

一种基于单目视觉的实时人机测距系统研究

张帆, 董秀成, 王艺

(西华大学 四川省信号与信息处理重点实验室, 成都 610039)

摘要: 为获取视觉交互系统中人机距离信息, 利用单摄像头检测人脸特征区域、完成人脸与摄像头距离测量。首先通过改进的 AdaBoost 算法快速检测并定位特征区域, 其次利用系统约束条件、摄像头标定原理和面积映射关系推导像距测量方程, 完成距离测量。通过规定距离实验, 证明了像距测量方程的可行性。复杂环境实验表明系统在复杂背景下有较好的鲁棒性, 但不同光照会影响测量精度。系统适应性实验验证了对多名待测者的普遍适用性, 但佩戴眼镜会影响测量的精度和实时性。测距系统在有效量程内有较高精度, 处理速度约为 5 fps, 满足精度与实时性要求。

关键词: 单目视觉; 距离测量; 特征提取; 像距建模; 图像处理

中图分类号: TP391.7 **文献标志码:** A **文章编号:** 1001-3695(2013)12-3866-04

doi:10.3969/j.issn.1001-3695.2013.12.091

Real-time face to camera distance measurement system based on monocular vision

ZHANG Fan, DONG Xiu-cheng, WANG Yi

(Key Laboratory of Signal & Information Processing in Sichuan Province, Xihua University, Chengdu 610039, China)

Abstract: To accomplish the face to camera distance measurement, this paper proposed a method by extracting characteristic region, calculating pixel-area, and constructing the pixel-distance mathematic model. To detect and locate the characteristic region fast, it improved the traditional AdaBoost algorithm. It derived a mathematic formula revealing the relationship between pixel-area and distance from the pinhole model, camera calibration and area mapping. The specific distance and system adaptive experiment verified the feasibility of proposed method and the general applicability of measuring system. The experiment results show that measuring system performance is robust under complicated environment, the accuracy is high within the range and the system operating frequency is about 5 frames per second.

Key words: monocular vision; distance measurement; features extraction; pixel-distance modeling; image processing

视觉测距作为计算机视觉领域基础技术具有重要的地位, 它被广泛应用于视觉导航、机器视觉定位、目标跟踪等研究中, 尤其在人机交互系统中具有重要的使用价值^[1]。单目视觉测量是指利用一台数码相机或摄像头拍摄单幅图像以完成距离测量工作。由传统的计算机立体视觉可知, 要恢复景物的三维信息, 需要从多个角度获取被测景物信息。对于部分视觉系统, 利用其本身存在的约束条件, 可以恢复图像的深度信息。

在基于人眼视线的人机交互系统中, 操作者与屏幕之间的距离是设计人机交互算法的必要参数。人机距离测量精度的高低直接影响交互效果的好坏。本文在设计距离测量算法时, 规定了两个约束条件: a) 屏幕与摄像头安置在同一平面; b) 人脸正对屏幕。在约束条件下可知, 当人脸离摄像头越近, 则人脸图像越大, 反之则越小^[2,3]。利用这一规律, 通过坐标变换找出图像像素与距离的关系, 建立对应的数学模型, 完成测距任务。测距算法流程如图 1 所示。

测距算法主要由四部分组成:

- a) 图像预处理, 降低输入图像分辨率, 完成图像增强处理。
- b) 特征提取与定位, 确定人脸双眼和鼻子的特征并定位,

构建人脸特征三角形, 计算像素面积。

c) 像距建模, 利用一次最优移动构建像素与距离的数学模型, 确定像距方程。

d) 误差修正, 利用多次测量和相关约束减少误差。

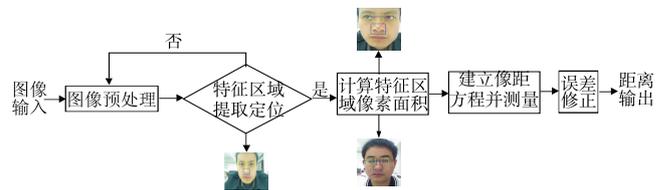


图1 人机测距算法流程

1 图像预处理

摄像头直接采集的图像分辨率为 640 × 480, 大分辨率图像虽然信息丰富, 易提取和定位特征点, 但处理时间长, 达不到实时性要求。为解决上述问题, 本文采用了金字塔降采样降低图像分辨率。图像金字塔是以多分辨率来解释图像的一种有效但概念简单的结构。一幅图像的金字塔是一系列以金字塔形状排列的分辨率逐步降低的图像集合^[4]。输入图像首先通

