

# 基于视觉注意的随机游走图像分割\*

王富治, 秦付军, 蒋代君, 宋昌林

(西华大学机械工程学院 成都 610039)

**摘要:**传统随机游走图像分割需要多次交互设置种子点以获得理想的分割结果。在视觉注意的基础上,提出了一种新的自动确定种子点的随机游走图像分割算法。首先对图像进行超像素分割,并生成概率边界图(PBM);然后基于 Itti 模型,通过视觉注意焦点的转移搜寻待分割的关键区域;为确定关键分割区域种子点,以当前注意焦点作为极点对概率边界图进行极坐标变换,在获得的极坐标概率边界图上建立关于焦点区域边界的能量函数,采用图论 max-flow min-cut 算法最小化能量函数检测焦点区域的最优边界,焦点区域边界内的超像素即为种子点;最后以超像素为节点构造图,在图上随机游走完成图像分割。在 Berkeley Segmentation Data Set 上的实验表明本文方法能有效分割复杂图像。

**关键词:** 图像分割;随机游走;视觉注意;概率边界图

中图分类号: TH74 文献标识码: A 国家标准学科分类代码: 510.4050

## Random walks for image segmentation based on visual attention

Wang Fuzhi, Qin Fujun, Jiang Daijun, Song Changlin

(School of Mechanical Engineering of Xihua University, Chengdu 610039, China)

**Abstract:** The traditional random walks based image segmentation algorithm requires setting seed points interactively to obtain the desired segmentation results. Based on visual attention, the paper proposes a new random walks based image segmentation algorithm with the seed points determined automatically. Firstly, the probability boundary map (PBM) is generated and the image is divided into super pixels. Then, the key segmentation region is searched by shifting visual attention focus with Itti model. In order to determine the seed points of the key segmentation region, the probabilistic boundary map is transformed into polar coordinates map taking the current focus of attention as the pole. The energy function about the boundary of the focal region is established on the obtained polar coordinate probabilistic boundary map. The energy function can be minimized by the max-flow min-cut algorithm, and the super pixels within the boundary of focal region are the seed points of segmentation region. Finally, super pixels of images are used as nodes to construct a graph, random walks algorithm is conducted on the graph to complete the image segmentation. The experiments on Berkeley Segmentation Data Set show that the proposed method is effective to complex images' segmentation.

**Keywords:** image segmentation; random walks; visual attention; probabilistic boundary map (PBM)

## 1 引 言

图像分割是机器视觉的基础任务,应用广泛。其困难在于本身属于底层视觉,却与图像理解、场景感知等高层视觉密切相关。近年来兴起的交互式图像分割算法,由于加入了用户的先验知识,较好消除了关于分割“一致性”的歧义,使得分割的性能得到了很大的提高。随机游

走分割算法由 Grady L 所提出,是当前研究和应用最为广泛的交互式分割算法之一,该算法的基本思想:首先由用户确定不同分割区域的种子点,然后将一幅图像视为以像素作为节点的无向图,进而将图像分割转化为图上随机游走者首次到达种子点的概率计算问题<sup>[1]</sup>。在 Grady L 工作的基础上,文献[2]通过对 Laplace 矩阵进行分解,提出了随机游走图像分割快速算法。传统随机游走分割方法主要适用于灰度渐变的图像,并且分割质量

和种子点的选择密切相关,往往需要多次交互设置种子点以获得理想的分割结果,限制了该类方法的使用。因而,如何实现基于随机游走的复杂图像自动分割,就成为一个值得注意的研究方向。这方面的工作首先也是由 Grady L 开展起来的。Grady L 提出了带先验模型的随机游走分割算法<sup>[3]</sup>,利用高斯核函数对图像灰度分布进行估计作为图像的先验知识,可以不需要交互输入种子点,实现了随机游走自动图像分割。但该方法仅仅适用于图像灰度具有明显分布规律的图像,对于复杂自然图像,效果不好。

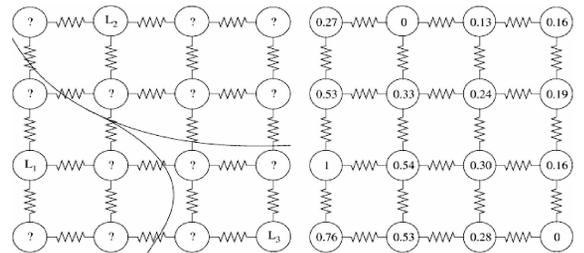
视觉注意机制不仅是一种高效的信息处理机制,还是一种重要的认知机制,与高层视觉与底层视觉任务都紧密相关;并且视觉注意焦点区域通常是分割目标。因此,把视觉注意机制与图像分割结合起来,就为解决图像分割问题提供了一种可行之道,成为当前研究的热点。目前该方向研究多集中于单视觉显著区域的检测与分割<sup>[4-6]</sup>,如文献在[4]提出了基于多分辨率框架的显著目标分割算法,文献[5]通过二值化显著图来初始化 Grab 算法,然后迭代运用 Grab 算法来优化分割结果,文献[6]提出了基于 Graph-cut 的显著区域检测方法。基于这方面的工作,目前已有少数文献把视觉注意机制引入随机游走分割算法中,如文献[7]利用视觉显著区域检测算法和仿射传播聚类算法提取目标和背景的种子点,然后利用随机游走实现了单个显著区域与图像背景比较准确的分割。文献[8]则主要讨论了随机游走算法中图的构造问题,进而实现了显著目标的检测。

文献[7]所提出的单显著区域种子点提取与分割方法(也称二值分割)与多显著区域的检测与分割相比,后者更富有意义。这是因为,根据 Salha A. A. 等人<sup>[9]</sup>和 Soyer C 等人<sup>[10]</sup>的研究,图像理解、场景感知并不需要图像或者场景的全部信息,而只需要通过视觉注意焦点的转移获取若干关键信息就足以进行感知、理解和识别。而这些通过视觉注意焦点转移获取的关键信息正是图像分割的目标。基于此,针对传统随机游走分割算法的不足,本文在视觉注意机制的基础上,提出一种新的随机游走多标签(多区域)自动图像分割算法。该方法的特点在于,通过视觉注意焦点的转移来获取分割区域,在极坐标概率边界图中建立能量函数检测焦点区域边界进而确定分割区域超像素种子点,最后利用传统的随机游走分割模型实现自动分割,这种以超像素为节点的图上游走大大提高了计算效率。本文的分割过程与生物视觉过程相似,分割区域种子点的确定属于高层视觉任务,随机游走分割属于底层视觉任务,底层视觉任务与高层视觉任务以视觉注意机制为纽带相连。因而本文方法的分割过程体现了高层视觉与底层视觉任务相结合的特点。实验证明了本文方法能有效分割复杂图像。

