

## 基于结构张量的 GrabCut 图像分割算法

张 勇<sup>1</sup>, 袁家政<sup>1,2</sup>, 刘宏哲<sup>1</sup>, 李 青<sup>1</sup>

(1. 北京联合大学 北京市信息服务工程重点实验室, 北京 100101; 2. 北京成像技术高精尖创新中心, 北京 100048)

**摘 要:** 传统的 GrabCut 图像分割方法大多基于图像的像素值建立图模型, 未考虑到彩色图像中丰富的纹理信息。为此, 提出一种新的 GrabCut 模型图像分割算法。对比基于结构张量的 GrabCut 分割方法和传统 GrabCut 分割方法的结果, 利用结构张量和像素值构建紧缩的结构张量。为提高计算的简洁性和高效性, 将 GrabCut 方法构建的混合高斯模型扩展到张量空间, 并用 Kullback-Leible 散度代替常用的黎曼度量。在合成纹理图像和自然图像上进行的实验结果表明, 与 Carsten Rother, GACWRF 等方法相比, 该算法具有更精确的分割效果, 不仅实现了纹理信息与颜色信息的无参融合, 而且提高了计算效率。

**关键词:** 图像分割; 结构张量; 图割模型; Kullback-Leible 散度; 混合高斯模型

**中文引用格式:** 张 勇, 袁家政, 刘宏哲, 等. 基于结构张量的 GrabCut 图像分割算法[J]. 计算机工程, 2017, 43(8): 258-265, 271.

**英文引用格式:** Zhang Yong, Yuan Jiazheng, Liu Hongzhe, et al. GrabCut Image Segmentation Algorithm Based on Structure Tensor[J]. Computer Engineering, 2017, 43(8): 258-265, 271.

## GrabCut Image Segmentation Algorithm Based on Structure Tensor

ZHANG Yong<sup>1</sup>, YUAN Jiazheng<sup>1,2</sup>, LIU Hongzhe<sup>1</sup>, LI Qing<sup>1</sup>

(1. Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, China;

2. Beijing High-tech Innovation Center of Imaging Technology, Beijing 100048, China)

**[Abstract]** Traditional GrabCut based image segmentation method is mainly based on the image pixel values to build a graph model, and does not take into account the rich texture of color image information. This paper presents an image segmentation algorithm based on GrabCut model, and contrasts results of Structure Tensor (ST) GrabCut segmentation method and traditional GrabCut segmentation method. The method uses the ST and the pixel values to construct the tight ST. For concise and efficient calculation, this paper extends Gaussian Mixture Model (GMM) built based on Grabcut method to tensor space, and uses Kullback-Leible (KL) divergence instead of the commonly used the Riemannian metric. Through a lot of experiments on synthetic texture images and natural images, results show that, compared with carstem Rother, GACWRF method the algorithm has more accurate segmentation effects, not only achieves the texture and color information parameter fusion, but also improves the computational efficiency.

**[Key words]** image segmentation; Structure Tensor (ST); GrabCut model; Kullback-Leible (KL) divergence; Gaussian Mixture Model (GMM)

**DOI:** 10.3969/j.issn.1000-3428.2017.08.044

### 0 概述

从复杂环境中分割出前景目标在计算机视觉中具有重要的现实意义, 尤其是从自然彩色图像中分割出前景目标, 具有更大的挑战性。彩色图像相比于灰度图像能够携带更多的信息<sup>[1]</sup>, 而这些信息一方面增加了分割的难度, 另一方面这些信息也可用于增强分析的过程, 提高分割的效果。针对彩色图像的分割已

经研究了多年, 尤其是最近几年, 前景目标分割在电影、电视、摄影出版物以及一些桌面应用程序上的应用, 使得彩色图像的分割受到了更多的关注<sup>[2]</sup>。

由于彩色图像中包含丰富的信息以及它们的不可预测性与复杂性, 一个通用的图像分割技术应该能够准确地定义所需的对象, 并应用最少的用户交互将其自动或半自动分割出来。现有的图像分割算法一般分为 3 个大类: 基于特征的, 基于区域的和基

**基金项目:** 国家自然科学基金(61271369, 61502036, 61571045); 国家科技支撑计划项目(2014BAK08B, 2015BAH55F03); 北京市自然科学基金(4152018, 4152016)。

**作者简介:** 张 勇(1990—), 男, 硕士研究生, 主研方向为图像处理; 袁家政, 教授、博士后; 刘宏哲, 教授、博士; 李 青, 讲师、博士。

**收稿日期:** 2016-06-28 **修回日期:** 2016-08-12 **E-mail:** 908715004@qq.com

于边界的。在一般情况下,每类方法都有自己的适用范围。基于特征的方法尝试去基于它们在特征空间上的位置分类像素点。常用的特征包括颜色、梯度、纹理、深度等。基于灰度阈值和基于距离信息的分类都属于基于特征的方法。这种方法有一个严重的缺点,由于不考虑相似点之间的空间连通性,一些在特征空间具有较高相似性但在空间上不具有连通性的像素点将被分为一类。基于区域的方法是基于特征方法的扩展,它不仅要保持特征空间的相似性,还要保持区域的连通性。基于区域的分割算法包括斑点着色、区域生长、区域合并、区域分割和智能涂料等。然而,这种方法有可能产生数量众多的零碎的离散区域。基于边界的方法不存在这个缺点,该方法是通过优化现有的边缘标准曲线去逼近真实的边界,提取出轮廓包围的对象或子区域。现有的基于边界的算法包括:边界追踪,动态规划,活动轮廓和智能剪刀等算法。但是由于边缘信息的局部性,如果想要获得所需的全局对象,需要大量的人工交互。为了能够运用更全面的信息,本文提出一种交互式的图像分割方法。

## 1 相关研究

近年来,基于图论的图像分割方法显现出了较高的准确性,这些算法最大的特点是将区域内分割能量函数与边缘正则化相结合。它们的共同策略是构造加权图,图的每个顶点对应一个图像像素或区域,每 2 个顶点之间连线的权重表示它们属于同一区域的概率的大小<sup>[3]</sup>。权重通常与提取的特征相关。随后,图被划分成多个最小化能量函数的组件。

