

一种基于视觉模型与连通域统计的阈值分割新算法

赵峙江¹, 张田文¹, 张志宏²

(1. 哈尔滨工业大学计算机科学与技术学院, 黑龙江哈尔滨 150001; 2. 哈尔滨工程大学计算机科学与技术学院, 黑龙江哈尔滨 150001)

摘要: 该文提出了一种新的基于视觉侧抑制模型和连通域统计的阈值图像分割算法。该方法模拟人眼处理图像机制, 建立侧抑制模型, 提取出图像中相对变化的信息, 并计算变化量矩阵的直方图, 然后对该直方图自下而上进行水平切割, 并计算割线上的连通域数量, 绘制连通域曲线图。仿真实验表明, 该方法具有很好的亮度、对比度不变性和线条连通性, 分割结果能更好的符合视觉感受。

关键词: 图像分割; 视觉模型; 侧抑制模型; 连通域

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2005)05-0793-05

A Novel Threshold Algorithm about Image Segmentation Based on Vision Model and Statistics of Consecutive Fields

ZHAO Shi-jiang¹, ZHANG Tian-wen¹, ZHANG Zhi-hong²

(1. Dept. of Computer Science & Technology, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China;

2. Dept. of Computer Science & Technology, Harbin Engineering University, Harbin, Heilongjiang 150001, China)

Abstract: This paper proposes a novel algorithm about image segmentation based on the vision lateral inhibition model and statistics of consecutive fields. The algorithm simulated the mechanism of the retina model of human eyes. First it picked up the variance information in the image, and then used level secant to cut the histogram from bottom to top. The algorithm counted the consecutive fields on every secant and compared each secant position and consecutive field numbers, and then made a consecutive curve. It can be proved that the secant corresponding to the curve's peak contains the boundary of the biggest consecutive field, which is the optimized threshold. The algorithm is very effective and robust. It is not only well adaptive to the change of brightness and contrast, and also can be applied to much complex background. It can improve the connectivity of strokes and lines. The result of segmentation is well coincident with the vision feeling of human eyes and the main reasons are discussed.

Key words: image segmentation; vision model; lateral inhibition; consecutive field

1 引言

计算机视觉中的图像理解包括目标检测、特征提取和目标识别等, 都依赖于图像分割的质量。到目前为止还不存在一种通用的方法, 也没有一个判断分割是否成功的客观标准, 因此图像分割技术被认为是计算机视觉研究中的一个瓶颈^[1-3]。阈值法是一种被广泛应用的简单有效的图像分割方法, 它用一个或几个阈值将图像的灰度级分为几个部分, 认为属于同一部分的像素是同一个物体。

在图像处理中, 有一类普遍存在的纹理图像, 如印刷文字, 手写文字, 视频文字, 工程图纸, 条码, 指纹, 掌纹, 印刷线路板等, 它们的特点是: (1) 目标图像由线条组成, 背景复杂; (2) 目标所占比例与背景相比相对较小, 且具有很宽的亮度、对比度范围; (3) 分割后的线条之间的连通关系对于目标识别

具有决定性作用。此类图像总体可分为简单图像和复杂图像两大类。简单图像中的对象种类较少, 像素在属性空间聚为少数几个子类, 直方图具有明显的峰谷特性。对此类图像进行阈值分割的方法可以大致分为两类: (1) 基于点的阈值法, 如二值扫描法、P-tile^[4]、双峰法^[5]、Otsu法^[6]等; (2) 基于区域的阈值法, 如边缘算子法^[7]、二叉树法^[8]、灰度共生矩阵法^[9]、二维直方图法、熵法^[10]等。但是由于目标质地、浓淡、采集设备噪声等影响, 这类方法分割结果的连通性不是很好, 存在线条缺损问题(如图1), 其中的噪声很容易使后续识别工作产生错误。另一类复杂图像是由于背景造成的, 其包含的对象数目和种类较多, 灰度层次比较丰富, 图像直方图峰谷特征不明显。对此类图像分割, 现有方法都存在一定的不足: (1) 若采用局部自适应阈值方法, 不仅难以自动选择合适的窗口尺寸, 而且容易破坏对象的完整性^[11]; (2) 采用区域生长方法则难以自