## 2 随机游走图像分割原理与本文算法框架

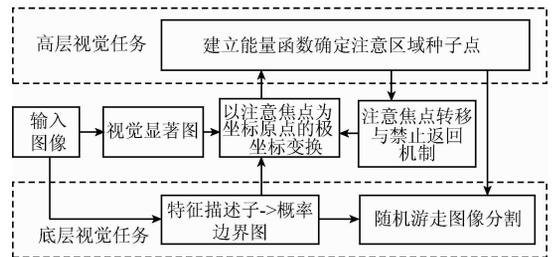
### 2.1 随机游走图像分割算法原理

如图 1(a)所示,随机游走分割算法将图像转化为由一定顶点、边和权值构成的带权无向图  $G(V, E, W)$ 。其中,  $V$  表示由图像像素、超像素或其他同质区域构成的顶点集;  $E$  表示连接图  $G$  顶点  $i$  和  $j$  的边  $e_{i,j}$  的集合;  $W$  表示相邻顶点间权值  $w_{i,j}$  的集合,  $w_{i,j}$  描述了图节点之间的相似性,也反映了游走时通过这条边的概率。随机游走图像分割算法的基本思想是,首先由用户标记分割区域种子点,有多少个分割区域,种子点就分为多少类。然后将属于本区域的种子点的概率记为 1,其他区域种子点的概率记为 0。而未标记顶点归属于各分割区域种子点的概率,可以等效为二维模拟电路各点电位的计算,在数学上是一个边界值为 0 和 1 的 Dirichlet 积分求解问题。最后根据所求得各顶点概率的大小对未标记点进行分类,从而实现图像分割。



(a) 随机游走分割算法原理

(a) Principle of random walks based image segmentation



(b) 本文算法框架

(b) The framework of the proposed algorithm

图 1 随机游走图像分割原理与本文算法框架

Fig. 1 Principle of random walks based image segmentation and the framework of the proposed algorithm

### 2.2 本文算法框架

由 2.1 节可见,传统的随机游走分割算法是基于种子点的分割算法,哪些区域需要分割,待分割区域种子点皆由人工指定,使用上不方便;另外,传统的随机游走分割算法,基于图像像素实现无向图节点的构建,随着节点数目的增多,算法的计算开销急剧增大。

为了解决传统随机游走图像分割的这两个缺点,本

文提出基于视觉注意的随机游走分割算法。该算法框图如图1(b)所示。该算法将传统的以像素为节点的图上随机游走分割改变为以超像素为节点的图上随机游走图像分割;对于最为重要的种子点确定问题,输入图像后,首先对图像进行超像素分割<sup>[11]</sup>,并生成显著图和概率边界图;接着通过视觉焦点的不断转移来搜寻分割区域;为了确定分割区域的种子点,首先以当前视觉注意焦点为原点对概率边界图进行极坐标变换,然后在所获得的极坐标概率边界图上建立关于焦点区域边界的能量函数,采用最大流最小割算法最小化能量函数检测焦点区域的最优边界,焦点区域边界内部的超像素即构成分割区域内部的种子点。获得种子点后,即可利用传统的随机游走分割模型实现自动分割。这里,分割区域种子点的确定属于高层视觉任务,而超像素分割、概率边界图以及随机游走图像分割属于底层视觉任务,因此,本文方法体现了高层视觉与底层视觉任务相结合的特点。

在算法框架描述中,涉及到两个概念:分割区域和焦点区域。更确切地说,分割区域分为实际分割区域和理想分割区域。理想分割区域是指期望的分割结果,也称为 Ground truth;实际的分割区域则是分割算法所得到的分割结果;焦点区域则是以注意焦点为原点对概率边界图进行极坐标变换后所确定的最优分割区域。从后文焦点区域的确定办法可知,焦点区域一般是理想分割区域的子集,若能整合多种信息得到准确的边界图,则焦点区域将逼近于理想分割区域。本文将理想分割区域称为 Ground truth,实际的分割区域简称为分割区域。

视觉注意是生物视觉信息处理和认知过程中一个非常重要的机制,是本文分割方法的基础。目前已提出了数十种视觉注意模型及算法,这其中,Itti 视觉注意模型具有里程碑的意义<sup>[12]</sup>。该模型属 Bottom-up 类型,通过线性滤波提取图像颜色特征、亮度特征和方向特征,将特征分别进行归一化整合后得到显著图,显著图产生了一系列的待注意目标,然后通过胜者为王机制和禁止返回机制进行转移,由此确定一系列焦点区域,即待分割的目标。在本文中即采用 Itti 视觉注意模型来产生注意图以及进行焦点转移。

### 3 分割区域种子点自动确定

Itti 视觉注意模型只能发现分割区域位置所在,而随机游走图像分割是基于种子点的图像分割方法,因而首先要确定各个分割区域内的种子点。为此,在图像超像素分割的基础上提出如下种子点确定方法。

#### 3.1 概率边界图

边界是指能区分不同分割区域的边缘。传统的边缘检测方法多以不同形式捕捉图像中的梯度变化信息,从而使得图像的纹理区域往往产生大量的无意义干扰性边缘。为了准确检测图像边界,近年来的趋势是综合图像多种特征信息进行检测。比较有代表性的方法是 Martin D. R. 等人<sup>[13]</sup>提出的基于多种视觉特征自然统计特性的概率边界图(probabilistic boundary map, PBM)检测算子(Pb 算子),该算子的整体思想是将图像转换到 LAB 空间,构建亮度 L、颜色 A、颜色 B、纹理等多个特征通道,然后采用梯度方法来表示图像特征。基于梯度的特征表示方法是,对图像的每个像素点,构建 8 个方向的半圆对掩膜,在每个半掩膜区域,使用直方图对其中每个像素的特征(亮度、颜色和纹理)进行统计,得到各类特征的统计分布,然后使用  $\chi^2$  距离计算中心像素点两个半圆形区域的差异作为特征输出。