GraphCut 算法是一种基于图论的交互式分割技术。GraphCut 由文献[4]提出和检验,该算法需要用户先标定出需要分割的对象,然后通过计算得到所有符合用户标定的分割结果,选取其中的最优值,实现图像的分割。该算法使得用户能够通过非常直观的交互得到满意的分割结果。这种方法在计算机视觉和图形领域被广泛研究,主要用于图像修复、二维和三维图像的分割等。GrabCut 算法是对 GraphCut 的一种扩展,它将 GraphCut 算法扩展到彩色图像,并且允许使用不完全的标注。GraphCut 算法的目标和背景的模式是灰度直方图,GrabCut 取代为各个通道的混合高斯模型(GMM),GraphCut 的能量最小化(分割)是一次达到的,而 GrabCut 取代为一个不断进行分割估计和模型参数学习的交互迭代过程<sup>[5]</sup>。GraphCut 需要用户指定目标和背景的一些种子点,但是 GrabCut 只需要提供背景区域的像素集即可,也就是说只需要框选目标,那么在方框外的像素全部当成背景,方框内的像素全部当成前景目标,这时就可以进行 GMM 建模,并且完成良好的分割<sup>[6-8]</sup>。这些发展使 GrabCut 可以更方便地进行图像编辑。

原始的 GrabCut 算法主要基于图像的像素值建

立 GMM, 在应用方面有很高的局限性。在文献[6,8-10]中,将图像分为目标与背景的过程主要基于图像价值的区域统计,然而对于图像价值的统计计算可能不足以对各区域进行判别。在一些情况下,纹理信息往往更适合作为一个判别特征<sup>[11]</sup>。在图像分割领域已经存在很多纹理特征描述方法,包括马尔可夫随机场<sup>[3,12-13]</sup>、多尺度的解决方法<sup>[5,14]</sup>、Gabor 小波滤波器等。1991 年,Bigun 等人提出了一种迭代方案<sup>[15]</sup>,即二阶矩阵结构张量(ST)被用来描述纹理,并分析了其基本特性,即仅产生 3 个特征通道,且每个通道比 Gabor 滤波能提供更多非常显著的信息,而且从中提取出图像的边界、拐角、纹理等重要信息。多种基于结构张量的图像分割算法也已被提出,包括活动轮廓模型、CV 模型等<sup>[16-17]</sup>。例如,文献[18]使用各通道的高斯近似的非线性化结构张量和加强的灰度直方图作为特征,将其带入到活动轮廓模型中实现图像的分割。此外,文献[19]的 GraphCut 技术被用来分割弥散张量核磁共振图像,并采用 KL 距离作为相异性度量,通过计算平均值来获得各个子张量的距离并以此作为终端之间的权重。然而,所有这些距离权重的计算量随着图像种子的增加而急剧增多。

文献[20]扩展这些基于张量的方法,通过在张量空间中引入黎曼几何,用于多对象与背景的图像分割。而且,交互式 GraphCut 技术也被用于多模态张量图像分割。为提高分割效率,一个  $3 \times 3$  的包含强度信息的扩展结构张量被构造,并作为 GraphCut 输入的数据。然而,在  $3 \times 3$  的结构张量中,颜色信息和纹理信息各占的比例是 50%,由于混合因子不能自适应调节,因此会存在很多的无用信息。此外,如果引入高阶的  $5 \times 5$  (考虑所有的颜色通道)的结构张量,则意味着必须在更高阶的空间实现能量的最小化,这将是极其困难的,并且可能得到多个局部极小的结果。选用 KL 散度量张量之间的距离,比文献[20]中所用的黎曼度量计算更简单,效率更高。在文献[17,19]中已经证明 KL 散度是一种有足够鲁棒性和判别力的低秩距离度量。

为了能够运用更全面的信息,并且保持结构张量的低阶,本文构造非线性的紧缩的结构张量,将其融入到 GrabCut 框架中,并进一步简化用户交互。

## 2 特征提取

本节将构造非线性的紧缩结构张量来提取图像的颜色和纹理特征。

### 2.1 结构张量

最初由 Di Zenzo 和 Förstner 先后于 1986 年和 1987 年提出了结构张量(ST)的定义,并将其应用于特征检测。然后,由 Weickert 等人将其应用在变分 PDE 中进行各向异性非线性正则。结构张量允许同

时进行方向估计和图像结构分析,在图像处理和计算机视觉中起着非常重要的作用。结构张量采用矩阵场表示,对其各通道进行滤波时允许结合通道内的局部邻域信息,因而可以避免梯度消去效应。例如,考虑一条细的线条(thin line),在它的某一边有一个正的梯度,在另一边有一个负的梯度。若直接对梯度进行平滑操作,会导致2个梯度相互消除,但对矩阵场数据进行平滑时,就会避免这种消去效应。通过对初始张量内积进行平滑滤波,可以获得更强的噪声鲁棒性。大多数存在于初始矩阵场中的噪声都能被平滑操作移除。结构张量结合了图像的局部方向信息,因而可以获取更多额外的结构信息。例如,可以提取图像的局部结构测度(包括边、角测度),能够自适应地区分平坦区域、结构区域等。

对于给定的一个2D图像 $I$ ,能够得到在像素点 $P_0$ 处的结构张量 $T$ ,并且考虑 $P_0$ 的附近区域 $R_{P_0}$ :

$$T_{(P_0)} = G_{R(P_0)}(DD)^T \quad (1)$$

其中, $G_{R(P_0)}$ 是区域的一个区域平均算子,以 $P_0$ 为中心; $D$ 表示区域 $R$ 中每一个点 $P$ 的梯度向量, $P \in R_{(P_0)}$ 。在图像 $I$ 像素点 $P$ 处的梯度算子 $D$ 可以表示为:

$$D(P) = \begin{bmatrix} \frac{\partial I(P)}{\partial_x} \\ \frac{\partial I(P)}{\partial_y} \end{bmatrix} = \begin{bmatrix} I_x(P) \\ I_y(P) \end{bmatrix} \quad (2)$$

其中, $I_x(P)$ 和 $I_y(P)$ 表示像素点 $P$ 在 $x$ 和 $y$ 方向的离散倒数。在最简单的情况 $G_{R(P_0)}$ 表示离散二项式或高斯滤波器<sup>[18]</sup>。关于 $G_{R(P_0)}$ 更多的信息将在后文介绍。

