

# 实时语义图像分割模型研究

刘尚旺<sup>1</sup>, 陈平<sup>1</sup>, 王统昊<sup>2</sup>

(1. 河南师范大学 计算机与信息工程学院; “智慧商务与物联网技术”河南省工程实验室, 河南 新乡 453007; 2. 郑州大学 软件与应用科技学院, 郑州 450002)

**摘要:** 为了实现快速语义图像分割, 提出一种简化整合模型. 首先, 对频域视觉注意模型 PQFT 的四元数图像虚部系数进行简化改进. 然后, 将改进 PQFT 模型的显著图与简化 PCNN 的内部活动项结合起来对显著目标区域进行粗略定位, 并以提出的显著目标区域均值的 3/2 倍进行精细分割. 最后, 根据“尺寸变化与否”准则判断输出正确的语义图像分割结果. 实验结果表明, 提出的整合模型具有实时性, 且取得的 AUC 值和 F 值较原 PQFT 模型分别提高了 29.9% 和 44.2%.

**关键词:** 语义图像分割; 频域视觉注意模型; 四元数图像; PCNN; AUC

**中图分类号:** TP391.41

**文献标志码:** A

语义图像分割技术是计算机视觉、图像理解和机器人视觉导航等领域中具有挑战性的研究课题. 语义图像分割旨在把图像分割为一个或几个感兴趣的区域, 以便进行深入的图像分析与理解<sup>[1]</sup>. 而人们在观察一幅图像时, 会选择性地将注意焦点放在自己感兴趣的区域, 这与语义图像分割的目的不谋而合. 视觉注意模型是为模拟人类视觉系统的注意机制而提出的模型, 它能获得图像中最容易引起人们注意的显著区域, 从而可应用于图像分割<sup>[2-5]</sup>. 视觉注意模型有空间域和频域视觉注意模型之分. 但大多数空间域视觉注意模型的结构复杂、需要设置的参数多、计算量比较大, 而频域视觉注意模型的结构较简单、需手工可调参数少、计算量小. 因此, 频域视觉注意模型在近年来得到了广泛关注.

2007年, Hou等提出首个频域视觉注意模型——谱残差(Spectral Residual, SR)<sup>[6]</sup>. 2008年, Guo等认为只利用频域中的相位谱信息同样能检测到图像的显著目标, 从而提出相位谱傅里叶变换模型(Phase Spectrum of Fourier Transform, PFT), 接着把其拓展为相位谱四元数傅里叶变换模型(Phase Spectrum of Quaternion Fourier Transform, PQFT)<sup>[7]</sup>. 2009年, Achanta等通过模拟生物视觉神经元的中央兴奋-周边抑制(center-surround, C-S)机制, 提出图像频率自动调节的显著性检测模型, 该模型能输出全分辨且边缘保持良好的显著图, 但是当图像的背景较复杂或者图像中的显著目标较大时, 该方法往往无能为力<sup>[8]</sup>. 为此, Achanta和Süsstrunk在2010年又提出最大对称环绕显著性检测方法(Maximum Symmetric Surround Saliency, MSSS)<sup>[9]</sup>. 2012年, Hou又提出了基于离散余弦变换(Discrete Cosine Transform, DCT)的图像签名显著性检测模型<sup>[10]</sup>. 2013年, Li等提出了超复数傅里叶变换模型(Hypercomplex Fourier Transform, HFT)<sup>[11]</sup>, 该模型本质上仍是对PQFT模型的改进. 通过上述文献分析以及本文的预实验验证, PQFT模型框架值得深入研究<sup>[2,5]</sup>, 其显著性检测结果与人类视觉系统比较接近.

另一方面, 作为第三代人工神经网络的主要代表, 脉冲耦合神经网络(Pulse Coupled Neural Network, PCNN)根据动物的大脑视觉皮层神经元的同步脉冲发放现象提出, 其在图像处理方面具有生物学依据的优势. 另外, PCNN属于单层神经网络, 不需任何训练即可实时处理图像. 由于PCNN图像分割结果与人类视

收稿日期: 2015-06-10; 修回日期: 2016-01-13.

基金项目: 国家自然科学基金(U1304607); 河南省高等学校重点科研项目(15A520080); 河南师范大学博士科研启动基金(qd12138).

第1作者简介(通信作者): 刘尚旺(1973-), 男, 河南新乡人, 河南师范大学副教授, 博士, 研究方向为生物图像处理, E-mail: shwl2012@hotmail.com.

觉系统特点比较吻合,PCNN 经常用于图像分割<sup>[12-15]</sup>.

