

# 自适应柱状全景图拼接

殷润民<sup>1)</sup> 李伯虎<sup>1)</sup> 柴旭东<sup>2)</sup>

<sup>1)</sup>(北京航空航天大学自动化科学与电气工程学院,北京 100083) <sup>2)</sup>(北京仿真中心,北京 100854)

**摘要** 为了实现高效高精度的全景图拼接,提出了一种面向柱状全景图的自适应拼接方法。该方法首先使用基于灰度边缘特征点的图像匹配来获得匹配坐标,并由匹配坐标唯一性计算焦距;然后在该焦距参数估计的基础上,采用新的自适应混合图像灰度匹配方法生成初期拼接图像,并根据其宽度以及相关判别条件进行焦距估计的调整,再重复迭代,以获得准确的焦距参数;最后利用基于角度1维投影的图像匹配方法进行匹配坐标的小范围校正,以及使用线性平滑算法生成最终柱状全景图。该法在精确焦距的基础上,结合了多种图像匹配算法,并根据自相关指数自适应选取结果,实现了水平和垂直两方向上的匹配定位和错位检测。实验证明,该方法不仅在真实图像上的可行,且具有较高的实时性、精度和鲁棒性。

**关键词** 柱状全景图 图像拼接 混合图像灰度匹配 模式识别 基于图像绘制

中图法分类号:TP391.41 文献标识码:A 文章编号:1006-8961(2008)06-1191-06

## Adaptive Cylindrical Panoramas Mosaic

YIN Run-min<sup>1)</sup>, LI Bo-hu<sup>1)</sup>, CHAI Xu-dong<sup>2)</sup>

<sup>1)</sup>(School of Automation Science and Electric Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083)

<sup>2)</sup>(Beijing Simulation Center, Beijing 100854)

**Abstract** In this paper, the problem of efficiently and precisely constructing the cylindrical panoramas view of a scene from an image sequence is considered. The proposed method firstly detects correspondences based on gray level edge feature so that the focal lens can be counted based on correspondences uniqueness. Then the focal lens will be directly used in following image matching processing with a novel adaptive mixed gray level image matching algorithm and outputs a rough panorama result. The width of it reacts to a judgment to adjust the focal lens, which induces an iterative computation to guarantee an exact focal lens. Finally an angle-oriented one-dimension projection template matching algorithm will be employed to modify the correspondences in narrow range and the final panorama is obtained by linear image blending. Adaptive using multiple image matching algorithms according to self-relation variant after precise focal lens searching, the method redraws the accurate overlapping image horizontally and vertically. Experimental results proved that the proposed approach are given works well over most image sequence capture by handholding camera, and is fast, accurate and robust enough for real time applications.

**Keywords** cylindrical panoramas, image mosaic, mixed gray level image matching, pattern recognition, image-based rendering

## 1 引言

柱状全景图拼接技术涉及到的问题包括图像全局/局部对齐、对齐图像结果的本地像素调整、对融

合像素区域的选择、图像融合技术、融合图像的光照补偿技术等等。由于对上述问题的理解和解决的方法不同,因而出现了许多不同的全景图拼接方法。其中 QuickTimeVR 系统<sup>[1]</sup>在图像拼接时,必须知道相邻帧间摄像机的旋转角。文献[2]方法是依据摄

基金项目:国家十五预研专题项目(413090303)

收稿日期:2006-09-01;改回日期:2006-12-26

第一作者简介:殷润民(1979~),男,北京航空航天大学自动化科学与电气工程学院博士研究生。主要研究方向包括:虚拟现实、基于图像可视化、协同可视化。E-mail:yinrunmin@163.com

像机绕轴  $360^\circ$  所拍摄的图像序列来求解摄像机参数,进而完成全景图拼接<sup>[2]</sup>,但其要求相邻帧间有  $2/3$  以上的重叠。文献[3]方法在保持摄像机光心基本不发生变化的前提下,先通过 8 个未知参数的矩阵来描述两幅图像之间的变换关系<sup>[3]</sup>,然后利用最小化算法估计的矩阵将各图像投影到某一参考图像的平面坐标系,最终利用拼接区域加权平均的方法进行全景图拼接。该法虽简单易用,且快速,但该算法有以下不足:(1)必须由用户给定参数初始值,且容易导致误匹配;(2)采用加权平均的方法不能保证无缝;(3)拼接区域宽度对拼接的效果影响明显。文献[4]提出了只对源图进行移动和旋转两种简单操作的复制映射方法(manifold projection),其虽然避免了由于使用其他源图对齐手段而减少了拼接结果的失真,但是它需基于源图,而且没有较大焦距变化的限制假设。