在此基础上,可以看出 $T_{(P_0)}$ 是一个正定的对称2D矩阵,能够描述 $P_0$ 周围范围像素点梯度分量的平均值。因此,结构张量 $T$ 不仅能够表达 $P_0$ 点的一些信息,而且能够表达 $P_0$ 周围像素点的一些信息。它也可以解释为区域 $R$ 的局部梯度方向一致性的一个衡量。另外还可以看出,如果计算出图像 $I$ 每一个像素点 $P$ 的结构张量 $T$ ,则 $T_{(P)}$ 能够传递每个像素点 $P$ 周围一些重叠区域的信息,即它包含图像纹理和局部曲率的信息。因此,结构张量(ST)是类似于Harris角点检测方法。然而,如果要同时传达图像的颜色信息,ST需要进一步加强颜色分量。将式(2)带入式(1)可以得到:

$$\begin{aligned} T &= G_R \left( \begin{bmatrix} I_x(P) \\ I_y(P) \end{bmatrix} \begin{bmatrix} I_x & I_y \end{bmatrix} \right) \\ &= G_R \left( \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix} \right) = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix} \end{aligned} \quad (3)$$

为简化计算,取均值滤波器。

## 2.2 扩展的结构张量

如前文所述,ST提供了2D图像局部区域的结构信息。然而,在许多应用中,将亮度信息或者颜色

信息与结构张量结合到一起使用。例如,文献[21]通过对ST和图像亮度信息添加适当的加权系数来进行立体匹配。文献[22]将ST与图像亮度或颜色信息结合到一起,通过能量的方式进行图像分割。在这种方法中,一个二维梯度向量 $D$ 被简单地扩展到一个三维向量 $E$ ,如式(4)所示。

$$E_{(P)}^T = [D_P^T \ I_P]^T = [I_x \ I_y \ I]^T \quad (4)$$

将式(4)带入到式(1),用 $E$ 取代 $D$ ,则能得到扩展的结构张量(EST)。

$$\begin{aligned} T_E &= G_R([EE^T]) = G_R \left( \begin{bmatrix} I_x \\ I_y \\ I \end{bmatrix} \begin{bmatrix} I_x & I_y & I \end{bmatrix} \right) \\ &= G_R \left( \begin{bmatrix} I_x^2 & I_x I_y & I_x I \\ I_y I_x & I_y^2 & I_y I \\ I_x I & I_y I & I^2 \end{bmatrix} \right) \end{aligned} \quad (5)$$

从式(5)可以看出,EST包含梯度的平均分量信息、沿边缘的平均平方亮度信息以及梯度分量与亮度混合信息。这些定义对于图像分割和一些其他的图像处理任务,如图像匹配、跟踪等,具有较强的判别力特征。

当考虑彩色图像时,每个像素点包含3个颜色通道的信息。在这种情况下,式(4)可够扩展到包含颜色分量:

$$F_{(P)}^T = [D_{(P)}^T I_{(P)}]^T = [\hat{I}_x \ \hat{I}_y \ I_R \ I_G \ I_B]^T \quad (6)$$

其中:

$$\hat{I} = \frac{1}{3}(I_R + I_G + I_B) \quad (7)$$

其中, $\hat{I}_x(P)$ 和 $\hat{I}_y(P)$ 分别表示 $\hat{I}$ 在 $P$ 点 $x$ 方向和 $y$ 方向的离散空间导数。将式(6)带入到式(1)中就能得到彩色图像的EST,它仍是一个正定的对称矩阵,包含 $n$ 个相互独立的分量。在一般情况下,对于包含15个分量的矢量,在它的外积矩阵将包含 $\frac{1}{2}n(n+1)$ 个独立分量。

## 2.3 紧缩的结构张量

尽管建立在式(6)基础上的EST能够提供足够的局部结构信息,但是在彩色图像的情况下,将需要计算15个独立分量,计算量太大。将PCA算法运用到式(6)的 $F$ 上,然后用得到的2个主要分量根据式(1)构建EST。更确切地说,对于每一个向量 $F$ ,它的PCA投影 $\hat{F}$ 为:

$$\hat{F} = A(F - \bar{F}) = [\hat{F}_1 \ \hat{F}_2]^T \quad (8)$$

其中, $A$ 表示PCA转换矩阵; $\bar{F}$ 表示所有向量 $F$ 的平均值; $\hat{F}_1$ 和 $\hat{F}_2$ 表示 $F$ 的2个最主要分量。因此,紧缩结构张量 $T_c$ 将与ST一样,只包含3个独立分量。用 $\hat{F}$ 替换 $D$ ,将式(4)带入到式(1)中可得到CST。

$$T_C = G_R([\hat{F}\hat{F}^T]) = G_R\left(\begin{bmatrix} \hat{F}_1\hat{F}_1 & \hat{F}_1\hat{F}_2 \\ \hat{F}_1\hat{F}_2 & \hat{F}_2\hat{F}_2 \end{bmatrix}\right) \quad (9)$$

PCA 的主要作用是将独立分量的数量有 15 个减少到 3 个,但是在一些图像区域,这样的减少会导致很高的信息丢失率。对于这种情况的一种解决方法是对主成分分析中的特征值进行检测,并且设定方差百分比的阈值。此外,本文的实验显示 PCA 不仅能够对 EST 进行压缩,而且能够剔除信息中的噪声。CST 的主分量示意图如图 1 所示。

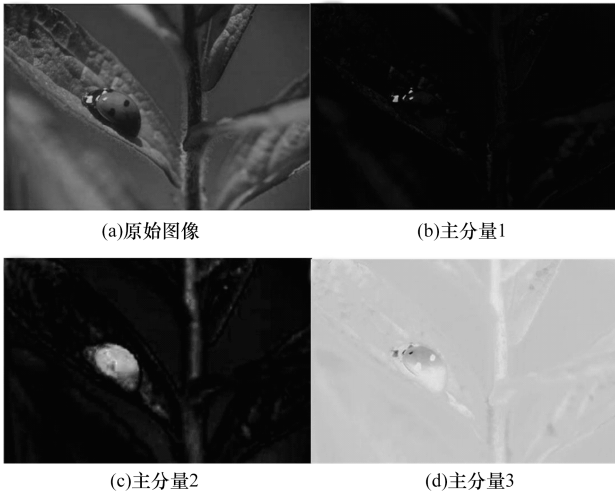


图 1 原始图像及其 CST 的 3 个主分量

### 2.4 非线性的紧缩结构张量

在 ST, EST 和 CST 的计算中,均值滤波函数  $G_R$  的选择是很重要的,如前文提出,如果重点考虑计算速度,可以使用简单的布尔滤波或高斯滤波函数,然而它们都属于各向同性的滤波器,不考虑区域的局部性质。因此,本文将使用一种非线性各向异波方法。在本文的框架中将使用 PM 方程的向量化形式作为滤波器<sup>[23]</sup>。

$$\partial_t I(P, t) = \text{div}(c(\|D_t(P)\|) \times D_t(P)) \quad (10)$$

其中,  $D_t(P)$  表示像素点  $P$  在时刻  $t$  的梯度;  $c$  表示函数  $D_t(P)$  参数的一个非线性控制函数。多种形式的函数  $c$  已经在文献中提出,在本文实验中使用 Tukey bi-weight 函数<sup>[24]</sup>。

$$c(x) = \frac{1}{2} \left(1 - \frac{x^2}{\sigma^2}\right)^2, |x| \leq \sigma \quad (11)$$

