

基于高斯尺度空间的尺度参数自适应算法研究^①

刘 立^② * ** 彭复员 * 田 岩 * 万亚平 **

(* 华中科技大学电信系 武汉 430076)

(** 南华大学计算机科学与技术学院 衡阳 421001)

摘 要 以屋脊型边缘和阶梯型边缘的特征作为图像的信息度量依据,研究了图像信息量随尺度参数变化而变化的规律,并基于视觉特征模型提出了一种高斯尺度参数的自适应算法。该算法通过使尺度空间信息量在每层均匀分布来实现尺度参数的自适应变化,避免了在尺度空间的构造过程中因计算过于复杂或因丢弃过多关键信息而造成失真过大的问题。实验结果显示,该算法能有效地抑制噪声,从而为图像的高层次处理如目标识别等奠定了基础。

关键词 高斯尺度空间, 尺度参数, 视觉特征, 特征点

0 引 言

近年来,多尺度技术在图像处理领域得到越来越多的关注,最典型的例子就是金字塔结构与小波分解。另一种重要的多尺度技术就是所谓的尺度空间(scale space)，“尺度空间”概念是著名的图像金字塔概念的新补充,最先在 Kelly 的图像处理中被采用,后来又以各种形式进行扩展。在现代的尺度空间表示中,每一层是前一层的模糊结果。通常每一层通过一个方差为 σ^2 的高斯函数卷积原图像而产生,而每一层的 σ 都不一样,这个方差就是“尺度参数”,尺度参数越大,图像越模糊,越多的细节被丢弃,有文献称这种细节为特征点,这些特征点对人眼视觉起关键作用,利用它可以有效地进行目标识别、特征提取等^[1,2]。目前大多数应用中采用的层间的尺度参数 σ 关系有的为等间隔或近似等间隔关系^[3,4],有的采用比例关系^[5],但图像的细节并不随尺度参数的变化而同步变化,采用简单的尺度参数处理会使某些层因尺度参数过小造成过多的细节冗余,或某些层会因尺度参数过大使细节丢失过多,使得分层效果不理想。

本文针对不同图像的多尺度空间信息量度量模型给出了尺度参数的自适应算法,并通过试验证明了它的正确性,从而为多尺度空间的图像处理提供了可靠的依据。最后还给出了等间隔方法与这种自

适应算法的实验比较。

1 尺度空间

1.1 定义

Lindeberg 在其文献中较规范地定义了尺度空间^[6,7]。每一层均有一个尺度参数代表该层尺度的特征,该尺度参数可以是离散的也可以是连续的。所有的尺度空间应该具有如下性质:

- (1) 所有信号应该在同一定义域中被定义;
- (2) 随着尺度参数的增大,输出图像的表示应该越来越模糊;
- (3) 一个信号在粗的层次中包含的细节要比细的层次中少,如果将局部极值看作平滑度的一种度量,那么随着尺度变模糊,极值是非增的,这种属性称为“尺度空间的因果关系”;
- (4) 所有的表示都是通过一个卷积核算子产生。

图像的这种变模糊是通过与尺度核卷积完成的,Koenderink 与 Lindeberg 的研究表明,唯一可能的尺度核就是高斯核^[8-10]。

1.2 高斯尺度核与高斯差函数

一幅图像的尺度空间被定义为函数 $L(x, y, \sigma)$,它是尺度变化的高斯函数 $G(x, y, \sigma)$ 与图像 $I(x, y)$ 的卷积,即

$$L(x, y, \sigma) = G(x, y, \sigma) \times I(x, y) \quad (1)$$

① 国家自然科学基金(60475024, 60872076)和航天技术创新基金(2006AA09Z203)资助。

② 男,1971 年生,博士;研究方向:图像处理,机器视觉;联系人,E-mail: 11710915@yahoo.com.cn
(收稿日期:2007-11-13)