收稿日期: 2013-01-09; 修回日期: 2013-03-04

作者简介: 张帆(1987-), 男, 四川宜宾人, 硕士研究生, 主要研究方向为计算机视觉、人机交互、模式识别(ybz912@163.com); 董秀成(1963-), 男, 教授, 硕士, 主要研究方向为基于视觉的智能控制与智能信息处理、复杂非线性控制系统建模与控制、工业自动化技术及系统、虚拟现实技术及应用。

过与 $G_{5 \times 5}$ 的高斯核进行卷积,然后通过拒绝偶数的行与列来下采样图像。通过一次降采样,原始输入图像分辨率变换为 320×240 。值得注意的是,金字塔的分辨率越低,伴随的细节就越少。经实验,若不使用降采样处理,系统完成一次测量的时间约为 800 ms,达不到实时性要求。若采用了一次降采样,处理时间仅需 230 ms,实时性要求获得满足。若进行两次降采样,图像运算时间虽然可以更短,但图像信息丢失太多,测量精度达不到测量要求,故本系统只采用一次图像金字塔降采样变换。

摄像头采集的人脸图像受光照、人脸肤色等环境因素影响。当环境光照过暗或摄像头曝光不足时,图像的灰度范围局限在低灰度区间,此时图像模糊不清,对比度不强,人脸特征不易提取。采用直方图均衡化方法可快速有效解决上述问题。

直方图均衡化的思想是把原始的直方图变换成均匀分布的形式,增加像素灰度值的动态范围,达到增强图像整体对比度的效果^[5]。如图 2 所示,经直方图均衡化后图像灰度级明显扩大,输入图像明显改善。



图2 均衡化前后图像对比

2 人脸特征点提取与定位

人脸特征点的提取与定位是整个测距系统的基础,此处特征点是指双眼和鼻子。通过定位特征点,构建特征三角形,计算其像素面积,为后续测距提供参数。本系统关注的是特征点像素面积,故特征点的绝对位置并非必需,且按常规方法精确定位眼鼻坐标耗时过多,系统实时性得不到满足。为此,本文改进了传统的 AdaBoost 算法,优化了眼、鼻特征集,训练了效率更高的分类器。改进的 AdaBoost 算法提取特征点耗时少、鲁棒性更高、相对定位更准确。

2.1 特征区域提取

采用级联 AdaBoost 算法对人脸特征进行提取是比较常用的一种方法^[6]。传统的 AdaBoost 算法存在训练速度缓慢、检测过分依赖训练样本的现象。本文在传统 AdaBoost 算法基础上,充分考虑双眼、鼻子的特征,通过简约特征数和引入样本扩张策略来提高训练速度和特征检测率。

样本的训练速度主要受制于训练集数量和特征数量^[7]。为保证最后强分类器的效果,训练集数量是不能减少的。因此,只能通过精简特征数量来提高训练速度。精简的主体思路是剔除那些对人眼和非人眼区分不敏感的特征。主要做法是在特征生成阶段剔除对人眼和非人眼区分贡献不大的特征以及在训练过程中减少矩形特征。

为提高特征的检测率,本文采用了样本扩张策略,使扩张后的样本能更为全面地区别特征与非特征。首先随机选取正负样本构成初始的训练库;然后利用训练库训练弱分类器,再由弱分类器组合成为强分类器;其次利用训练好的强分类器对不同条件下的特征进行检测,将无特征图像中的误检测图像作为非特征样本进行样本扩充,同时也对有特征图像中的漏检测图像进行样本扩充^[8];最后对新训练的样本库再次训练,可得