对于  $\sigma$  的确定,如式(12)所示<sup>[24]</sup>。

$$\sigma = 1.4826 \cdot \text{med}(\|\nabla I - \text{med}(\|\nabla \text{---} I\|)\|) \quad (12)$$

其中,  $\text{med}(\cdot)$  表示中值函数。上述各向异性滤波器的一个缺点是,它的迭代过程不能使用并行执行,在本文的框架中,10 次~30 次迭代一般能够满足要求。经过上述过程,得到图像  $I$  的非线性紧缩结构张量 (Nonlinear Compact Structure Tensor, NCST)。

### 3 GrabCut 模型图像分割算法

在一些情况下,仅依靠颜色信息并不能将感兴趣区域分割出来,这时如果能够有效地融合纹理信息,将极大地提升图像分割方法的性能。现在已经有很好很多的算法,能够将颜色与纹理融合到一起来进行合成纹理图像分割或自然图像分割。本节将详细介绍如何将 NCST 融合到 GrabCut 模型中,实现图像的分割。GrabCut 算法首先需要用户简单交互选定前景和背景样本,对前景、背景区域建立 GMM,并利用 k-means 算法初始化 GMM,分别计算节点到前景或背景的距离和相邻节点之间的距离,然后获得分割能量权重,对未知区域构建 s-t 网络图,接着采用最大流-最小割算法对其进行切分。GrabCut 算法的分割过程是通过迭代不断更新、修正 GMM 参数,保证  $E$  收敛于最小值,最终实现图像的分割<sup>[6-7]</sup>。本文尝试将 NCST 融入 GrabCut 的迭代分割过程中,并对 GrabCut 算法的初始化、距离测量和能量构建等方面进行了改进。实验结果显示,该算法不仅提高了分割的精度,而且有助于减少下一步的用户编辑。

#### 3.1 距离测量

从图像中提取颜色特征和纹理特征后,必须选择合适的距离度量,从而有效地区分这些特征,则对于图像分割的准确性来说是至关重要的。对于一幅给定的图像  $I$ ,提取它的非线性紧缩结构张量  $T_C$ :

$$T_C = \begin{bmatrix} \hat{D}_{xx} & \hat{D}_{xy} \\ \hat{D}_{xy} & \hat{D}_{yy} \end{bmatrix} \quad (13)$$

其中,符号“ $\hat{\cdot}$ ”表示对应的分量已经进行了非线性化扩散。张量空间分析中的一个关键因素是张量距离范数的选择,它需要能够度量张量之间与均值张量的相似性与差异性。文献[14,16]将张量压缩为一个标量,但是丢失了很多重要的信息,而且辨识度严重下降。文献[17]采用了 Frobenius 范数,但是它依赖于一个非常严格的假设。文献[25-26]考虑了张量空间的黎曼结构,给出了更为准确的张量统计估计。在本文的方法中,为了更容易地实现和更高效的计算,采用了对称 KL 散度,在低维时它与黎曼空间分布十分接近,而且在文献[17,19]中已经被证明有足够高的鲁棒性和辨识度。它自然地遵从扩散物理现象的规律,以局部高斯分布的协方差矩阵解释了张量的对称正定性,然后定义了基于信息理论概念的相异性度量。虽然 KL 散度与黎曼度量相比会丢失一些辨识度,但是它仍具有仿射不变性,并且具有封闭式计算和计算简易的优势。对于给定图像  $I$  上的点  $m$  和  $n$ ,可以定义 NCST 之间的 KL 距离为:

$$\text{dis}(T_m, T_n) = \sqrt{\frac{1}{4}(\text{tr}(T_m^{-1}T_n + T_n^{-1}T_m) - 4)} \quad (14)$$

其中,函数  $\text{tr}(\cdot)$  表示矩阵迹算子,其他符号按照 NCST 中的定义。

### 3.2 初始聚类

在原始的 GrabCut 算法中,用户通过直接框选目标来得到一个初始的 trimap  $T$ ,即方框外的像素全部作为背景像素  $T_B$ ,而方框内  $T_U$  的像素全部作为“可能是目标”的像素。对  $T_B$  内的每一像素  $n$ ,初始化像素  $n$  的标签  $\alpha_n = 0$ ,即为背景像素;而对  $T_U$  内的每个像素  $n$ ,初始化像素  $n$  的标签  $\alpha_n = 1$ ,即作为“可能是目标”的像素。经过上面的步骤,就可以分别得到属于目标( $\alpha_n = 1$ )的一些像素,剩下的为属于背景( $\alpha_n = 0$ )的像素,通过这个像素可估计目标和背景的 GMM。可以通过 k-mean 算法分别把属于目标和背景的像素聚类为  $K$  类,即 GMM 中的  $K$  个高斯模型,这时 GMM 中每个高斯模型就具有了一些像素样本集,它的参数均值和协方差就可以通过 RGB 值估计得到,而该高斯分量的权值可以通过属于该高斯分量的像素个数与总的像素个数的比值来确定。

而在 NCST 空间,为了建立 GMM,在初始的背景和前景区域被人工确定后,需要根据纹理特征将它们聚类为  $K_T$  类。将要得到一个( $2 \times K_T$ )的高斯分量。本文实验显示,初始聚类的结果对 GrabCut 迭代分割的效率和精度有很大影响。紧密且分离良好的集群算法有助于提高 GrabCut 迭代分割的效率和精度。初始聚类越好,迭代分割越精确,迭代次数越少。