为了减少先验条件,同时降低源图像拍摄条件对拼接结果的影响。本文提出了一种自适应柱状全景图拼接方法。其具体步骤如下:(1)使用拍摄设备进行场景采样,将连续的  $360^\circ$  环视场景先分割为有一定重合区域的离散图像;(2)借助已有的离散图像对拍摄的焦距参数进行估计;(3)在焦距估计的基础上,通过对离散图像进行圆柱面投影来生成柱状离散图像;(4)通过对相邻柱状离散图像进行图像匹配来确定相邻两幅图像对应点的坐标;(5)按照匹配点坐标进行图像拼接,并对重合区域进行无缝混合。

采用不同的图像匹配方法会直接影响图像的拼接效果。传统的通过平方差和(sum of square differences, SSD)或者归一化交叉相关性(normalized cross-correlation, NCC)模板的匹配算法,会因为视点和亮度变化而性能下降,甚至无法正常工作<sup>[5]</sup>。基于列灰度比值匹配的全景图拼接方法<sup>[6]</sup>和基于水平灰度 1 维投影匹配的全景图拼接方法<sup>[7]</sup>是直接利用图像垂直和水平方向上的灰度信息来进行拼接,虽然其执行速度较快,但是由于它利用的图像特征太少,且在匹配源的选取上,对源图的要求较高,又不支持垂直或水平方向变化均匀的图像,因此在使用上存在较大缺陷。基于图像特征匹配的全景图拼接方法<sup>[8]</sup>,由于是建立在特征检测算法适合、特征的数目和位置都相同的假设上,而且会耗费大量计算资源,因此其应用性有所局限。

鉴于这些方法存在的实现速度和计算量、准确

性和适应性之间的矛盾,本文提出的自适应柱状全景图拼接方法,在图像匹配上采用了多种图像匹配算法,其不仅可以保证较高的运行速度,同时在精度和鲁棒性上也能得到相应的满足。

## 2 柱状投影与焦距计算

柱状投影与焦距计算采用小孔成像原理。假设拍摄的焦距为  $f$ ,不失一般性,可以假设以小孔中心为原点的空间坐标系里的  $Z$  轴通过图像中心(图 1)。

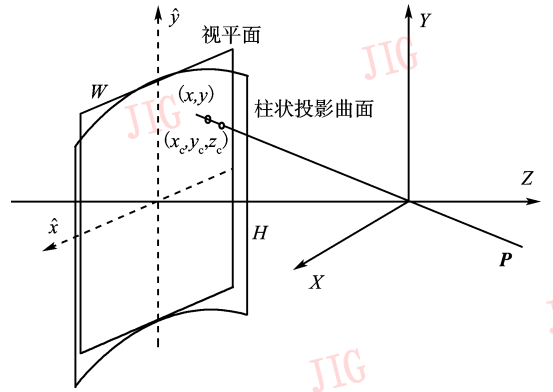


图 1 柱状投影原理图

Fig. 1 Cylindrical projection diagram

根据图像的宽度  $W$  和高度  $H$ ,可先将视平面中任一点的坐标  $(x, y)$  转换为在圆柱面上的投影坐标  $(x_c, y_c, z_c)$ ,然后可将其投影转换到 2 维矩形平面上的坐标  $(\hat{x}, \hat{y})$ 。其中,2 维矩形平面以左下角为坐标原点,  $x = 0$  处为起始展开点;圆柱曲面按顺时针方向旋转:

$$\begin{cases} x_c = \frac{f}{\sqrt{\left(x - \frac{W}{2}\right)^2 + f^2}} \left(x - \frac{W}{2}\right) \\ y_c = \frac{f}{\sqrt{\left(x - \frac{W}{2}\right)^2 + f^2}} \left(y - \frac{H}{2}\right) \\ z_c = \frac{f^2}{\sqrt{\left(x - \frac{W}{2}\right)^2 + f^2}} \end{cases}$$

$$\begin{cases} \hat{x} = f \arctan\left(\frac{x - \frac{W}{2}}{f}\right) + f \arctan\left(\frac{W}{2f}\right) \\ \hat{y} = \frac{f}{\sqrt{\left(x - \frac{W}{2}\right)^2 + f^2}} \left(y - \frac{H}{2}\right) + \frac{H}{2} \end{cases}$$

