

基于KFCM和改进分水岭算法的猪肉背最长肌分割技术*

伍学千 廖宜涛 樊玉霞 成芳

(浙江大学生物系统工程与食品科学学院, 杭州 310029)

【摘要】 提出一种利用核模糊C均值聚类(KFCM)和改进分水岭算法分割猪肉眼肌切面图像中背最长肌区域的方法。该算法对经中值滤波去噪后图像的R分量利用最大方差自适应阈值(OTSU)去除背景,再采用KFCM提取出肌肉组织,然后进行空洞填充,最后由改进的分水岭算法分割出背最长肌区域。利用该算法对采集的60幅猪肉眼肌图像进行处理,分割正确率为86.67%;与传统的形态学算法相比,该算法能真实、完整地恢复出背最长肌区域。结果表明:该算法能有效地分割出猪肉眼肌图像中的背最长肌区域,与改进前分水岭算法相比,能避免背最长肌区域出现欠分割。

关键词: 无损检测 图像分割 猪肉 背最长肌 核模糊C均值聚类 分水岭算法

中图分类号: TP391.41; TS251 **文献标识码:** A **文章编号:** 1000-1298(2010)01-0172-05

Segmentation of Pork Longissimus Dorsi Based on KFCM Clustering and Improved Watershed Algorithm

Wu Xueqian Liao Yitao Fan Yuxia Cheng Fang

(College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310029, China)

Abstract

A method for automatic segmentation of pork longissimus dorsi muscle (MLD) region from rib-eye image was developed using KFCM clustering and improved watershed algorithm. Firstly, median filter and OTSU were used to remove noise and background. And then, kernel fuzzy C-means clustering (KFCM) was applied to remove fat pixels. Finally, hole filling operation and improved watershed algorithm were employed to segment the area of MLD. Sixty samples were used to test the performance of the proposed method. The success rate of segmentation was 86.67%. Compared with the traditional morphology and watershed methods, it is proved that the developed method could segment MLD region perfectly and avoid under-segmentation effectively.

Key words Nondestructive detection, Image segmentation, Pork, Longissimus dorsi, Kernel fuzzy C-means clustering, Watershed algorithm

引言

利用计算机视觉技术对猪肉品质进行评价时,猪肉第1~4肋骨中段的眼肌切面中背最长肌区域的面积、颜色和大理石纹分布等特征是评价的重要指标。提取出眼肌图像中的背最长肌区域,是实现基于计算机视觉技术的猪肉无损检测的关键步骤之

一。国内外对背最长肌区域分割的研究主要集中在牛肉上。McDonald和Chen^[1-2]利用数学形态学方法进行了牛肉背最长肌区域提取;Gerrard等^[3]以快速结构标记算法分析出的背最长肌结构特征作为形态学运算的结构元素提取出了背最长肌;赵杰文等^[4]先采用固定的圆形结构元素去掉细小连接部分,然后通过一个不断扩大的圆结构元素进行腐蚀

操作,直至图像中只有一个区域为止,用最后一次腐蚀的结构元素对腐蚀结果进行膨胀操作从而提取出背最长肌;Subbiah^[5]根据背最长肌的凸形特征,利用凸包生成算法生成凸包,判断凸包中目标区域的面积比率,通过反复迭代运算直到满足设定条件来提取出背最长肌区域;于铂等^[6]采用基于颜色和纹理特征的区域增长算法对猪肉背最长肌进行了分割;赵杰文^[7]等根据牛胴体眼肌切面图像的颜色特征,依次在 Ohta 和 RGB 颜色空间下对眼肌切面图像进行阈值分割,然后进行填充、腐蚀、面积阈值分割和膨胀等运算来分割出背最长肌区域。这些算法都能对背最长肌区域进行分割,但存在结构元素复杂、时间复杂度高、粘连严重时难以正确分割等制约因素;而且对于猪肉,其背最长肌与周围相连接的肌肉组织颜色差别更小,以上方法不适合,需要研究一种分割猪肉背最长肌区域的新算法。

本文提出一种结合核模糊 C 均值聚类和改进分水岭算法来提取猪肉图像中背最长肌区域的新方法,即先通过核模糊 C 均值聚类算法对肌肉和脂肪进行初分割,再利用改进的分水岭算法去除背最长肌与周围肌肉之间的连接提取出背最长肌。