目前已经有很多优秀的初始聚类的算法,比如 K-means、模糊 c-means 等。但是由于提取的纹理特征在 NCST 空间中被认为是矩阵值而不是向量值。定义区域间的基于矩阵值的协方差矩阵的高斯近似是很困难的一件事,所以引入了 KL 度量而不是运用黎曼流体去定义距离。在文献[27]的启发下,本文对 k-means 算法进行了扩展,使用式(14)作为距离度量的选择,使用式(15)来计算聚类中心。

### 3.3 能量函数

原始的 GrabCut 的概率密度函数是由颜色空间定义的,对于每个 GMM,它包含  $K_C$  个分量,每个分量包含 3 个参数:均值向量,对称的正定的全协方差矩阵和一个正的混合加权系数,更多的细节可参考文献[12]。本文在 NCST 空间中构建了 GMM,因此将纹理信息整合到 GrabCut 框架中,并且保持了框架的可扩展性。需要注意的是,基于纹理构建的 GMM 与基于颜色构建的 GMM 相同,它们都是基于初始聚类的结果或从迭代分割的上一步中得到。利用 KL 距离定义在区域  $R$  上的平均张量场  $\Gamma$ :

$$\bar{\mathbf{M}}_r(T, R) = \{\bar{\mathbf{M}}_r(T_1, R), \bar{\mathbf{M}}_r(T_2, R), \dots, \bar{\mathbf{M}}_r(T_n, R)\} \quad (15)$$

其中:

$$\{\bar{\mathbf{M}}_r(T_n, R) = \sqrt{\mathbf{B}_n^{-1}} [\sqrt{\mathbf{B}_n} \mathbf{A}_n \sqrt{\mathbf{B}_n}] \sqrt{\mathbf{B}_n^{-1}}\}_{n=0,1,\dots,N}$$

其中,  $\mathbf{A}_n = \int_R \mathbf{T}_n(x) dx$ ;  $\mathbf{B}_n = \int_R \mathbf{T}_n^{-1}(x) dx$ ;  $N$  为区域  $R$  中点的个数。然而,如上一部分提到,在张量空间构建协方差矩阵是很困难的。为了解决这个问题,文献[28]在黎曼空间定义距离的方法,在张量空间定义了基于 KL 距离在区域  $\Omega_T$  的随机变量的方差  $\sigma_T^2$ :

$$\sigma_T^2 = \frac{1}{|\Omega_T|} \int_{\Omega_T} \text{dis}_T^2(\Gamma_x, \bar{\mathbf{M}}_T) dx \quad (16)$$

因此,在 NCST 空间的每个 GMM 模型可以被认为是一个一般的混合高斯密度模型,在 NCST 空间中包含  $K_T$  个分量,类似于一阶高维的情况。对于每一个分量,它也包含 3 个参数:均值  $\bar{\mathbf{M}}_T$  (一个  $2 \times 2$  的矩阵),方差  $\sigma_T^2$  (正值),权重  $\pi_T$  (正值)。其中,  $\pi_T$  的值可以很容易地通过计算此分量包含像素点的个数与前景或背景像素点的总数的比值得到。

大量的实验结果证明,原始 GrabCut 算法的 GMM 近似在颜色空间中具有很高的鲁棒性。当它与 NCST 空间相结合,不仅保持了原有的鲁棒性,而且能够处理更大范围的图像。在 NCST 空间中,原始的 GrabCut 算法的能量构建被更新为:

$$E(\alpha) = \varepsilon_T(\alpha) \quad (17)$$

其中,  $\alpha$  表示指定的标签,当像素点属于背景时取 0,当像素点属于前景时取 1。  $\varepsilon_T(\alpha)$  由下式得到:

$$\begin{aligned} \varepsilon_T(\alpha) &= \sum_{U \ni u} -\text{lb} \sum_{j=1}^{K_r} \left\{ \frac{\pi_r(a_u, j)}{2\pi\alpha_r^2(a_u, j)} \exp\left(-\frac{\text{dis}_r^2(\Gamma_u, \mathbf{M}_r(a_u, j))}{2\alpha_r^2(a_u, j)}\right) \right\} \\ &+ \sum_{N \ni (m, n)} [a_m \neq a_n] \{ \pi_r \text{dis}^{-1}(m, n) \exp(-\beta_r \text{dis}_r^2(\Gamma_m, \Gamma_n)) + \tau \} \end{aligned} \quad (18)$$

其中,当  $\alpha_u = 1$  时,  $(\alpha_u, j)$  表示前景 GMM 的第  $j$  个分量;当  $\alpha_u = 0$  时,  $(\alpha_u, j)$  表示背景 GMM 的第  $j$  个分量;  $U$  表示初始时人为圈定的矩形区域,  $u$  为区域  $U$  内的像素点,  $N$  表示设定的邻近像素点的对数,  $\tau$  表示新添加的去噪常数。在分割过程中可以自适应的更新  $\beta_T$  的值:

$$\beta_T = \left( 2 \frac{\sum_{N \ni (m, n)} \text{dis}_T^2(\Gamma_m, \Gamma_n)}{|N|} \right)^{-1} \quad (19)$$

其中,  $|N|$  表示设定的邻近像素点对数  $N$  的数值。

### 3.4 迭代收敛准则

当迭代分割产生全局最优分割时,能量函数  $E(\alpha)$  也将达到最小值,但是在迭代分割的过程中,收敛在何时发生,最简单的标准是检测像素点的标签在迭代前后是否发生变化。但是在迭代过程中像素点的标签可能产生震荡变化,所以对于标签变化的检测往往是对计算资源的浪费。在本文中,一个自适应的准则可以被定义在基于纹理空间的前景与背景的 KL 距离上。迭代过程在满足下述公式时自动终止:

$$\| \mathbf{KL}_A - \mathbf{KL}_{A-1} \|^2 \leq \sigma \| \mathbf{KL}_1 - \mathbf{KL}_D \|^2 \quad (20)$$

其中,  $\mathbf{KL}_A$  表示第  $A$  次迭代时又得到的张量;  $\mathbf{KL}_0$  表示最初由用户放置的矩形确定的前景与背景之间的

KL 距离;  $\sigma$  是一个系数用来控制收敛速度和分割的准确度。很明显  $\Lambda \geq 2$ , 满足式(20)的  $\Lambda$  值表示了迭代的次数。对于颜色空间 2 个 GMM 模型的 KL 距离可以由下式计算:

$$KL_C(N_{C,k}^F \parallel N_{C,i}^B) = \frac{1}{2} \left( \ln \frac{\sum_{C,i}^B}{\sum_{C,k}^F} + \text{tr}((\sum_{C,i}^B)^{-1} \sum_{C,k}^F) + (\sum_{C,i}^B)^{-1} (\mu_{C,k}^F - \mu_{C,i}^B) \right) \quad (21)$$

