

自适应 Mean Shift 算法的彩色图像平滑与分割算法

王晏¹ 孙怡¹

摘要 采用 Mean shift 算法对图像进行平滑和分割处理时, 带宽和采样点权重的选择直接影响平滑和分割的效果. 带宽分为空域带宽和值域带宽. 本文根据图像颜色分布的丰富程度定义了自适应空域带宽. 在此基础上, 通过最小化局部方差函数和最大化频域结构相似度函数获得自适应值域带宽. 此外, 通过定义采样点权重, 克服了图像过平滑问题. 通过随机选取大量的图像进行实验, 结果表明运用本文所选择的带宽和权重, 可以得到正确的图像区域分割结果.

关键词 Mean shift, 带宽, 权重, 频域结构相似度, 图像分割

DOI 10.3724/SP.J.1004.2010.01637

Adaptive Mean Shift Based Image Smoothing and Segmentation

WANG Yan¹ SUN Yi¹

Abstract Bandwidths and weights of sampling points are two key points in mean shift based image smoothing and segmentation. Bandwidths indicate spatial bandwidths and range bandwidths. Adaptive spatial bandwidths are defined according to color distribution of the image. Then, adaptive range bandwidths are obtained by minimizing the local variance function and maximizing the frequency structural similarity function. Additionally, weights of sampling points are defined to overcome over smoothness. Experimental results prove that the correct segmented regions are obtained by using the proposed bandwidths and weights.

Key words Mean shift, bandwidth, weights, frequency domain-based structural similarity, image segmentation

彩色图像分割是机器视觉领域中一个重要而具有难度的研究方向, 也是目标检测的基础. 在目标检测中, 需要首先将图像分割成不同的区域, 待检测的目标通常可以认为是这些区域的某种组合. 过细的分割会增加检测的难度和复杂度, 过粗的分割会导致错误的检测结果. 本文的目的是在运用 Mean shift 算法平滑和分割图像时, 折中考虑图像的平滑效果和失真程度, 使分割的区域不会太细或太粗, 从而能正确表达待检测目标的整体性. Mean shift 算法的优点是其算法结构简单, 不需要先验知识, 运行速度快等. Han 等^[1] 将 Mean shift 算法与双边滤波等算法进行了性能及运行时间的对比, 从而进一步说明了 Mean shift 算法在图像处理方面的优越性. 在运用 Mean shift 算法进行图像平滑与分割时, 核函数带宽是一个重要的参数, 它不但决定了参与迭代的采样点数量, 而且还会影响算法的收敛速度和准确性. 带宽一般分为两种: 1) 全局固定带宽, 是指在整个 Mean shift 算法的迭代过程中, 带宽值不变. 确定全局带宽的方法一般有两种: 一种是根据经验人为确定^[2-6], 但这种方法制约了 Mean shift

算法的自动化过程, 并且不能适应图像的多变性和复杂性. 另一种是根据整体数据估计最优全局带宽. Hong 等^[7] 运用 plug-in 规则估计全局带宽, 并将所估计的带宽乘以一个比例系数, 但当采样点包含多种模态时, 一般很难计算出全局最优带宽. Gondra 等^[8] 运用示例学习的方法训练得出全局带宽, 但所选学习样本的种类和数量会直接影响平滑和分割结果. Zheng 等^[9] 通过设置带宽的变化步长选取不同的带宽进行分析和比较, 从而得出较为理想的带宽, 但这种方法实际上也是人为参与的, 并且不同的带宽初值和变化步长对分割的结果会有直接影响. 2) 局部自适应带宽, 是指根据数据或其特征, 对密度大的区域采用小带宽, 而对密度小的区域采用大带宽. 由于全局固定带宽的局限性, 一些学者对局部自适应带宽作了进一步的研究. Comaniciu 等^[10-11] 提出了两种局部自适应带宽选择方法并用于图像的平滑与分割中, 其中, 一种方法是局部自适应带宽反比于局部估计的平方根. 另一种方法是通过求取局部均值向量模的最大值得到最优局部自适应带宽. Jiménez-Alaniz 等^[12] 也将局部自适应带宽设定为反比于局部估计的平方根, 并且给出了比例系数的求解方法. Mayer 等^[13-14] 将局部自适应带宽定义为 K 近邻采样点与被平滑点差的范数, 文中的近邻系数是根据经验选取的, 实际上, 近邻系数的选取会直接影响平滑及分割结果. 在上述方法中, 还存在一些问题需要进一步研究: 1) 没有讨论采样点权重对

收稿日期 2010-02-08 录用日期 2010-06-08
Manuscript received February 8, 2010; accepted June 8, 2010
“十一五”国家科技支撑计划课题 (2006BAK02B01-20) 资助
Supported by the National Key Technology Research and Development Program in the “Eleventh Five-Year Plan” Period (2006BAK02B01-20)

1. 大连理工大学信息与通信工程学院 大连 116024
1. School of Information and Communication Engineering,
Dalian University of Technology, Dalian 116024

平滑及分割结果的影响; 2) 上述自适应带宽方法并没有考虑人眼的主观视觉特性. 鉴于此, 本文根据图像的颜色信息分布, 定义了权重函数, 求取了自适应空域带宽, 并且通过最小化局部方差和最大化图像频域结构相似度确定了值域带宽, 从而将带宽的选择同图像颜色信息、评价图像质量的频域结构相似度有机结合起来, 得到符合人眼感知的平滑和分割结果.

