

基于分割线的漫画帧识别与自动排序

焦利敏 何中市 李 佳

(重庆大学计算机学院 重庆 400030)

摘 要 分析了漫画构成文法,提出一种漫画结构定义法;分析了漫画帧排布特点和分治思想的关联性,提出基于二叉树结构的漫画帧迭代识别排序法;分析了漫画分割线特点,提出一种基于分割比的分割线选择算法。本方法可以对漫画帧进行自动识别、抽取与排序,帮助减少手动分割的工作量。分割后的漫画帧可以放在屏幕较小的手机或 PDA 上阅读,使漫画阅读方便、简单、生动。实验证明该方法对大多数漫画帧的识别分割和故事情节的先后顺序排列具有较好的适应性与实时性。

关键词 漫画帧识别,分割线选择,漫画文法,直线检测,霍夫变换

中图分类号 TP391 **文献标识码** A

Frame Recognition and Sort of Comic Pages Based on Split Line

JIAO Li-min HE Zhong-shi LI Jia

(College of Computer Science, Chongqing University, Chongqing 400030, China)

Abstract After analyzing comic grammar, this paper presented one kind of definition for comic structure. According to the relevance between features of frames layout and Divide-and-Conquer, iterative segmentation algorithm based on binary tree was proposed. After analyzing comic split line, split line choosing algorithm based on split ratio was proposed. The proposed method could decompose one comic page into multiple comic frames and sort them in order automatically, so it can reduce the amount of manual segmentation work. Comic frames after division can be placed on the smaller screen of cell phone screen or PDA, one by one in order to allow simple and vivid reading even through screen with very limited size. Experiments show that for most comics this method can identify each frame and then arranged them in chronological order according to the story and that it has better performance in both accuracy and real time.

Keywords Frame recognition, Split line choosing, Comic grammar, Line detection, Hough transform

1 引言

电子漫画不但具有电子书方便经济、自动排好序的优点,而且具有动感。随着智能手机等电子设备的普及,电子漫画逐渐发展并流行起来,尤其在青少年消费市场中非常受欢迎。对漫画图像的研究也相继出现,主要有以下方面:

漫画帧的识别、抽取与排序的研究;漫画中内容的识别:如人物识别方法^[1];漫画搜索算法的研究^[2];漫画本身的研究,如漫画聊天内容的产生方法^[3]。本文主要的研究内容为漫画帧的识别、抽取与排序。

漫画帧的识别排序问题即是如何把一页漫画图片分割成更小的漫画帧并且按照故事情节的顺序排序。这样在手机或 PDA 等屏幕较小的设备上也能有良好的阅读效果,且随着上下键翻动产生生动漫画效果。

主要的排序方法有基于分割线角度和基于迭代法两种:

1) 基于分割线角度的排序方法。如 Masashi YAMADA 等人提出一种基于分割线之间的角度的规则对漫画帧进行排序的方法^[4],即通过分割线的角度范围确定分割线两侧漫画帧的顺序,根据对话框相对位置对漫画帧中对话框内容进行

排序。其主要着重于排序结果的正确性,用漫画帧之间的白色区域来分割连在一起的漫画帧。基于分割线角度的方法充分结合漫画文法处理,对漫画帧的形状没有限制,排序结果的正确率较高,但不能保证完全正确。

2) 基于迭代的排序方法。如 Takamasa Tanaka 等人提出一种用树结构来表示漫画帧排布过程的方法,即通过迭代的方式对漫画帧进行排序^[5],先把一页漫画分成几个部分,再对每部分细分。水平或竖直分割线的确定仍然基于分割线角度的判断。这种方法对形状为四边形的漫画帧排序基本都是正确的。

已有学者在漫画帧识别方面取得一定成果,主要的漫画帧识别方法包括以下 3 种:

1) 基于空白位置识别法。通过检测漫画帧之间的空白位置把每一帧识别出来^[4,6]。方法简单,对漫画帧的形状无限制。但对于两帧有重叠或两帧之间只有分割线、没有空白位置的情况无效。对前景内容超出漫画帧的情况效果不好。

2) 基于分割线识别法。通过检测漫画帧之间的分割线来识别漫画帧^[5]。对两帧间没有空白位置的情况仍然有效。对漫画帧的形状有限制,一般为四边形,对话框超出漫画帧与漫