其中  $N_{C,k}^F, N_{C,i}^B$  分别表示前景 GMM 的第  $k$  个分量和背景 GMM 的第  $i$  个分量。张量空间的 KL 距离  $KL_r$  与  $KL_c$  的计算是类似的,但是基于 NCST 空间的 GMM 是一阶高维的情况,不能直接套用式(20)。这种状况是广义高斯分布的一种特殊情况,它的形状参数是固定的,近似等于  $2^{[29]}$ , 可得到在 NCST 空间的近似的 KL 距离:

$$KL_r(N_{r,k}^F \parallel N_{r,i}^B) = \frac{1}{2} \left( \ln \frac{(\sigma_{r,1}^B)^2}{(\sigma_{r,x}^F)^2} + \ln \frac{(\sigma_{r,x}^F)^2}{(\sigma_{r,1}^B)^2} - 1 \right) \quad (22)$$

实验结果表明,本文策略不仅可以有效地缩短迭代的时间,而且具有较好的分割精度。

### 4 实验结果与分析

通过对大量自然场景图像的分割来测试本文提出方法的性能。

#### 4.1 合成纹理图像分割

图 2、图 3 分别展示了对灰度合成纹理图像和彩色合成纹理图像的处理,图像来自于 Brodatz 纹理库,实验结果显示了本文提出的 NCST 具有强大纹理提取能力。

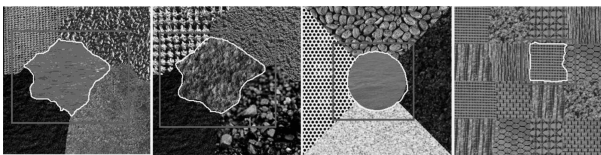


图 2 灰度合成纹理图像分割

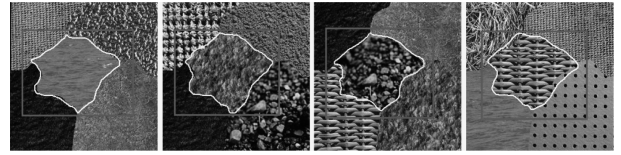


图 3 彩色合成纹理图像分割

#### 4.2 人工交互对比

本文将 NCST 融入到 GrabCut 算法中,不仅提高了分割的精度,而且能够有效地减少后续的人工编辑。图 4 展示了 GrabCut 算法与本文所提出的算法分割过程的人工交互。其中,第 1 列与第 3 列为 GrabCut 算法的分割过程,第 2 列与第 4 列为本文提出算法的分割过程。从第 2 行初始迭代的结果可以看出,本文提出的算法在初始迭代中实现了比 GrabCut 更为精确的结果,所以在第 3 行的人工交互过程中,本文提出算法的交互更为简洁。

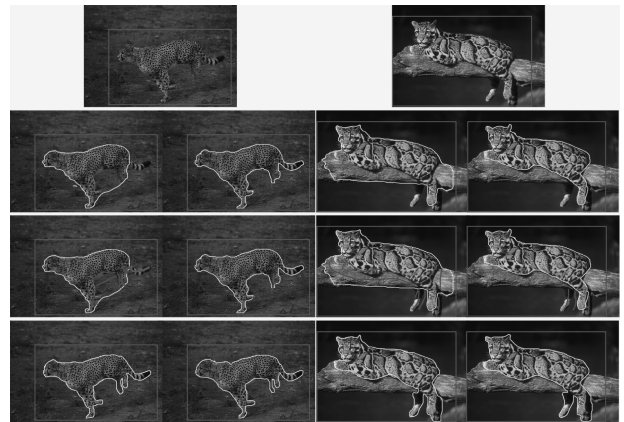


图 4 人工交互对比

#### 4.3 自然图像处理

本文提出的方法面向的对象是复杂环境中的目标提取。图 5(a)展示了本文提出的算法对一幅图像的迭代分割过程,图 5(b)和图 5(c)分别是其对应的 KL 距离与能量  $\epsilon_r(\alpha)$ 。

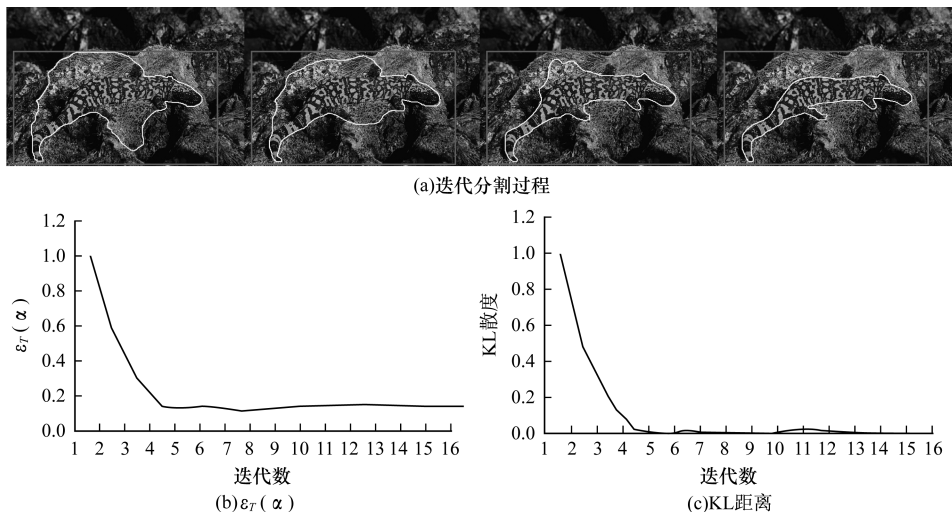


图 5 迭代分割过程展示

为进一步验证本文方法的有效性和准确性,在 BSDS500 自然图像库上,将本文方法、Carsten Rother 方法<sup>[30]</sup>、GACWRF 方法<sup>[31]</sup>、Gobar 与 GrabCut 结合