1 Mean shift 算法

Mean shift 算法的基本原理如式 (1) 所示:

$$m(x) = \frac{\sum_{i=1}^n G\left(\frac{x_i - x}{w}\right) \omega(i)x_i}{\sum_{i=1}^n G\left(\frac{x_i - x}{w}\right) \omega(i)} \quad (1)$$

x 表示被平滑点的像素值. x_i ($i = 1, 2, \dots, n$) 表示以被平滑点为中心, 边长为 $2r$ 的正方形区域内的像素点的值, 正方形区域内的所有像素点称为采样点, r 称为空域带宽. $G(x)$ 为核函数, 一般有两种常用的核函数: 单位核函数和高斯核函数. w 为值域带宽. 本文采用的核函数为高斯函数, w 越大, 高斯函数曲线越平坦, 对图像的平滑效果越强. $\omega(x_i)$ 为赋予每一个采样点的权重. 本文在 Lab 颜色空间对图像三个分量 l, a, b 分别运用 Mean shift 算法进行平滑, 具体步骤如下:

步骤 1. 给定初始条件, 包括: 起始像素点, 每个采样点的权重 $\omega(x_i)$, 核函数 $G(x)$, 误差 ε , 空域带宽 r , 值域带宽 w . 其中, 初始点从图像左上端的第一个像素点开始, 采用高斯函数为核函数. 设定一个误差极小值 $\varepsilon = 0.01$. 采样点权重、空域带宽和值域带宽的确定将在第 2 节中详细叙述.

步骤 2. 根据式 (1) 计算 $m(x)$.

步骤 3. 如果 $|m(x) - x| > \varepsilon$, 将 $m(x)$ 赋予 x , 即令 $x = m(x)$, 并且返回到步骤 2. 反之, 结束该像素点的迭代, 指向下一个像素点, 并且返回到步骤 2, 直至整个图像遍历结束. 为了更准确地表达两点之间的颜色差异, 本文对式 (1) 进行了调整, 用采样点与被平滑点之间的色差 $((l_i - l)^2 + (a_i - a)^2 + (b_i - b)^2)^{1/2}$ 替换式 (1) 中的 $x_i - x$ 项, 即

$$m(x) = \frac{\sum_{i=1}^n G\left(\frac{\sqrt{(l_i - l)^2 + (a_i - a)^2 + (b_i - b)^2}}{w}\right) \omega(i)x_i}{\sum_{i=1}^n G\left(\frac{\sqrt{(l_i - l)^2 + (a_i - a)^2 + (b_i - b)^2}}{w}\right) \omega(i)} \quad (2)$$

其中, l_i, a_i, b_i 为采样点在 Lab 颜色空间的分量值.

2 自适应 Mean shift 算法

在运用 Mean shift 算法进行图像平滑和分割时, 存在三个问题需要分析: 1) 怎样选取采样点权重; 2) 怎样自适应选取空域带宽; 3) 怎样自适应选取值域带宽. 现有的基于 Mean shift 的图像平滑和分割算法中, 一些研究者并没有考虑采样点权重的影响, 并且很多学者都会根据经验采用人为固定的空域带宽或值域带宽^[2-6]. 而对于自适应空域带宽和自适应值域带宽的分析, 通常采用数据优化的方法选取自适应带宽, 而并没有考虑人眼的视觉特性. 针对以上问题, 本文首先根据图像的颜色分布信息定义了采样点的权重函数, 确定了自适应空域带宽. 然后建立评价值域带宽选取的两个目标函数: 局部方差和频域结构相似度, 通过最大化目标函数求取自适应值域带宽, 从而对图像进行有效的平滑和分割.

2.1 采样点权重的设置

在式 (1) 中, $\omega(x_i)$ 为赋予每一个采样点的权重, 权重不同, 表示每个采样点在平滑过程中的重要性不同. 在彩色图像平滑和分割中, 采样点颜色与被平滑点颜色相差越大, 选取的权重系数应该越小, 因此, 本文定义了如式 (3)~(6) 所示的权重系数:

$$\omega(x_i) = \omega_h(x_i)\omega_s(x_i)\omega_v(x_i) \quad (3)$$

$$\omega_h(x_i) = \left(1 - \frac{|h_i - h_0|}{\max(|h_k - h_0|)_{k=1,2,\dots,n}}\right) \quad (4)$$

$$\omega_s(x_i) = 1 - |s_i - s_0| \quad (5)$$

$$\omega_v(x_i) = \left(1 - \frac{|v_i - v_0|}{\max(|v_k - v_0|)_{k=1,2,\dots,n}}\right) \quad (6)$$

其中, 以被平滑点为中心, 边长为 $2r$ 的正方形区域内的所有像素点作为采样点, r 称为空域带宽. h_0, s_0, v_0 分别表示被平滑点在 HSV 颜色空间的三个分量, h_i, s_i, v_i 分别表示采样点在 HSV 颜色空间的三个分量, n 表示采样点的数量, $(|h_k - h_0|)_{k=1,2,\dots,n}$ 表示所有采样点与被平滑点的 H 分量差的绝对值, $(|h_k - h_0|)_{k=1,2,\dots,n}$ 表示所有采样点与被平滑点的 V 分量差的绝对值. 当采用过大的值域带宽时, 相应的高斯核函数半径增大, 图像平滑效果增强, 平滑图像与原始图像的颜色偏差增大, 从而导致图像模糊. 为了克服这一现象, 本文定义了采样点权重, 使得采样点颜色与被平滑点颜色相差越大时, 权重越小, 从而减小了平滑图像与原始图像的颜色偏差, 有效解决了过大值域带宽导致图像模糊的问题. 由于分别在 HSV 颜色空间和 Lab 颜色空间计算的权重系数所对应的平滑结果差别很小, 而 HSV 颜色空间比 Lab 颜色空间更符合人眼的感知特性, 因此本文

采用在 Lab 颜色空间进行图像平滑, 在 HSV 颜色空间计算采样点权重.