1 试验材料与装置

试样为不同猪胴体第 1~4 根肋骨中段的眼肌肉,肉样厚度约为 30 mm,共 60 个,分 10 d 购买,购买后置于 22℃ 实验室环境下试验。试验装置如图 1 所示,图像由工业数码相机(MVC3000-B 型,北京微视图像公司)经过 12 mm 光学镜头(Pentax)进行采集,相机分辨率为 1 280 × 1 024,以 24 bits 的位图格式存储。试验显微镜镜头固定焦距,调节光圈固定照明强度参数,保证图像的一致性;所有图像处理算法在 Visual studio 2005 上实现。

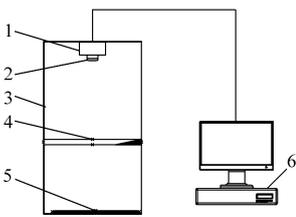


图 1 计算机视觉检测系统示意图

Fig.1 Sketch map of the computer vision detection system

2 图像处理方法

利用图 1 所示的试验装置采集的样本图像如图 2 所示。

2.1 预处理

受到采集装置、环境等因素的影响,采集的猪肉眼肌切面图像需进行预处理。中值滤波能有效地去

除噪声且较好地保护图像细节^[8];猪肉彩色图像中背景为黑色,脂肪为白色,肌肉为红色, R 、 G 、 B 直方图中 R 直方图出现双峰现象,采用最大方差自适应阈值法^[9](OTSU)在 R 分量上进行阈值分割可以有效地去除背景。

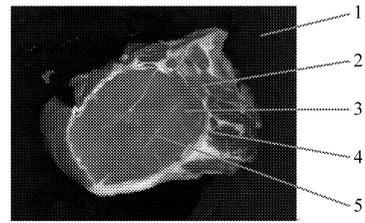


图 2 采集的猪肉眼肌切面图像

Fig.2 Collected image of pork rib-eye

1. 背景区域
2. 周围肌肉组织
3. 背最长肌区域
4. 肌间脂肪
5. 肌内脂肪

2.2 核模糊 C 均值聚类算法

预处理后的猪肉图像各分量直方图没有明显的双峰现象,用常归的阈值分割算法很难找到合适的阈值,基于模糊理论的核模糊 C 均值聚类算法(KFCM)可以有效地提高图像分割的精度。KFCM 的基本思想是利用一个非线性映射 Φ 将输入空间 $X = \{x_1, x_2, \dots, x_n\}$ 映射到特征空间 F ,然后在 F 中进行聚类^[10-12],其目标函数为

$$J_{\text{KFCM}}(U, V) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m (K(x_k, x_k) - 2K(x_k, v_i) + K(v_i, v_i)) \quad (1)$$

式中 U —— $c \times n$ 的隶属度矩阵

V ——聚类中心矩阵

m ——权重系数

$K(x, y)$ ——核函数

实现 KFCM 算法的具体步骤如下^[10-13]:

(1) 初始化聚类数目 c 、权重系数 m 和允许误差 E_{\max} , 总迭代循环次数 l (l 从 1 开始)。

(2) 用 $[0, 1]$ 上的随机数初始化隶属矩阵 U , 使其满足 $\sum_{i=1}^c u_{ik} = 1$ 。

(3) $l = l + 1; i = 1, 2, \dots, c; k = 1, 2, \dots, n$ 。聚类中心计算公式为

$$v_i = \frac{\sum_{k=1}^n u_{ik}^m K(x_k, v_i) x_k}{\sum_{k=1}^n u_{ik}^m K(x_k, v_i)} \quad (2)$$

(4) 根据式(1)计算相邻两次目标函数的改变量 $e = |J^{(l)} - J^{(l+1)}|$, 如果 $e < E_{\max}$, 则退出, 否则根据式(3)更新隶属度矩阵并转到步骤(3)。

$$u_{ik} = \frac{[1/(K(x_k, x_k) + K(c_i, c_i) - 2K(x_k, c_i))]^{1/(m-1)}}{\sum_{j=1}^c [1/(K(x_k, x_k) + K(c_j, c_j) - 2K(x_k, c_j))]^{1/(m-1)}} \quad (3)$$