综上所述,为了进行快速语义图像分割,在深入研究 PQFT 和 PCNN 模型的基础上,本文提出改进 PQFT 与简化 PCNN 整合模型,期望在保证实时语义图像分割的同时,取得较好的语义图像分割效果.改进 PQFT 与简化 PCNN 整合模型的整体工作流程图,如图 1 所示.

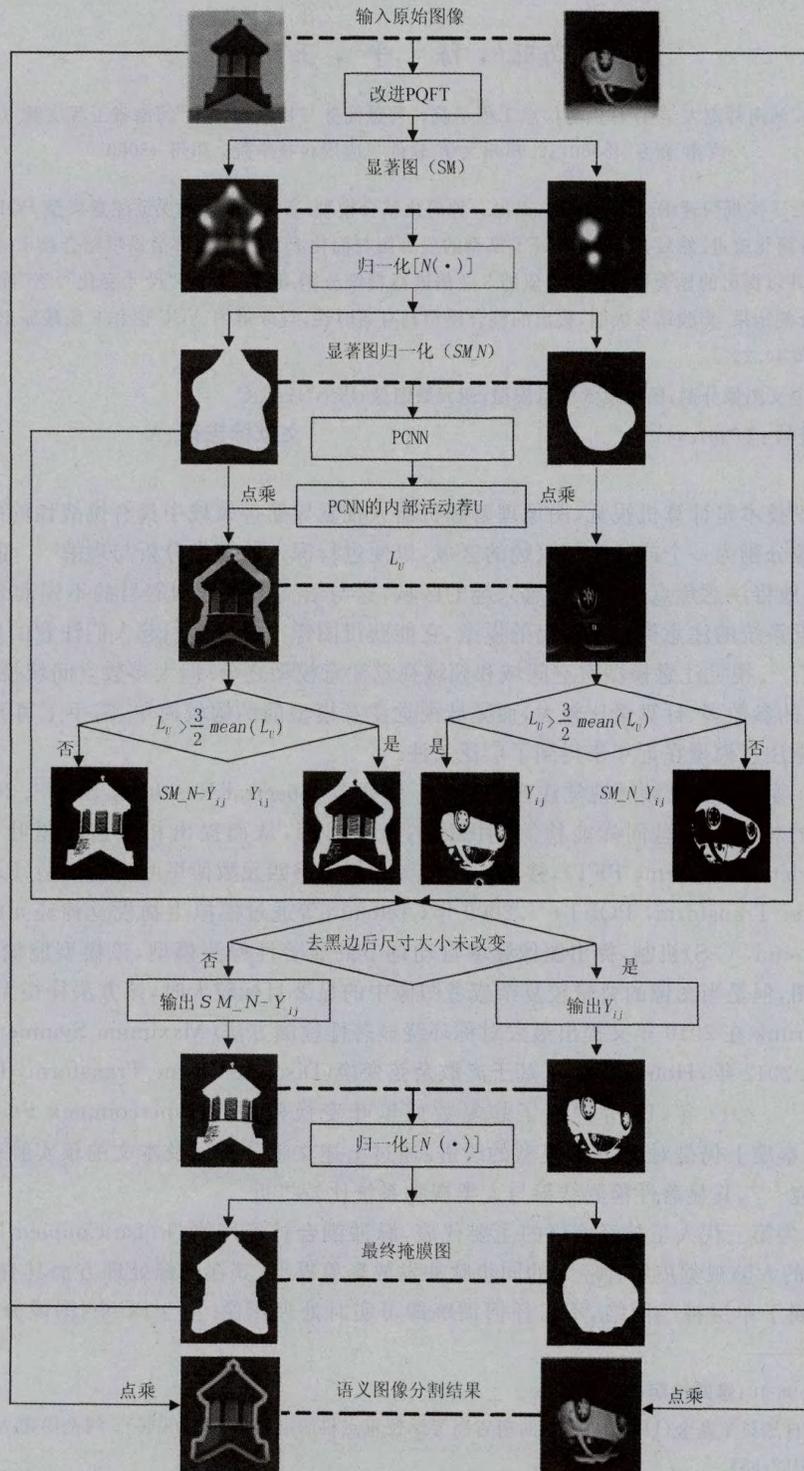


图1 改进PQFT与简化PCNN整合模型的整体工作流程图

从图1可以看出,本文提出的快速语义图像分割模型中的主要步骤方法有:改进PQFT模型,改进PQFT模型与简化PCNN模型的整合,以及正确语义图像分割结果的自动判断与输出。

## 1 改进PQFT模型

### 1.1 PQFT模型

PQFT模型不仅考虑了图像的颜色、亮度特征,还将视频图像帧与帧之间的运动信息考虑在内,来构建显著图。PQFT模型显著图的输出分为两个阶段,第一个阶段是利用输入图像的4种特征信息将其表示成四元数图像;第2个阶段是对四元数图像进行傅里叶变换及其相关处理。