2.2 自适应空域带宽

在运用 Mean shift 算法对图像平滑时, 是采用一个边长为 $2r$ 的正方形窗遍历图像中的每一个像素, 因此, 采用不同的空域带宽, 参与运算的采样点则不同, 从而导致平滑及分割的结果也不同. 空域带宽越大, 参与运算的采样点越多, 被平滑的像素值距离原始像素值的偏离也越大, 从而会导致图像失真. 但如果空域带宽过小, 参与运算的采样点过少, 又达不到有效平滑的目的. 因此, 本文采用的方法如图 1 所示. 给定一个空域带宽初值 r_0 , 以步长 $step$ 逐渐增大空域带宽, 本文设 $r_0 = 5$, $step = 1$, 即 $r_i = r_{i-1} + 1$, 当满足停止条件 $\sum \sigma_j < 0.7 \times n_i$ ($j = 1, 2, \dots, n_i$), 即在空域带宽 r_i 内统计所有与被平滑点颜色相似的采样点数量, 当该数量小于全部采样点数量 n_i 的 70% 时, 对应的即为所求的空域带宽, 记为 r . 其中, 采样点颜色与被平滑点颜色相似的条件为 $|h_j - h_0| \leq 5$. h_j 为采样点颜色的 H 分量值, h_0 为被平滑点颜色的 H 分量值, σ_j 为符号函数, 满足条件 $|h_j - h_0| \leq 5$ 时, $\sigma_j = 1$. 反之, $\sigma_j = 0$. 以上的空域带宽的选取方法保证了在图像信息比较丰富, 即细节或边缘信息比较多时, 选用比较小的空域带宽. 而在图像细节或边缘信息不丰富, 图像比较平坦时, 选用比较大的空域带宽, 从而可以在一定程度上保证在平滑图像时尽可能不丢失细节信息.

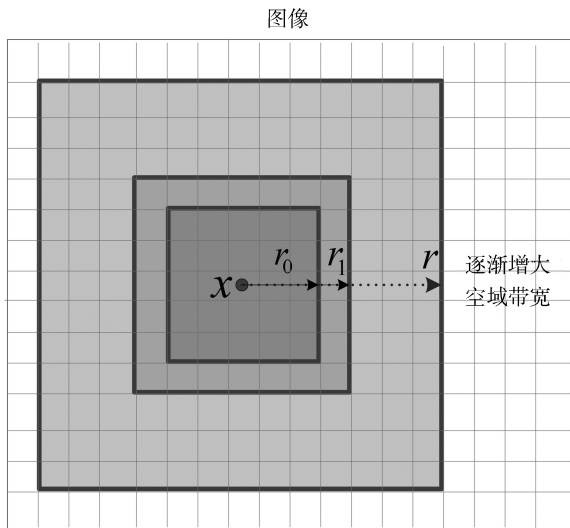


图 1 确定空域带宽示意图

Fig. 1 The illustration of calculating spatial bandwidth

2.3 自适应值域带宽

在运用式 (1) 对图像进行平滑时, 值域带宽是一个重要参数, 一般值域带宽的确定方法有两种: 全

局最优固定值域带宽和自适应值域带宽. 但无论哪种方法, 都是基于数据本身的, 而没有考虑到人眼的视觉特性, 即人对图像的感受. 因此, 本文将图像质量的客观评价方法引入到带宽的选择中, 使平滑图像更接近于人眼的主观感觉. 为了更接近主观评价, 本文根据 Wang 等^[15] 提出的结构相似度 ($ssim$) 并且参考李航等^[16] 提出的频域结构相似度, 给出了本文所采用的频域结构相似度函数:

$$FSSIM = [l(I_0, I_d)]^\alpha [c(I_0, I_d)]^\beta [s_F(I_0, I_d)]^\gamma \quad (7)$$

$$l(I_0, I_d) = \frac{2u_{I_0}u_{I_d} + C_1}{u_{I_0}^2 + u_{I_d}^2 + C_1} \quad (8)$$

$$c(I_0, I_d) = \frac{2\sigma_{I_0}\sigma_{I_d} + C_2}{\sigma_{I_0}^2 + \sigma_{I_d}^2 + C_2} \quad (9)$$

$$s_F(I_0, I_d) = \frac{\sigma_{F_0F_d} + C_3}{\sigma_{F_0}\sigma_{F_d} + C_3} \quad (10)$$

其中, I_0 和 I_d 分别为参考图像和失真图像, $l(I_0, I_d)$, $c(I_0, I_d)$ 和 $s_F(I_0, I_d)$ 分别为 I_0 和 I_d 的亮度函数、对比度函数和结构度函数. $\alpha > 0$, $\beta > 0$ 和 $\gamma > 0$ 分别为 $l(I_0, I_d)$, $c(I_0, I_d)$ 和 $s_F(I_0, I_d)$ 的权重, u_{I_0} , u_{I_d} 分别为 I_0 和 I_d 的均值, σ_{I_0} , σ_{I_d} 分别为 I_0 和 I_d 的标准差, $\sigma_{I_0I_d}$ 为 I_0 和 I_d 的协方差, C_1, C_2 和 C_3 是为了防止当分母接近零时产生不稳定现象而设置的常数. F_0, F_d 分别为 I_0 和 I_d 的离散余弦变换图像.