通过上述步骤获得一个像素点的多特征表示后,Pb 算子采用二项 Logistic 回归模型对 200 幅训练图像进行统计学习,模型参数通过极大似然法进行估计后,由该模型即可获得图像每一像素点属于边界的概率,生成概率边界图。

#### 3.2 基于概率边界图的焦点区域确定

概率边界图仅仅反映了某像素属于边界的概率。而对于随机游走分割算法来说,需要的是属于某区域的种子点,要求它们不应“越界”,即属于各自的焦点区域。种子点“不越界”,从最优化的观点来看,也即找到一条最优的封闭曲线,该曲线包含该区域的视觉注意焦点,并将属于当前焦点区域的种子点与不属于当前焦点区域的种子点分割。该封闭区域应尽可能逼近理想的焦点区域边界。而愈逼近焦点区域理想边界的封闭曲线,其搜索代价愈高。为保证不同的封闭曲线搜索代价相差不大,对图像以当前视觉注意焦点为原点进行极坐标变换,在极坐标空间里进行搜索<sup>[6]</sup>。这样做的好处是显然的,同圆心而不同大小的圆,经极坐标变换后,在极坐标空间为长度相等的两条直线,搜索代价相同。

设  $I_E^0$  表示直角坐标中的概率边界图,  $F(x_o, y_o)$  为当前焦点(直角坐标),以该焦点为坐标原点(也即极点)进行极坐标变换,所得的极坐标概率边界图记为  $I_E^r$ 。设  $P$  表示极坐标边界图中所有像素的集合,  $L = \{0, 1\}$  为极坐标边界图中每个像素可能标记的集合,假定像素  $p$  的标记  $l_p = 0$  表示该像素处于焦点区域封闭边界的内部,  $l_p = 1$  表示该像素处于焦点区域封闭边界外部,那么寻找焦点区域最优封闭边界的问题等价于,在概率边界图上寻求从集合  $P$  到集合  $L$  的标号函数,该标号函数满足如式(1)。

$$Q(f) = \sum_{p \in P} U_p(l_p) + \lambda \sum_{(p,q) \in \Omega} V_{p,q} \cdot \delta(l_p, l_q),$$

$$V_{p,q} = \begin{cases} \eta \exp\left(-\frac{I_E^{\text{pol}}(r_p, \theta_p) + I_E^{\text{pol}}(r_q, \theta_q)}{2}\right), & \frac{I_E^{\text{pol}}(r_p, \theta_p) + I_E^{\text{pol}}(r_q, \theta_q)}{2} \neq 0 \\ k, & \text{其他} \end{cases} \quad (1)$$

式中:  $\delta(l_p, l_q)$  为 Kronecker 符号。  $U_p(l_p)$  为能量函数数据项。关于数据项, 只有两种情况是确定的, 即对于极坐标图像第 1 列来说, 其对应于极点, 必处于焦点区域内, 因而对于  $\forall p \in (0, 0^\circ \leq \theta \leq 360^\circ)$ , 有  $l_p = 0$ ,  $U_p(l_p = 0) = 0, U_p(l_p = 1) = Cst$ ; 而对于最后一列来说, 其必位于焦点区域封闭边界外部, 因而对于  $\forall p \in (r_{max}, 0^\circ \leq \theta \leq 360^\circ)$  必有  $l_p = 1, U_p(l_p = 0) = Cst, U_p(l_p = 1) = 0$ 。而对于其他坐标位置像素, 没有任何关于数据项的知识, 因此统一规定为 0。

式(1)所规定的能量函数, 可用图论中的最大流-最小割法<sup>[14]</sup>进行求解。其中极坐标概率边界图像素之间的连接权值按下式确定:

$$w_{p,q} = \exp(-\beta(I_E^{pol}(r_p, \theta_p) - I_E^{pol}(r_q, \theta_q))^2) \quad (2)$$

在极坐标下获得了焦点区域封闭边界的最优解后, 对该边界像素坐标进行逆极坐标变换, 即得直角坐标下的焦点区域封闭边界的最优解。

从上述焦点区域的确定方法可知, 经  $Q(f)$  最小化所获得的焦点区域, 由于仅仅利用了颜色梯度, 亮度梯度和纹理梯度等等信息, 往往收敛于  $Q(f)$  一个局部极小值, 因而通常是分割区域的一个子集。

图 2 所示为焦点区域的确定过程。图 2(a) 所示为待处理图像, 图 2(b) 中的符号“x”表示视觉注意焦点, 该焦点的位置由 Itti 视觉注意模型获得。以该焦点为坐标原点, 对概率边界图 2(c) 进行极坐标变换, 得到极坐标概率边界图如图 2(d) 所示, 根据式(1)所得到的最优边界如图 2(e) 所示, 对图 2(e) 进行逆极坐标变换得图 2(f), 得到一封闭区域, 即当前焦点区域。实验证明, 采用式(1)绝大多数情况下都能提取到准确的焦点区域, 如图 3 所示。

