及强光、弱光、车厢反光和不均匀光照导致,得到的边缘二值图效果欠佳,以致不能正确定位,此外,后续行扫描定位过程中阈值的设定没有充分考虑不同车型的影响,如C64K车厢上斜槽的边缘。本文定位失败的原因,一是由于现场采集的图像存在严重反光、污损、缺失等现象;二是由显著性模型得到的车号显著图,只能得到车号的可能性区域,不够精准。总之,文献[2]和文献[3]的算法非常依赖于采集图像的质量,不能克服不同光照的影响,只能定位车号区域满足一定大小、位置的车厢图像,而且对于不同车型的鲁棒性不好;本文算法对于采集的图像没有苛刻的要求,更能适应不同光照下各类车厢的车号定位,因而具有更好的鲁棒性和普适性,但是本文定位算法

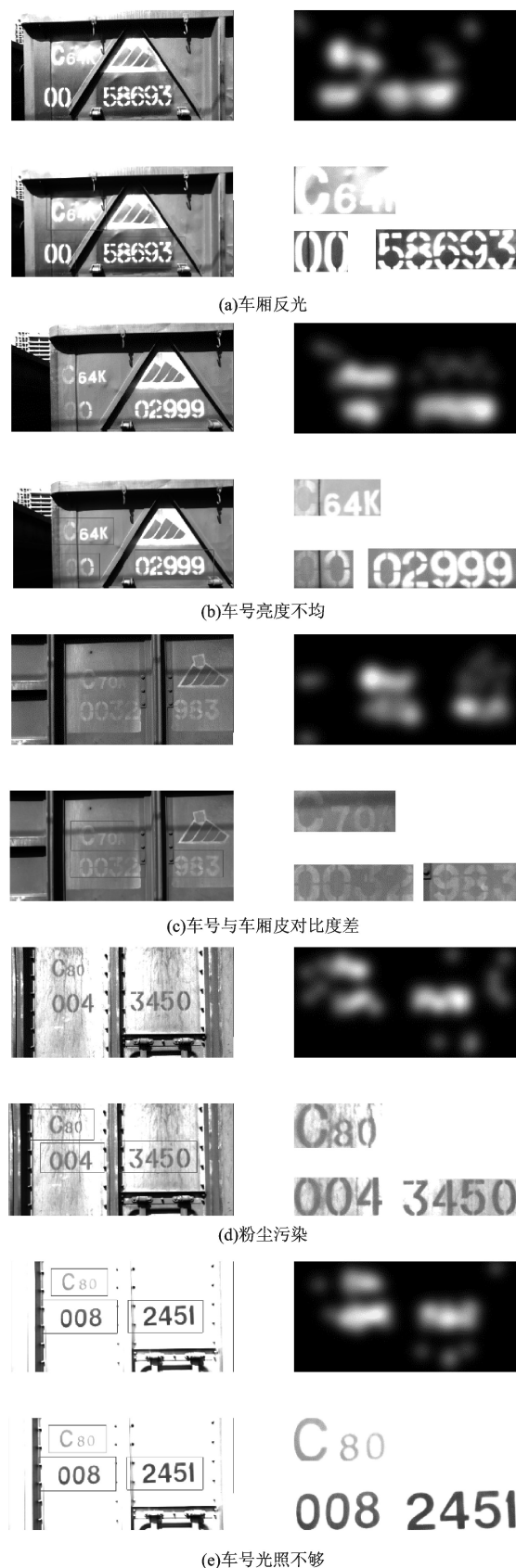


图6 质量较差车厢图像的定位结果

先要建立视觉注意力模型,过程复杂、耗时较长,而且注意力模型利用特定位置采集的车厢图像构建

的,不能很好的适用于任意车厢图像的车号定位,因而具有一定的局限性。

图6显示的是四组利用本文方法能够准确定位而文献[2-3]定位错误的图像,每组图像由4幅图组成,依次是:原图、车号显著图、车号定位图以及车号图像。由图6可以看出,本文算法能够较好地解决反光、强光、光照不均、遮挡以及车号边缘模糊等情况下的车号定位。

4 结束语

本文提出了一种基于视觉注意力机制的列车车号定位的新方法。该方法利用眼动仪获取真实眼动数据,提取车厢图像的显著特征,进而建立视觉注意力模型,利用该模型预测车号显著区域,并分析定位车号区域。实验证明本文方法具有较好的鲁棒性和普适性,更能适应不同光照下各类车厢的车号定位。然而,本文方法过程较为复杂、耗时较长。

参考文献:

- [1] 耿青云,孙泳江. TMIS中央系统整体架构概述[J]. 铁路计算机应用, 2005, 14(7): 39-41.
- [2] 杨绍华,李建勇,王恒,等. 基于BP神经网络的货车车号识别方法研究[J]. 计算机应用, 2007, 16(12): 4-7.
- [3] 赵入宾. 铁路货车车号识别的算法研究[D]. 河北: 河北工业大学, 2010.
- [4] 王少杰,朱志刚,石定机,等. 货运列车车型车号自动分割和识别算法[J]. 模式识别与人工智能, 1998, 11(3): 328-334.
- [5] 沈兰荪,张菁,李晓光. 图像检索与压缩域处理技术的研究[M]. 北京: 人民邮电出版社, 2008.
- [6] Zhai Y, Shah M. Visual attention detection in video sequences using spatiotemporal cues[C]// ACM Multimedia. Santa Barbara, 2006: 815-824.

作者简介:

葛莉女, (1989-), 硕士研究生. 研究方向为图像处理、视线跟踪。

(上接第33页)

- [6] 刘道华,原思聪,兰洋,等. 混沌映射的粒子群优化方法[J]. 西安电子科技大学学报, 2010, 37(4): 764-768.
- [7] 栾丽君,谭立静,牛奔. 一种基于粒子群优化算法和差分进化算法的新型混合全局优化算法[J]. 信息与控制, 2007, 36(6): 708-714.
- [8] 阳春华,钱晓山,桂卫华. 一种混沌差分进化和粒子群优化混合算法[J]. 计算机应用研究, 2011, 28(2): 439-441.
- [9] Ou C M. Design of block ciphers by simple chaotic functions [J]. IEEE Computational Intelligence Magazine, 2008, 3(2): 54-59.
- [10] 杨妍,陈如清,俞金寿. 差分进化粒子群混合优化算法的研究与应用[J]. 计算机工程与应用, 2010, 46(25): 238-241.
- [11] Storm R. On the usage of deferential evolution for function optimization[C]//Proc of Biennial Conference

of the North American Fuzzy Information Processing Society. Piscataway: IEEE Press, 1996: 519-523.

- [12] 吴昱,李元香,徐星. 基于群智能的新型反向混合差分进化算法[J]. 小型微型计算机系统, 2009, 5(4): 903-907.

作者简介:

匡芳君女, (1976-), 博士, 副教授. 研究方向为模式识别、智能信息处理技术及应用、信息安全等。

张思扬男, (1973-), 硕士, 讲师. 研究方向为人工智能、故障诊断、电子技术等。

金忠男, (1961-), 博士, 教授, 博士生导师. 研究方向为模式识别、机器学习、计算机视觉等。

徐蔚鸿男, (1963-), 博士, 教授, 博士生导师. 研究方向为人工智能、模式识别、软件工程等。