KFCM算法中聚类数目 c 、权重系数 m 、核函数 $K(x, y)$ 是影响分割效果的关键参数,其中 m 是控制算法柔性的参数, m 过大则聚类效果会很差, m 过小算法会接近硬聚类算法,这些参数的选择需根据实际试验效果确定。

2.3 改进的分水岭算法

分水岭算法常被用来分割图像中相粘连区域,其最简单的实现方法是采用数学形态学中的腐蚀、膨胀运算,即先用极限腐蚀标记所有区域,然后用条件膨胀从极限腐蚀得到的核心区域膨胀恢复得到无粘连的区域^[14-15]。

传统的极限腐蚀是保留每个连通区域消失之前的最后一步作为其“种子”,但存在缺点:当相粘连区域面积相差很大时由于极限腐蚀得到的种子区域面积都较小,经条件膨胀后大面积区域出现欠分割、小面积区域出现过分割。为解决上述问题,提出了

改进的极限腐蚀方法,即在迭代腐蚀 N 次之后判断图像中剩余连通区域的个数,如果为 1 则停止腐蚀。

设原始输入的二值图像为 I ,第 k 次腐蚀图像为 O_k , O_k 中连通区域个数为 N_k ,腐蚀结果元素为 B ,则改进的极限腐蚀算法步骤如下:①初始化: $k = 0$, $O_k = I$ 。②)迭代腐蚀: $k = k + 1$, $O_k = O_{k-1} \ominus B$ 。③)如果 $N_k = 1$ 且 $k \geq N$,则退出,否则继续往下执行。④)如果 O_{k-1} 中某个连通区域在 O_k 中不存在,则在 O_k 中恢复该连通区域,否则转到步骤②。

经上述极限腐蚀后将每个区域分离成种子小区域之后再反复使用条件膨胀即可恢复得到无粘连的原始区域。

结构元素 B 和迭代腐蚀终止次数 N 是正确分离粘连区域的关键参数,需要根据实际试验效果确定。

3 图像处理结果与分析

3.1 基于 KFCM 和改进分水岭的背最长肌分割

采用上文介绍的 KFCM 和改进分水岭算法对猪肉图像中背最长肌区域进行分割的效果如图 3 所示。

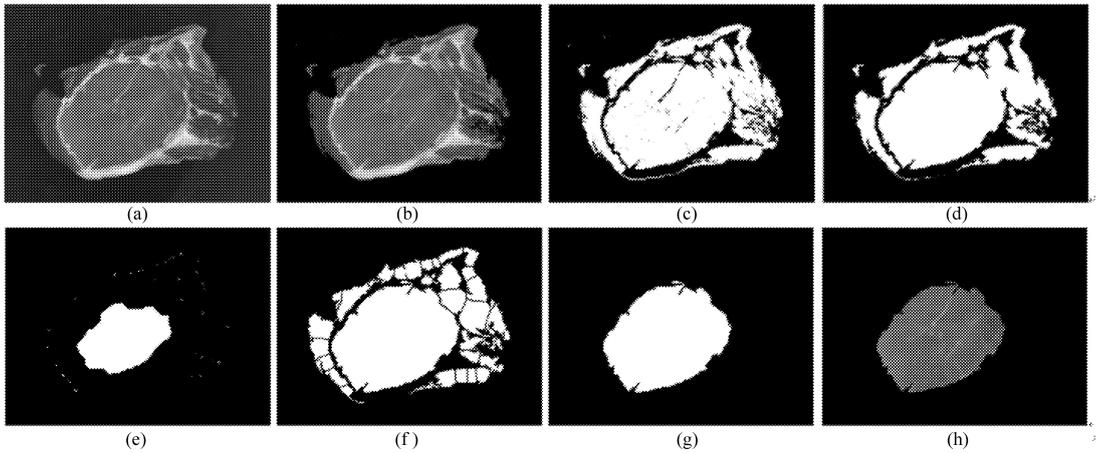


图 3 猪肉图像中背最长肌分割过程

Fig. 3 Segment process of longissimus dorsi in pork image

(a) 原图像 (b) 预处理后 (c) KFCM 分割 (d) 空洞填充 (e) 改进极限腐蚀后 (f) 条件膨胀后
(g) 提取最大区域 (h) (a)与(g)作“与”运算