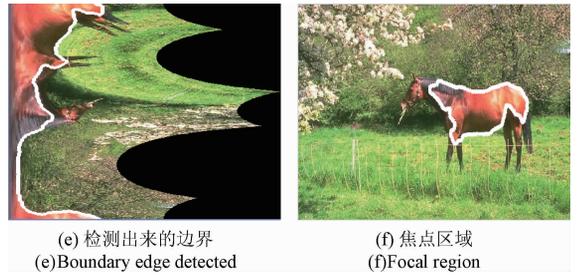
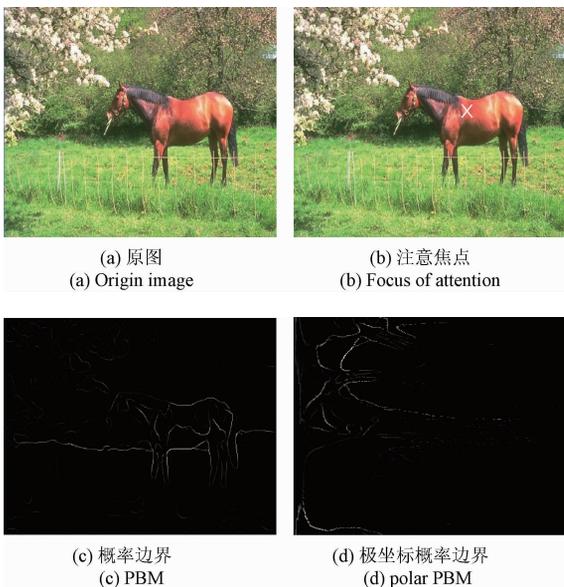


图 2 焦点区域确定

Fig. 2 The determination of focal region

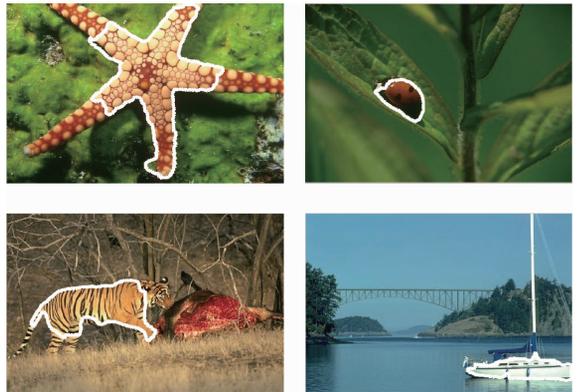


图 3 焦点区域检测例子

Fig. 3 Examples of detection of focal region

从图 2 和 3 可知, 焦点区域通常是分割区域的子集, 为了获得完整的分割区域, 还需要通过焦点区域种子点, 采用基于随机游走的方法分割完成分割。

### 3.3 焦点区域内超像素种子点确定

焦点区域由视觉注意焦点的转移而产生, 有多少个视觉注意焦点, 就有多少个焦点区域; 通过能量函数最小化确定焦点区域后, 区域内有多少个超像素分割区域, 就有多少个种子点。但焦点区域与分割区域之间并非完全对应, 同一分割区域, 可能存在多个视觉注意焦点, 由此产生多个焦点区域, 但这些焦点区域既然同属于一个分割区域, 它们的特征必然具有相似性, 由此表现为这些焦点区域相互交叉, 甚至相同。反之, 若两焦点区域虽然相交, 但是交叉面积较小(实验证明这种情况出现的可能性较小), 如果出现此种情况, 则这两个焦点区域是否属于同一分割区域, 还需根据两区域之间的相似性进行处理。

设有焦点区域  $A_i, A_j, \rho(A_i, A_j)$  为焦点区域  $A_i, A_j$  之间基于 Bhattacharyya 系数的相似性度量<sup>[15]</sup>。Bhattacharyya 系数是基于直方图的相似性统计量, 直方图计算方法是, 将图像区域 R、G、B 三个颜色通道都分成 16 个等级, 即在  $16 \times 16 \times 16 = 4\ 096$  维的空间中计算直方图。设  $H_i, H_j$  分别为  $A_i, A_j$  区域的规范化直方图, 则二区域之间的 Bhattacharyya 系数为:

$$\rho(A_i, A_j) = \sum_{u=1}^{4096} \sqrt{p_u q_u} \quad (3)$$

式中:  $p_u$  表示直方图  $H_i$  的第  $u$  维特征,  $q_u$  表示直方图  $H_j$  的第  $u$  维特征。Bhattacharyya 系数实际上等于计算单位向量  $(\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_m})^T$  与  $(\sqrt{q_1}, \sqrt{q_2}, \dots, \sqrt{q_m})^T$  之间的余弦距离。

设焦点区域  $A_i, A_j$  相互交叉,  $A_i \cap A_j = S$ , 若  $\rho(A_i, A_j) > T_{h_i}$ , 则可将  $A_i, A_j$  直接合并; 若  $\rho(A_i, A_j) < T_{h_i}$ , 则不能合并, 且将  $S$  从  $A_i, A_j$  中的任一方去除。经过调整后的焦点区域数目即为最终分割区域数目。焦点区域内部的超像素即为所属分割区域的种子点。对焦点区域进行处理后, 对焦点区域所属的超像素种子点也需进行重标记。

经过上述方法处理, 在大多数情况下都能获得准确的种子点。由于自然图像的复杂性, 在某些情况下可能产生对焦点区域的错误检测, 如图 4(a) 所示。图 4(a) 用白线框住部分为采用式(1)检测到的焦点区域  $foc1$  和  $foc2$ , 可见  $foc1$  和  $foc2$  包含了其他区域的种子点, 经随机游走后会产生错误分割。在此情况下, 为了获得准确的检测结果, 需要采用更强的相似性判断准则。为此可以采用基于密度的聚类(density-based spatial clustering of applications with noise, DBSCAN)算法<sup>[16-17]</sup>对焦点区域超像素进行聚类处理。超像素之间的邻接关系可直接从 SLIC 超像素分割后所得到的邻接矩阵得到, 而超像素之间的相似关系既需考虑到不同超像素之间颜色的相似性, 还需考虑到不同超像素之间纹理特征的相似性。颜色的相似性可根据图像 Lab 颜色空间中的距离(或 Bhattacharyya 系数)计算; 而对于纹理的相似性来说, 可利用概率边界图的纹理特征, 该纹理特征定义为 64 个纹元在超像素分割区域内统计值。设两个超像素内对纹理基元的直方图统计量分别为  $h_i(k)$  和  $h_j(k)$ , 则两个超像素纹理特征之间的相似性定义为<sup>[16]</sup>:

$$\chi^2(h_i, h_j) = \frac{1}{2} \sum_k \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \quad (4)$$