#### 1.1.1 建立输入图像的四元数图像

令 $t$ 时刻输入的视频帧为 $F(t)$ ,  $t = 1, \dots, N$ ,  $N$ 为视频帧的总数。每帧包含红、绿、蓝三个颜色通道信息,分别记为: $r(t)$ ,  $g(t)$ ,  $b(t)$ 。则,4个广义调制颜色通道的计算,如式(1)~(4)所示。

$$R(t) = r(t) - \frac{g(t) + b(t)}{2}, \quad (1)$$

$$G(t) = g(t) - \frac{r(t) + b(t)}{2}, \quad (2)$$

$$B(t) = b(t) - \frac{r(t) + g(t)}{2}, \quad (3)$$

$$Y(t) = \frac{r(t) + g(t)}{2} - \frac{|r(t) - g(t)|}{2} - b(t). \quad (4)$$

在人的某些神经元感受视野里,神经元受到一种颜色的刺激而兴奋,但却会被另一种颜色抑制。人的视觉皮层中有4组颜色拮抗对,它们分别为红/绿,绿/红,蓝/黄,黄/蓝。故,两个颜色特征的计算公式分别为:

$$RG(t) = R(t) - G(t), \quad (5)$$

$$BY(t) = B(t) - Y(t). \quad (6)$$

从而,亮度特征和运动特征,可分别由(7)式和(8)式计算得到:

$$I(t) = (r(t) + g(t) + b(t))/3, \quad (7)$$

$$M(t) = I(t) - I(t - \tau), \quad (8)$$

式中, $\tau$ 代表时延,通常取值为3。

至此,得到了输入图像的4个特征信息:两个颜色特征、一个亮度特征和一个运动特征。于是,原始输入图像的四元数形式可表示为:

$$q(t) = M(t) + RG(t)\mu_1 + BY(t)\mu_2 + I(t)\mu_3, \quad (9)$$

式中, $\mu_i$ 为单位纯四元数, $i = 1, 2, 3$ ,它满足 $\mu_i^2 = -1$ ,  $\mu_1 \perp \mu_2$ ,  $\mu_2 \perp \mu_3$ ,  $\mu_1 \perp \mu_3$ ,  $\mu_3 = \mu_1\mu_2$ 。

进一步,将 $q(t)$ 表示成symplectic形式:

$$q(t) = f_1(t) + f_2(t)\mu_2, \quad (10)$$

式中,

$$f_1(t) = M(t) + RG(t)\mu_1, \quad (11)$$

$$f_2(t) = BY(t) + I(t)\mu_1. \quad (12)$$

#### 1.1.2 四元数图像的傅里叶变换

四元数图像 $q(n, m, t)$ 的傅里叶变换,如(13)式和(14)式所示。

$$Q[u, v] = F_1[u, v] + F_2[u, v]\mu_2, \quad (13)$$

$$F_i[u, v] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-\mu_i 2\pi \left( \frac{u}{M} + \frac{v}{N} \right)} f_i(n, m), \quad (14)$$

式中, $(n, m)$ 和 $(u, v)$ 表示图像像素分别在时间域和频率域对应的位置; $f_i$ ,  $i \in \{1, 2\}$ 则由(11)式和(12)式计算所得。

(14) 式的逆变换为:

$$f_i[n, m] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-\mu_1 2\pi \left( \left( \frac{m}{M} \right) + \left( \frac{n}{N} \right) \right)} F_i[u, v]. \quad (15)$$

对应地, 四元数图像傅里叶变换的极坐标形式为:

$$Q(t) = \|Q(t)\| e^{\Phi(t)} \quad (16)$$

式中,  $\mu$  是单位纯 4 元数.

此时, 令  $\|\Phi(t)\| = 1$ , 以起到抑制频域中的幅度谱冗余信息的作用; 并且  $Q(t)$  中包含有频域的全部相位谱信息.

可利用(15)式来计算  $Q(t)$  的逆变换  $q'(t)$ , 如(17)式所示,

$$q'(t) = a(t) + b(t)\mu_1 + c(t)\mu_2 + d(t)\mu_3. \quad (17)$$

则根据 PQFT 方法的思想, 空间域中显著图  $SM(t)$  的计算公式为:

$$SM(t) = g * \|q'(t)\|^2, \quad (18)$$

式中,  $g$  表示  $\sigma = 8$  的二维高斯滤波器.