根据圆柱曲面投影推导,可以断定交叠图像在满足拍摄水平线重合的情况下,匹配点距离投影后曲面的水平中心线相等,即

$$\frac{f}{\sqrt{\left(x_A - \frac{W}{2}\right)^2 + f^2}} \left(y_A - \frac{H}{2}\right) = \frac{f}{\sqrt{\left(x_B - \frac{W}{2}\right)^2 + f^2}} \left(y_B - \frac{H}{2}\right)$$

其中,  $(x_A, y_A)$  和  $(x_B, y_B)$  分别是相邻图像的匹配点  $A, B$  的坐标,最后可以获得

$$f = \frac{\sqrt{\left(x_A - \frac{W}{2}\right)^2 \left(y_B - \frac{H}{2}\right)^2 - \left(x_B - \frac{W}{2}\right)^2 \left(y_A - \frac{H}{2}\right)^2}}{\left(y_A - \frac{H}{2}\right)^2 - \left(y_B - \frac{H}{2}\right)^2}$$

需要注意的是,应该将匹配点放置在与水平中心线不同距离的位置。

获得匹配点坐标的算法主要包括以下两部分:(1)通过灰度差统计信息进行边缘分割<sup>[9]</sup>,并选择阈值范围内分割区域数量最多的作为模板;(2)进行边缘模板的匹配(下一节详述)。由于采用的是以像素点为最小单位,其在计算焦距时会有量化误差,因此需要参考多个匹配点的参数进行计算,才可获得其统计平均。为了提高估计精度,匹配点位置还需尽量离水平中心线较远。尽管如此,焦距计算仍然需要满足判别条件:  $2\pi f = W_s$ 。其中  $W_s$  指的是序列离散图像经柱状投影后拼接所得结果的水平宽度。因此其迭代过程如下:

- (1) 寻找多位置匹配点;
- (2) 计算焦距  $f$ ;
- (3) 利用混合图像匹配进行图像拼接结果的图像宽度累加;
- (4) 若  $2\pi f - W_s$  大于允许误差  $\varepsilon$ , 则令  $f = \left(f + \frac{W_s}{2\pi}\right)/2$ , 重复步骤(3)。否则确定  $f$ 。

### 3 自适应混合图像匹配方法

传统的基于灰度模板的 SSD 算法是直接利用方形  $n \times n$  大小的模板内的累计灰度差作为匹配程度的判别指标,虽然该法简单,但却无法实现精

确匹配,而且焦距  $f$  的不精确会导致柱状投影产生一定图像变形,从而极大影响模板灰度信息的匹配程度。为了降低这种影响,本文提出了使用平均灰度列作为匹配元素的算法(简称为列匹配)。由于按照柱状投影算法,如果投影点离图像中心越远,则变形程度越大,因此灰度列中每点灰度是一定范围内灰度的平均值,灰度计算范围可按离水平中心线距离的远近进行相应扩散(图 2)。除此之外,为了降低图像整体的亮度差异,列内各点的灰度需要进行归一化处理,即令取值范围边长为  $\lambda_h d_h + \lambda_v d_v$ , 其中  $d_h$  和  $d_v$  分别代表参与计算的像素点至水平中心线和垂直中心线的距离,  $\lambda_h$  和  $\lambda_v$  分别对应其相关系数,通常可以取  $\lambda_h = \frac{2}{5W}$ ,  $\lambda_v = \frac{2}{3H}$ 。