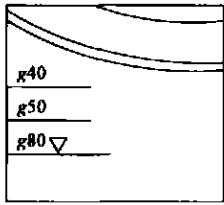


图1 用传统方法分割后的图像线划有缺损

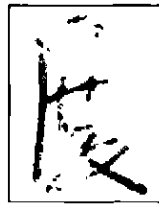


图2 手写文字图像中各线划灰度不均匀

动选取种子,尤其是当光照不均匀时,“分开-合并”分析方法受限于矩形网格,存在“歧义区”问题,容易破坏对象的完整性。

通过实验,我们发现线条缺损的根本原因是图像线条灰度不均匀、图像录入噪声过大造成的,没有一个阈值可以很好地分割一个线条或标注中的所有像素(如图2)。另外,每个线段的最佳阈值均有差别,阈值选取很难兼顾图像中所有内容,所以直接进行阈值分割不能取得理想效果。另外,由于摄像采集系统受光照的影响,图像的亮度、对比度变化范围很大,造成传统方法的适应困难。为解决这些问题,本文通过模拟人眼处理图像的机制,提出了一种新的基于视觉侧抑制模型与连通域统计的阈值分割算法。实验证明,该方法具有很好的亮度、对比度不变性和线条连通性,分割结果能更好的符合视觉感受。

2 阈值分割算法

2.1 侧抑制模型

生理学研究表明,侧抑制是生物神经网络中广泛存在的现象。当一个神经元受到刺激兴奋后,它不但加强自身,而且带动周围的神经元得以相应加强,同时抑制距离较远的神经元,这就是所谓“加强中心而抑制周围”的现象。进一步研究表明,侧抑制的作用半径大约是三倍的水平细胞大小,衰减率与其通过的面积成正比^[12,13]。视网膜神经元之间的侧抑制作用主要有以下几种功能:(1)可以突出边缘,增加反差(如马赫带效应);(2)高通滤波器,将很大的输入变化范围压缩到网络本身的动态范围之内,具有明显的亮度适应作用;(3)可以对图像的细微间断处进行拟合,具有明显的聚类作用(如 Hermann 网格幻觉现象^[14])。

为模拟视网膜的侧抑制作用,我们考虑一个的二维阵列网络,用表示位置的神经元,每个神经元只与其相邻神经元相连接,侧抑制半径为3。模仿生理模型,我们得到神经元兴奋后对其邻域所产生的侧抑制模型,如式(1):

$$\begin{aligned} \Delta E = & \sigma U_{ij} + \sigma k_1 \sum_{m=i-1, m \neq n}^{i+1} \sum_{n=j-1, n \neq j}^{j+1} U_{mn} - (\sigma k_2 \sum_{m=i-2}^{i+2} U_{m(j-2)} + \sigma k_2 \sum_{m=i-2}^{i+2} U_{m(j+2)}) + \sigma k_2 \sum_{n=j-1}^{j+1} U_{(i-2)n} + \sigma k_2 \sum_{n=j-1}^{j+1} U_{(i+2)n}) \\ & - (\sigma k_3 \sum_{m=i-3}^{i+3} U_{m(j-3)} + \sigma k_3 \sum_{m=i-3}^{i+3} U_{m(j+3)}) + \sigma k_3 \sum_{n=j-2}^{j+2} U_{(i-3)n} \\ & + \sigma k_3 \sum_{n=j-2}^{j+2} U_{(i+3)n} \end{aligned} \quad (1)$$

其中, ΔE 为神经元 S_{ij} 兴奋后产生的侧抑制变化总量, σ 为神

经元 S_{ij} 的兴奋增强系数, k_p ($p = 1, 2, 3$) 为衰减系数, U_{ij} 为神经元自身兴奋值。

由式(1)整理合并得式(2),

$$\Delta U_{ij} = \sigma U_{ij} + \sigma k_1 \sum_{m, n \text{ 为 } (i, j) \text{ 的 } 8 \text{ 邻域}} - \sigma k_2 \sum_{m, n \text{ 为 } (i, j) \text{ 的 } 16 \text{ 邻域}} - \sigma k_3 \sum_{m, n \text{ 为 } (i, j) \text{ 的 } 24 \text{ 邻域}} \quad (2)$$

其中, ΔU_{ij} 为神经元 S_{ij} 的兴奋值变化量。

另外,

$$k_i = \beta_i / s_i \quad (3)$$

其中, β_i 为距离衰减系数, s_i 为通过的面积。

将式(3)代入式(2),

$$\begin{aligned} \Delta U_{ij} = & \sigma (U_{ij} + \frac{\beta_1}{s_1} \sum_{m, n \text{ 为 } (i, j) \text{ 的 } 8 \text{ 邻域}} U_{mn} - \frac{\beta_2}{s_2} \sum_{m, n \text{ 为 } (i, j) \text{ 的 } 16 \text{ 邻域}} U_{mn} \\ & - \frac{\beta_3}{s_3} \sum_{m, n \text{ 为 } (i, j) \text{ 的 } 24 \text{ 邻域}} U_{mn}) \end{aligned} \quad (4)$$