焦利敏(1989—),女,硕士生,主要研究方向为图像处理,E-mail:jiaolimin@cqu.edu.cn;何中市(1965—),男,博士后,教授,CCF 会员,主要研究方向为图像处理、自然语言处理;李 佳(1971—),男,博士后,教授,主要研究方向为图像处理、异步电路原理。

画分割线被截断的情况效果不好。

3) 基于连通区域识别法。通过检测漫画图像中的联通区域来识别漫画帧^[7,8], 认为联通的一块区域为一帧。对漫画帧的形状无限制, 附加检测分割线的不连接点可以解决部分对话框超出漫画帧的情况, 正确率较高。但对两帧有重叠的情况无效, 对内容超出漫画帧的情况效果不好。

针对传统规整的漫画帧分割排序问题已经取得了很好的效果, 但怎样有效地对不规则漫画进行分割排序仍然需要进一步研究。现有方法没有对对话框超出漫画帧情况下的漫画帧分割取得理想的研究成果, 不能保证分割后不破坏漫画帧内容, 且大多数方法不可以进行实时处理。

本文分析了漫画文法并提出一种漫画结构定义法, 结合人脑识别机制和漫画帧排布特点, 提出二叉树迭代排序法对漫画帧排序。根据漫画帧之间分割线所具有的特点提出一种按分割比对漫画帧分割线进行识别的方法。本文竖直与水平分割线的判断没有基于角度, 而是根据分割线与边缘的交点位置确定, 方法简单, 正确率高。本方案对分割线被对话框截断的情况下亦能执行, 且具有实时性。

2 基于二叉树的漫画文法分析

2.1 漫画结构分析

漫画书由若干页组成, 每页由若干漫画帧组成。每帧都是一个由情景画面、气泡状的对话框和边缘线组成的四边形。连在一起的几帧漫画组成的四边形本文定义为漫画段。如图 1 为一页漫画, 小方格为漫画帧, 下面框起来的部分和剩余的部分分别为两个漫画段。漫画页或漫画帧亦属于一个漫画段。漫画段之间的分割线就是相邻的边缘线。



图 1 漫画页

漫画的阅读顺序与平时阅读习惯一致。如在中国习惯从上到下、从左到右阅读, 中国漫画也是按照这个顺序阅读。又例如在日本习惯按照从上到下、从右到左的顺序阅读。本文采用日本漫画的阅读规则进行举例, 本方法同样可以用于中国漫画, 只需要改变漫画段在二叉树中存储的左右位置即可得到正确结果。

2.2 人脑识别规律

在了解漫画帧排序规则的基础上, 人看漫画时很容易判断漫画帧的排序, 这是人的视觉系统与大脑识别机制相结合的结果。观察事物时, 人的视觉首先注意到的是最鲜艳、最明显、位置趋于中心的部分。同样人在看一幅漫画时, 首先注意

到的分割线是最明显最居中的分割线, 然后通过这条分割线把一幅漫画分为两部分, 接着再对每一部分分别按照同样的规则区分出来, 直到每一帧都被识别出来为止。如图 2(a) 所示, 中间分割线①首先被看到。在先上后下规则的基础上, 分割线上面一部分受到关注。同样首先注意到上面部分最长的分割线②, 上部分的漫画段分为左右两部分。依次类推把整页漫画分割成一帧一帧的方格。

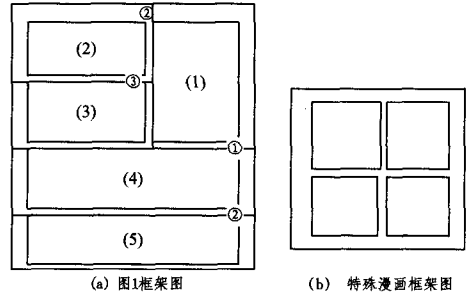


图 2 漫画框架图

2.3 二叉树与漫画帧关系提取

用二叉树表示漫画分割过程这种分而治之的思想符合漫画文法规则, 并且与人判断漫画帧排序过程具有相同的原理。一棵二叉树对应一页漫画的分割过程。根节点表示一整页漫画, 漫画首先被分为两部分, 按照阅读顺序分别放在根节点的左右子树中。对每棵子树继续分割, 直到全部分割完成。分割完成的漫画帧保存在二叉树的叶子节点。图 2(a) 是图 1 的框架图, 对应的二叉树结构图如图 3。图 2(b) 漫画中同时有一条以上明显分割线的情况, 本文不作考虑。

