

基于疲劳驾驶的人眼定位方法研究

杜永昂, 杨耀权, 金玥佟

(华北电力大学 控制与计算机学院, 河北 保定 071003)

摘要: 人眼定位是疲劳驾驶的研究关键。由粗到精, 先后进行了人脸检测、瞳孔定位。针对现有方向梯度直方图人脸检测算法泛化能力不佳的问题, 提出了一种基于信息熵加权的 HOG 特征提取算法, 该算法将待分类的人脸特征进行信息熵阈值加权, 形成新的 HOG 特征, 然后通过支持向量机进行分类; 针对现有瞳孔定位算法准确率不高的问题, 提出了多算法协同工作的瞳孔定位方法, 以自商图为基准寻找二值分割点, 实现了眼部区域光照不变性, 以灰度积分投影为依据实现了瞳孔精确定位。实验结果表明, 该文提出的人脸检测算法在 CelebA 验证数据集准确率可达到 98.26%, 较传统识别方法有更高的准确率; 而瞳孔定位算法也可达到令人满意的精确度, 提高了瞳孔定位的准确性。

关键词: 瞳孔定位; 信息熵; 自商图像; 光照不变性

中图分类号: TP391.41

文献标识码: A

DOI: 10.19358/j.issn.2096-5133.2020.06.004

引用格式: 杜永昂, 杨耀权, 金玥佟. 基于疲劳驾驶的人眼定位方法研究[J]. 信息技术与网络安全, 2020, 39(6): 19-23, 30.

Research on eye localization method based on fatigue driving

Du Yongang, Yang Yaoquan, Jin Yuetong

(School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: Eye positioning is the key of fatigue driving research. In this paper, from rough to fine, face detection and pupil positioning were carried out successively. To solve the problem of poor generalization ability of existing face detection algorithms based on histogram of directional gradient, this paper proposes a HOG feature extraction algorithm based on information entropy weighting. This algorithm can calculate the threshold weight of face features to be classified to form a new HOG feature, and then classify it by support vector machine. Aiming at the problem that the accuracy of pupil positioning is not high, this paper proposes a method of pupil positioning that works in collaboration with multiple algorithms, searches for binary segmentation points based on the self-quotient chart, realizes the illumination invariance of the eye region, and realizes the accurate pupil positioning based on the gray integral projection. Experimental results show that the face detection algorithm proposed in this paper can achieve 98.26% accuracy in CelebA validation data set, which is higher than the traditional recognition method. The pupil positioning algorithm can also achieve satisfactory accuracy, and improve the accuracy of pupil positioning.

Key words: pupil orientation; information entropy; self quotient image; illumination invariance

0 引言

随着社会经济的发展, 我国汽车保有量逐年增加, 交通安全形势愈发严峻。在所有交通事故的成因中, 由疲劳驾驶造成的交通事故占总交通事故的 20% 左右, 占特大交通事故 40% 以上^[1]。目前, 疲劳驾驶的检测方法主要分为三大类: 基于生理指标的检测方法、基于驾驶员行为特性分析的检测方法和基于面部表情识别的检测方法。基于生理指标的检

测方法采用接触式测量方式, 一般通过测试驾驶人员的生理信号来推测驾驶人的疲劳状态。文献[2]使用一种自制驾驶员脑电图信号检测装置, 采用快速傅里叶变换提取脑电图功率谱密度, 当分析脑电图功率谱密度发现疲劳时, 采取一种自动减速的车辆速度控制策略。这种接触式生理参数测试方法通常需要被测者佩戴相应的装置(如电极片等), 会给驾驶行为造成极大干扰, 不适合实际行车环境下的