式中:  $k$  为纹元序号,  $k = 1 \sim 64$ 。为确定  $h_i(k)$ , 首先确定超像素中心, 并以该中心建立适当大小的窗口  $W$ , 然后通过滤波器组对图像滤波生成纹元分配图  $T, 1 < T(i, j) < 64$ , 最后以纹元直方图量化此区域, 表示该区域里某纹元出现的频率。直方图统计量  $h_i(k)$  按下式确定<sup>[13]</sup>:

$$h_i(k) = M_w * [T = k] \quad (5)$$

式中:  $M_w$  表示窗口  $W$  的卷积掩膜。上述两种相似性归一化后, 分别记为  $S_c$  和  $S_r$ , 则总的相似性  $S$  可整合为:

$$S = \omega S_c + (1 - \omega) S_r \quad (6)$$

式中: 参数  $0 < \omega < 1$ , 反映了两种相似性的不同影响。聚类时, 当邻域超像素  $p$  与中心超像素  $s$  之间的相似性大于某阈值时, 则  $p$  可保留, 与中心超像素  $s$  聚为一类, 并作已聚类标记, 否则不能聚类。重复上述过程, 直至所

有的邻域超像素检测完毕; 然后对未标记的超像素从新开始执行上述过程, 直至检测完焦点区域所有像素。该聚类过程不需要事先指定类别数目。聚类完成后, 按上文方法对聚类区域进行合并处理, 得到修正后的准确焦点区域和种子点。

图 4(a) 聚类后的结果如图 4(b) 所示,  $foc1$  聚类为  $reg1, reg2$  两个子区域,  $foc2$  聚类为  $reg3, reg4$  和  $reg5$  三个子区域。即  $reg1$  和  $reg3$  为单独的焦点区域, 而  $reg2, reg4$  和  $reg5$  合并为一新的焦点区域。当多个相似的子区域合并为一新的焦点区域时, 也可仅仅选取其中一个子区域作为焦点区域。

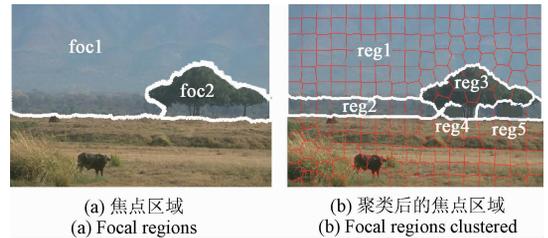
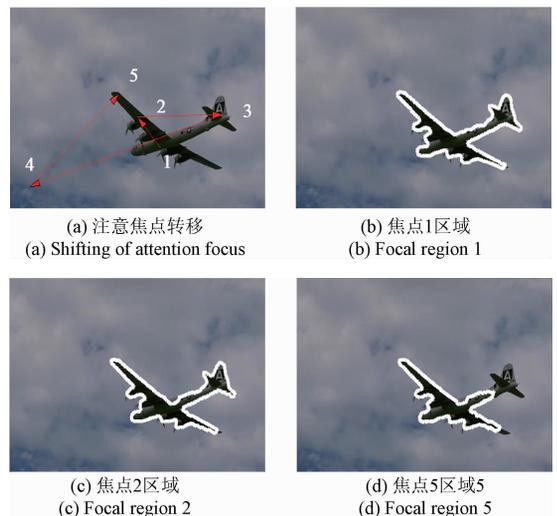


图 4 焦点区域的聚类处理

Fig. 4 Clustering of focal regions

图 5 所示为焦点区域种子点的确定过程。其中图 5(a) 表示根据 Itti 视觉注意模型得到的视觉注意焦点及其转移过程, 从该图可知, 有 5 个注意焦点。图 5(b) ~ (d) 分别表示视觉注意焦点 1、焦点 2 和焦点 5 所对应的焦点区域, 焦点 3 区域与焦点 1 的区域类似。这 4 个焦点区域为重合和包含关系, 根据前述焦点区域调整方法可知, 该图实际上可标记为 2 个区域, 即焦点区域 1、2、3 和 5 标记为区域 1, 焦点区域 4 标记为区域 2。图 5(e) 所示为超像素分割结果, 图 5(f) 给出了重标记后焦点区域 1 白色边界包围的超像素, 即焦点区域 1 的种子点。



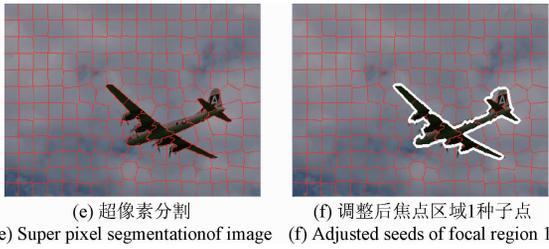


图5 焦点区域种子点确定

Fig.5 Determination of seeds of focal region

## 4 基于超像素种子的随机游走图像分割

随机游走分割就是确定其余未标记超像素到达各个种子点的概率值,并将该超像素归并到概率最大值最大的种子点所属的区域中。为确定未标记超像素到达各个种子点的概率值,随机游走分割算法将图像视为无向加权图。传统的随机游走分割算法以像素作为图节点,随着节点数目的增多,计算开销急剧加大。本文则以超像素作为图节点,使得节点数目大大减小,计算效率显著提高;而各节点之间的连接权值,按 Bhattacharyya 系数进行计算。Bhattacharyya 系数是基于直方图的相似性计算方法,考虑到不相似的两个超像素可能具有相似的直方图,为降低这种情况的发生概率,规定只有相邻的超像素才按 Bhattacharyya 系数计算连接权值,否则为 0。

一个随机游走者首次到达一个种子点的概率恰好等于种子点边界条件下的 Dirichlet 问题的解<sup>[1]</sup>。为实现 Dirichlet 问题的求解,设以待分割图像超像素为节点所构成的图为  $G$ ,首先定义图  $G$  的 Laplace 矩阵:

$$L_{s_i, s_j} = \begin{cases} d_{s_i}, & v_{s_i} = v_{s_j} \\ -\omega_{s_i, s_j}, & v_{s_i} \text{ 和 } v_{s_j} \text{ 相邻} \\ 0, & \text{其他} \end{cases} \quad (7)$$