到稳定性更好的强分类器。

本文利用上述方法改进了传统 AdaBoost 算法,生成了特征提取率更高的分类器。特征检测效果如图 3 所示。

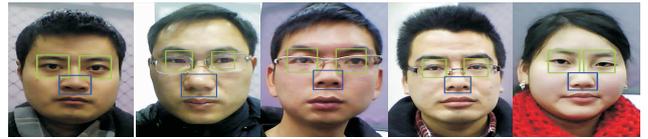


图3 特征区域检测

2.2 特征区域定位

在正面人脸特征约束条件下,双眼特征框应保持平行,且双眼的瞳孔位置应在框内中心处^[9,10]。为了精确定位双眼的相对位置,限定单眼的最大尺寸为 25×25 像素,且在第一次搜索到人眼后,将搜索范围扩大 5%,并连续搜索五次,记录每一次搜索框的中心点坐标。若五次搜索都存在,则判定人眼特征存在,且人眼区域特征最终的中心点坐标定义为五次搜索坐标的平均值^[11]。对于鼻子的定位,本文直接采用改进的 AdaBoost 方法进行特征提取,并选取定位框的中心点坐标作为鼻子的像素坐标。经实验,该方法对鼻子特征的提取有较高的鲁棒性,定位精度高。眼鼻定位及人脸特征三角形如图 4 所示。

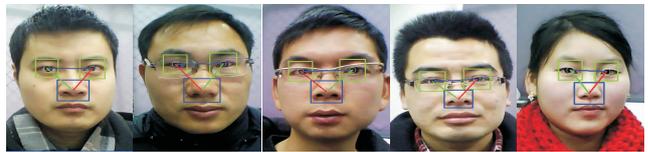


图4 人脸特征区域定位

3 测量原理

3.1 针孔成像模型

本文采用高清网络摄像头采集图像,将人脸图像投影到摄像机二维平面^[12]。如图 5 所示,由针孔成像模型定义,物平面点 P 通过摄像机坐标系将像投影到摄像机像平面 p' ,由比例关系可知:

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (1)$$

其中: (x, y) 为 p' 点在图像坐标系下的物理坐标, (X_c, Y_c, Z_c) 为 P 点在摄像机坐标系下的坐标。

图像在计算机内部以数字形式存储,在图像坐标系下,建立数字图像的像素坐标系 (u, v) ,图像平面的中心点坐标为 (u_0, v_0) 。由摄像机标定知识可以得到以下公式:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} d_x & 0 & -u_0 d_x \\ 0 & d_y & -v_0 d_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (2)$$

其中: d_x, d_y 为单位物理尺寸上的像素点个数。摄像机坐标系和世界坐标系存在如下关系:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (3)$$

其中: (X_w, Y_w, Z_w) 为点 P 在世界坐标系下的坐标; R 为 3×3 的正交旋转矩阵; T 为 3×1 的平移向量。将式(2)(3)代入式(1)则可以得出物体在世界坐标系下和图像像素坐标系下的

关系:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & u_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (4)$$

其中: $\alpha_x = f/d_x, \alpha_y = f/d_y$ 。根据上述分析,可以从世界坐标系下的点直接求出图像像素坐标系下点的坐标。

3.2 像距建模

由上文可知,本文所采用的人脸特征三角形为两眼的瞳孔中心点及鼻尖所构成的特征三角形 $\triangle ABC$,如图6所示。

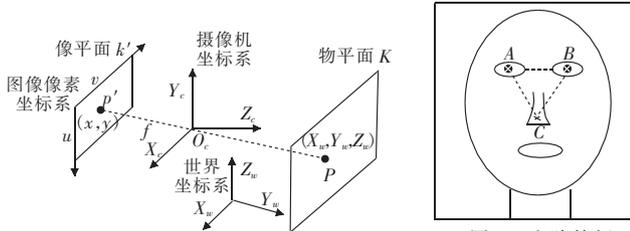


图5 针孔成像模型

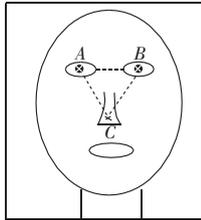


图6 人脸特征三角形示意图

利用上述已构建的世界坐标与图像像素坐标的关系,将世界坐标系与摄像机坐标系按图7方式建立^[13],则可得 $R = I, T = [0 \ 0 \ d]$,代入式(4)可得

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x X_w + u_0 Z_w + u_0 d \\ \alpha_y Y_w + v_0 Z_w + v_0 d \\ Z_w + d \end{bmatrix} \quad (5)$$