从理论上分析, 值域带宽越大, 图像平滑程度越高, 但过度的平滑会使图像出现模糊失真现象, 从而失去了平滑的意义. 因此, 平滑图像需要考虑两个方面的问题: 1) 图像越平滑越好, 可以用图像方差表征图像的平滑程度, 即方差越小, 图像平滑程度越好. 2) 图像平滑的同时应该符合人眼的视觉特点, 即保证平滑图像不失真, 可以运用平滑图像与原始图像的频域结构相似度表征图像的失真程度, 即频域结构相似度越小, 图像失真越严重. 理论上, 值域带宽越大, 图像平滑程度越高, 图像方差越小, 而平滑图像与原始图像的频域结构相似度就会越低. 因此, 在图像平滑的过程中, 应该使局部方差尽可能小而图像的频域结构相似度尽可能大. 本文联合考虑了这两方面的内容, 建立了如下两个目标函数:

$$F_1(w) = e(w) = \frac{\sum_{i=1}^n (x_i - \bar{x}(w))^2}{n} \quad (11)$$

$$F_2(w) = FSSIM(I_0, I_d) \quad (12)$$

式 (11) 中 $e(w)$ 表示采样点的方差, x_i 为采样点像素值, n 为采样点数量, $\bar{x}(w)$ 表示采样点均值, 如式 (13) 所示:

$$\bar{x}(w) = \sum_{i=1}^n x_i \quad (13)$$

理论上, 值域带宽越大, 图像平滑程度越高, 图像方差越小, 图像的方差从某种程度上可以表征值域带宽的选取情况. 式 (12) 中, $F_2(w)$ 表示空域带宽内平滑图像与原始图像的频域结构相似度, I_d 表示平滑图像, I_0 表示原始图像. 值域带宽越大, 图像平滑程度越高, 平滑图像与原始图像的颜色偏差越大, 对应的频域结构相似度就会越低. 因此, 在图像的平滑过程中, 应该使局部方差尽可能小而图像的频域结构相似度尽可能大, 通过最小化方差函数 $F_1(w)$ 和最大化频域结构相似度函数 $F_2(w)$, 得到一个折中的值域带宽, 此时, 既达到图像平滑的效果, 又能保证图像不失真. 本文进行最小化 $F_1(w)$ 和最大化 $F_2(w)$ 时所采用的具体方法是根据实验经验设定一个值域带宽的搜索范围 $[0.01W_x, 0.5W_x]$, 运用折半查找的方法寻找满足目标函数的值域带宽. 根据实验经验, W_x 取值为采样点与被平滑点差的平均值, 表示如下:

$$W(x) = \frac{\sum_{i=1}^n (x_i - x)}{n} \quad (14)$$

其中, x_i, x 和 n 的含义同上.

3 实验结果及分析

为了验证本文带宽分析方法的准确性和有效性, 随机选取了 378 幅图像进行实验, 算法采用 Matlab 语言编程实现. 计算频域结构相似度时所采用的参数是参考叶盛楠等^[17] 所采用的参数, 分别为: $\alpha = \beta = \gamma = 1, C_1 = 3, C_2 = 30, C_3 = 15$. 实验内容包括: 1) 采样点权重对平滑结果的影响; 2) 值域带宽与频域结构相似度关系分析; 3) 自适应 Mean shift 算法的图像平滑及分割结果; 4) 分割结果误差分析. 在实验 1 和实验 2 中, 采用大小为 100 像素 \times 140 像素的图像. 为了更清晰地显示实验结果, 其他实验均采用大小为 315 像素 \times 400 像素和 400 像素 \times 315 像素的图像.

1) 采样点权重对平滑结果的影响

本文定义采样点权重函数的基本思想是采样点的颜色与被平滑点颜色相差越大, 权重越小, 即采样点在运算中的重要性越小. 因此, 即使值域带宽选取得比较大, 也可以通过权重对采样点的重要性加以控制, 有效克服大的值域带宽导致图像模糊失真的缺点. 为此, 本文分别作了两组实验, 如图 2 所示, 其中, 图 2(a) 为原图, 图 2(b) 是采样点权重为 1, 即 $\omega(x_i) = 1$, 值域带宽为 $0.5W_x$ 时的平滑结果, 图 2(c) 是采用本文方法设置权重, 值域带宽为 $0.5W_x$

时的平滑结果.

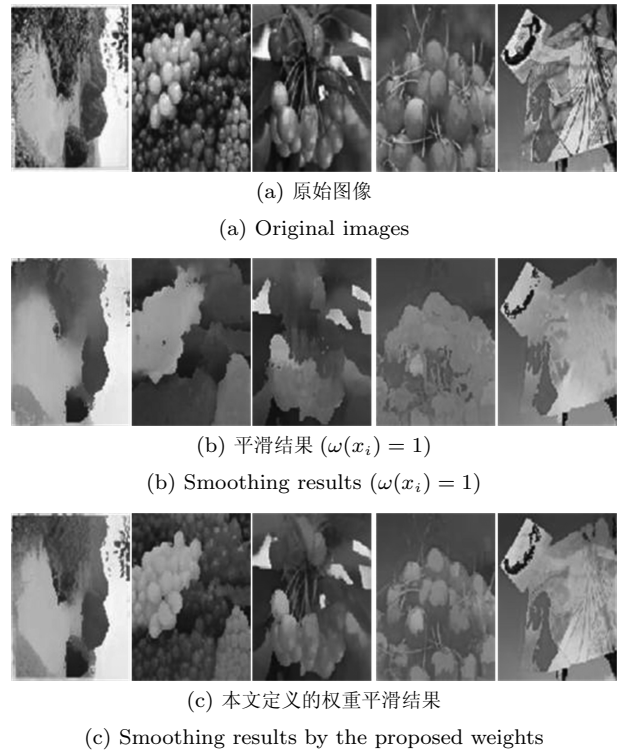


图 2 权重对平滑结果的影响

Fig. 2 The effect of weights on smoothing results

从图中可以看出, 不考虑权重, 即 $\omega(x_i) = 1$ 时, 当值域带宽比较大时, 容易造成过平滑, 导致图像模糊. 而采用本文设置权重的方法, 可以通过权重对采样点加以控制, 从而避免了图像严重失真的现象.