整理得

$$\Delta U_{ij} = \sigma (U_{ij} + \beta_1 U_8 - \beta_2 U_{16} - \beta_3 U_{24}) \quad (5)$$

其中, U_8 、 U_{16} 和 U_{24} 分别为神经元 S_{ij} 的 8 邻域、16 邻域和 24 邻域的平均兴奋值。

式(5)表明了神经元 S_{ij} 的兴奋值变化量与其邻域神经元平均兴奋值的关系。像素 S_{ij} 如果距离亮点(兴奋值较高的神经元)很近,则可以得到较大提升,如果较远,则会受到很大抑制,具有明显的空间整合作用。图3是原图像的灰度直方图和归一化后的侧抑制矩阵的灰度直方图的比较,我们可以看出侧抑制矩阵的直方图密度更大,分布更均匀,另外由于侧抑制矩阵提取的是原图像的相对变化量,所以其直方图具有很好的亮度、对比度适应性。

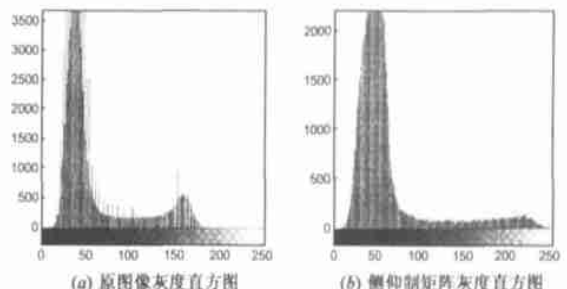


图3 原图像灰度直方图与侧抑制矩阵灰度直方图比较

2.2 最佳阈值搜索算法

为获得分割结果,需要对侧抑制矩阵求取最佳阈值。最佳阈值是侧抑制矩阵灰度直方图峰脚下的谷底。由于灰度分布的参差性,寻找谷底是困难的。本文提出一种快速搜索谷底的方法,通过水平割线在直方图中自下而上的运动,获得一系列割线连通域,通过对连通域分裂和合并的分析,得到最佳阈值。该方法不仅可以分割双峰直方图,对于单峰或多峰也可以获得较好的结果。

我们考察一个侧抑制矩阵直方图,当切割线由下至上移动时,会产生连通域分裂、缩小等几种情况,如图4所示:

(1)一个大连通域分裂为多个小连通域,如图4(a);

(2)连通区域面积变小(或保持不变),如图4(b);

(3) 连通区域消失,如图 4(c).

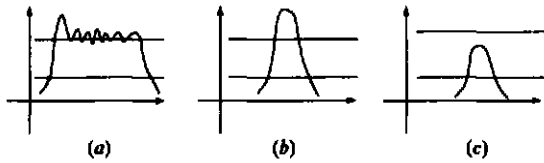


图 4 水平切割线切割侧抑制矩阵直方图

通过分析切割线上移过程中的连通域变化情况,我们根据连通域的定义和集合运算性质很容易证明如下结论:

- (1) 高位置切割线的连通域必然为低位置切割线连通域的子集.
- (2) 如果连通域数目为定值,则切割线位置在峰腰,如图 5(a)所示.
- (3) 如果连通域数目突然变大而后恢复原值,则切割线位置在峰脚,如图 5(b)所示.
- (4) 如果连通域数目突然变大而后稳定值大于原值,则切割线位置在谷底,如图 5(c)所示.
- (5) 如果连通域数目突然变大而后稳定值小于原值,则切割线位置在峰顶,如图 5(d)所示.