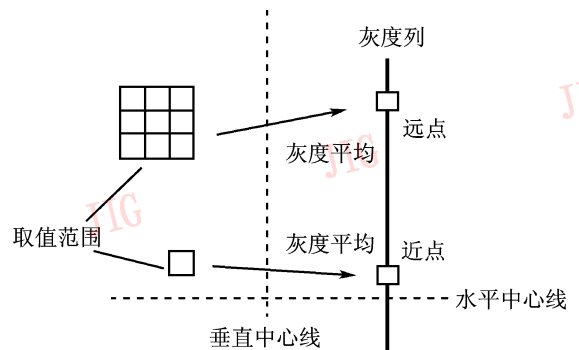


图 2 列匹配原理图

Fig. 2 Column pixels match diagram

由于列匹配算法使用的是列模板,该法对拼接图像垂直方向上匹配的检测能力比水平方向强,而且相邻图像的重合区域未知会导致无法将该法运用在灰度行上,为此本文引入了基于边缘统计直方图的匹配算法,以增强对水平方向的考虑,该法简称为行匹配。要获得准确唯一的边缘信息,通常需要花费很高的计算量,而使用灰度差统计得到的一级准边缘<sup>[9]</sup>实际上已经能够得到足够的边缘信息供匹配使用,且运算简单快速(图 3)。行匹配是对一定列宽的各列边缘点数进行直方图统计,即先选择变化较明显的一段作为匹配源;然后与相邻图像做同样操作后获得的直方图统计进行比对;最后选择最小差异者作为匹配结果。由于相邻图像的重合区域有限,因此不能对所有的边缘统计进行统一的归一化处理,而只是对其中选为模板和匹配对象的行段进行归一化处理。

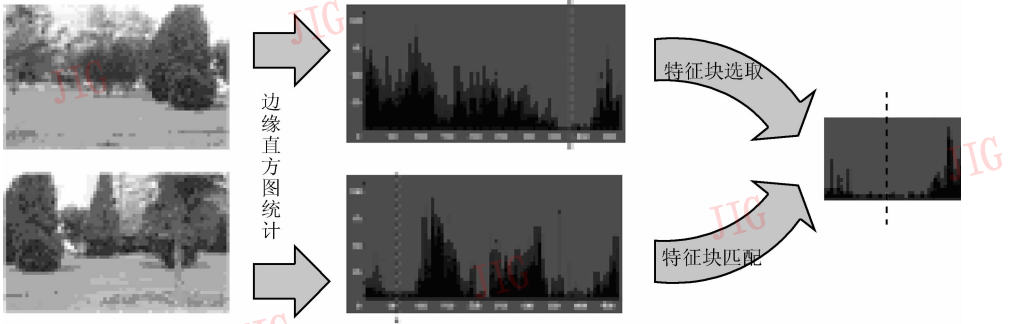


图 3 行匹配原理图

Fig. 3 Row pixels match diagram

由于列匹配和行匹配使用的匹配信息是多像素信息的集合,因此不能保证得到准确的匹配结果。本文在进行焦距计算时,采用的是基于边缘分割 1 维角度投影的模板匹配算法(简称精匹配),可以将其作为进一步的精确匹配方法。精匹配首先需要进行之前经灰度差统计过滤后图像的分割区域检测,并寻找区域数量最多的检测块中心来充当模板中心。拥有了模板中心点坐标  $(x_c, y_c)$  后,再根据模板半径,就可以按角度分割来对模板和目的图像分别进行灰度 1 维投影映射计算(图 4)。当其坐标  $(x_c + r\cos\theta, y_c + r\sin\theta)$  不为整数时,则使用线性插值:

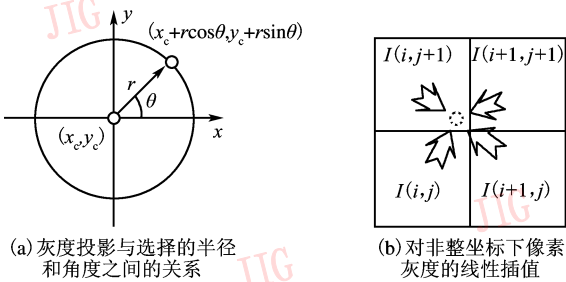


图 4 基于灰度 1 维投影的精匹配原理图

Fig. 4 Precise image match diagram based on 1-dimension gray level projection

接着需要在目的图像中不同计算区域的中心位置(边缘点)上进行相同的 1 维灰度映射运算,以获得对应的  $f(r, \hat{\theta})$ , 判别是否匹配的工作可以描述为变形 AD:

$$D(r, \hat{\theta}) = \sum_{0^\circ \leq \theta, \hat{\theta} < 360^\circ} |f(r, \hat{\theta}) - \omega(r, \theta)|$$

$$0^\circ \leq \theta, \hat{\theta} < 360^\circ, \hat{\theta} = \theta + \Delta\theta$$