上述方法得到的显著图不仅考虑了颜色、亮度特征, 还考虑了运动特征. 输入图像被表示为四元数图像, 使得各个特征被并行处理, 降低了算法的复杂度, 实现了 PQFT 的实时运算, PQFT 模型还具有不依赖任何参数的选取和先验知识信息的优点. 若将运动通道设置为 0, 即可用于静态自然图像的处理. 从 PQFT 模型的预实验显著图结果来看, PQFT 模型能较完整地检测出输入图像中的显著目标; 但是, 检测出的显著目标对象不够平滑, 另外不少背景点也被检测出来. 针对该缺点, 本文对其进行改进.

## 1.2 改进 PQFT 模型

PQFT 模型之所以检测出大量的背景像素点, 究其原因在于该模型将全部相位谱信息毫无选择地保留下来, 导致过多冗余低频信息参与视觉计算. 因此, 本文从减少图像冗余低频信息的角度出发, 对 PQFT 模型进行改进. 四元数图像傅里叶变换的结构框架值得借鉴, 而决定该框架显著性检测性能好坏的因素在于其四元数图像 3 个虚部系数的选择<sup>[5]</sup>. 所以, 本文具体从这 3 个特征系数的选择上对 PQFT 模型进行改进.

依据人类视觉系统中神经元的 C-S 机制, 本文在 CIE Lab 颜色空间上, 选择去掉冗余低频信息的  $L$ ,  $a$  和  $b$  这 3 个通道作为四元数图像的虚部系数. 这 3 个特征通道的计算, 如(19)~(21)式所示.

$$X_1 = L - L_1, \quad (19)$$

$$X_2 = a - a_1, \quad (20)$$

$$X_3 = b - b_1, \quad (21)$$

式中,  $X_1, X_2$  和  $X_3$  分别用来代替 PQFT 模型中  $RG(t), BY(t)$  以及  $I(t)$ ;  $L, a$  和  $b$  是 CIE Lab 颜色空间 3 个特征通道; 若  $X_1, X_2$  和  $X_3$  为负值, 则将其设为 0. 因为对于真实的自然图像来讲, 值为负的像素点表示它是亮度较低或者颜色较暗的像素点, 一般不为人类视觉注意;  $L_1, a_1$  和  $b_1$  是对应特征通道像素点周围区域的均值. 对于如何确定周围区域大小, 文中简单地利用将图像的宽和高与 PCNN 链接强度系数  $\beta$  相乘的方法; 其中,  $\beta$  值的计算, 如(22)式所示.

$$\beta = 1 / \sqrt{STD_{ij} + 1}, \quad (22)$$

式中,  $STD_{ij}$  表示像素点  $I(i, j)$  的灰度值与其 8-邻域灰度值的标准方差. 这样, 标准方差越大, 说明周边像素对当前像素点的影响越小, 从而其周围区域也就越小, 计算出来的周围区域均值就愈科学合理, 反之亦然.

上述四元数虚部系数和周边区域的确定方法不仅可去掉图像中部分冗余低频信息, 而且还可突出图像边缘等高频信息, 亦即还有效利用了少量频域中幅度谱的显著信息; 这就达到了改进 PQFT 模型在均匀检测到整个显著目标的同时, 能够有效去除背景冗余低频信息的预期目标.

## 2 改进 PQFT 与简化 PCNN 整合模型

### 2.1 简化 PCNN

鉴于 PCNN 的基本原型有很多局限性, 如手工可调参数多、参数的选择和系数的确定无规律、网络常量

系数难以确定等,实际图像处理时,经常将其进行简化. 本文根据语义图像分割的需要,采用如下简化 PCNN 模型<sup>[16-17]</sup>,如(23)~(27)式所示.

$$F_{ij}[n] = I_{ij}[n], \tag{23}$$

$$L_{ij}[n] = \sum_{k, l \in N(i, j)} W I_{kl}, \tag{24}$$

$$U_{ij}[n] = F_{ij}[n] \{1 + \beta L_{ij}[n]\}, \tag{25}$$

$$T_{ij}[n] = \begin{cases} T_{ij}[n-1], & t = n-1, \\ \exp[-\alpha_T] T_{ij}[n-1], & n-1 < t \leq n, \end{cases} \tag{26}$$

$$Y_{ij}[n] = \begin{cases} 1, & U_{ij}[n] > T_{ij}[n-1], \\ 0, & \text{其他}. \end{cases} \tag{27}$$

式中,  $n$  是迭代次数;  $t$  为运行时间,以迭代次数为单位;  $N(i, j)$  代表神经元  $(i, j)$  的邻域;  $F_{ij}[n]$  表示在  $(i, j)$  位置的神经元第  $n$  次迭代的主输入;  $I_{ij}[n]$  是外界输入的激励信号,也即  $(i, j)$  位置的图像像素的灰度值;  $L_{ij}[n]$  为链接输入;  $U_{ij}[n]$  是 PCNN 的内部活动项;  $T_{ij}[n]$  为动态阈值;  $Y_{ij}[n]$  为 PCNN 的神经元输出;  $\beta$  是链接权值;  $W$  为链接强度矩阵;  $\alpha_T$  是阈值衰减常数.