式中:  $d_{s_i}$  为极点  $S_i$  的度,定义为所有与极点  $S_i$  相连的极点之间的权值之和。

假设最终的焦点区域数为  $K$ ,区域标记变量为  $t, 1 \leq t \leq K$ 。对于某一分割区域来说,图中所有节点分为两类:标记点集合(所有焦点区域内种子点集合)  $V_M$  和未标记点集合  $V_U, V_M \cup V_U = V$  且  $V_M \cap V_U = \Phi$ 。因此,根据节点所属的不同集合, Laplace 矩阵可以写为:

$$L = \begin{bmatrix} L_M & B \\ B^T & L_U \end{bmatrix} \quad (8)$$

再设节点  $v_{s_i}$  到达标记为  $t$  的种子点的概率为  $m_{s_i}^t$ ,对当前焦点区域种子点定义标记函数  $Q(v_{s_i})$ ,从而对所有  $V_M$  中的节点有:

$$m_{s_i}^t = \begin{cases} 1, & Q(v_{s_i}) = t \\ 0, & \text{其他} \end{cases} \quad (9)$$

而对于未标记  $V_U$  中的节点对标记为  $t$  的种子点的概率,根据 Dirichlet 边界条件,可根据下式求解:

$$L_U X = -B^T M \quad (10)$$

最后可根据  $L_U$  中各未标记点到各超像素种子点概率大小决定其归属于哪一分割区域,从而实现图像分割。如果未标记点到达所有种子点的概率都很小,说明这些未标记点与已给定种子点的相似性很小,此时,这些未标记点可单独归为一区域。

## 5 实验

为了说明本文方法的性能,将本文方法与传统交互式图像分割进行对比;其次,本文是一种多标签分割算法,因此,还将本文算法与常规的 Ncut 多标签分割算法<sup>[16]</sup>进行了比较。实验是在 Berkeley 标准图像库上进行的。

### 5.1 实验结果与分析

实验分割结果如图 6 所示;表 1 所示为本文算法与 NCUT 算法以及经典的随机游走分割算法采用概率 rand 指数(probabiliditic rand index, PRI)<sup>[18]</sup>、信息变化指数(variation of information, VOI)<sup>[19]</sup>和全局一致程度(global consistency error, GCE)<sup>[20]</sup>指标进行定量评价的结果。其中, PRI 是指算法产生的分割结果与给定标准分割结果标记相一致性的概率度量, PRI 的取值范围为区间  $[0, 1]$ , 对应的值越大越好; GCE 是衡量一个分割结果可以被看做另一个分割结果的子集程度,取值范围为区间  $[0, 1]$ , 值越小越好; VOI 表示像素点从一个聚类变化到另一个聚类丢失或获得了多少信息, VOI 取值区间是  $[0, \infty)$ , 取值越小越好。

图 6(a)所示的 4 幅图分别为所选图像根据 Itti 视觉注意模型得到的视觉注意焦点及其转移过程;图 6(b)所示为每个注意焦点相对应焦点区域,已根据 3.2 节的所介绍的方法进行了处理(为便于比较,此处显示未聚类处理结果),如 159091 号图按 Itti 模型由 5 个注意焦点,其中第 1 个注意焦点区域和第 2 个注意焦点区域几乎重合,因而被合并,因此,该图焦点区域数目为 4。图 6(c)所示为 Ground truth,从 Berkeley 数据库提供的 4~7 个观察者的分割结果中选择了一种作为作为基准;图 6(d)所示为常规多标签分割算法 NCUT 的分割结果;图 6(e)所示为传统随机游走分割算法的分割结果;图 6(f)所示为本文算法分割结果。NCUT 算法需要人为输入分割区域数目,而本文算法的分割区域数目由算法自动确定。为了便于比较,每幅图像 NCUT 算法的分割数目按本文算法取为一致。传统随机游走分割法的分割区域数也跟本文算法取为一致。