其中,  $r$  是模板半径,  $\hat{\theta}$  是判定是否匹配的起始角度,  $\theta$  是参考图像的起始角。  $\Delta\theta$  是  $\hat{\theta}$  与  $\theta$  间的角度差。考虑到图像偏转角度较小,可直接取  $\Delta\theta = 0$ 。由于精匹配需要较大的计算量,因此需要降低其执行范围,可令其在列匹配结果  $(\hat{x}_v, \hat{y}_v)$  和行匹配结果  $(\hat{x}_h, \hat{y}_h)$  重合的较小区域范围内工作。若源图像在垂直或水平方向上存在相似内容时,则列匹配和行匹配可能会出现一定的误差。当  $(\hat{x}_v, \hat{y}_v)$  与  $(\hat{x}_h, \hat{y}_h)$  之间的误差大于阈值  $\hat{\epsilon}$ , 则需要进行自适应匹配结果选择。由于列匹配和行匹配最终都是通过计算最小 1 维模板的差异来判断匹配程度,因此可以通过模板源在本图像内的匹配结果来计算自相关指数  $G = \frac{\min(D_{h,v})}{N}$ , 以用于匹配可能性判断。其中  $D_{h,v}$  是列匹配或行匹配中计算差异的总和,  $N$  是匹配模板的像素点数目,自相关指数  $G$  是对模板在本图像中匹配的最小差异和平均到单个像素上的数值。当行自相关指数  $G_h$  小于列自相关指数  $G_v$ , 则可以认为行匹配出现误匹配的可能性较大,故可选择列匹配结果作为精匹配范围的选择依据。反之则选择行匹配结果。

### 4 实验结果

本文对使用真实的手持照相机拍摄的图像进行了实验。在第 4、5 张图像的匹配过程中,由于出现了相似的树林和黄土路,致使列匹配和行匹配的结果差异较大(图 5)。本文通过计算自相关指数最终选择了使用行匹配的结果,从匹配的结果(见图 5(d))来看,该自适应拼接方法具有较高的精度和鲁棒性。随后给出了全景图的匹配结果(见图 6),该

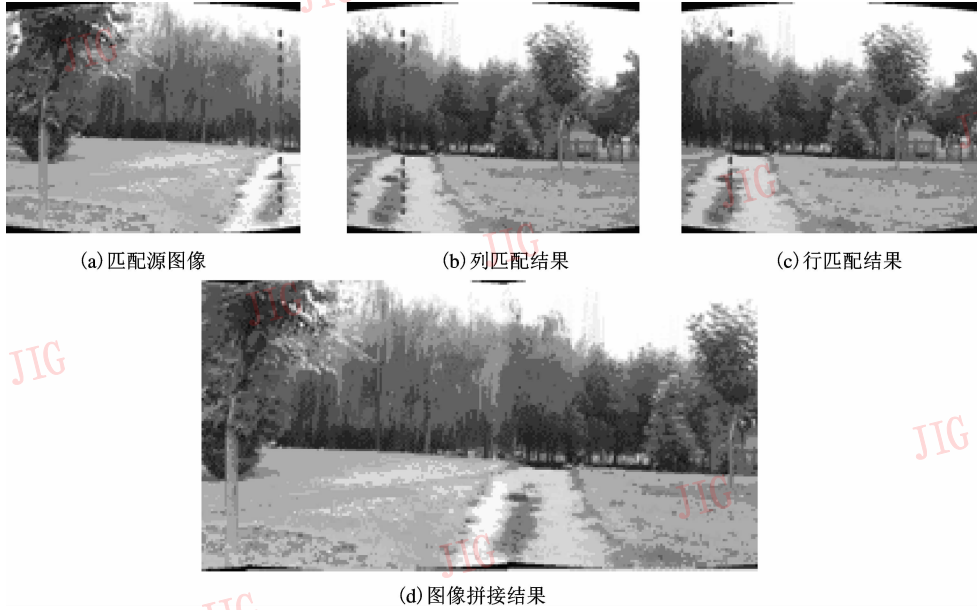


图 5 自适应拼接结果(匹配结果通过虚线表示)

Fig. 5 Adaptive mosaic result (correspondences highlight with dash line)