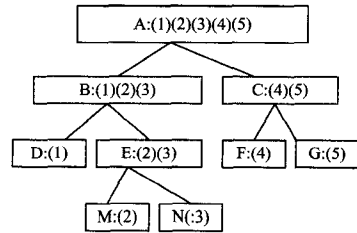


图 3 漫画帧的二叉树结构图

采用二叉树数据结构表示漫画分割过程, 可对漫画帧直接排序。所有叶子为最终分割后的漫画帧, 且已经从左到右排好顺序。其他节点表示分割过程产生的漫画段。按照日本漫画排布规则, 图 3 中二叉树的所有叶子结点 D、M、N、F、G 为最终的分割结果, 与图 2 漫画帧的阅读顺序相对应。

3 迭代分割排序与分割线识别

3.1 漫画结构的定义

通过研究漫画文法, 为了更好地分析漫画结构, 本文提出了一种漫画定义法。

每个漫画页、漫画段、漫画帧都是四边形。定义这样一个四边形为 Coma。

定义 1 (Coma) 一个四边形的结构体, 其有 4 个元素, 分别为四边形的 4 个顶点, 规定按顺时针排序, 分别为左上、右上、右下、左下。表示为:

$$Coma\{LT, RT, RB, LB\}$$

定义 2 Split Line, SPL: 分割线。把一个 Coma 分成两

部分的直线。表示为： $SPL\{FP, LP, SS, SR\}$ 。

定义 3 First Point, FP:分割线与 Coma 相对两条边的起始交点; last Point(LP):分割线与 Coma 相对两条边的结束交点。

定义 4 Split Status, SS:分割线类别。有 3 个值:V 表示竖立式的分割线, H 表示横卧式的分割线, N 表示无效分割线。

V 分割线是把 Coma 分成左右两部分的分割线,即与线段 LT 到 RT、RB 到 LB 都有交点,如图 4 所示。H 分割线是把 Coma 分成上下两部分的分割线,即与线段 LB 到 LT、RT 到 RB 有交点,如图 5 所示。分割线需符合以下情况,否则定义为无效分割线 N:

(1)分割线只与 Coma 中相对的两条边相交,与其他边没有交点。

(2)保证分割前后每个漫画段、漫画帧为四边形。

(3)Coma 边缘线与分割线的交点把边缘线分成的两段之比的值范围在 $[0.17, 6]$ 。

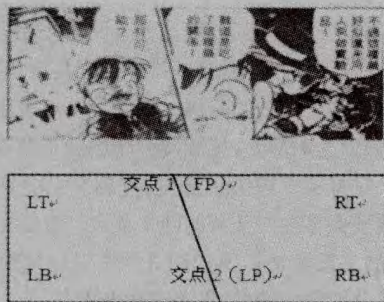


图 4 V 分割线图(下)与示例图(上)

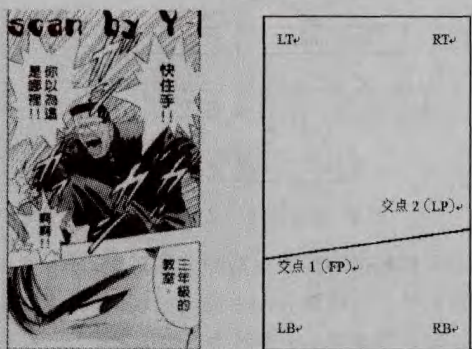


图 5 H 分割线图(右)与示例(左)

定义 5 Split Ratio, SR:分割比。分割线的实际长度与分割线的理论长度之比。

分割线的理论距离为 FP 到 LP 的距离。实际漫画中分割线常常在中间被对话框或者情景画面截断,实际长度在这种情况小于 FP 到 LP 的距离。所以 $0 < SR \leq 1$ 。

3.2 迭代分割排序法

本文首先对漫画图片做预处理工作,包括二值化^[9]、轮廓检测、细化,然后对图片进行直线检测。直线检测方法为 L. A. F. Fernandes 和 M. M. Oliveira 提出的通过投票方案进行实时直线检测的方法^[10]。该方法是在直线段检测方法霍夫变换^[11,12]上进行改进的,执行效率高,对漫画图片适应性好且检测出的直线已按权重排好序。直线检测后用二叉树构造算法对漫画进行分割排序,分割过程中需要按权重由高到低