应用。基于驾驶人行为特性的疲劳检测方法受道路环境、行驶速度、个人习惯、操作技能等的影响,并不适用我国复杂多变的道路情形。得益于机器视觉等人工智能的发展,基于面部表情的疲劳驾驶检测方法因其应用成本较低、无需干扰驾驶人员,已经成为该领域的研究热点。文献[3]基于开源库提取的人脸轮廓点计算眼和嘴的纵横比,生成眯眼和哈欠特征,基于线性拟合趋势提取法生成车辆操纵活跃度特征,然后采用改进的随机森林模型对疲劳状态进行识别。文献[4]提出了一种采用级联的序列级和帧级两层检测模型,分别处理持续睁眼和眨眼的视频序列,实现由粗到精渐进地检测眨眼的方法,通过随机回归森林计算眼睛开合度,最终完成眨眼过程。文献[5]针对复杂光照和姿态变化造成疲劳驾驶的原因,提出了实时增强约束局部模型的多信息疲劳检测方法。文献[6]提出比例缩减区域(PRA)的方法快速定位人眼,其原理为对面部区域反复迭代缩减,最后得到对人眼区域的定位。因此,构建一种可靠快速的人眼定位技术是对驾驶人员疲劳检测的关键。由于光照、道路两侧树木及楼房阴影等原因的影响,直接在原始图像中进行人眼定位存在较大困难,而且准确率较低,因此,人眼定位一般分为如下步骤:第一步,在原始图像中定位出人脸区域;第二步,在人脸图像中粗定位出人眼区域;第三步,对人眼区域进行精确定位。

考虑到不同特征维度对人脸的贡献程度不同,本文提出了一种基于信息熵加权的方向梯度直方图(Histogram of Oriented Gradients, HOG)与支持向量机(Support Vector Machine, SVM)相结合的人脸检测算法,可以有效提取人脸特征,并准确完成人脸定位。在完成人脸检测之后,以人脸区域为母版,先采用 OpenCV 开源实现对人眼的粗定位;然后采用自商图消除光照影响,遍历寻找最优二值分割点;最后用二值分割图像进行积分投影,实现瞳孔的精确定位。该算法解决了复杂光照下的瞳孔定位问题。

1 人脸检测算法

梯度直方图 HOG 是在 2005 年 CVPR 会议上由 Dalal 提出的特征提取算法^[7]。近年来, HOG 特征提取算法越来越多地受到国内外学者的青睐,积极探索其潜能及应用范围,但在对 HOG 特征的相关研究

中,大部分学者只是将 HOG 算子应用到人脸全局中或只是简单地进行全局特征和局部特征的级联。文献[8]提出的信息熵加权的 HOG 特征提取算法将人脸分割成不同块,在不同块中进行加权来提高不同部位的识别率。

本文提出一种基于信息熵加权的 HOG 特征提取算法。首先计算每幅图片的信息熵;然后提取每幅图片的原有 HOG 特征;最后对原有提取的 HOG 特征实行阈值分割,将符合条件的特征点加权信息熵形成新的 HOG 特征。算法采用 SVM 作为分类器,将加权后的 HOG 特征作为 SVM 的输入特征进行分类。算法步骤如下:

(1) 人脸图像预处理

对检测到的图像进行灰度变换,单通道图片相较于 RGB 三通道图片信息量更少,可以更加快速地完成训练。

$$\text{Gray} = 0.299R + 0.587G + 0.114B \quad (1)$$

(2) 图片 HOG 特征提取

本文将人脸图像矩阵 I 的维数设定为 $I_x \times I_y$, 其中 x, y 均为 64; 每个细胞单元尺寸设定为 $(8, 8)$; 每 2×2 个细胞单元组成一个块(block); 每个细胞单元分为 9 个方向块; 滑动步长设定为一个胞元长度。检测窗口在横向滑动 7 个块, 在纵向滑动 7 个块, 共提取 1 764 个特征点。

(3) 信息熵加权到特征点

由于 HOG 特征点只在人脸轮廓线或鼻子、嘴巴等区域有较为明显的特征, 而在人脸大部分较为平滑的区域特征点不明显, 故将信息熵加权到该区域。设 $D_{N \times M}$ 为特征点向量, 其中 $N=1, M=1\ 764$, 则特征点均值为:

$$\text{Mean} = \frac{\sum_{i=1}^N \sum_{j=1}^M x_{i,j}}{\text{len}(D_{N \times M})} \quad (2)$$

其中 $\text{len}(D_{N \times M})$ 为特征点向量长度。

加权特征点计算为:

$$\begin{cases} D_{N \times M}(i) \times E[I(x_i)], & \text{Mean} \geq D_{N \times M}(i) \\ \text{其他}, & \text{Mean} < D_{N \times M}(i) \end{cases} \quad (3)$$