### 2.2 改进 PQFT 与简化 PCNN 整合

改进 PQFT 与简化 PCNN 整合进行快速语义图像分割的模型框架图,如图 2 所示.

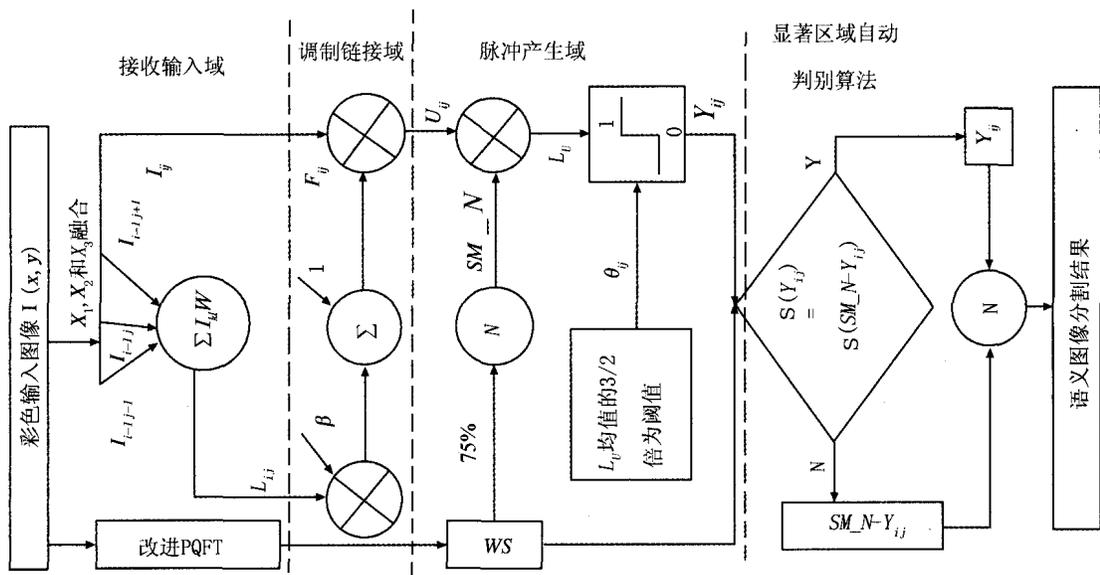


图2 改进PQFT与简化PCNN整合模型的框架图

从图 2 可以看出,该整合模型由 4 部分构成:接收输入域、调制链接域、脉冲产生域和显著区域自动判别算法.

#### 2.2.1 接收输入域

对于给定彩色输入图像,将  $X_1$ 、 $X_2$  和  $X_3$  的特征融合图  $(I_{ij})$  作为 PCNN 的输入图像,这样不仅可以增强输入图像中的显著对象,而且可以提高 PCNN 的抗噪能力. PCNN 的主输入  $(F)$  和链接输入  $(L)$  分别用来接收外界刺激和局部刺激,分别如(28)式和(29)式所示.

$$F_{ij}[1] = I_{ij}, \tag{28}$$

$$L_{ij}[1] = \sum_{k, l \in N(i, j)} I_{kl} w, \tag{29}$$

式中,  $[1]$  表示 PCNN 只需要迭代 1 次;  $N(i, j)$  表示外界输入刺激  $(I_{ij})$  的 8-邻域;  $W$  是链接权重矩阵,利用它能将外界刺激的 8-邻域  $(I_{kl})$  像素值按权重大小链接起来,  $W$  可根据图像当前像素到其 8-邻域边界上像素的距离自动设置,如(30)式所示,

$$W_{ijkl} = 1/[(i-k)^2 + (j-l)^2]. \quad (30)$$

从(30)式可以看出,距离当前像素点越远的像素点的权重值越小,越近的像素点的权重值越大,这完全符合生物视觉神经元的中央兴奋一周边抑制机制.按照(30)式,就8-邻域而言,链接权重矩阵  $W = [0.5 \ 1 \ 0.5; 1 \ 0 \ 1; 0.5 \ 1 \ 0.5]$ .

### 2.2.2 调制链接域

在调制链接域,对被接收进来的主输入和链接输入进行全局调制耦合,如(31)式所示.

$$U_{ij}[1] = F_{ij}[1]\{1 + \beta L_{ij}[1]\}, \quad (31)$$

式中, $U$ 表示PCNN的内部活动项; $\beta$ 代表神经元之间的链接强度系数,可由(22)式自适应设置,其值越大,表示在8-邻域内的神经元对其中心神经元影响越大,反之亦然.