的顺序选出符合定义的最佳分割线,把漫画段分为两部分,对每一部分按照同样的方法迭代分割并排序,直到不可再分割为止。

3.2.1 二叉树构造

构造一棵以 Coma 为节点的的二叉树来记录漫画分割过程并存放分割结果。二叉树构造算法描述如下:

名称:二叉树构造算法 ContributeTree(Coma)

输入:Coma 结构体

输出:一棵二叉树

首次输入一页漫画时,定义 Coma{LT, RT, RB, LB} 作为二叉树的根节点。漫画页 4 个顶点作为 Coma 的元素{LT, RT, RB, LB}

//调用 3.2.2 节 ChooseSPL(Coma)分割线选取算法,选取当前 Coma 最佳分割线。

SPL ← ChooseSPL(Coma)

//如分割线选取算法返回值 SPL 为无效分割线,停止。

If SPL. SS == N

Stop

End if

If SPL. SS == V

//分割线右边的部分作为一个新的 Coma 结构赋值给当前 Coma 的左子树;左边部分赋值给右子树

Coma. LeftChild ← Coma{ FP, RT, RB, LP};

Coma. RightChild ← Coma{ LT, FP, LP, RB};

//构造左子树和右子树

ContributeTree(Coma. LeftChild);

ContributeTree(Coma. RightChild);

End if

If SPL. SS == H

//分割线上边的部分作为一个新的 Coma 结构赋值给当前 Coma 的左子树;下边部分赋值给右子树

Coma. LeftChild ← Coma{ RT, RB, LP, FP};

Coma. RightChild ← Coma{ FP, LP, RB, LB};

//构造左子树和右子树

ContributeTree(Coma. LeftChild);

ContributeTree(Coma. RightChild);

End if

3.2.2 分割线识别

二叉树构造中分割线的选取调用分割线选择取法 ChooseSPL(Coma)。分割线选择的正确性与高效性直接决定了漫画帧识别与自动排序方法的好坏。

为了不损坏漫画内容,需找出一条当前 Coma 的最佳分割线并且在分割线附近找出一个合适的分割位置。如图 6 所示,最佳分割线为中间位置的两条边缘线之一,为了尽量确保漫画内容的完整性,应该在最佳分割线旁边的空白地方,即边缘线的中间位置将其分割开。



图 6 框架图(左)与示例图(右)

下面详细阐述分割线选择过程。表 1 是分割线选取过程用到的所有阈值和参数。

表 1 参数表

名称	值	意义
FRR	0.05	前景区域与整页漫画面积之比最小值
FCR	0.25	前景图像像素与其面积上所有像素数量之比最小值
MSR	0.08	前景区域与整幅漫画分别的长度、高度之比最小值
SPW	8	分割位置离分割线的最大宽度的像素值
R0	0.55	分割线与其他直线无交点时分割比最小值
R1	0.6	分割线与其他直线有一个交点时分割比最小值
R2	0.75	分割线与其他直线有两个交点时分割比最小值

分割线选取算法 ChooceSPL(Coma)步骤描述如下,直线已按权重从高到低排列。

名称:分割线选取算法 ChooceSPL(Coma)

输入:Coma,漫画页中按权重降序排序的未使用的直线

输出:SPL(最佳分割线)

Step1:

定义一条无效分割线:SPL. SS←N

如果当前 Coma 满足以下条件之一,则认为当前 Coma 不可分,返回无效分割线:

- (1)当前 Coma 中最大连通前景区域面积与整页漫画面积的比值小于 FRR;
- (2)当前 Coma 中前景区域的长度、高度与整幅漫画的长度、高度的比小于 MSR;
- (3)当前 Coma 中前景图像实际像素数量与前景区域面积占有像素数量之比小于 FCR。

Step2:

选取 Coma 内部直线:如果所有未使用的直线落入 Coma 内部的像素数量均小于直线像素总数量,则认为 Coma 内部无分割线,返回无效分割线。

Step3:

对 Coma 内部这些备选分割线按其权重的降序进行筛选,依次做如下操作:

检查当前直线是否符合分割线 SPL 定义,并确定分割线状态。如果分割线 SPL 不符合定义或者 SPL. SS=N 则选取下一条 Coma 内部直线进行判断,否则认为本条直线为分割线,执行以下步骤:

在分割线邻近的 2SPW 个像素点宽度范围,与分割线平行的一组直线内,按离分割线由近到远的顺序选出符合以下条件之一的直线(如图 7 所示)作为分割位置,并确定分割线为最佳分割线。如果没有符合条件的分割位置,选取下一条 Coma 内部直线判断。