从而有

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/(2\sigma^2)} \quad (2)$$

要保留尺度空间中的信息,可通过高斯-拉普拉斯算子(LOG)来实现^[8],也可以用高斯差函数(difference-of-Gaussian, DOG)与图像卷积的方式来近似表示,高斯差函数 $D(x, y, \sigma)$ 从两个相邻尺度之差计算得出:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) \times I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (3)$$

2 基于视觉特征的信息度量

图像的特征有统计特征,也有视觉特征。本文采用视觉特征来度量信息量的大小。

在 Marr 的视觉理论中的初始图像是由图像的亮度变化中直接推导提取出来的第一个表象,由图像要素组成的基元在很大程度上直接反映了物理实在,因此,初始简图中的要素数量就可以作为一个十分好的信息度量标准。初始简图中的要素有边缘、棒、斑点和端点,它们是通过检测图像中灰度变化的检测和定位得到的。这些基元就构成了图像的视觉特征。初始简图的生成是比较复杂的,可以使用所谓的特征点来描述,而图像的信息量主要由两部分特征点组成,即局部非平凡极点和非平凡拐点(图像局部灰度变化最剧烈的像素点),其中第一类点对应屋脊型边缘(ramp edge)的特征,第二类对应阶梯型边缘(step edge)的特征。由于人的视觉对这两类特征点比较敏感,因此它们可构成图像信息的主要内容。直接采用文献[11-13]有关图像信息度量与特征点检测的结论,定义在尺度 σ 下的离散图像信息量为两类特征点的和。

第一类特征点集合:

$$\begin{aligned} \Phi_1(f) &= \{(m, n): 1 < m < M, 1 < n < N; \\ &(f(m, n) - f(m-1, n))(f(m, n) - f(m+1, n)) > 0; \\ &(f(m, n) - f(m, n-1))(f(m, n) - f(m, n+1)) > 0; \\ &(f(m, n) - f(m-1, n-1))(f(m, n) - f(m+1, n-1)) > 0; \\ &(f(m, n) - f(m+1, n+1))(f(m, n) - f(m+1, n+1)) > 0;\} \end{aligned} \quad (4)$$

第二类特征点集合:

$$\begin{aligned} \Phi_2(f) &= \{(m, n): 2 < m < M-1, 2 < n < N-1; \\ &(f(m+1, n) - f(m-1, n))(f(m+2, n) - f(m, n)) \times \\ &(f(m+1, n) - f(m-1, n))(f(m, n) - f(m-2, n)) > 0; \\ &(f(m, n+1) - f(m, n-1))(f(m, n+2) - f(m, n)) \times \\ &(f(m, n+1) - f(m, n-1))(f(m, n) - f(m, n-2)) > 0; \\ &(f(m+1, n+1) - f(m-1, n-1))(f(m+2, n+2) - f(m, n) \\ &\times (f(m+1, n+1) - f(m-1, n-1))(f(m, n) \\ &- f(m-2, n-2)) > 0; \\ &(f(m+1, n-1) - f(m-1, n+1))(f(m+2, n-2) - f(m, n)) \\ &\times (f(m+1, n-1) - f(m-1, n+1))(f(m, n) \\ &- f(m-2, n+2)) > 0;\} \end{aligned} \quad (5)$$

其中, $f(m, n)$ 表示图像 f 在像素点 (m, n) 处的灰度值。图像 $f(m, n)$ 的信息量等于:

$$I_p = |\Phi_1(f)| + |\Phi_2(f)|^{[11-13]}$$

通过式(4),(5)得到的信息量称为候选信息点,需要进一步的确认。因为在对比度低的地方(该处对噪声敏感)有可能产生伪信息点,需要过滤掉。在式(4)中增加阈值 c ,可以有效地抑制噪声,即:

$$\begin{aligned} \Phi_1(f) &= \{(m, n): 1 < m < M, 1 < n < N; \\ &(f(m, n) - f(m-1, n))(f(m, n) - f(m+1, n)) > c; \\ &(f(m, n) - f(m, n-1))(f(m, n) - f(m, n+1)) > c; \\ &(f(m, n) - f(m-1, n-1))(f(m, n) - f(m+1, n-1)) > c; \\ &(f(m, n) - f(m+1, n+1))(f(m, n) - f(m+1, n+1)) > c;\} \end{aligned} \quad (6)$$