处理步骤如下:

(1)预处理。对采集的原始图像(图 3a),采用中值滤波和 OTSU 去背景预处理后得到结果如图 3b所示,从图中可以看出原彩色图像的噪声和背景被有效去除。

(2)利用 KFCM 分割对预处理后的图像进行脂肪与肌肉像素分割。以数据矩阵 $X = \{x_1, x_2, \dots, x_n\} = \{[R_1, G_1, B_1], [R_2, G_2, B_2], \dots, [R_n, G_n, B_n]\}$ 作为原始输入,预处理后的图像包括肌肉和脂肪两类,聚类数目 $c = 2$;权重系数 m 从 $\{2, 3, \dots, 10\}$ 中进行选择,处理结果表明 $m = 8$ 效果最佳;然后分别采

用线性核、多项式核和高斯核 3 种核函数进行分割效果对比试验,以确定最优核函数。各核函数定义如下^[12]:

$$\text{线性核 } K(x, y) = xy$$

$$\text{多项式核 } K(x, y) = (xy + 1)^q$$

$$\text{高斯核 } K(x, y) = \exp(-\|x - y\|^2 / (2\sigma^2))$$

其中多项式核中的 $q = 3$,高斯核的 $\sigma = 2^8$,允许误差 $E_{\max} = 10^{-5}$,试验得到的分割效果如图 4 所示。

以背最长肌内脂肪像素提取的好坏作为评价标准,由图 4 可知,多项式核的分割效果最差,线性核其次,高斯核最好,因此本文采用高斯核进行肌肉

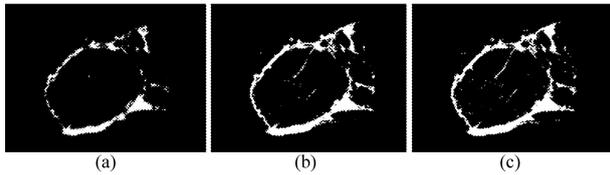


图 4 不同核函数分割结果

Fig. 4 Segmentation results with different kernel functions

(a) 多项式核分割 (b) 线性核分割 (c) 高斯核分割

与脂肪的分割。

(3)改进分水岭算法去除粘连。经 KFCM 分割后背最长肌内脂肪形成黑色空洞,利用空洞填充操作处理后的结果如图 3d 所示,可以看出背最长肌与周围肌肉组织出现粘连,采用改进分水岭算法进行粘连区域分割。选用 5×5 的圆形结构元素,对 N 取不同值进行分割试验,结果如表 1 所示。

表 1 不同迭代腐蚀次数的背最长肌分割结果

Tab. 1 Segmentation results of longissimus dorsi with different number of iterative erosion

N	8	9	10~25	26	27
分割正确率/%	76.67	83.33	86.67	81.67	75.00

由表 1 可知, $10 \leq N \leq 25$ 时分割性能最优,分割正确率为 86.67%, $N > 25$ 或 $N < 10$ 分割性能逐渐减弱。因为 N 取值过小时,对于粘连程度宽度较大的情况,相粘连区域极限腐蚀后还未来得及分离,导致背最长肌出现过分割;而 N 取值过大时,会与改进前分水岭算法一样背最长肌出现欠分割。 $N = 10$ 的处理结果如图 3e 和 3f,可以看出背最长肌与周围肌肉组织被较好地分离;最后找出图 3f 中面积最大的区域,再与原图 3a 进行与运算即可得到背最长肌区域,结果分别见图 3g 和 3h 所示。

出现错误分割的 8 幅图像,其中有 5 幅图是因为在取样时厚度不均而引起的背最长肌区域表面光照不均匀,或猪肉表面渗水情况严重而产生反光现象,致使一部分背最长肌像素点表现出与脂肪像素点接近的颜色特征信息,导致 KFCM 聚类时出现误判;另外 3 幅图像的背最长肌区域与周围肌肉粘连程度过大,如采集的图 5a 经 KFCM 处理后得图 5b,改进的分水岭算法无法实现正确分割。

3.2 传统方法的背最长肌分割

利用传统形态学方法进行分割:先进行有限次腐蚀使图中只存在一个区域,结果见图 6a;然后再进行相同次数的膨胀操作,结果见图 6b,可以看出提取到的区域并不是真实、完整的背最长肌区域,其原因是形态学开运算并不是可逆的,开运算处理之后只有那些在附近存在完整结构元素的像素点才会被保留,其他像素点都会被消除,从而处理得到的图