的方法<sup>[32]</sup>进行实验对比分析,实验结果如图 6 所示。需要注意的是,这里利用相同的矩形框框住前景目标。



图 6 图像分割结果对比

从图 6 可以看出,本文提出的方法能够分割出更多的细节,达到更高的分割精度。为了进一步对本文所使用的 4 种对比方法的效率进行有效性分析,本文在 BSD500 数据库上随机选取 20 幅图片,采用数据库中的 PRI 概率随机检索进行准确率度量。将数据库中的人工分割图像作为地面真实的参考,量化的结果如表 1 所示。由表 1 中平均 PRI 准确率可见,本文方法的分割准确率最高,表明本文方法分割的结果更加接近于地面真实分割结果,也可以说明本文采用的 NCST 相比于 Gobar 滤波,具有更强的纹理提取能力。同时,从分割算法的时间消耗可见,本文方法分割过程的收敛速度快于 GACWRF 方法和 Gabor GrabCut 方法,原因在于本文方法对

k-means 初始化进行了改进,同时也对 GrabCut 算法进行了适用于 NCST 的改进,提高了收敛速度。但是,其时间分割代价则略高于 Carsten Rother。从整体上考虑,本文方法的分割性能优于 Carsten Rother 方法、GACWRF 方法和 Gabor GrabCut 等方法。

表 1 PRI 准确率度量

对比方法	平均 PRI 准确率	平均耗时/s
Carten Rother 方法	0.895	6
GACWRF 方法	0.923	12
Gabor GrabCut 方法	0.937	10
本文方法	0.973	8

综上所述,本文方法具有分割效果好、准确性高等特点。实验结果表明,该方法能够较好地完成图像的分割,显示出基于结构张量的纹理图像分割方法的潜在应用前景。

## 5 结束语

本文通过结合 NCST 与 GrabCut 模型,提出一种交互式的图像分割方法。利用 NCST 来实现颜色信息与纹理信息的融合,不仅简化了计算,而且解决了纹理信息与颜色信息融合时遇到的参数问题。将 KL 散度代替黎曼度量来计算 NCST 空间的距离,提高了计算的简洁性与高效性。大量合成纹理图像与自然图像的实验结果表明,本文提出的方法相比于其他 3 种方法具有更高的分割精度。下一步将对 NCST 的非线性化扩散进行改进,从而提高计算速度。

### 参考文献

- [1] Cheng H D, Jiang X H, Sun Y, et al. Color Image Segmentation: Advances and Prospects [J]. *Pattern Recognition*, 2001, 34(12): 2259-2281.
- [2] 吕骁博,郭耀,陈向群. 基于分割的数字图像云存储机制[J]. *计算机研究与发展*, 2014, 51(5): 1129-1135.
- [3] 刘磊,石志国,宿浩茹,等. 基于高阶马尔可夫随机场的图像分割[J]. *计算机研究与发展*, 2013, 50(9): 1933-1942.
- [4] Lombaert H, Sun Y, Grady L, et al. A Multilevel Banded Graph Cuts Method for Fast Image Segmentation[C]// *Proceedings of IEEE ICCV '05*. Washington D. C., USA: IEEE Press, 2005: 259-265.
- [5] Eccleston D, Chan B, Holt G, et al. Multiple Resolution Segmentation of Textured Images[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1991, 13(2): 99-113.
- [6] Faltn P, Nguyen P A, Chaisaowong K, et al. Interactive 3D Segmentation of Pleural Thickenings Simultaneously at Different Points of Time Using Graph Cut [M]. Berlin, Germany: Springer, 2014: 378-383.
- [7] Ciesielski K C, Udupa J K, Falcão A X, et al. Fuzzy Connectedness Image Segmentation in Graph Cut Formulation: A Linear-time Algorithm and a Comparative Analysis [J]. *Journal of Mathematical Imaging & Vision*, 2012, 44(3): 375-398.
- [8] Ciesielski K C, Udupa J K, Falcão A X, et al. Fuzzy Connectedness Image Segmentation in Graph Cut Formulation: A Linear-time Algorithm and a Comparative Analysis [J]. *Journal of Mathematical Imaging & Vision*, 2012, 44(3): 375-398.
- [9] Li Y, Li X. A Background Correction Method Based on Lazy Snapping [C]// *Proceedings of the 7th IEEE International Conference on Image and Graphics*. Washington D. C., USA: IEEE Press, 2013: 144-148.
- [10] Rother C, Kolmogorov V, Blake A. GrabCut: Interactive Foreground Extraction Using Iterated Graph Cuts [J]. *ACM Transactions on Graphics*, 2004, 23(3): 307-312.
- [11] Newbould R D, Nicholas R, Thomas C L, et al. AgeIndependently Affects Myelin Integrity as Detected by Magnetization Transfer Magnetic Resonance Imaging in Multiple Sclerosis [J]. *NeuroImage: Clinical*, 2014, 4: 641-648.
- [12] Cross G R, Jain A K. Markov Random Field Texture Models [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1983, 5(1): 25-39.
- [13] Na K O, Kim C. Unsupervised Texture Segmentation of Natural Scene Images Using Region-based Markov Random Field [J]. *Journal of Signal Processing Systems*, 2016, 83(3): 423-436.
- [14] Fourati W A, Bouhlel M S. Trabecular Bone Image Segmentation Using Wavelet and Marker-controlled Watershed Transformation [J]. *Chinese Journal of Engineering*, 2014, 2014: 1-10.
- [15] Bigun J, Granlund G H, Wiklund J. Multidimensional Orientation Estimation with Applications to Texture Analysis and Optical Flow [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1991, 13(8): 775-790.
- [16] Rousson M, Brox T, Deriche R. Active Unsupervised Texture Segmentation on a Diffusion Based Feature Space [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. [S. l.]: IEEE Computer Society, 2003: 699-704.
- [17] Wang Z, Vemuri B C. Tensor Field Segmentation Using Region Based Active Contour Model [C]// *Proceedings of ECCV '04*. Prague, Czech Republic: [s. n.], 2004: 304-315.
- [18] Cyganek B. Visual Pattern Recognition Framework Based on the Best Rank Tensor Decomposition [M]. Berlin, Germany: Springer, 2015.
- [19] Weldelessie Y T, Hamameh G. DT-MRISegmentation Using Graph Cuts [J]. *Proceedings of SPIE*, 2007(6512): 1-9.
- [20] Malcolm J, Rathi Y, Tannenbaum A. A Graph Cut Approach to Image Segmentation in Tensor Space [EB/OL]. (2007-10-21). <http://www.docin.com/p-109227654.html>.
- [21] Cyganek B. Object Detection and Recognition in Digital Images [M]. [S. l.]: John Wiley & Sons, 2013.
- [22] Luis-García R D, Deriche R, Rousson M, et al. Tensor Processing for Texture and Colour Segmentation [C]// *Proceedings of Scandinavian Conference on Image Analysis*. Berlin, Germany: Springer-Verlag, 2005: 1117-1127.
- [23] Perona P, Malik J. Scale-space and Edge Detection Using Anisotropic Diffusion [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1990, 12(7): 629-639.
- [24] Andrew A M. Geometric Partial Differential Equations and Image Analysis [J]. *Kybernetes*, 2002, 31(2).