- (1)与其他直线无交点,且分割线的分割比大于 R0;
- (2)与其他直线有一个交点,且分割线的分割比大于 R1;
- (3)与其他直线有两个交点,且分割线的分割比大于 R2。

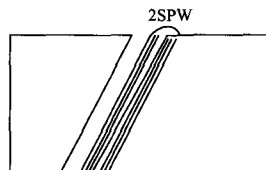


图 7 分割位置确定过程

4 实验与分析

实验通过测试漫画帧识别的成功率和效率来证明方法的有效性。实验环境为 Visual Studio 2008,语言为 C++。实验数据来自一些网站,为了贴近实际应用以及保留数据的完整性,图片的像素没有完全统一,且只去掉了一些空白页图片,并没有去掉漫画图片中标题页等情况,通过判断是否可分

本方法对标题页可以进行自动处理。

图 8 直观显示了漫画帧识别与排序结果。漫画帧识别结果用线条框起来的部分,序号为排序结果。



图 8 分割排序结果

首先选取漫画龙珠的第 42 卷和海贼王第 1 卷进行实验,将结果与 Kohei Arai 所提方法的实验结果^[10]进行对比,见表 2。本文方法中海贼王的实验数据较对比方法的实验数据多 13 张图片的原因是本文方法没有去掉标题页。可见本文方法在两组实验数据中成功率均高于对比方法,尤其是龙珠 42 卷高于对比方法 5.07%。

表 2 Kohei Arai 方法与本文方法的对比结果

数据集	Kohei Arai			本文		
	总	成功	成功率	总	成功	成功率
龙珠.42 卷	237	218	91.98%	237	230	97.05%
海贼王.1 卷	191	171	89.53%	204	191	93.63%

另外还选取网球王子第 16 话到 22 话分割排序,实验结果见表 3,其中列出了龙珠和海贼王实验结果的详细数据。

表 3 实验结果

数据集	总数	成功	失败	过分	少分	成功率
网球王子	106	93	13	5	7	87.74%
龙珠	237	230	7	1	5	97.05%
海贼王	204	191	13	7	6	93.63%
总张数	547	514	33	13	18	93.97%

实验统计平均效率为 0.215 张/s。失败情况包括 3 种:过分,一帧被分为更小的单位,这是由于情景画面中存在干扰直线而造成的,如图 9 中(a)第③、④应为一帧;另一种情况:少分,连在一起的几帧没有分开,这是由于帧太小或分割线被覆盖过多造成的,如图 9 中(b)第③帧应被分为两帧。第 3 种为错分。



图 9 分割结果展示

ACM SIGGRAPH/EUROGRAPHICS conference on Graphics hardware(HWWS'04). ACM, New York, NY, 2004; 85-94

- [21] Mochocki B, Lahiri K, Cadambi S, et al. Signature-based workload estimation for mobile 3d graphics[C]//Proceedings of the 43rd Annual Conference on Design Automation (DAC'06). ACM, New York, NY, 2006; 592-597
- [22] Lee N B-G, et al. A low-power handheld GPU using logarithmic arithmetic and tripleDVFS power domains[C]//Proceedings of the 22nd ACM SIGGRAPH/EUROGRAPHICS Symposium on Graphics hardware(GH'07). 2007; 73-80
- [23] Igehy H, Eldridge M, Proudfoot K. Prefetching in a texture cache architecture[C]//HWWS'98; Proceedings of the ACM SIGGRAPH/ EUROGRAPHICS Workshop on Graphics Hardware. ACM, New York, NY, 1998; 133-142
- [24] Powell M, Yang S-H, Falsafi B, et al. Gated-vdd: A circuit technique to reduce leakage in deep-sub micron cache memories[C]//Proceedings of the International Symposium on Low Power Electronics and Design (ISLPED'00). ACM, New York, NY, 2000; 90-95