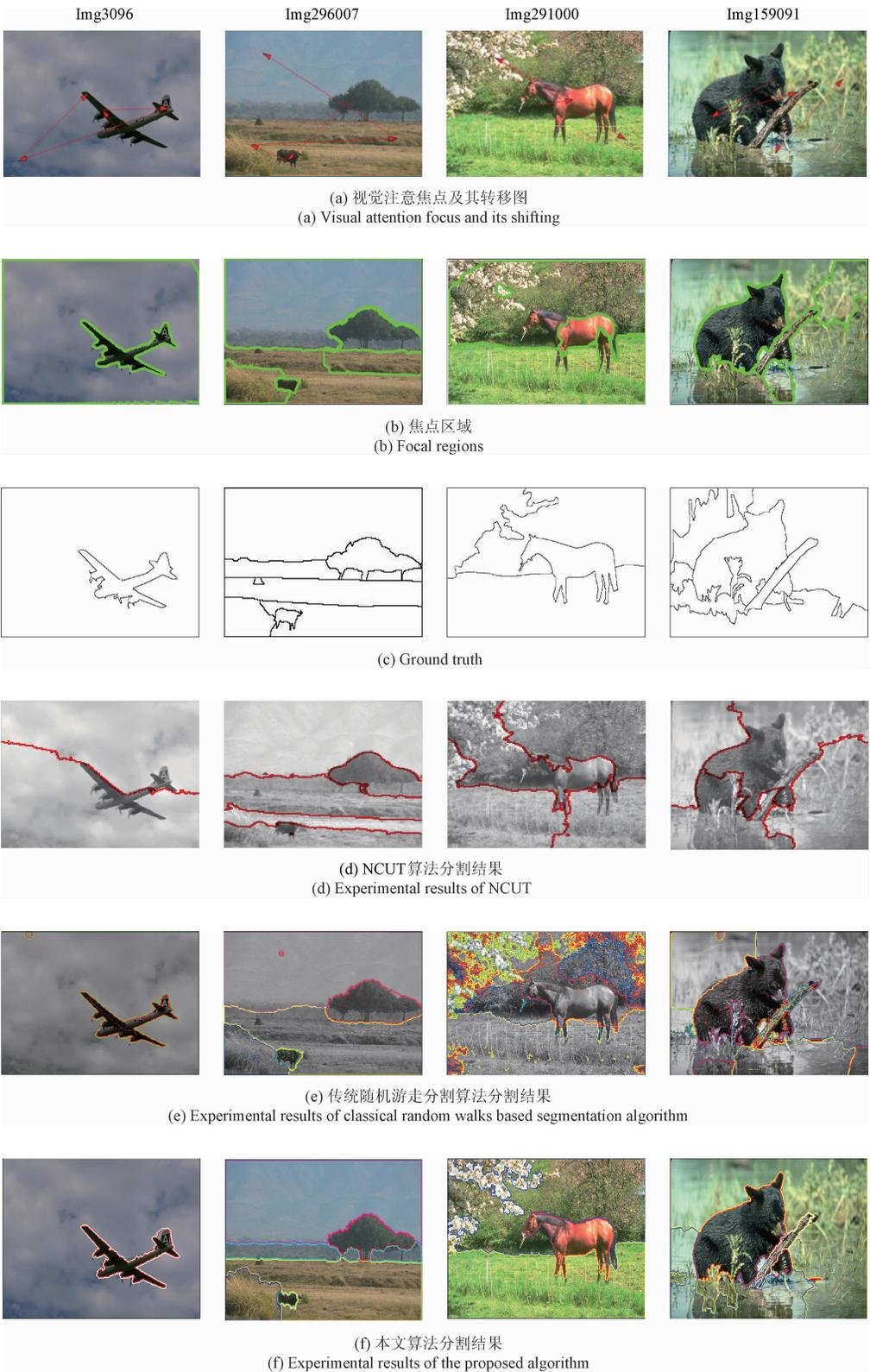


图6 实验结果

Fig. 6 Experimental results

从图6可知,本文算法分割性能明显好于NCut算法和传统的随机游走算法。从图6(d)可以看到,NCut算

法很容易出现过分割,欠分割乃至误分割的现象,难以获得理想的分割结果。因而其 PRI 指标偏低,而 GCE 和 VOI 指标偏高。而由图 6(e) 可知,传统的随机游走图像分割算法,主要适用于图像区域灰度变化平滑的图像,对于区域特性比较简单的图像,分割效果较好,如图 6(e) \_3096 所示,其但对于复杂图像,则难以获得有效分割,如图 6(e) \_291000 和图 6(e) \_159091 所示等。传统随机游走图像分割算法的特性可以从其所采用的高斯权函数得到解释。而本文通过视觉注意机制获取关键分割区域,建立能量函数检测焦点区域,以焦点区域内的超像素为种子点,以超像素之间 Bhattacharyya 系数作为权值进行图上游走实现分割,更能反映复杂图像不同区域的特性,因而能取得更好的分割效果,如图 6(f) 所示。比较图 6(d) ~ (f) 与表 1 可知,图 6 的分割结果与采用 PRI, GCE 和 VOI 指标进行定量评价的结果是一致的。本文方法能取得良好分割效果的内在原因是,焦点区域是通过  $P_b$  算子检测概率边界图获得的,而  $P_b$  算子采用 Logist c 回归模型对 200 幅训练图像进行统计学习,模型参数通过极大似然法进行估计得到,利用了更多的关于图像的知识。由此说明,图像分割是一个复杂的过程,其分割成功与否,往往依赖于和待分割图像的相关知识储备是否充分。

## 5.2 讨论

本文的分割过程包含了 3 个关键环节,即视觉注意模型、焦点区域确定和图网络节点特征,利用这些方法,本文建立了一个以随机游走为基础的分割框架。从 Berkeley 标准图像库中选了大量图像进行试验,试验结果表明所提出的方法大都能取得较为理想的分割效果,但由于本文所采用的  $P_b$  算子仅仅利用了图像的亮度、颜色和纹理的梯度信息,对部分复杂图像,如 296007 号、159091 号、106024 号图像、181079 号图像、130026 图像等,由于图像本身的复杂性,采用式(1)获得的焦点区域存在干扰(如图 6B\_296007),需要采用 3.3 节中的聚类方法处理后,才能获得较为理想的分割效果(见图 6F\_296007),增加了计算消耗。而生物视觉系统往往综合利用单目视觉信息、立体视觉信息与运动信息乃至更多的背景知识来感知外部世界,可以获得复杂区域准确度非常高的边界,从而可以实现对复杂自然图像的准确分割。因此,解决问题的最根本方法还在于更充分地利用多种信息,如在 3.1 节生成概率边界图的过程中,加入深度信息或运动信息,则该问题就可以得到彻底解决。这说明本文方法具有很强的扩充性。

表 1 本文方法与其他方法性能比较

Table 1 Performance comparison of the proposed algorithm with other algorithms

图像	传统随机游走算法			NCUT			本文方法		
	PRI	GCE	VOI	PRI	GCE	VOI	PRI	GCE	VOI
3 096	0.850 5	0.181 2	0.587 2	0.307 2	0.651 4	3.624 7	0.938 6	0.042 1	0.383 1
29 6007	0.662 5	0.361 3	2.513 6	0.742 5	0.353 2	2.626 5	0.801 5	0.158 6	1.683 4
291 000	0.365 3	0.723 5	3.846 5	0.684 7	0.254 6	2.735 9	0.854 5	0.229 5	1.084 6
159 091	0.562 3	0.214 6	2.589 5	0.754 6	0.268 4	1.675 6	0.828 7	0.246 8	1.108 5
130 026	0.647 8	0.423 4	2.674 2	0.656 3	0.584 5	3.354 6	0.835 4	0.147 8	1.024 0
86 016	0.885 6	0.219 8	0.646 5	0.478 9	0.345 7	2.456 1	0.923 4	0.043 6	0.898 5
38 092	0.556 8	0.293 2	3.132 1	0.514 6	0.456 2	2.746 3	0.806 4	0.265 9	1.738 5
197 017	0.564 5	0.584 7	3.456 2	0.632 4	0.341 8	2.224 5	0.789 2	0.325 1	1.234 5
106 024	0.647 5	0.465 9	3.345 6	0.348 7	0.476 3	3.554 6	0.746 2	0.278 8	2.365 4