## 5 结束语

为更好地实现对运动阴影的去除和检测,本文结合传统码本原理和运动阴影的特性,对传统码本方法进行改进。首先在 YCbCr 颜色空间中建立背景码本用于检测包含运动目标前景和运动阴影的运动区域;然后改进码本用于运动阴影的建模,实现亮度范围阈值和颜色距离阈值的自适应获取,同时提出根据码字的匹配次数、标示位和时间来更新运动阴影码本的方法。实验结果表明,本文算法可有效提高运动阴影的检测率和识别率,并且具有较好的稳定性。尽管本文取得了较好的实验结果,但是有大量的阈值参数需要在实验过程中人工取值,因此,下一步将对算法进行改进,减少人工取值的参数数量。

### 参考文献

- [1] 刘 雪,常发亮,王华杰. 运动目标检测中的阴影去除方法[J]. 微处理机,2008,29(5):116-117.
- [2] 苑 颖,罗青山,曾贵华. 视频监控系统中阴影去除方法的研究[J]. 计算机工程与应用,2008,44(36):187-188.
- [3] Lee B E, Nguyen T B, Chung S T. An Efficient Cast Shadow Removal for Motion Segmentation [C]//Proceedings of WSEAS International Conference on Signal Processing, Computational Geometry and Artificial Vision. Stevens Point, USA: World Scientific and Engineering Academy and Society, 2009:83-87.
- [4] Sanin A, Sanderson C, Lovell B C. Improved Shadow Removal for Robust Person Tracking in Surveillance Scenarios [C]//Proceedings of the 20th International Conference on Pattern Recognition. Washington D. C., USA: IEEE Press, 2010:141-144.
- [5] Sanin A, Sanderson C, Lovell B C. Shadow Detection: A Survey and Comparative Evaluation of Recent Methods [J]. Pattern Recognition, 2012, 45(4):1684-1695.
- [6] 张红颖,李 鸿,孙毅刚. 基于混合高斯模型的阴影去除算法[J]. 计算机应用,2013,33(1):31-34.
- [7] Kim K, Chalidabhongse T H, Harwood D, et al. Real-time Foreground-Background Segmentation Using Codebook Model [J]. Real-Time Imaging, 2005, 11(3):172-185.
- [8] 张华南,冯 全,杨 梅,等. 基于 YCbCr 颜色空间的叶片阴影检测与去除[J]. 计算机系统应用,2015,24(11):262-265.
- [9] 王 雯,陈 丽,李 晨,等. YUV 空间下基于码本模型的视频运动目标检测方法[J]. 武汉大学学报(工学版),2015,48(3):412-416.
- [10] Kim K, Chalidabhongse T H, Harwood D, et al. Background Modeling and Subtraction by Codebook Construction [C]//Proceedings of International Conference on Image Processing. Washington D. C., USA: IEEE Press, 2004:3061-3064.
- [11] Ting S U N, Hsu S C, Huang C L. Hybrid Codebook Model for Foreground Object Segmentation and Shadow/Highlight Removal [J]. Journal of Information Science and Engineering, 2014, 30(6):1965-1984.
- [12] 付 萍,方 帅,徐心和,等. 视频监控系统中运动目标检测的阴影去除方法[J]. 计算机工程,2007,33(10):22-24.
- [13] Benedek C, Szirányi T. Study on Color Space Selection for Detecting Cast Shadows in Video Surveillance [J]. International Journal of Imaging Systems & Technology, 2007, 17(3):190-201.
- [14] Al-Najdawi N, Bez H E, Singhai J, et al. A Survey of Cast Shadow Detection Algorithms [J]. Pattern Recognition Letters, 2012, 33(6):752-764.
- [15] 杨 丹,余孟泽. 车辆视频检测及阴影去除[J]. 计算机工程与设计,2011,32(6):2072-2074.
- [16] Khronos B. OpenCL: Parallel Computing for Heterogeneous Devices [C]//Proceedings of 2009 IEEE Hot Chips 21 Symposium. Washington D. C., USA: IEEE Press, 2010:1-56.
- [17] Prati A, Mikic I, Trivedi M M, et al. Detecting Moving Shadows: Algorithms and Evaluation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2003, 25(7):918-923.
- [25] Lenglet C, Rousson M, Deriche R, et al. Statistics on the Manifold of Multivariate Normal Distributions: Theory and Application to Diffusion Tensor MRI Processing [J]. Journal of Mathematical Imaging & Vision, 2006, 25(3):423-444.
- [26] Pennec X, Fillard P, Ayache N. A Riemannian Framework for Tensor Computing [J]. International Journal of Computer Vision, 2006, 66(1):41-66.
- [27] Li Y, Sun J, Tang C K, et al. Lazy Snapping [J]. ACM Transactions on Graphics, 2004, 23(3):303-308.
- [28] Fletcher P T, Joshi S. Principal Geodesic Analysis on Symmetric Spaces: Statistics of Diffusion Tensors [M]. Berlin, Germany: Springer, 2004:87-98.
- [29] Do M N, Vetterli M. Wavelet-based Texture Retrieval Using Generalized Gaussian Density and Kullback-leibler Distance [J]. IEEE Transactions on Image Processing, 2002, 11(2):146-158.
- [30] Rother C, Kolmogorov V, Blake A. GrabCut: Interactive Foreground Extraction Using Iterated Graph Cuts [J]. ACM Transactions on Graphics, 2004, 23(3):307-312.
- [31] Tao W. Iterative Narrowband-based Graph Cuts Optimization for Geodesic Active Contours with Region Forces [J]. IEEE Transactions on Image Processing, 2012, 21(1):284-296.
- [32] 杨章静,钱建军,黄 璞,等. 基于 Gabor 变换的 GrabCut 纹理图像分割[J]. 控制与决策,2016,31(1):149-154.

编辑 金胡考

(上接第 265 页)

编辑 索书志