其中 $D_{N \times M}(i)$ 为第 i 个特征点向量, $E[I]$ 为图片信息熵。本文阈值点的选择在实验部分有详细的实验数据作为参考。

(4) 对信息熵加权后的 HOG 特征点进行 SVM 分类。

2 人眼定位

2.1 基于自商图的局部纹理建模

自商图的计算公式为:

$$R = \frac{I}{\hat{I}} = \frac{I}{F \otimes I} \quad (4)$$

其中, R 为自商图, I 为灰度图, F 为滤波核。

根据自商图的计算公式, 将采集的原始图像转变为自商图, 消除眼睛区域的光照变化, 以更好实现二值化阈值分割。将人眼区域灰度图进行低通滤波, 遍历灰度图与低通滤波图像中的所有像素点, “点对点”地作商运算, 并将计算结果归一化到 $[0 \sim 255]$ 灰度范围。

2.2 积分投影定位人眼

一般而言, 脸部图像中眼睛区域与周围区域相比较暗, 利用该特征, 常采用灰度积分投影法来定位眼睛, 根据图像的波谷和波峰的分布情况来定位眼睛区域^[9]。

假设图像大小为 $m \times n$, $G(x, y)$ 为像素点 (x, y) 的灰度值, 其在水平方向的积分投影为:

$$H(y) = \sum_{x=x_1}^{x_2} G(x, y) \quad (5)$$

在图像 $[y_1, y_2]$ 区域的垂直积分投影为:

$$V(x) = \sum_{y=y_1}^{y_2} G(x, y) \quad (6)$$

3 本文瞳孔定位算法

本文提出了由粗到精的瞳孔检测方法, 以自商图为基准寻找二值分割点, 实现了眼部区域光照不变性, 以灰度积分投影为依据实现了瞳孔精确定位。传统灰度积分投影是直接对原始图像进行灰度化, 然后将灰度图进行二值化, 最后将二值化图像进行水平投影与垂直投影。此方法虽然计算效率高, 但是容易受光照及阴影所影响, 在实际应用中存在一定的缺陷。本文引入自商图, 用以消除光照的影响, 同时可以保留积分投影的运算效率, 满足了实际应用。瞳孔定位算法主要步骤为:

(1) 应用 OpenCV 得到左、右眼睛的大体区域。

(2) 对眼睛区域进行自商图变化, 结果如图 1(b) 所示, 本文滤波核采用高斯滤波。

(3) 从图 1(b) 不难看出, 眼部灰度最小值在瞳孔附近, 故遍历眼部区域自商图, 将灰度值进行排序, 得到灰度最小值。设 $F_{N \times M}$ 为自商图, N 、 M 分别为该图行列数, 则最小值为:

$$T = \text{Min}F(i, j), i \in N; j \in M \quad (7)$$

(4) 以 T 为分割点做灰度阈值变换, 将其转换为二值图像:

$$f(x) = \begin{cases} 0, & x \leq T \\ 255, & x > T \end{cases} \quad (8)$$

(5) 对二值图像进行水平、垂直积分投影。

(6) 得到瞳孔定位的横纵坐标。

如图 1 所示, (a) 为原始图像人脸检测区域; (b) 为图像自商图; (c) 为二值化图像。可以看出基于自商图的眼部细节更加突出。本数据来源 LFW 公开数据集, 数据编号为 Zico_0002。



图 1 原始图像、自商图及二值图像对比结果

本文提出的瞳孔定位算法流程图如图 2 所示。

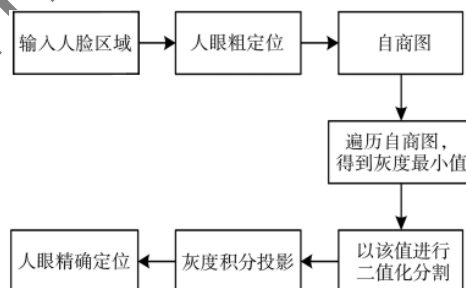


图 2 瞳孔定位算法流程图

4 实验结果及分析

通过 Python 工具平台进行实验, 实验运行环境为 InterCore I5 处理器, 8 GB 运行内存, Windows 10 操作系统。人脸检测数据集采用香港中文大学公布的大型人脸识别数据集 CelebA, 该数据集包含 20 万张人脸图片, 人脸属性有 40 多种, 在表情、光照及姿态等方面更权威。瞳孔定位实验数据集采用 HELEN 系列数据集, 该数据集包含 348 幅图片。实验识别分类器采用支持向量机分类器。