### 2.2.3 脉冲产生域

视觉注意模型输出的显著图,常对其进行等级二值化<sup>[7]</sup>,对模型得到的显著图,根据其显著性大小顺序,取其前75%的像素作为二值图像分割结果,反复进行提取和覆盖,用以拟合返回抑制机制.本文为了增强图像分割的泛化性能,用显著图的统计直方图的峰谷点数来表示给定图像中的显著对象个数,而峰谷之间的像素点数目占的比重作为返回抑制机制选择像素数的依据.

首先,根据上述机制对改进PQFT模型得到的显著图进行显著像素点的截取和输出,并进行归一化( $N(\cdot)$ )操作,生成二值图像  $SM\_N$ .

然后,将  $U$  和  $SM\_N$  进行点乘得到显著目标对象的粗略定位结果.

$$L_U = U_{ij}[1] \cdot SM\_N, \quad (32)$$

式中, $L_U$ 是改进PQFT模型在PCNN内部活动项中对显著目标对象的粗分割结果.

最后,利用  $L_U$  均值的  $3/2$  倍作为阈值来对  $L_U$  进行精细分割,如(33)式所示.

$$Y_{ij}[1] = \begin{cases} 1, & L_U > \frac{3}{2}M(L_U), \\ 0, & \text{其他.} \end{cases} \quad (33)$$

式中, $M(L_U)$ 表示  $L_U$  的均值.

### 2.2.4 显著区域自动判别算法

精细分割把  $L_U$  分成了  $Y_{ij}$  和  $SM\_N - Y_{ij}$  两幅不同图像,如图1所示.

如图1所示,为了确定选择哪一个图像作为最终正确的语义图像分割结果,文中根据“尺寸变化与否”准则来自动判断并输出正确的语义图像分割结果:利用一个两层循环嵌套程序实现自动去除  $SM\_N$  及其对应  $Y_{ij}$  的黑色边缘部分,然后判断去掉黑色边缘部分后二者的尺寸大小是否相等.若相等,则所需输出的正确语义图像分割结果为  $SM\_N - Y_{ij}$ ;否则,正确语义图像分割结果为  $Y_{ij}$ .

## 3 实验结果与分析

在整合模型的验证试验中,采用的测试数据集为 Achanta 等在2009年提供的包含1000张图像的图像集<sup>[9]</sup>.在性能指标的验证试验中,采用最常用的受试者工作特征曲线(Receiver operating characteristics, ROC)下的面积(Area under curve, AUC)、召回率(Recall)、查准率(Precision)以及  $F$  值等数字化指标.其中,AUC的值越大表明其模型性能越好, $F$  值是召回率和查准率的综合体现,表明图像分割结果的一致性或光滑性.

### 3.1 视觉效果实验结果

改进PQFT模型与简化PCNN整合模型和其他相关模型的对比实验结果,如图3所示.

从图3的实验结果来看,本文整合模型能够忽略背景点,并且能较准确地分割出图像中的显著对象.虽然分割出的显著目标的形状特征存在过分割现象,但与其他模型相比,语义分割结果和手工标注的 Ground-truth 图最为接近.究其原因,整合模型能够准确地分割出显著目标对象应归功于:PCNN良好的图像分割能力、四元数图像的虚部系数和均值阈值的合理选择.