2) 值域带宽与频域结构相似度关系分析

从理论上分析, 值域带宽越大, 图像平滑程度越高, 平滑结果与参考图像的频域结构相似度就会越低, 本文将频域结构相似度与图像方差结合起来, 建立了选择值域带宽的目标函数, 从而使所选择带宽对应的平滑和分割图像更符合人眼的视觉特性, 为此, 需要实验验证以下两个方面的内容: a) 图像是否符合频域结构相似度随值域带宽增大而减小的变化规律; b) 当频域结构相似度下降到什么程度时, 平滑图像会严重失真, 从而导致分割错误. 首先, 我们将固定空域带宽设置为 24, 分别采用不同的值域带宽对图像进行了 Mean shift 平滑, 并且绘制了频域结构相似度随值域带宽的变化曲线, 随机选取其中 5 幅图像的频域结构相似度变化曲线, 如图 3 所示. 其中, 横坐标表示的是不同的值域带宽值, 纵坐标表示的是相应的频域结构相似度. 可以看出, 所实验的图像都符合频域结构相似度随值域带宽的增加而减小的规律.

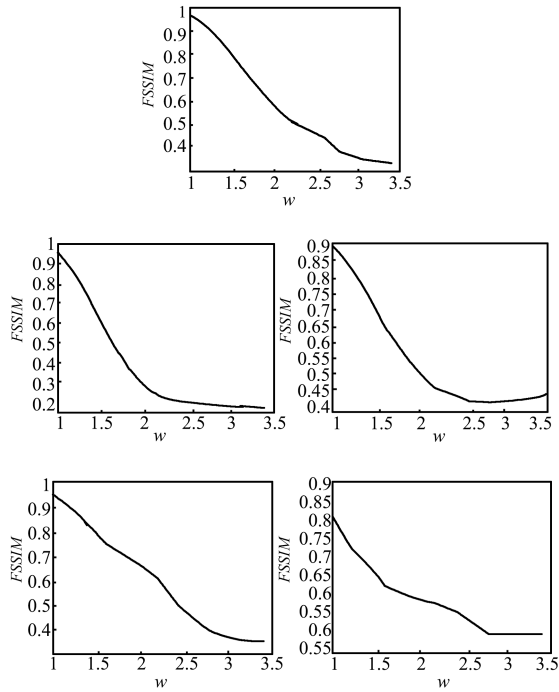


图3 频域结构相似度随值域带宽变化曲线
(图像大小为 100×140 , 单位: 像素, 空域带宽为 24)
Fig.3 The changing curves of *FSSIM* and range bandwidth (image size: 100 pixels \times 140 pixels, range bandwidth: 24)

其次, 为了分析不同频域结构相似度对应的平滑和分割结果, 本文选择了值域带宽分别为 $0.05W_x$, $0.13W_x$, $0.29W_x$, 对应的频域结构相似度分别为 0.93, 0.75, 0.65 时的平滑和分割结果, 如图 4 所示。从图中可以看出, 当选取较小的值域带宽 $0.05W_x$, 其频域结构相似度对应为 0.93 时, 图像平滑效果不明显, 分割结果容易出现过分割现象。当增大值域带宽 ($0.13W_x$), 相应的频域结构相似度降低为 0.75 时, 平滑效果增强, 可以得到相对较好的分割结果。当值域带宽继续增加为 $0.29W_x$ 时, 相应的频域结构相似度下降为 0.65, 此时, 平滑结果存在一定程度的模糊, 从而导致分割结果局部不够准确。根据 378 幅平滑图像的实验结果分析, 无论是平滑还是分割, 所选择的值域带宽所对应的频域结构相似度都不应该低于 0.6。

3) 自适应 Mean shift 算法的图像平滑及分割结果

为了验证本文自适应 Mean shift 算法的有效性, 本文实现了 plug-in^[18] 规则下最优带宽 Mean shift 算法的图像平滑与分割, 并且将之与本文的自适应 Mean shift 算法的平滑和分割结果进行了对比分析, 如图 5 所示。其中, 图 5(a) 是原始图像, 图 5(b) 和 5(c) 分别是 plug-in 规则下 Mean shift 算法的平滑结果和分割结果, 图 5(d) 和 5(e) 分别是本

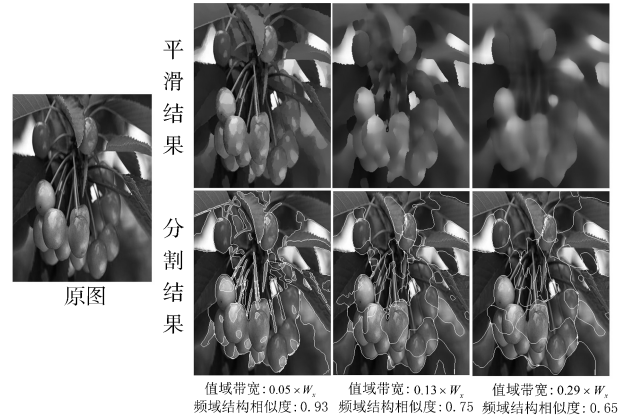
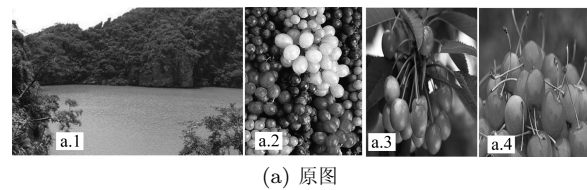
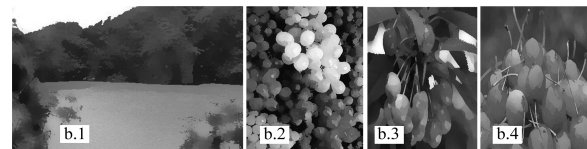