图 6 柱状全景图拼接结果

Fig. 6 Cylindrical panoramas mosaic result

匹配结果可以体现其可行性。

为了提高运行速度,并且考虑到焦距计算中所允许的误差,在进行焦距迭代时,可以无需进行精匹配,同时可令行匹配和列匹配的范围分别为半边图像(在图像拍摄时可以很简单地满足该条件)。具体在计算焦距时可以分别只对图像离中心较远的右上部分和右下部分进行两次模板源选取,而在匹配时则只在左上部分和左下部分进行模板匹配。该方法运算快的原因是:(1)由于在整个方法里的列匹配和行匹配使用的是 1 维加减运算,计算量很小,且无需复杂的前期工作;(2)1 维准边缘检测算法也只是进行简单的灰度差计算;(3)需要进行较复杂的线性插值工作的精匹配只是在以已有结果坐标为中心的  $5 \times 5$  小范围内进行。实验证明,该方法具有一定的实时性。表 1 给出了在 Pentium 1.7G、512MB 内存的计算机里几种柱状全景图拼接使用的时间比较:

表 1 使用时间比较

Tab. 1 Comparison of using time

使用方法	13 幅 $300 \times 400$ 大小的图像的 实验时间 (ms)	18 幅 $640 \times 480$ 大小的 图像的实验时间 (ms)
本文方法	6 051. 053	19 817. 109
文献[7]的方法	2 763. 506	6 099. 531
文献[6]的方法	11 054. 027	24 369. 333

## 5 结 论

现有的柱状全景图的拼接,大部分都是基于焦距已知或直接通过手动拼接获得的连续图像宽度来进行焦距估计的。本文提出了一种自适应柱状全景图的拼接方法,它无需事先知道拍摄焦距,而是先通过精确定位相邻图像的匹配点坐标,然后按照匹配点对应柱状投影时的与水平中心线垂直距离相等的关系通过计算得到;同时通过对焦距进行估计与在其基础上产生的全景图宽度是圆半径和周长关系的

原则来迭代获得较精确的焦距参数。在拼接过程中, 单种图像匹配方法通常无法满足实时性、准确性或适用性。自适应柱状全景图拼接方法, 由于混合采用了不同的图像匹配方法因此能自适应选择最佳匹配结果, 并最终通过精匹配校准来确保全景图拼接结果的精确和鲁棒性。由于使用的核心算法计算复杂度较小, 因此运算时间较少。实验证明, 该方法具有一定的实时性。

### 参考文献 (References)

- 1 Chen S. QuickTime VR: an image-based approach to virtual environment navigation[J]. Computer Graphics, 1995, **29**(4): 29 ~ 38.
- 2 Memillan L, Bishop G. Plenoptic modeling: an image-based rendering system[J]. Computer Graphics, 1995, **29**(4): 39 ~ 46.
- 3 Szeliski R, Shum H Y. Creating full view panoramic image mosaics and environment maps[A]. In: Proceedings of ACM SIGGRAPH97 [C], Los Angeles, CA, USA, 1997: 251 ~ 257.
- 4 Peleg S, Herman J. Panoramic mosaics by manifold projection[A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], San Juan, Puerto Rico, 1997: 338 ~ 343.
- 5 Berg A, Malik J. Geometric blur for template matching [A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], Kauai, HI, USA, 2001: 607 ~ 614.
- 6 Zhong Li, Hu Xiao-feng. Overlapping image mosaic [J]. Journal of Image and Graphics, 1998, **3**(5): 367 ~ 370. [钟力, 胡晓峰. 重叠图像拼接算法 [J]. 中国图象图形学报, 1998, **3**(5): 367 ~ 370.]
- 7 Ping Jie, Yin Run-min. An algorithm for fast creating panoramic image and its implementation [J]. Microcomputer Applications, 2006, **27**(1): 59 ~ 62. [平洁, 殷润民. 一种全景图快速生成算法及其实现 [J]. 微计算机应用, 2006, **27**(1): 59 ~ 62.]
- 8 Yasushi K, Kenichi K. Robust image matching preserving global consistency [J]. Joho Shori Gakkai Kenkyu Hokoku, 2003, (2): 171 ~ 178.
- 9 Yin Run-min, Ping Jie, Chai Xu-dong, *et al.* An edge detector based on gray level difference analysis [J]. Computer Engineering, 2006, **18**(3): 201 ~ 203. [殷润民, 平洁, 柴旭东等. 一种基于灰度差统计的边缘检测方法 [J]. 计算机工程, 2006, **18**(3): 201 ~ 203.]