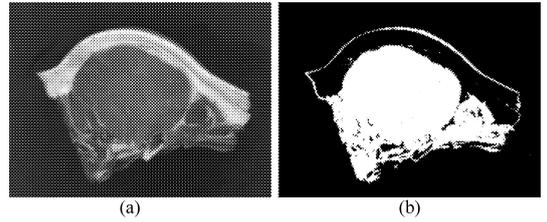


图 5 粘连程度过大时分割结果

Fig. 5 Segmentation result of the large degree connection

(a) 原图像 (b) KFCM 分割

像相对于原图像是有损的,因此经过相同腐蚀次数的膨胀运算之后并不能恢复出完整的原始区域。

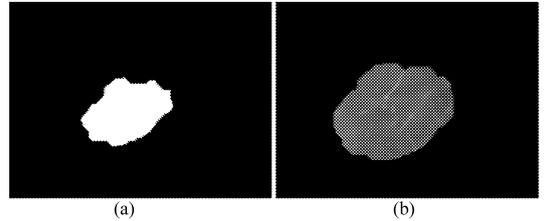


图 6 传统形态学方法分割结果

Fig. 6 Segmentation results by using traditional morphological method

(a) 腐蚀到只剩下一个区域 (b) 进行相同次数的膨胀

而分水岭算法并不是简单的保留最大目标区域,然后进行膨胀,而是先进行极限腐蚀,并保留区域中所有区域的种子区域,再以所有种子区域为基点进行有限性的膨胀。但使用传统的分水岭算法对相粘连区域面积相差很大的情况进行处理时容易出现欠分割。如图 7a 和 7b 所示,背最长肌区域出现欠分割,其原因是:背最长肌区域面积较相粘连的肌肉区域大很多,在使用传统的极限腐蚀方法处理后每个区域都收缩成一个点或无内点的小区域,条件膨胀时背最长肌与周围相粘连肌肉的种子区域在相同的结构元素作用下每次向外扩张的范围也基本一致,从而导致得到的分水线出现在背最长肌内部。

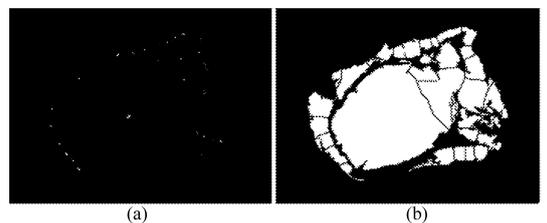


图 7 改进之前分水岭算法分割结果

Fig. 7 Segmentation results by using original watershed algorithm

(a) 改进前极限腐蚀后 (b) 条件膨胀后

4 结论

(1)在利用核模糊 C 均值聚类算法对肌肉和脂肪像素进行分割时,选择参数 $\sigma = 2^8$ 的高斯核函数、权重系数 $m = 8$ 进行聚类可以得到较好的分割效

果。

(2)将传统分水岭算法中的极限腐蚀操作迭代次数限定为10可以有效地解决当粘连区域面积相差很大时出现的大区域欠分割、小区域过分割的问题;同时在腐蚀和膨胀等形态学算法中选用 5×5 圆形结构元素,可以减少在形态学运算中造成图像失