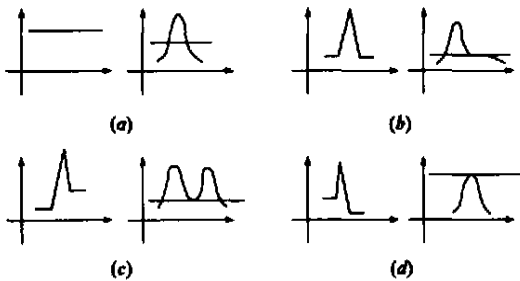


图 5 连通域数目变化与切割线位置的关系

通过实验我们得到不同图像最佳阈值的位置特点:

- (1) 对于普通的纹理图像,其侧抑制直方图具有双峰性质,最佳阈值在谷底;
- (2) 对于局部阈值分割产生的线条和文字很少甚至不存在的区域,其侧抑制直方图具有明显的单峰性质,最佳阈值在峰脚处;
- (3) 对于退化严重的图像,其侧抑制直方图具有多峰性质.由于线条和文字相对背景所占比例较小,并且侧抑制矩阵直方图线条灰度与背景相距较远,所以最佳阈值在最低的谷底.

综合上述讨论,我们首先将侧抑制矩阵直方图(图 6(a))自下而上以一定的步长进行水平切割,计算割线上的连通域数目,然后结果以割线的位置为横坐标、连通域的数目为纵坐标绘成连通域曲线,如图 6(b)所示.由于文字相对于背景像

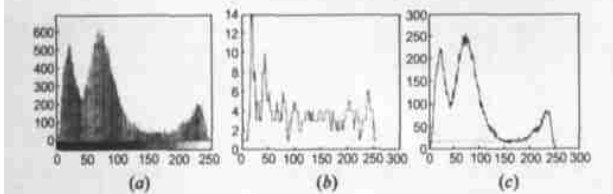


图 6 切割线的连通域数量曲线

素较少,所以切割的终止条件为割线下的面积为总面积的一半.根据前述连通域曲线与割线位置的对应关系及最佳阈值位置的分析,我们不难证明连通域曲线最大峰值所对应的割线是包含阈值的最佳割线如图 6(c)所示.最后,搜索最佳割线包含的最大连通域,其边界所对应的灰度值即为最佳分割阈值.(注:亮目标对应右边界,暗目标对应左边界)

2.3 整体算法

全部算法流程如下:

- ```

{
Step 1. 计算图像的侧抑制矩阵;
Step 2. 将侧抑制矩阵归一化;
Step 3. 计算侧抑制矩阵的直方图;
Step 4. 归一化直方图;
Step 5. 以一定步长自下而上进行水平切割;
Step 6. 计算割线的连通域值并保存;
Step 7. 判断割线下部的直方图面积,若不足总面积一半,重复 step 5;
Step 8. 绘制连同域曲线;
Step 9. 计算连同域曲线最大值,得到其对应的割线,称为最优切割线;
Step 10. 计算最优割线中面积最大的连通域的标值;
Step 11. 确定面积最大连通域边界所对应的灰度值;
Step 12. 较大的灰度值(暗背景亮目标)或较小的灰度值(亮背景暗目标)即为最佳分割阈值.
}

```

### 3 仿真实验

#### 3.1 实验

采用第 2 节中描述的算法对各类图像进行处理,图 7 至图 12 是运用本文方法与类间最大方差法进行比较结果.实验中,根据式(5),取  $\beta_i = 1/2^i$ , ( $i = 1, 2, 3$ ),由于侧抑制矩阵归一化后参与运算,所以的取值对结果不会有影响,这里取常数  $\sigma = 0.2$ ,由此得到邻域半径为 3 的侧抑制模型矩阵,见式 6.

$$B = \begin{bmatrix} -0.001 & -0.001 & -0.001 & -0.001 & -0.001 & -0.001 & -0.001 \\ -0.001 & -0.003 & -0.003 & -0.003 & -0.003 & -0.003 & -0.001 \\ -0.001 & -0.003 & -0.025 & -0.025 & -0.025 & -0.003 & -0.001 \\ -0.001 & -0.003 & -0.025 & -0.200 & -0.025 & -0.003 & -0.001 \\ -0.001 & -0.003 & -0.025 & -0.025 & -0.025 & -0.003 & -0.001 \\ -0.001 & -0.003 & -0.003 & -0.003 & -0.003 & -0.003 & -0.001 \\ -0.001 & -0.001 & -0.001 & -0.001 & -0.001 & -0.001 & -0.001 \end{bmatrix} \quad (6)$$



图 7 手写文字处理结果

### 3.2 实验结果分析

(1) 由图 7 可以看出由于质量较差的图像背景不均匀, 采用通常的类间最大方差法分割的结果含有很大噪声, 说明类间最大方差法对背景噪声很敏感. 而采用本文方法分割的结果比较清晰饱满, 光滑连通性好, 基本没有线条缺损的现象, 说明该方法具有抑制背景噪声的能力, 可以获得更好的分割效果. 这是因为侧抑制算法同时具有图像增强和图像聚类的作用, 图像经过处理之后, 不但线条与背景灰度差别更大, 而且线条灰度在侧抑制处理过程中得到补偿, 从而更有效地抑制噪声, 更容易获得最佳分割结果, 这为后续的认识工作打下了坚实的基础.