本文考虑人脸特征近似在一个平面,故 $Z_w = 0$,进一步化简式(5)可得

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{d} \begin{bmatrix} \alpha_x X_w + u_0 d \\ \alpha_y Y_w + v_0 d \\ d \end{bmatrix} \quad (6)$$

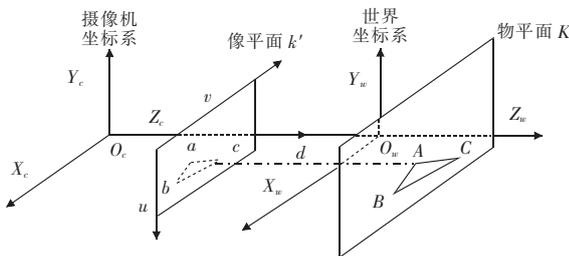


图7 坐标系变换

如图7所示,定义世界坐标系下三角形 $\triangle ABC$ 面积为 $S_{\triangle ABC}$,投影在像平面内的三角形 $\triangle abc$ 面积为 $S_{\triangle abc}$ 。物平面内 A, B, C 三点坐标分别为 $(X_A, Y_A), (X_B, Y_B), (X_C, Y_C)$,像平面内 a, b, c 三点的像素坐标为 $(u_a, v_a), (u_b, v_b), (u_c, v_c)$ 。

由三角形面积公式的向量表达式可知: $S_{\triangle ABC} = (1/2) \times |AB \times AC|$,将其转换为坐标表达式为

$$S_{\triangle ABC} = (1/2) |(X_B - X_A)(Y_C - Y_A) - (X_C - X_A)(Y_B - Y_A)| \quad (7)$$

同理可得

$$S_{\triangle abc} = (1/2) |(u_b - u_a)(v_c - v_a) - (u_c - u_a)(v_b - v_a)| \quad (8)$$

将式(6)代入式(8),可得

$$S_{\triangle abc} = (\alpha_x \alpha_y) / d^2 \cdot S_{\triangle ABC} \quad (9)$$

$$d = \sqrt{\alpha_x \alpha_y S_{\triangle ABC} S_{\triangle abc}^{-0.5}} \quad (10)$$

从式(10)中可以看出, α_x, α_y 为摄像机内参数,经过摄像机标定可以求得^[14], $S_{\triangle ABC}$ 为人脸特征三角形的物理面积,其大小

只与待测者有关,且不会随距离的变化而变化。至此,建立了一个物理距离随像素面积变化的数学模型,即像距测量方程。

4 实验结果分析

本文测量在 CPU 为 i3-2100,主频为 3.1 GHz,内存 DDR 2 GB 的 PC 机上进行。利用 500 万像素 Logitech 快看太空版网络摄像头和 VS2010 编程平台对测量数据进行了分析。每幅输入图像的分辨率为 640×480 ,系统处理速度为 5 fps。为减少因单次获取特征点像素面积不准确而引入的测量误差,将同一距离的像素面积求和三次再取其平均值。在求和过程中,为防止因特征点误识别而带来的像素面积突变,约束前后两次像素面积差的绝对值不得大于 350。测量系统界面如图8所示。

为验证系统测量的准确性、复杂环境的适应性和使用的普遍性,进行以下实验:

实验1 为验证测距方程的可行性,选择一名固定待测者#1,在 20~80 cm 的范围内以 5 cm 为移动步长进行规定距离测量。待测者#1 的像距方程为: $d = 3154.79 \times S_{\triangle abc}^{-0.5}$ 。测量准确率如图9所示,测量结果如表1所示。