4.1 基于 CelebA 库人脸检测实验

本文选用 CelebA 人脸库编号为 1~200 的图片作为训练样本, 选用编号为 201~350 的图片作为测试样本。正、负训练集样本共为 400 张; 正、负测试集样本共为 300 张。首先对特征点阈值进行选择,

通过不同阈值点对比选择出最佳的阈值点。

在 CelebA 人脸库中对于不同阈值共进行了 4 次实验,每一次对人脸特征点进行不同的阈值划分,通过对 4 次阈值划分所得的实验结果进行比较,选取最佳阈值点。得到的最佳实验结果即为本文方法的最终实验结果。

令 $\text{Mean1}=0.8\times\text{Mean}$; $\text{Mean2}=1.25\times\text{Mean}$; $\text{Mean3}=1.5\times\text{Mean}$ 。表 1 给出了几种阈值的识别结果。识别结果用百分比表示。不同阈值识别率比较如图 3 所示。

表 1 不同阈值所得识别结果

阈值点划分	Mean1	Mean	Mean2	Mean3
准确率	96.95	98.26	98.25	97.82
误检率	3.05	1.74	1.75	2.18

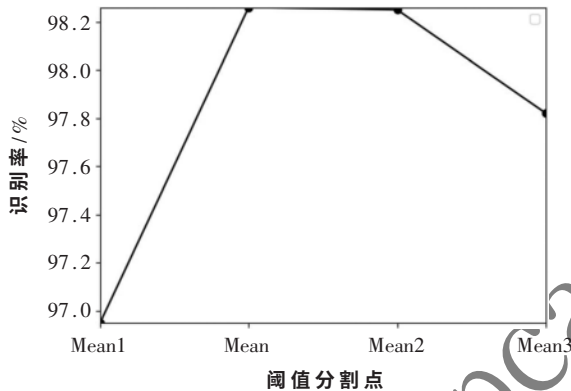


图 3 不同阈值识别率比较

根据不同阈值仿真结果可以看出以 Mean 为阈值点识别效果最佳,Mean2 次之,然后依次为 Mean3、Mean1。如果 Mean1 作为阈值划分点,信息熵只能加权到小部分特征点,不能很好表示全局特征点,导致识别率较低;Mean2、Mean3 则是因为信息熵加权到了太多的特征点,造成了大量噪声,给分类带来了干扰,导致识别率有所下降。

在选择好最佳阈值点后,用原始 HOG+SVM 人脸检测与本文人脸检测算法进行实验对比,实验结果如表 2 所示。

从不同算法的对比仿真结果中可以看出,本文

表 2 不同算法识别率比较

算法	原 HOG+SVM	本文算法
准确率	97.39	98.26
误检率	2.61	1.74

提出的人脸检测算法识别率高于原始 HOG+SVM 人脸检测算法。原始 HOG+SVM 算法主要依赖于全局特征,有部分关键特征点并不能很好地体现出来,而本文所提出的信息熵加权法则对这些特征点进行信息熵加权,从而提高了识别率。

此外,本文还在 HELEN_1 公开数据集对原始 HOG+SVM 人脸检测算法与本文提出的人脸检测算法进行了仿真实验,并选择部分样本图片作为展示结果,以便更加直观地体现该算法性能。其中数据类型分别为脸部右侧旋转、脸部左侧旋转、光照明亮、光照较暗、光影对比强烈、侧脸、阴影强烈、佩戴眼镜、抬头、低头等类型。数据编号分别为 10697993_1、13789611_1、16542667_1、103236168_1、137346980_1、170508172_1、143588896_1、151902153_1、167629013_1、213033315_1。仿真结果如图 4、图 5 所示。