(2) 图 8、9 显示出本文方法具有很好的亮度和对比度适应性. 这是因为侧抑制处理结果矩阵只包含变化量信息, 将原图像的很大的灰度范围映射至一个相对较小的范围内, 与灰度的绝对值没有直接关系, 主要体现灰度之间的对比关系, 从而有利于突出特征. 这与人的视觉神经机制是一致的.



图 8 改变亮度的图纸分割结果



图 9 改变对比度的文字图像分割结果

(3) 图 10、11 显示出本方法对少量线条甚至是无线条的图像区域也可获得比较理想的分割结果. 这种能力对于分块

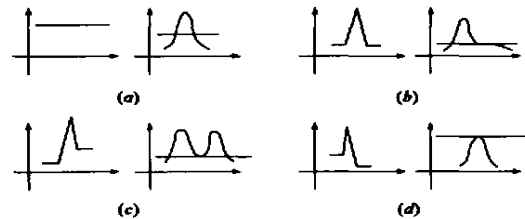


图 5 连通域数目变化与切割线位置的关系

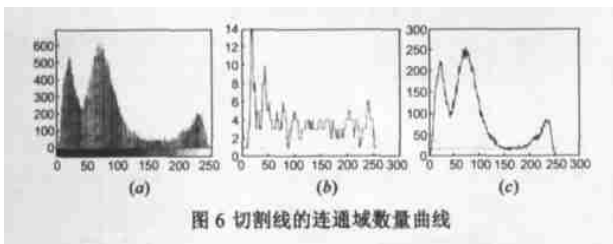


图 6 切割线的连通域数量曲线

自适应阈值分割尤其重要. 因为无论如何分块, 也不可能避免线条很少的区域出现, 这些区域通常很难进行分割. 本文方法取得较好的效果是因为侧抑制矩阵灰度直方图比原图的直方图更加细密、规整, 对它进行分割可以更容易找到最佳阈值. 对于少量线条或无线条的图像的情况, 本文方法也可以进行稳定的分割, 主要是因为线条与文字像素所占面积很少, 直方图为明显的单峰. 连通域数目最大的切割在峰脚, 即使对于极特殊的连通域曲线恒为 1 的情况, 由于由左至右搜索, 切割线也可以切割在峰脚, 从而可以搜索到最佳分割阈值. 由以上分析, 本算法用于局部自适应多阈值分割可以获得满意的结果, 具有较好的窗口适应性.

(4) 由图 12 可以看出, 对于具有复杂背景的图像, 本文方法也可以给出有效的分割结果. 由于侧抑制运算同时具有锐化和聚类的作用, 所以可以增大目标线条和背景的距离, 使多峰直方图的目标与背景之间的谷底更宽阔, 连通域曲线峰值更大, 搜索到的最佳阈值分割效果更好, 体现了本文方法具有更强的适应复杂背景的能力.

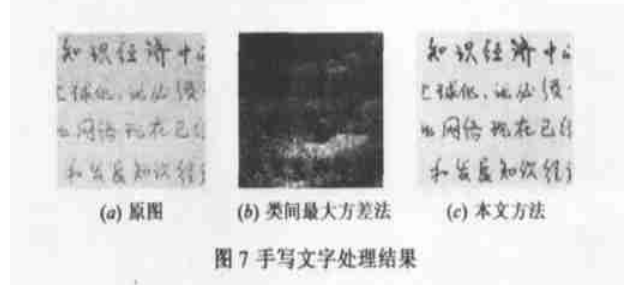


图 7 手写文字处理结果

### 4 结束语

本文详细论述了基于视觉模型与连通域统计的阈值分割方法. 该方法模拟生物侧抑制机理, 建立了视觉侧抑制模型, 设计了最佳阈值快速搜索算法. 仿真实验结果表明, 该方法具有很好的亮度和对比度适应性, 可以有效的抑制背景噪声, 具有突出目标特征和适应复杂背景的能力, 更容易获得最佳分割结果. 运用该算法进行图像分割能更好的符合视觉感受, 这为后续的认识工作打下了坚实的基础, 大大扩展了它的应用范围.