当 $c = 0$ 时就是文献[13]介绍的求信息量的方法,一般取 $c = 0.03$ 比较合适^[5]。采用对 Lena 图像进行每层尺度间隔为 0.4 的平滑,得到 10 幅图像如图 1 所示。



图 1 尺度参数在 0.4~4 等间隔变化的 Lena 图像

从左到右依次为 Lena 图像原图和尺度参数从 0.4 到 4 以间隔为 0.4 的 11 幅 Lena 图像,可以发现前三幅图像没有明显的变化,从第 4 幅图像开始出

现明显的模糊,最后四幅图像的模糊程度又没有明显的差别,采用上面所描述的信息两度量所得到的每幅图像信息量数据如表 1 所示。

表 1 Lena 图像的不同尺度下信息度量值

尺度参数	0	0.4	0.8	1.2	1.6	2	2.4	2.8	3.2	3.6	4
$ \varphi_1(f) + \varphi_2(f) $	600	593	300	91	44	16	8	7	6	4	2

表 1 与图 2 验证了 Lena 系列图像的变化。随着尺度参数的增加,视觉信息量先是平缓地递减,然后是剧烈递减,最后又是平缓地递减。从 DOG 图像也可以看出每层图像信息量的丢失,如图 3 所示。

从图 3 可以看出,信息量的损失在每个层次之间不一样,开始信息量损失少,反映平滑程度低,在中间信息损失多,最后信息量损失又减少,进一步验证了信息量度量的合理性。

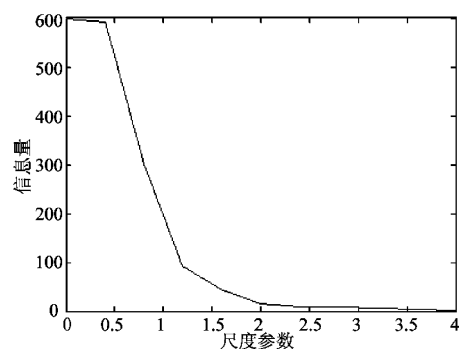


图 2 图像信息量与尺度参数的变化曲线

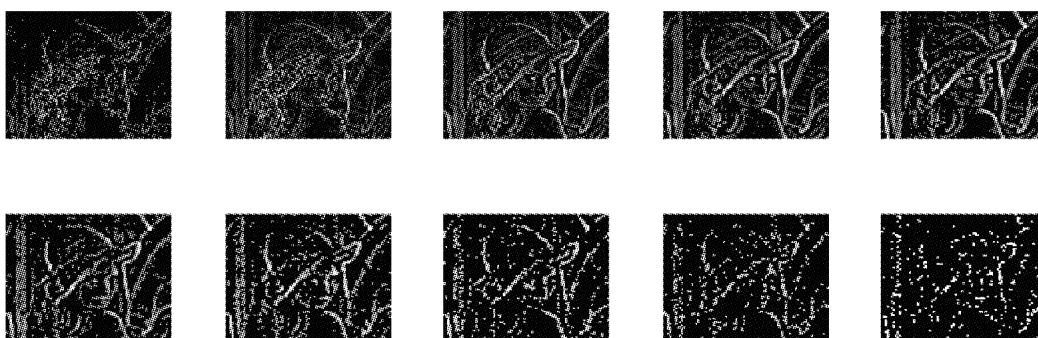


图 3 尺度参数 0.4~4 等间隔变化的 Lena DOG 图像

3 基于高斯尺度空间的尺度参数自适应算法

3.1 尺度参数自适应算法概述

尺度参数自适应算法的主要步骤如下:

- (1) 利用公式(5)、(6)求得原图像的信息量,设为 inf_0 ;
- (2) 确定合适的阈值 ϵ ;
- (3) 如果 $inf_0 < \epsilon$, 退出;
- (4) 采用较小的等间隔尺度对原图像进行平滑,得到一系列的平滑图像 f_i ;
- (5) 利用公式(5)、(6)求得每幅图像 f_i 的信息量,设为 inf_i ;
- (6) 分两种情况重新建立高斯尺度空间 f_k :

1) 若相邻两个层的图像信息量小于 ϵ , 将尺度参数调整到较大值,直到找到满足相邻两层图像信息量 $\geq \epsilon$ 的最小值,即相邻两层图像之间的信息量 $inf_k - inf_{k-1} \geq \epsilon, f_k = \min\{f_i \mid inf_i - inf_{i-1} \geq \epsilon\}$, 若 $inf_k < \epsilon, f_i \rightarrow f_{i+1}$, 返回(6);

2) 当相邻两个层的图像信息量大于 ϵ 时重新调整尺度参数 σ_i 为 $\frac{\sigma_i}{2}$, 返回(1), 循环直到生成新的高斯尺度空间,若 $inf_k < \epsilon, f_i \rightarrow f_{i+1}$, 返回(6);

3) 完成遍历所有层。

(7) 利用相邻高斯图像之差建立高斯差(DOG)尺度空间。

3.2 阈值 ϵ 的确定

阈值 ϵ 的确定是算法的关键,这里分两种情况讨论(设原图像信息量为 N):

(1) 当分层层数不确定时,选择 ϵ 既要满足层间保存的信息量足够多而分层层数不能过多,同时又要满足层间保存的信息不能太多使得图像细节不能充分表达在尺度空间里。由于要满足生物视觉模型,对 50 幅不同类型的图像进行了试验,设 STD 为图像灰度数据的方差,实验表明当取 $\epsilon = 6\text{STD}$ 时,既能满足人的视觉分辨足够的细节,也能限制分层

数目不至于过多。

(2) 当分层层数确定时,则 $\epsilon = \frac{N}{\text{分层层数}}$ 。

以 Lena 图像为例,若将 ϵ 定为 80,则尺度参数调整如下:

0 0.6 0.8 1 1.2 1.4 1.6 2 2.4 2.8 4

调整以后得到的 DOG 图像如图 4 所示:

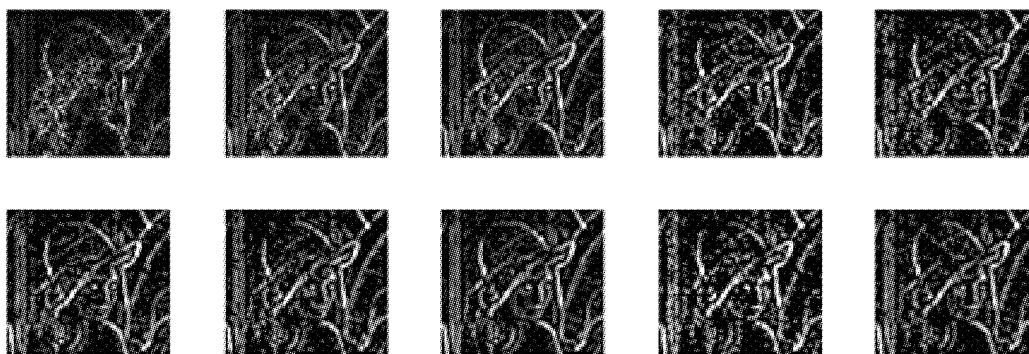


图 4 尺度参数自适应变化的 Lena DOG 图像

图 4 反映出的每层图像间信息量的丢失呈现均匀的特点。

4 图像多尺度表示的处理

目前图像的多尺度的表示方法广泛应用在机器视觉定位与识别、机器人导航、路径规划等研究中^[14,15],利用尺度空间的带通性能也可以较好地实现图像去噪等应用。

假设原图为 $I(x, y)$, $L(x, y, \sigma)$ 是采用尺度为 σ 的高斯平滑图像, $D(x, y, \sigma)$ 是图像高斯差,代表层之间信息的丢失,由式(1)、(2)、(3)得:

$$I(x, y) = L_n(x, y, \sigma_n) + D_1(x, y, \sigma_1) + D_2(x, y, \sigma_2) + \cdots + D_n(x, y, \sigma_n) \quad (7)$$

其中 $L_n(x, y, \sigma_n)$ 是最后一层的平滑图像。去噪过程往往伴随着图像的模糊,如果噪声集中在某些尺度下就可以将该层高斯差图像丢失,则应在实现图像去噪的同时保存图像的细节。

由式(7)可知,丢弃某些噪声集中的 $D_i(x, y, \sigma_i)$, 可用得到的 $I'(x, y)$ 近似表示原图像,实现图像的去噪。由于等间隔尺度空间中各层包含的细节不均匀,若不能充分地将有用的细节信息与噪声分开,则有可能在去噪的同时丢失有用信息。图 5 反映了图像去噪的过程,第 1 幅图像是被随机噪声和椒盐噪声污染的 Lena 图像,第 2 幅是采用中值滤波方法得到的图像,第 3 幅为采用等间隔尺度滤波的图像,最后一幅为本文推荐的自适应尺度参数滤波方法的图像。

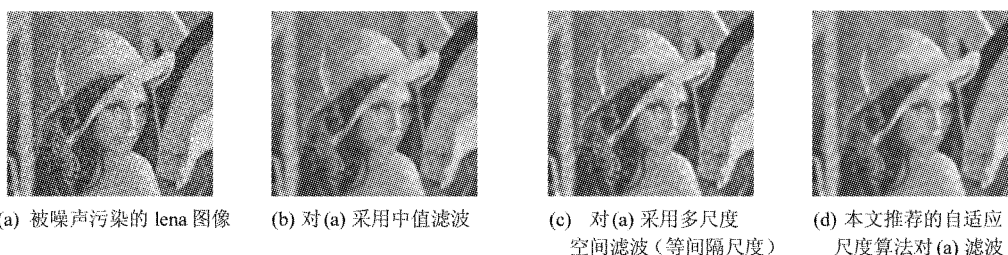


图 5 自适应尺度算法和倍增尺度的图像去噪比较

为了度量算法对去噪的效果,本文参考了标准差和峰值信噪比(PSNR)这两个参数,其中

$$PSNR = 10 \lg \frac{255^2 MN}{\sum_{i=1}^M \sum_{j=1}^N |R(i,j) - \hat{X}(i,j)|^2} \quad (8)$$

表 2 去噪实验数据

比较内容	原图	噪声图像	中值滤波	等间隔尺度	自适应算法
标准方差	40.1037	43.1412	40.8481	42.2427	40.6658
峰值信噪比(dB)		30.8231	33.5413	31.8754	33.9781

利用图像的高斯多尺度表示可以有效地实现图像去噪,而且尺度参数自适应的去噪效果好于尺度参数倍增变化的去噪效果,通过将平滑图像和部分高斯差图像融合还可以实现图像的压缩。

5 结 论

在对传统的图像信息度量方法改进的基础上提出一种高斯尺度参数的自适应算法,该算法产生的高斯尺度空间能在各层之间均匀地表示图像的信息量。最后给出了在图像去噪方面的应用实例。实验结果表明,多尺度的表示方法能有效地实现图像去噪,而且在通常情况下基于信息量均匀变化的分层处理效果优于基于尺度参数等间隔的分层处理效果。