- [25] Kaxiras S, Hu Z, Martonosi M. Cache decay: exploiting generational behavior to reduce cache leakage power[C]//Proceedings of the 28th Annual International Symposium on Computer Architecture(ISCA'01). ACM, New York, NY, 2001; 240-251
- [26] Flautner K, Kim N S, Martin S, et al. Drowsy caches: Simple techniques for reducing leakage power[C]//Proceedings of the 29th Annual International Symposium on Computer Architecture(ISCA'02). IEEE Computer Society, 2002; 148-157
- [27] Youssef A, Anis M, Elmasry M. Dynamic standby prediction for leakage tolerant microprocessor functional units[C]//Proceedings of the 39th Annual IEEE/ACM International Symposium on Microarchitecture. IEEE Computer Society, 2006; 371-384
- [28] 杨毅, 郭立, 史鸿声, 等. 面向移动设备的 3D 图形处理器设计[J]. 小型微型计算机系统, 2009, 30(8): 1668-1672
- [29] 韩俊刚, 蒋林, 杜慧敏, 等. 一种图形加速器和着色器的体系结构[J]. 计算机辅助设计与图形学学报, 2010, 23(3): 363-372
- [30] 韩俊刚, 刘有耀, 张晓. 图形处理器的历史现状与发展趋势[J]. 西安邮电学院学报, 2011, 16(3): 61-64

(上接第 195 页)

由结果可见, 实验结果受到几个因素的影响。如漫画风格、图片质量。同时, 过分和少分相互影响。如果把 3 个分割比值调大些, 对分割线的要求将更严格, 出现过分的情况会随之减少, 但少分情况由于分割线要求的升高会增加, 反之亦然。所以根据大多数漫画文法综合考虑, 选取过分和少分总和较少的值。

结束语 本文分析漫画后提出一种基于二叉树的漫画帧分割排序法和基于分割比的分割线选取方法对漫画帧进行分割排序。其虽尚存在不足之处, 但总体分割成功率较高, 有以下几点:

(1) 本方法对漫画帧的排序没有基于分割线角度来判断, 而是基于分割线与边缘线交点的特点, 对漫画帧形状唯一的限制是四边形, 与漫画文法相契合。

(2) 通过设定分割比, 对分割线不完整的情况放松条件, 使情景画面和对话框覆盖分割线的情况得到解决。

(3) 直线检测是基于细化后的轮廓线上进行的, 提高了原本非常耗费的直线检测的效率。

(4) 对分割线与前景图像相交情况做了判断, 从而出现了很少的错分与过分图片。

本文只对一般漫画进行分割, 对于嵌套或重叠漫画帧内容的无损分割是本文以后的研究内容。

参 考 文 献

- [1] Yamada T, Watanabe T. Extraction of Person Objects from Japanese. Four-Scenes Comics [J]. IEIC Technical Report, 1999, 99(47): 152-159
- [2] Hoashi K, Ono C, Ishii D, et al. Automatic Preview Generation of Comic Episodes for Digitized Comic Search [C]//Proceedings

of the 19th ACM international conference on Multimedia. 2011; 1489-1492

- [3] Kurlander D, Skelly T, Salesin D. Comic Chat [C]// Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. 1996; 225-236
- [4] Yamada M, Budiarto R, Miyazaki S. Comic image decomposition for reading comics on cellular phones [J]. IEICE transactions on information and systems, 2004, 87(6): 1370-1376
- [5] Tanaka T T, Shoji K, Toyama F, et al. Layout analysis of tree-structured scene frames in comic images [C]// Proceedings of IJCAL 2007; 2885-2890
- [6] Chan C, Leung H, Komura T. Automatic panel extraction of color comic images [J]. Advances in Multimedia Information Processing-PCM 2007, 2007, 4810; 775-784
- [7] Arai K, Herman T. Method for automatic e-comic scene frame extraction for reading comic on mobile devices[C]//Information Technology; New Generations (ITNG), 2010 Seventh International Conference on. IEEE, 2010; 370-375
- [8] In Y, Oie T, Higuchi M, et al. Fast frame decomposition and sorting by contour tracing for mobile phone comic images [J]. International journal of systems applications, engineering and development. 2011, 5(2): 216-223
- [9] Nobuyuki O. A threshold selection method from gray-level histograms [J]. IEEE transactions on systems, 1979, 9(1): 62-66
- [10] Fernandes L A F, Oliveira M M. Real-time line detection through an improved Hough transform voting scheme[J]. Pattern Recognition, 2008, 41(1): 299-314
- [11] 郭斯, 霍文娟, 唐求, 等. 结合 Hough 变化与改进最小二乘法的直线检测[J]. 计算机科学, 2012, 39(4): 196-200
- [12] Illingworth J, Kittler J. A survey of Hough transform [J]. Computer Vision Graphics Image Processing, 1988, 44(1): 87-116