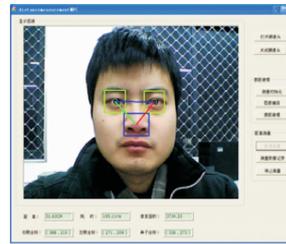


图8 系统测量界面

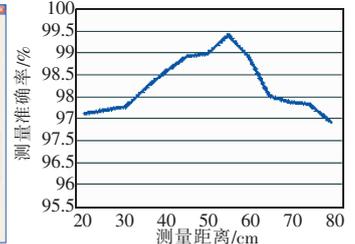


图9 待测者#1 规定距离测量准确率

表1 待测者#1 测量结果

序号	真实距离/cm	像素面积	测量距离/cm	绝对误差/cm	准确率/%	单次耗时/ms
1	20	23729.32	20.48	0.48	97.60	233.07
2	25	16688.52	24.42	0.58	97.68	223.45
3	30	11571.31	29.33	0.67	97.76	229.66
4	35	8422.72	34.38	0.62	98.21	221.32
5	40	6399.38	39.44	0.56	98.59	231.95
6	45	5021.33	44.52	0.48	98.93	236.62
7	50	4062.81	49.49	0.51	98.99	229.32
8	55	3330.17	54.67	0.33	99.40	239.46
9	60	2825.13	59.35	0.65	98.92	234.96
10	65	2452.33	63.71	1.29	98.01	242.82
11	70	2119.83	68.52	1.48	97.89	233.56
12	75	1694.51	76.64	1.64	97.82	231.32
13	80	1639.07	77.92	2.08	97.41	242.17

通过实验可以发现,由像距数学模型推导建立的理想测距方程作用于待测者#1,在所测范围内最大绝对值误差为 2.08 cm,测量准确率高于 97.41%,测量精度较高,测距理论得以验证。

实验2 为检验测量系统在不同光照条件下的测量性能,选取#1 待测者在三种不同光照条件下进行实验,记测量准确率为 ρ 。

条件1 白天自然条件,待测者上方有两处白炽灯源,背景无强光源干扰,实验过程光照充足。

条件2 白天自然条件,待测者上方无白炽灯源,背景无强光源干扰,实验过程光照强度中等。

条件3 白天自然条件,待测者上方有两处白炽灯源,背景有强光源干扰,实验过程光照强。

条件 4 夜间自然条件,待测者上方有两处白炽灯源,实验过程光照强度较弱。测量结果如表 2 所示。

表 2 四种不同条件测量准确率对比

测量条件	规定距离测量准确率 $\rho/\%$			有效量程/ cm ($\rho \geq 95\%$)	单次平均 耗时/ms
	30 cm	55 cm	70 cm		
条件 1	97.54	99.31	97.32	18 ~ 86	233
条件 2	97.11	98.92	97.04	21 ~ 81	241
条件 3	95.12	96.45	93.23	27 ~ 66	272
条件 4	96.29	97.94	95.21	24 ~ 73	253

通过实验 2 可以发现,系统的测量性能整体表现稳定,在各种测试条件下都完成了测距任务,且获得较高的测量准确率,随外界光照条件的改变而发生较大变化。在光照充足且无其他光源干扰的条件下,测量性能最好;当系统受外界光源干扰时,测量精度和量程都受较大影响,测量实时性也降低。

实验 3 为检验测量系统对不同待测者的适应性,选择 6 名待测者,分别标记为待测者#1 ~ #6,检验在 25 ~ 80 cm 范围内测量的精度,以验证测量系统对不同待测者的适应性和使用的普遍性。六名待测者如图 10 所示。



图 10 六名待测者人脸图像

为保证测量的可靠性,本文对测量环境进行了约束,测量实验都规定在实验 2 中条件 1 下进行。其中前三个待测者没有佩戴眼镜,后三个待测者佩戴有框眼镜。测量结果如图 11 所示。

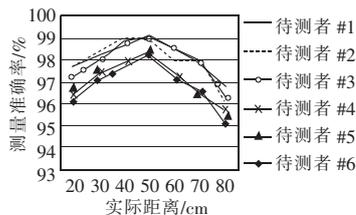


图 11 不同待测者测量准确率对比图

测量系统准确地捕捉了不同待测者的人脸特征信息,完成测距任务,使用普遍性较好。分析测量结果可知,佩戴眼镜的待测者测量准确率略低于未佩戴眼镜的测量者。但对于所有待测者来说,测量精度都保持在 95% 以上。

5 结束语

本文提出一种基于人脸特征区域像素面积的单目人机测

距新方法。该方法结构简单,测量时不需要附加设备,测量具有较好的精度且能够满足实时性。系统首先通过图像预处理,将输入图像处理为满足系统要求的图像,其次通过精简特征数量和样本扩张策略改进传统的 AdaBoost 算法快速检测并定位特征区域,构建特征三角形。最后利用系统约束条件、摄像机标定原理和面积映射关系推导像素面积与人机距离的数学关系,建立像距测量方程,完成距离测量。通过规定距离实验、复杂环境实验和系统适应性实验,验证了测量系统的可行性、稳定性和普遍适用性。实验结果表明,系统测量精度较高,处理速度约为 5 fps,满足精度与实时性要求。