图4 不同频域结构相似度对应的平滑及分割结果
Fig.4 Smoothing results for different *FSSIM*

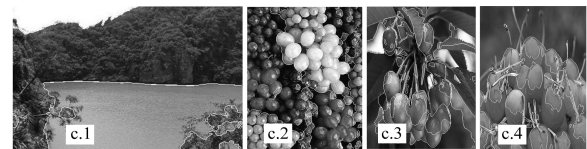


(a) Original images



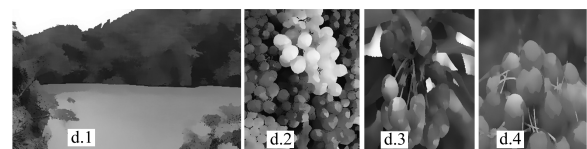
(b) plug-in 规则下的 Mean shift 算法平滑结果

(b) Smoothing results of mean shift algorithm with plug-in rule



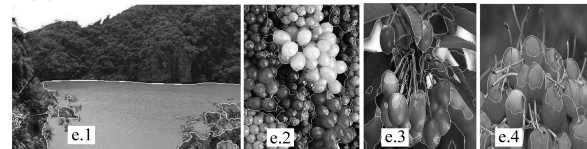
(c) plug-in 规则下的 Mean shift 算法分割结果

(c) Segmenting results of mean shift algorithm with plug-in rule



(d) 自适应 Mean shift 算法的平滑结果

(d) Smoothing results of adaptive mean shift algorithm



(e) 自适应 Mean shift 算法的分割结果

(e) Segmenting results of adaptive mean shift algorithm

图5 自适应 Mean shift 算法的平滑及分割结果
Fig.5 Smoothing and segmenting results of adaptive mean shift algorithm

文算法的平滑结果和分割结果. 采用 plug-in 规则的基本思想是通过折中选取估计的偏差和方差, 使估计的均方误差最小, 对应的最优带宽是在最小均方误差意义下的最优, 而非主观观察下的最优. 因此, plug-in 规则下的带宽对应的平滑结果容易出现欠平滑的现象, 对应的分割结果容易出现过分割. 本文将带宽的选择同图像颜色信息、图像的局部方差以及评价图像质量的频域结构相似度有机结合起来, 既可以有效平滑图像, 又可以获得符合人眼视觉特点的平滑和分割结果. 对比图 5(b) 和 5(d), 可以明显看出, 图 5(b) 对应的 plug-in 规则下的 Mean shift 算法的平滑效果明显弱于图 5(d) 所示的自适应 Mean shift 算法的平滑效果. 相应的图 5(c) 所示的 plug-in 规则下 Mean shift 算法的分割结果中, 局部出现过分割现象, 而图 5(e) 所示的自适应 Mean shift 算法的分割结果相对而言, 更符合人眼的视觉特性. 如图 5(a.3) 的水果部分, plug-in 规则下 Mean shift 算法的分割结果 (图 5(c.3)) 出现了分割过细的现象, 而本文自适应 Mean shift 算法分割结果 (图 5(e.3)) 则将水果作为整体分割出来.

为了进一步分析本文提出的自适应带宽与 plug-in 规则下带宽的不同, 在图 6 中比较了以 A 为左上顶点的 3×3 邻域内的自适应带宽数值和 plug-in 规则下带宽数值. 可以看出, 本文计算的带宽略大于运用 plug-in 规则计算的带宽, 但带宽变化趋势并没有改变, 因此, 本文自适应 Mean shift 算法不容易出现欠平滑和过分割的现象.

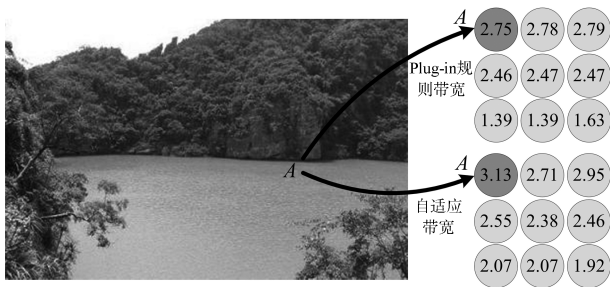


图 6 plug-in 规则带宽与自适应带宽
Fig. 6 Plug-in bandwidths and adaptive bandwidths

4) 分割结果误差分析

在分割的过程中, 不同颜色区域边界肯定会存在颜色重叠和颜色交叉的现象. 为了分析颜色重叠和颜色交叉所造成的分割误差, 本文以人工分割结果作为标准图像, 计算自适应 Mean shift 算法的分割结果与标准图像误差, 如图 7 所示. 在图 7 中, 图 7(a) 为原始图像, 图 7(b) 和 7(c) 分别为手工分割某一区域的结果及其二值图像, 图 7(d) 和 7(e) 分别为本文算法分割某一区域的结果及其二值图像, 图 7(f) 为误差二值图像, 即本文算法分割结果图

7(e) 与标准图像图 7(c) 的不重叠区域. 计算分割误差的公式如下:

$$err = \frac{num(I_e)}{num(I_s)} \tag{15}$$

其中, $num(I_e)$ 表示如图 7(f) 所示的分割误差区域的总像素数, $num(I_s)$ 为标准分割图像的总像素数.