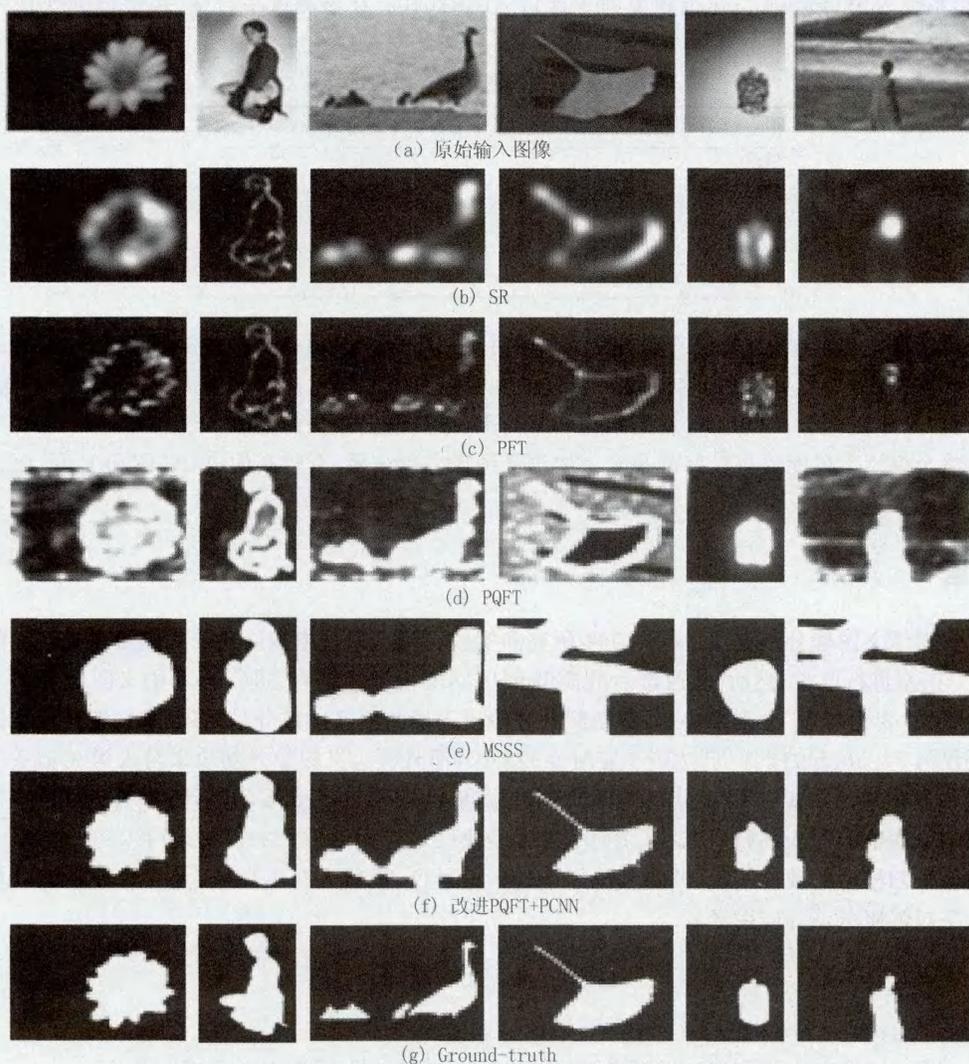


图3 相关模型语义图像分割结果与标注图的比较

### 3.2 性能指标实验结果

相关模型的平均性能指标的对比,如表 1 所示.

表 1 改进 PQFT 模型与 PCNN 整合模型的性能评价指标

模型	性能指标			
	平均 AUC 值及其均方差( $\sigma$ )	平均召回率/%	平均查找率/%	平均 F 值及其均方差( $\sigma$ )
SR	0.6249(0.077 50)	44.15	31.49	0.42(0.021)
PFT	0.5958(0.0822)	87.39	34.63	0.42(0.018)
MSSS	0.6785(0.0710)	86.36	47.61	0.59(0.016)
PQFT	0.6230(0.0753)	87.9	42.57	0.43(0.013)
本文整合模型	0.8095(0.0704)	87.79	62.97	0.62(0.012)

从表 1 可以看出,本文方法的 AUC 值、平均召回率、平均查找率以及  $F$  值均取得最大值.本文整合模型和原 PQFT 模型相比,其 AUC 值提高了 29.9%,均方差反而降低了 6.5%,这表明整合模型具有较好的语义图像分割能力;整合模型的  $F$  值比 PQFT 提高了 44.2%,均方差也有所降低,这意味着整合模型的图像分割结果有较好的一致性和光滑性.这些性能指标实验结果,也从侧面验证了视觉效果实验结果的正确性.

### 3.3 运行时间

测试平台:图像数据集为 Achanta 中 1000 幅图像(图像分辨率为  $400 \times 500$ );CPU 为 Intel Core 2;内

存:4 G;操作系统为 Windows 7;运行环境为 MATLAB R2010a. 在该测试平台下,相关模型处理一幅图像的平均时间,如表 2 所示.

表 2 相关模型的平均运行时间比较

模型	平均时间/ms
SR	24.49
PFT	22.53
MSSS	3936.28
PQFT	24.66
改进 PQFT+PCNN	33.90

从表 2 可以看出,整合模型处理一幅图像的平均时间相比 MSSS 模型减少很多,仅为 33.90 ms,完全能够满足实时处理的要求.究其原因,这应归因于整合模型中的频域视觉注意模型 PQFT 模型、简化 PCNN、均值阈值以及“尺寸变化与否”准则等均是结构非常简单、运算速度极快的缘故.

本文的整合模型不仅能取得较好的视觉上的语义图像分割结果、有较高的 AUC 值和 F 值,而且完全能够满足实时性的要求.