参考文献:

- [1] 黄桂平,李广云,王保丰.单目视觉测量技术研究[J].计量学报,2004,25(4):314-317.
- [2] HSU Chen-chien, LU Ming-chih, WANG Wei-yen. Distance measurement based on pixel variation of CCD images[J]. ISA Trans, 2009, 48(1):389-395.
- [3] LU Ming-chih, WANG Wei-yen, CHU Chun-yen. Image-based distance and area measuring systems[J]. IEEE Sensors Journal, 2006, 6(2):495-503.
- [4] 王伟,黄非非,李见为.采用 LBP 金字塔的人脸描述与识别[J].计算机辅助设计与图形学学报,2009,21(1):94-100.
- [5] 陈文飞,廖斌,许雪峰,等.基于 Piecewise 直方图均衡化的图像增强方法[J].通信学报,2011,32(9):153-158.
- [6] VIOLA P, JONES M. Robust real-time face detection[J]. International Journal of Computer Vision, 2004, 57(2):137-154.
- [7] 张彦峰,何佩琨.一种改进的 AdaBoost 算法——M-Asy AdaBoost [J].北京理工大学学报,2011,31(1):64-68.
- [8] 甘玲,朱江,苗东.扩展 Haar 特征检测人眼的方法[J].电子科技大学学报,2010,39(2):247-250.
- [9] 杜志军,王阳生.正面人脸图像中眼睛的定位算法[J].计算机辅助设计与图形学学报,2009,21(6):763-769.
- [10] 艾娟,姚丹,郭跃飞.基于块的眼睛定位方法[J].中国图象图形学报,2007,12(10):1841-1844.
- [11] 郝明刚,董秀成,黄亚勤.一种精确的人眼瞳孔定位算法[J].计算机工程,2012,38(8):141-143.
- [12] 敬泽,薛方正,李祖枢.基于单目视觉的空间目标位置测量[J].传感器与微系统,2011,30(3):125-127.
- [13] 朝廷祥,张志胜,戴敏.用于目标测距的单目视觉测量方法[J].光学精密工程,2011,19(5):1111-1117.
- [14] ZHANG Zheng-you. A flexible new technique for camera calibration [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(11):1330-1334.

(上接第 3809 页)

- [5] ALIZADEH J, MIRGHADRI A. A new distinguisher for CubeHash-8/b and CubeHash-15/b compression functions[J]. IJCSI International Journal of Computer Science Issues, 2011, 8(5):184-192.
- [6] MORAWIECKI P, PIEPRZYK J, SREBRNY R. Rotational cryptanalysis of round-reduced KECCAK [EB/OL]. (2012-12-18). http://eprint.iacr.org/2012/546.
- [7] WATANABE D, HATANO Y. Higher order differential attack on reduced round Luffa [R/OB]. (2010-11-19). http://eprint.iacr.org/2010/589.
- [8] KHOVRATOVICH D, NAYA-PLASENCIA M, ROCK A, et al. Cryptanalysis of Luffa v2 components [C]//Proc of the 17th International Conference on Selected Areas in Cryptography. Berlin: Springer-Verlag, 2011:388-409.
- [9] JIA Ke-ting, DESMEDT Y, HAN Li-dong, et al. Pseudo-cryptanalysis of Luffa [R/OB]. (2009-05-19). http://eprint.iacr.org/2009/224.
- [10] PRENEEL B, YOSHIDA H, WATANABE D. Finding collisions for reduced Luffa-256 v2 [C]//Proc of the 16th Australasian Conference on Information Security and Privacy. 2011:423-427.
- [11] AUMASSON J P, MEIER W. Zero-sum distinguishers for reduced Keccak-f and for the core functions of Luffa and Hamsi [EB/OL]. (2009-09-09). http://131002.net/data/papers/AM09.pdf.
- [12] KIRCANSKI A, YOUSSEF A M. Boomerang and slide-rotational analysis of the SM3 hash function [R/OB]. (2012-05-15). http://eprint.iacr.org/2012/274.