该算法在对多幅图像进行信息量的度量时具有一定的时间复杂度,考虑信息度量数据类似高斯分布,可以通过高斯曲线的某个参数拟合该特征曲线,从而快速地得到一系列非线性增长的尺度参数;同时如何找到合适的表达视觉特征曲线的参考模型,并将它应用到目标识别中将是下一步工作的重点。

参考文献

- [1] Kim J, Sikora T. Gaussian scale-space dense disparity estimation with anisotropic disparity-field diffusion. In: Proceedings of the 5th International Conference on 3-D Digital Imaging and Modeling, Washington, Ottawa, Canada, 2005. 556-563
- [2] Dork G, Schmid C. Selection of scale-invariant parts for object class recognition. In: Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV'03), Nice, France, 2003
- [3] Mikolajczyk K, Schmid C. Scale and affine invariant interest

其中: R 为无噪声的参考图像; $\hat{X}$ 为处理后得到的恢复图像。实验数据如表 2 所示。

可以看出,文中推荐的算法对图像滤波效果最好,在噪声能量密集在某个尺度空间的情况下,采用推荐的算法效果更好。

- point detectors. *International Journal of Computer Vision*, 2004, 60(1): 63-86
- [4] Dufournaud Y, Schmid C, Horaud R. Matching images with different resolutions. In: Proceedings of Computer Vision and Pattern Recognition, (CVPR'00), South Carolina, USA, 2000. 612-618
- [5] Lowe D. Object recognition from local scale-invariant features. In: Proceedings of the International Conference on Computer Vision, (ICCV'99), Kerkira, Greece, 1999. 1150-1157
- [6] Cunha A, Teixeira R, Velho L. Discrete scale spaces via heat equation. In: Proceedings of the Brazilian Symposium on Computer Graphics and Image Processing, Florianopolis, Brazil, 2001. 68-75
- [7] Lindeberg tony. Scale-space: a framework for handling image structures at multiple scales. In: Proceedings of the Conference European Organization for Nuclear Research School of Computing. Egmond aan zee, The Netherlands, 1996. 8-21
- [8] Lindeberg tony. Scale-space for discrete signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, (12): 234-254
- [9] Lindeberg tony. Scale-space theory: A basic tool for analyzing structures at different scales. *Journal of Applied Statistics*, 1994, (2): 224-270
- [10] Wesley 等著, 崔锦实等译. 机器视觉教程. 北京: 机械工业出版社, 2005. 64-79
- [11] Koenderink J J. The structure of images. *Biological Cybernetics*, 1984, (50): 363-396
- [12] Marr David. Vision. San Francisco CA, USA: Freeman Publishers, 1982
- [13] 王郑耀等. 基于视觉特征的尺度空间信息量度量. 中国图像图形学报, 2005, (7): 922-928
- [14] Stephen Se, Lowe. D, Little J. Global localization using distinctive visual features. *Intelligent Robots and System*, 2002, (1): 226-231
- [15] Stephen Se, Lowe D, Little J. Vision-based mobile robot lo-

calization and mapping using scale-invariant features.

Robotics and Automation, 2001, (2):2051-2058

An adaptive algorithm of scale parameter based on Gaussian scale-space

Liu Li^{* **}, Peng Fuyuan^{*}, Tian Yan^{*}, Wan Yaping^{**}

(^{*} Department of Telecommunication, Huazhong Science and Technology University, Wuhan 430076)

(^{**} Department of Computer Science and Technology, NanHua University, Hengyang 421001)

Abstract

Regarding ramp edge and step edge as the measurement of image information, the image information change law with the scale parameter is first studied. Then a new adaptive algorithm of scale parameter with the module of visual characters is proposed. The information of the image is uniformly distributed among every layer in this algorithm. The method can avoid the problem of complicated computation or over-distortion due to losing too much key information. The experimental results show that the algorithm can suppress the noise effectively and provide the basis for higher-level image processing technologies such as object recognition.

Key words: Gaussian scale-space, scale parameter, visual characters, feature point