## 4 结 语

本文对实时语义图像分割方法进行了初步研究和实践:首先,通过模拟简化的人类视觉神经元的 C-S 机制,对 PQFT 模型进行改进;然后,将改进后的模型和 PCNN 模型相结合,进行自动语义图像分割.最后,通过所提出的“尺寸变化与否”准则来快速、自动判断并输出正确的语义图像分割结果.本文整合模型的语义图像分割速度仅为 33.90 ms/图像,已达到了实时性的要求,而且其语义图像分割效果与人类视觉系统更加吻合:AUC 值和 F 值较原 PQFT 模型分别提高了 29.9%和 44.2%.更为重要的是,随着本文对该整合模型精细分割中的均值阈值的深入细致研究,本文模型各个模块已均能用硬件来加以实现.接下来,将重点研究该快速语义图像分割模型的硬件实现,并将其应用于行车记录仪、机器人自主导航、无人驾驶系统以及视频监控系统的数字矩阵服务器等应用场景.

## 参 考 文 献

- [1] 肖小龙. 基于区域的图像分割及视觉显著性检测算法研究[D]. 上海:华东理工大学,2014.
- [2] 郭文涛,王文剑,白雪飞. 基于视觉注意的 SVM 彩色图像分割方法[J]. 计算机工程与应用,2011,47(36):174-176.
- [3] Zhang L, Shen Y, Li H. VSI: a visual saliency-induced index for perceptual image quality assessment[J]. IEEE Transactions on Image Processing,2014,23(10):4270-4281.
- [4] Roux L, Buzsáki G. Tasks for inhibitory interneurons in intact brain circuits[J]. Neuropharmacology,2015,88:10-23.
- [5] 梁丹. 基于视觉注意机制及区域生长的图像分割方法研究[D]. 杭州:浙江大学,2013.
- [6] HOU X D, ZHANG L Q. Saliency detection: a spectral residual approach[C]. 2007 IEEE Conference on Computer Vision and Pattern Recognition,2007.
- [7] GUO C L, ZHANG L M. A novel multi-resolution spatiotemporal saliency detection model and its applications in image and video compression[J]. IEEE Transactions on Image Process,2010,19(1):185-198.
- [8] ACHANTA R, ESTRADA F, WILSON P, et al. Frequency tuned salient region detection[C]. Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, Miami,2009.
- [9] ACHANTA R, SÜSSTRUNK S. Saliency detection using maximum symmetric surround[C]. Proceeding of IEEE International Conference on Image Processing, Hong Kong,2010.
- [10] HOU X D, HAREL J, Koch C. Image signature: highlight sparse salient regions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,34(1):194-201.
- [11] LI J, LEVINE M, An X J, et al. Visual saliency based on scale space analysis in the frequency domain[J]. IEEE Transactions Pattern Analysis and Machine Intelligence,2013,35(4):996-1010.
- [12] 沈艳,张晓明,韩凯歌,等. PCNN 图像分割技术研究[J]. 现代电子技术,2014,37(2):38-41.
- [13] 邓翔宇,马义德. 基于 PCNN 的自动图像分割[J]. 兰州工业学院学报,2014,21(3):1-7.
- [14] 黄国峰. 基于脉冲耦合神经网络的图像分割算法研究[D]. 西安:西安电子科技大学,2013.

- [15] 李海燕,张榆锋,施心凌,等. 基于灰度迭代阈值脉冲耦合神经网络的图像分割[J]. 计算机应用,2011,31(10):2753-2756.
- [16] Hout M C, Goldinger S D. Target templates: the precision of mental representations affects attentional guidance and decision-making in visual search[J]. *Attention Perception & Psychophysics*,2015,77(1):128-149.
- [17] 刘尚旺. 基于改进选择性视觉注意模型的语义图像分割研究[D]. 杨凌:西北农林科技大学,2012.

## Research On Real-time Semantic Image Segmentation Model

LIU Shangwang<sup>1</sup>, CHEN Ping<sup>1</sup>, WANG Tonghao<sup>2</sup>

(1. College of Computer and Information Engineering; Engineering Lab of Intelligence Business & Internet of Things, Henan Normal University, Xixiang 453007, China; 2. College of Software, Zhengzhou University, Zhengzhou 450002, China)

**Abstract:** Aiming at addressing fast semantic image segmentation, a simple integrated model was proposed. Firstly, the PQFT model, a frequency-domain visual attention model, was improved by improving imaginary coefficients of its quaternion image. Then, the saliency map of the improved PQFT model was integrated with the inner activity of a simplified PCNN to locate the raw salient region, and the detected salient object was segmented perfectly according to the proposed 3/2 times mean-value threshold method. At last, the accurate semantic image segmentation result was output according to the size-changing rule. The experimental results show that the fast semantic image segmentation model proposed is of real-time, and its AUC and F values have been increased 29.9% and 44.2%, respectively.

**Keywords:** semantic image segmentation; frequency-domain visual attention model; quaternion image; PCNN; AUC