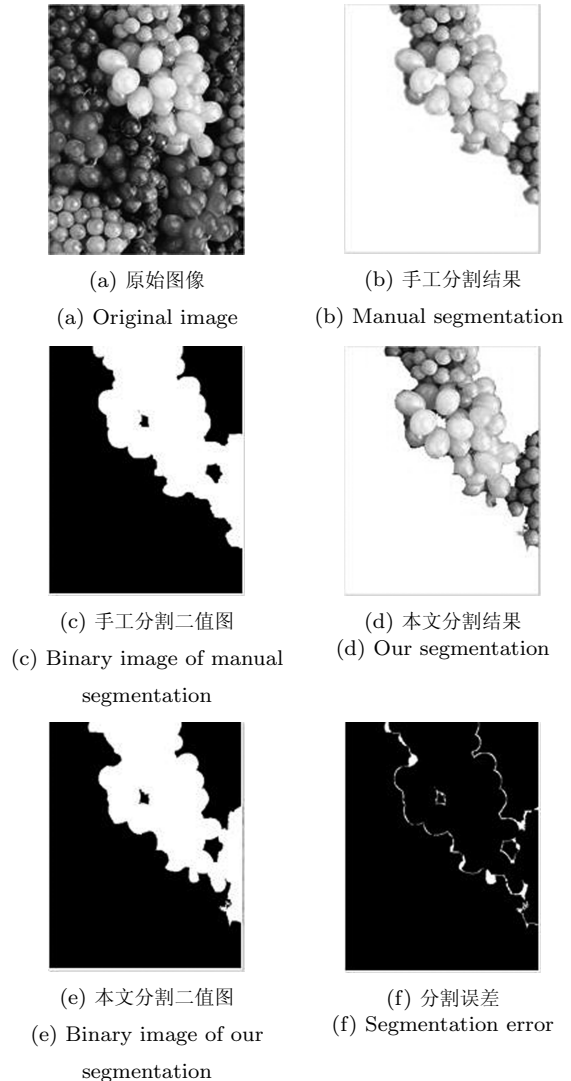
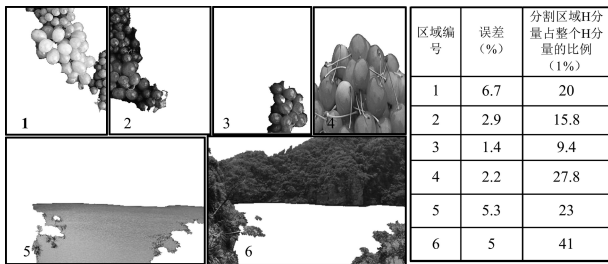


图 7 自适应 Mean shift 算法分割误差
Fig. 7 The errors of adaptive mean shift algorithm

根据式 (15), 本文对多幅图像进行了误差计算, 并且计算了所分割区域在整个颜色空间的比例, 将部分结果显示于图 8. 在图 8 中, 1~6 号图像是运用本文算法分割的某一区域图像, 右边表格显示的是对应区域的分割误差和所占颜色空间比例, 其中, 颜色空间比例是指分割区域的 H 分量占整个 H 分量的比例. 可以看出, 所有图像分割误差均小于 7%, 实际上, 此误差并不能绝对说明分割效果的好

坏, 如图 7 所示, 分割结果与标准图像并没有发生大的改变或错误分割。



区域编号	误差 (%)	分割区域中分量占整个H分量的比例 (%)
1	6.7	20
2	2.9	15.8
3	1.4	9.4
4	2.2	27.8
5	5.3	23
6	5	41

图 8 分割结果误差

Fig. 8 Segmentation errors

为了说明本文算法对所分割区域占视场比例的适应性, 本文计算了分割结果中部分区域所占视场的比例, 如图 9 所示, 可以看出, 即使目标占视场比例为 0.22% 时, 也可以得到正确的分割结果。但通过实验验证, 并不是所有小目标都可以进行正确分割, 对于部分边界模糊的小目标, 会出现漏分割的情况, 如图 9 中矩形框内的区域。实际上, 很难找到一个确定的比例和颜色差值来界定不能正确分割的小目标范围。

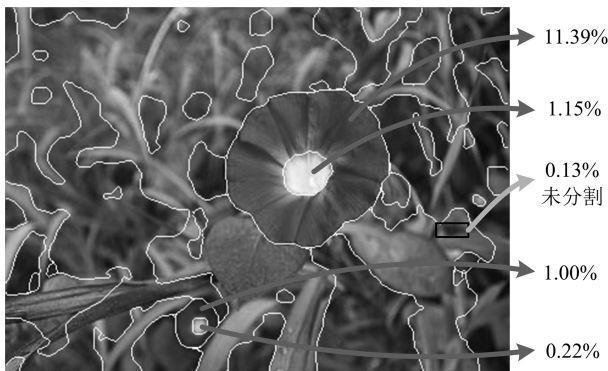


图 9 不同比例目标分割结果

Fig. 9 Object segmentation of different scales

4 结论

本文对 Mean shift 算法用于平滑和分割图像时做了以下三个方面的分析与改进: 1) 根据图像的颜色信息, 设置了采样点权重, 从而克服了大的值域带宽导致图像模糊失真的现象; 2) 根据图像的颜色分布信息自适应选取了空域带宽, 在图像信息丰富的区域采用较小的空域带宽, 而在图像信息相对不丰富的区域采用较大的空域带宽; 3) 将评价图像质量的客观标准即频域结构相似度, 与图像方差结合起来, 建立了选择值域带宽的目标函数, 从而使所选择带宽对应的平滑和分割图像更符合人眼的视觉特性。本文中自适应带宽的选择方法克服了人为设定带宽