真,保持原始目标的边缘信息且使经分水岭得到的分水线较平滑,较为真实地恢复出目标区域。

(3)迭代腐蚀终止次数 N 的取值是改进分水岭算法的一个至关重要的参数, N 值过小则在处理粘连宽度较大的情况时容易出现过分割, N 值过大则容易出现欠分割。

参 考 文 献

- McDonald T P, Chen Y R. Separating connected muscle tissue in images of beef carcass ribeyes[J]. Transactions of the ASAE, 1990, 33(6): 2 059 ~ 2 065.
- McDonald T P, Chen Y R. Visual characterization of marling in beef ribeyes and its relationship to taste parameters[J]. Transactions of the ASAE, 1991, 34(6): 2 499 ~ 2 504.
- Gerrard D E, Gao X, Tan J. Beef marling and color score determination by image processing[J]. Journal of Food Science, 1996, 61(1): 145 ~ 148.
- 赵杰文, 刘木华, 张海东. 基于数学形态学的牛肉图像中背长肌分割和大理石纹提取技术研究[J]. 农业工程学报, 2004, 20(1): 143 ~ 146.
Zhao Jiewen, Liu Muhua, Zhang Haidong. Segmentation of longissimus dorsi and marbling in ribeye imaging based on mathematical morphology[J]. Transactions of the Chinese Society of Agricultural Engineering, 2004, 20(1): 143 ~ 146. (in Chinese)
- Subbiah J, Ray N, Kranzler G A, et al. Computer vision segmentation of the longissimus dorsi for beef quality grading[J]. Transactions of the ASAE, 2004, 47(4): 1 261 ~ 1 268.
- 于铂, 郑丽敏, 田丽军. 基于颜色和纹理特征提取彩色图像的有意义区域[J]. 计算机工程, 2006, 32(3): 206 ~ 211.
Yu Bo, Zheng Limin, Tian Lijun. Image segmentation for meaningful region based on color and texture[J]. Computer Engineering, 2006, 32(3): 206 ~ 211. (in Chinese)
- 赵杰文, 方明, 刘木华, 等. 基于Ohta和RGB颜色空间牛胴体背最长肌的分割[J]. 江苏大学学报: 自然科学版, 2006, 27(3): 189 ~ 192.
Zhao Jiewen, Fang Ming, Liu Muhua, et al. Segmentation of longissimus dorsi in beef carcass based on Ohta and RGB color systems[J]. Journal of Jiangsu University: Natural Science Edition, 2006, 27(3): 189 ~ 192. (in Chinese)
- 贾渊, 姬长英. 采用改进的自适应模糊多级中值滤波算法去除牛肉图像斑点噪声[J]. 农业工程学报, 2005, 21(2): 102 ~ 105.
Jia Yuan, Ji Changying. Removing the speckle noises in beef image using an improved adaptive fuzzy multilevel median filter algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering, 2005, 21(2): 102 ~ 105. (in Chinese)
- Otsu N. A threshold selection method from gray level histograms[J]. IEEE Transactions on System Man and Cybernetics, 1979, 9(1): 62 ~ 66.
- Du Chengjin, Sun Dawen, Patrick Jachman, et al. Development of a hybrid image processing algorithm for automatic evaluation of intramuscular fat content in beef M. longissimus dorsi[J]. Meat Science, 2008, 80(4): 1 231 ~ 1 237.
- 张莉, 周伟达, 焦李成. 核聚类算法[J]. 计算机学报, 2002, 25(6): 587 ~ 590.
Zhang Li, Zhou Weida, Jiao Licheng. Kernel clustering algorithm[J]. Chinese Journal of Computers, 2002, 25(6): 587 ~ 590. (in Chinese)
- 王守志, 何东健, 李文, 等. 基于核K-均值聚类算法的植物叶部病害识别[J]. 农业机械学报, 2009, 40(3): 152 ~ 155.
Wang Shouzhi, He Dongjian, Li Wen, et al. Plant leaf disease recognition based on kernel K - means clustering algorithm [J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(3): 152 ~ 155. (in Chinese)
- 刘木华, 赵杰文, 张海东. 基于模糊C均值聚类的牛肉图像中脂肪和肌肉区域分割技术[J]. 农业工程学报, 2004, 20(2): 161 ~ 163.
Liu Muhua, Zhao Jiewen, Zhang Haidong. Segmentation of fat and lean meat in beef images based on fuzzy C-means clustering [J]. Transactions of the Chinese Society of Agricultural Engineering, 2004, 20(2): 161 ~ 163. (in Chinese)
- 张艳诚, 毛罕平, 胡波, 等. 作物病害图像中重叠病斑分离算法[J]. 农业机械学报, 2008, 39(2): 112 ~ 115.
Zhang Yancheng, Mao Hanping, Hu Bo, et al. Separate algorithm for overlapping spots in crop disease image [J]. Transactions of the Chinese Society for Agricultural Machinery, 2008, 39(2): 112 ~ 115. (in Chinese)
- 陆宗骐, 金登男. Visual C++ .NET 图像处理编程[M]. 北京: 清华大学出版社, 2006.