图 4 原始 HOG+SVM 人脸检测算法结果



图 5 本文人脸检测算法结果

通过对 HELEN_1 公开数据集验证可知,如图 4 所示,HOG+SVM 人脸检测算法对于脸部左侧旋转(13789611_1)、阴影强烈(143588896_1)、抬头幅度过大(167629013_1)等情况检测结果均不佳,出现漏检等情况。本文提出的人脸检测算法,如图 5 所示,无论对于脸部右侧旋转、左侧旋转,还是抬头等均可达到满意的检测效果,但是对于阴影强烈(143588896_1)的情况因为训练样本不足还有待增强。由仿真实验结果可知,本文提出的人脸检测算法比传统人脸检测算法精度更高、鲁棒性更强。

4.2 瞳孔定位实验

本文提出的基于自商图先粗后精的瞳孔定位算法,充分考虑了光照、阴影等环境,该算法对光影及噪声干扰等有较强适应性。一般积分投影法在眼睛定位时,采用的图像预处理方法为自适应阈值分割法,以该法获取二值图像。虽然该分割法在眼部区域光照较为明亮时会得到较好的效果,但当面部阴影较为严重时,使得整个眼部都会出现二值为黑的情况,如图 1(c)所示。文献[10]提出了一种基于肤色和积分投影结合的眼部定位方法,该方法在水平投影和垂直投影方向均出现了多个波峰、波谷的情况,不易得到准确的瞳孔坐标值。本文以眼部最小灰度值为分割点,可以较为理想地得到瞳孔横纵坐标值。

本文随机选取 50 张实际行车环境图片用于瞳孔定位实验,首先通过上文提到的自商图处理眼部区域,之后通过对每幅图片遍历求取灰度最小值,以该灰度值进行图像二值化,最后进行灰度积分投影,得到准确的瞳孔定位坐标。

表 3 给出了本文与传统积分投影瞳孔定位的识别结果。识别结果用百分比表示。

表 3 本文与传统积分投影瞳孔定位算法比较

算法	传统定位方法	本文算法
准确率	81.15	88.16
误检率	18.85	11.84

本文算法充分考虑了眼部区域结构性特点,即瞳孔附近灰度值最低的特性,并引入自商图消除光影的影响。由仿真实验结果可知,本文提出的瞳孔定位算法比传统算法泛化能力更强,准确度更高。

为了更加直观地展现该算法的优势,本文从样本中选取部分图片作为展示结果。其中数据类型分别为脸部右侧旋转、脸部左侧旋转、抬头、表情(打哈欠)等类型。数据帧数分别为 1316、1139、469、446、66、26。仿真结果如图 6、图 7 所示。



图 6 基于传统积分投影瞳孔定位



图 7 基于自商图二值分割点瞳孔定位

通过数据仿真结果可知,如图 6 所示,基于传统积分投影法对瞳孔定位时,如果头部出现偏转、抬头时将会出现误检等情况,如编号 1316、1139、469、446 所示,在实际应用中检测精度较低。本文提出的瞳孔定位算法,如图 7 所示,对于头部左右旋转、抬头等情况都能较为精准地定位出瞳孔位置。相较于传统的人眼定位方法,本文提出的瞳孔定位方法在准确度方面得到了较好提升。

5 结束语

人眼定位是疲劳驾驶检测的关键环节,针对现有方向梯度直方图人脸检测算法泛化能力不佳的问题,本文提出了一种基于信息熵加权的 HOG 特征提取算法,该算法将待分类的人脸特征进行阈值加权,提高了原有人脸检测的识别率。针对现有瞳孔定位准确率不高的问题,本文提出了多算法协同工作的瞳孔定位方法,以自商图为基准寻找二值分割点,实现了眼部区域光照不变性,以灰度积分投影为依据实现了瞳孔精确定位。仿真实验结果表明,本文提出的人脸检测算法比传统人脸检测算法鲁棒性更高、泛化能力更强;本文提出的瞳孔定位算法相较于传统方法,定位精度更高,可以为疲劳驾驶检测研究提供更加科学、可靠的数据。

参考文献

- [1] 刘文玲,钱晓飞,裴军.基于关联规则的公交事故受伤情况预测研究[J].控制工程,2016,23(9): 1448-1453.
- [2] ZHANG Z T, LUO D Y, RASIM Y, et al. A vehicle active safety model: vehicle speed control based on driver vigilance detection using wearable EEG and sparse representation[J]. Sensors, 2016, 16(2): 242.
- [3] 吴士力,唐振民,刘永.多特征融合的随机森林疲劳驾驶识别算法[J/OL].计算机工程与应用: 1-8 [2020-04-27]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20190916.1034.010.html>.
- [4] 高宁,王兴元,王秀坤.基于眼动序列分析的眨眼检测[J].计算机工程与应用,2019,55(8): 40-47, 73.