方法不能自动化平滑和分割图像的缺点。同时, 将带宽的选择同图像颜色信息、评价图像质量的频域结构相似度有机结合起来, 从而得到符合人眼感知特性的平滑和分割结果。本文的主要工作是运用改进的自适应 Mean shift 算法将彩色图像分割成不同区域, 这些区域不会被分割得过细或过粗, 从而能够正确表达待检测目标的整体性。作者下一步的工作重点是基于这些区域的特征和区域之间的关系实现目标检测任务。

References

- Han H, Sohn K. Automatic illumination and color compensation using mean shift and sigma filter. *IEEE Transactions on Consumer Electronics*, 2009, **55**(3): 978–986
- Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 603–619
- Tao W B, Jin H, Zhang Y M. Color image segmentation based on mean shift and normalized cuts. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2007, **37**(5): 1382–1389
- Park A, Kim J, Min S, Yun S, Jung K. Graph cuts-based automatic color image segmentation using mean shift analysis. In: *Proceedings of the Conference on Digital Image Computing: Techniques and Applications*. Canberra, Australia: IEEE, 2008. 564–571
- Li P F, Wang S R, Jing J F. The segmentation in textile printing image based on mean shift. In: *Proceedings of the 10th Conference on Computer-Aided Industrial Design and Conceptual Design*. Wenzhou, China: IEEE, 2009. 1528–1532
- Wang Z Y, Boesch R, Waser L, Ginzler C. Mean shift segmentation applied to ADS40 data for automatic forest detection. In: *Proceedings of the International Conference on Information and Automation*. Zhuhai, China: IEEE, 2009. 1099–1103
- Hong Y P, Yi J Q, Zhao D B. Improved mean shift segmentation approach for natural images. *Applied Mathematics and Computation*, 2007, **185**(2): 940–952
- Gondra I, Xu T. Adaptive mean shift-based image segmentation using multiple instance learning. In: *Proceedings of the 3rd International Conference on Digital Information Management*. London, UK: IEEE, 2008. 716–721
- Zheng L Y, Zhang J T, Wang Q Y. Mean-shift-based color segmentation of images containing green vegetation. *Computers and Electronics in Agriculture*, 2009, **65**(1): 93–98
- Comaniciu D, Ramesh V, Meer P. The variable bandwidth mean shift and data-drive scale selection. In: *Proceedings of the 8th International Conference on Computer Vision*. Vancouver, Canada: IEEE, 2001. 438–445

- 11 Comaniciu D. An algorithm for data-drive bandwidth selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(2): 281–288
- 12 Jiménez-Alaniz J R, Pohl-Alfaro M, Medina-Banuelos V, Yanez-Suarez O. Segmenting brain MRI using adaptive mean shift. In: Proceedings of the 28th IEEE Conference on Engineering in Medicine and Biology Society. New York, USA: IEEE, 2006. 3114–3117
- 13 Mayer A, Greenspan H. Segmentation of brain MRI by adaptive mean shift. In: Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro. Arlington, USA: IEEE, 2006. 319–322
- 14 Mayer A, Greenspan H. An adaptive mean-shift framework for MRI brain segmentation. *IEEE Transactions on Medical Imaging*, 2009, **28**(8): 1238–1250
- 15 Wang Z, Bovik A C, Sheikh H R, Simoncelli E P. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, **13**(4): 600–612
- 16 Li Hang, Lu Yang, Cui Hui-Juan, Tang Kun. Image quality assessment based on frequency domain based structural similarities. *Journal of Tsinghua University (Science and Technology)*, 2009, **49**(4): 559–562
(李航, 路羊, 崔慧娟, 唐昆. 基于频域的结构相似度的图像质量评价方法. 清华大学学报(自然科学版), 2009, **49**(4): 559–562)
- 17 Ye Sheng-Nan, Su Kai-Na, Xiao Chuang-Bo, Duan Juan. Image quality assessment based on structural information extraction. *Acta Electronics Sinica*, 2008, **36**(5): 856–861
(叶盛楠, 苏开娜, 肖创柏, 段娟. 基于结构信息提取得图像质量评价. 电子学报, 2008, **36**(5): 856–861)
- 18 Hardle W, Muller M, Sperlich S, Werwatz A. Nonparametric and semi parametric models [Online], available: <http://fedc.wiwi.hu-berlin.de/xplere/ebooks/html/spm/>, February 4, 2010



王 晏 大连理工大学博士研究生. 1994 年获北方工业大学学士学位. 主要研究方向为图像处理和模式识别. E-mail: yanzi_7656@yahoo.com.cn (WANG Yan Ph.D. candidate at Dalian University of Technology. She received her bachelor degree from North China University of Technology in 1994.

Her research interest covers image processing and pattern recognition.)



孙 怡 大连理工大学教授. 分别于 1989 和 2002 年获得大连理工大学学士学位和博士学位. 主要研究方向为成像与图像处理, 移动通信技术. 本文通信作者. E-mail: lslwf@dlut.edu.cn (SUN Yi Professor at Dalian University of Technology. She received her bachelor and Ph.D. degrees from

Dalian University of Technology in 1989 and 2002, respectively. Her research interest covers image processing, pattern recognition, and mobile communication. Corresponding author of this paper.)