#### 参考文献:

[ 1 ] 王爱民, 沈兰荪. 图像分割研究综述[J]. 测控技术, 2000, 5: 1-5.  
 [ 2 ] 韩思奇, 王蕾. 图像分割的阈值法综述[J]. 系统工程与电子技术, 2002, 6: 91-102.  
 [ 3 ] 邢延超, 谈正. 基于多阈值融合的图像分割[J]. 计算机学报, 2004, 2: 252-256.  
 [ 4 ] Doyle W. Operation Useful for Similarity Invariant, Pattern Recognition [J]. J Assoc Comput March, 1962, 9: 259-267.  
 [ 5 ] Lee S, Chung S. A comparative performance study of several global thresholding techniques for segmentation[J]. Computer Vision, Graphics, and Image Processing, 1990, 52: 171-190.  
 [ 6 ] Ostu N. A threshold selection method from gray level histograms[J].

- IEEE Trans on System, Man, and Cybernetics, 1979, 9(1):62-66.
- [ 7 ] Mason D, Lauder I J, Rutoritz D, et al. Measurement of G bands in human chromosomes[ J]. IEEE Trans on Pattern Anal Mach Intell, 1975, 7: 155- 164.
- [ 8 ] Wu A Y, Rosenfeld A. Threshold selection using quadtree [ J]. IEEE Trans on Pattern Anal Mach Intell, 1982, 4: 90- 94.
- [ 9 ] Ahuja N, Rosenfeld A. A note on the use of second order gray-level statistics for threshold selection[ J]. IEEE Trans on Systems, Man and Cybernetics, 1978, 8: 895- 899.
- [ 10 ] Abutaleb A S. Automatic thresholding of gray-level picture using two dimensional entropies[ J]. Pattern Recognition, 1989, 47: 22- 32.
- [ 11 ] Yen J C, Chang F J, Chang S. A new criterion for automatic multilevel thresholding[ J]. IEEE Trans on Image Processing, 1995, 4: 370- 378.
- [ 12 ] De Vries S H, Baylor D A. Mosaic arrangement of ganglion cell receptive fields in rabbit retina[ J]. J Neurophysiol, 1997, 78: 2048- 2060.
- [ 13 ] Krantz J H. A computational model of the retina Presented at the 30th annual meeting of the Society for Computers in Psychology [ J]. New Orleans, 2000, LA.
- [ 14 ] Lee S, Chung S. A comparative performance study of several global thresholding techniques for segmentation[ J]. Computer Vision, Graphics, and Image Processing, 1990, 52: 171- 190.
- [ 15 ] 张永平, 郑南宁, 雷雅莉. 具有侧抑制机制的视觉系统模型及在图像边缘提取中的应用[ J]. 电子学报, 1998, 8: 99- 101.
- [ 16 ] 陈卉, 欧阳楷. 用于图像增强的侧抑制网络模型的仿真比较[ J]. 系统仿真学报, 2003, 1: 100- 103.
- [ 17 ] 冯晓毅, 刘诗滨. 图像边缘增强的多分辨率侧抑制网络方法[ J]. 计算技术与自动化, 2003, 6: 97- 99.
- [ 18 ] 杨亚军, 李言俊, 张红林, 胡战虎. 循环侧抑制网络抑制系数的选取及在图像预处理中的应用[ J]. 弹箭与制导学报, 1998, 3: 15- 17.
- [ 19 ] 邢晓正, 戴国平, 周绍祥. 基于视觉侧抑制模型的图像增强算法[ J]. 测试技术学报, 1996, 2- 3: 523- 526.
- [ 20 ] 王怀野, 李言俊. 侧抑制网络在彩色图像边缘增强中的应用[ J]. 计算机工程与应用. 2003, 26: 55- 57.

#### 作者简介:



赵峙江 男, 1973 年生于黑龙江, 博士研究生, 主要研究方向为计算机视觉、图像处理、模式识别等. E-mail: zzh\_hrb@sohu.com.



张田文 男, 1940 年生于黑龙江, 教授, 博士生导师, 主要研究方向为图像信息理论、模式识别、计算机视觉、虚拟现实、人工生命, 近年来, 先后负责国家自然科学基金项目 3 项, 863 计划项目 3 项, 国家教委博士点基金项目 2 项; 获航天工业总公司科技进步奖 3 项, 在国内外学术刊物和会议上发表论文几十篇.

张志宏 女, 1974 年生于黑龙江, 博士研究生, 主要研究方向为图像处理、虚拟现实与可视化技术等.