(下转第 30 页)

- [11] CHENG Y, SHEN S Q, HE Z J, et al. Agreement-based joint training for bidirectional attention-based neural machine translation[C]. Proceedings of the 25th International Joint Conference on Artificial Intelligence. Menlo Park, CA: AAAI, 2016: 2761-2767.
- [12] CAI G, LI H. Joint attention LSTM network for aspect-level sentiment analysis[C]. Proceedings of the 24th China Conf on Information Retrieval (CCIR 2018). Los Alamitos, CA: IEEE Computer Society, 2018: 147-157.
- [13] YI Q, LIU J, ZHANG G, et al. Aspect-level sentiment classification with conv-attention mechanism[C]. Proceedings of the 25th Int Conference Los Alamitos. CA: IEEE Computer Society, 2018: 231-243.
- [14] YIN W P, SCHÜTZE H, XIANG B, et al. ABCNN: attention-based convolutional neural network for modeling sentence pairs[J]. Association for Computational Linguistics, 2016, 4: 259-272.
- [15] WANG X Y, XU G L, ZHANG J Y, et al. Syntax-directed hybrid attention network for aspect-level sentiment analysis[J]. IEEE Access, 2019, 7: 5014-5025.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]. Proceedings of the 31st Conference on NIPS2017. Cambridge, MA: MIT Press, 2017: 6000-6010.
- [17] KIM Y. Convolutional neural networks for sentence classification[C]. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar, 2014: 1746-1751.
- [18] YIN W P, SCHUTZE H, XIANG B, et al. Attention-based convolutional neural network for modeling sentence pairs[J]. Transaction of the Association for Computational Linguistics, 2016, 4(11): 259-272.
- [19] LIU J M, ZHANG Y. Attention modeling for targeted sentiment[C]. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, 2017, 2: 572-577.
- [20] WANG X, LIU Y C, SUN C J, et al. Predicting polarities of Tweets by composing word embeddings with long short-term memory[C]. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Beijing, China, 2015: 1343-1353.
- [21] TANG D Y, QIN B, FENG X C, et al. Effective LSTMs for target dependent sentiment classification[C]. Proceedings of the 26th International Conference on Computational Linguistics, 2016: 3298-3307.

(收稿日期: 2020-04-17)

作者简介:

宋婷(1984-), 女, 硕士, 中级实验师, 主要研究方向: 人工智能与数据挖掘。

陈战伟(1984-), 男, 硕士, 高级工程师, 主要研究方向: 人工智能与数据挖掘。

(上接第 23 页)

- [5] 李长勇, 吴金强, 房爱青. 基于多信息的疲劳状态识别方法[J]. 激光与光电子学进展, 2018, 55(10): 239-245.
- [6] 贺德强, 刘卫, 卢凯, 等. 基于机车车载视频序列的人眼疲劳快速定位方法研究[J]. 铁道科学与工程学报, 2018, 15(9): 2359-2366.
- [7] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005: 886-893.
- [8] 林克正, 张元铭, 李昊天. 信息熵加权的 HOG 特征提取算法研究[J]. 计算机工程与应用, 2020, 56(6): 147-152.

- [9] 侯向丹, 赵丹, 刘洪普, 等. 基于积分投影和差分投影的人眼定位[J]. 计算机工程与科学, 2017, 39(3): 534-539.

- [10] 荔小虎, 唐晶磊. 基于肤色匹配和积分投影结合的人眼定位方法研究[J]. 自动化应用, 2018(12): 59-61.

(收稿日期: 2020-04-27)

作者简介:

杜永昂(1993-), 通信作者, 男, 硕士研究生, 主要研究方向: 数字图像处理。E-mail: dya531500@163.com。

杨耀权(1962-), 男, 教授, 主要研究方向: 智能测试技术、数字图像处理。

金玥佟(1994-), 女, 硕士研究生, 主要研究方向: 数字图像处理。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所