## 6 结 论

图像分割是与图像理解、场景感知等高层视觉任务紧密相连的关键性基础任务。所有的图像分割算法,都是在解决如下两个问题:需要分割的区域和如何实现分

割。本文根据图像理解、场景感知并不需要图像或者场景的全部信息,而只需要通过视觉注意焦点的转移获取若干关键信息就足以进行感知、理解和识别的原理,在视觉注意的基础上,提出了一种新的多标签随机游走自动图像分割方法。该方法通过视觉注意焦点的转移,来搜寻待分割的关键区域;区域分割的实现乃通过在极坐标

概率边界图中建立能量函数以确定分割区域种子点,最后利用传统的随机游走分割模型实现自动分割。实验证明,本文的分割方法能有效分割复杂自然图像,具有高层视觉与底层视觉相统一的特点。

## 参考文献

- [ 1 ] GRADY L. Random walks for image segmentation [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(11) :1768-1783.
- [ 2 ] ANDREWS S, TANG L, HAMARNEH G, et al. Fast random walker image registration using precomputation [ C ]. Proceedings of the 2011th International Symposium on Biomedical Imaging, 2011 :409-416.
- [ 3 ] GRADY L. Multi-label random walker image segmentation using prior models [ C ]. IEEE Computer Society Conference on Computer Vision and Patter Recognition, 2011 :763-770.
- [ 4 ] FU Y, CHENG J, LI Z, et al. Saliency cuts: An automatic approach to object segmentation [ C ]. IEEE Computer Society International Conference on Pattern Recognition, ICPR, 2008 :1-4.
- [ 5 ] CHENG M M, ZHANG G X, MITRA N J, et al. Global contrast based salient region detection [ C ]. IEEE Computer Society International Conference on Pattern Recognition, 2011 :409-416.
- [ 6 ] MISHRA A K, ALOIMONOS Y, CHEONG L F, et al. Active visual segmentation [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(4) :639-653.
- [ 7 ] QIN CH CH, ZHANG G P, ZHOU Y C, et al. Integration of the saliency-based seed extraction and random walks for image segmentation [ J ]. Neurocomputing, 2014, 129(4) :378-391.
- [ 8 ] 胡正平, 孟鹏权. 基于混合图上随机游走的视觉注意显著目标检测模型 [ J ]. 仪器仪表学报, 2011, 32(7) : 1613-1620.  
HU ZH P, MENG P Q. Visual attention salient object detection model based on random walk in a hybrid graph [ J ]. Chinese Journal of Scientific Instrument, 2011, 32(7) :1613-1620.
- [ 9 ] SALAH A A, ALPAYDIN E, AKARUN L. A selective attention-based method for visual pattern recognition with application to handwritten digit recognition and face recognition [ J ]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(3) :429-425.
- [ 10 ] SOYER C, BOZMA H, ISTEANOPULOS Y. Attentional sequence-based recognition Markovian and evidential reasoning [ J ]. IEEE Transactions on Systems Man & Cybernetics, Part B Cybernetics, 2003, 33 ( 6 ) : 937-950.
- [ 11 ] ACHANTA R, SMITH K, LUCCHI A, et al. SLIC superpixel [ R ]. EPFL Technical Report, 2010.
- [ 12 ] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis [ J ]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1998, 20(11) :1254-1259.
- [ 13 ] MARTIN D R, FOWLKES C C, MALIK J. Learning to detect natural image boundaries using local brightness, color and texture cues [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26 ( 1 ) : 530-549.
- [ 14 ] MISHRA A, ALOIMONOS Y, FAH C L. Active segmentation with fixation [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(4) : 639-653.
- [ 15 ] 冯祖仁, 吕娜, 李良福. 基于最大后验概率的图像匹配相似性指标研究 [ J ]. 自动化学报, 2007, 33(1) : 1-8.  
FENG Z R, LU N, LI L F. Research on image matching similarity criterion based on maximum posterior probability [ J ]. Acta Automatic Sinica, 2007, 33 ( 1 ) : 1-8.
- [ 16 ] 南柄飞, 穆志纯. 基于 SLIC0 融合纹理信息的超像素分割方法 [ J ]. 仪器仪表学报, 2014, 35(3) :527-534.  
NAN B F, MU ZH CH. SLIC0-based superpixel segmentation method with texture fusion [ J ]. Chinese Journal of Scientific Instrument, 2014, 35(3) :527-534.
- [ 17 ] SHI J B, MALIK J. Normalized cuts and image segmentation [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8) : 888-905.
- [ 18 ] UNNIKRIISHNAN R, PANTOFARU C, HEBERT M. A measure for objective evaluation of image segmentation algorithms [ C ]. IEEE Computer Society International Conference on Pattern Recognition, 2005 : 34.
- [ 19 ] MARTIN D, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics [ C ]. IEEE International Conference on Computer Vision. Vancouver, 2001 :416-423.
- [ 20 ] MARTIN D, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring

ecological statistics [C]. IEEE International Conference on Computer Vision, 2001:416-423.

- [21] MEILA M. Comparing clusterings: An axiomatic view[C]. Proceedings of the 22th International Conference on Machine Learning, 2005:577-584.
- [22] ACHANTA R, SHAJI A, SMITH K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11):2274-2281.

## 作者简介



**王富治**, 1998年于四川工业学院本科毕业, 分别在2005年和2010年于电子科技大学获得硕士学位和博士学位, 现为西华大学讲师, 主要研究方向为视觉感知与视觉控制技术。

E-mail: redfires\_789@163.com

**Wang Fuzhi** received his B. Sc. degree from Sichuan Institute of Technology in 1998, and received his M. Sc. and Ph. D degrees both from UESTC in 2005 and 2010, respectively. Currently he is a lecturer in Xihua University. His main research interests include vision perception and vision control technology.



**蒋代君**, 分别在1998年和2004年于东北大学获得学士学位和硕士学位, 2007年于四川大学获得博士学位, 现为西华大学副教授, 主要研究方向为数字图像处理。

E-mail: jiangdaijunlw@sina.com

**Jiang Daijun** received her B. Sc. and M. Sc. degrees both from NEU in 1998 and 2004, respectively, and received her Ph. D. degree in 2010 from Sichuan University. Currently, she is an associate professor in Xihua University. Her main research interest is digital image processing.