

## 一种新的像素级多聚焦图像融合算法

吴艳<sup>①②</sup> 刘重阳<sup>②</sup> 廖桂生<sup>①</sup>

<sup>①</sup>(西安电子科技大学雷达信号处理国家重点实验室 西安 710071)

<sup>②</sup>(西安电子科技大学电子工程学院 西安 710071)

**摘要:** 该文在小波变换的基础上提出了一种将一维自组织特征映射(SOFM)网络和进化策略相结合的多聚焦图像融合算法。该方法对不同聚焦点的图像进行冗余小波分解,再分别将其各方向、各尺度的高频信息进行叠加,并在高频信息叠加层上提取反映图像清晰度差异的归一化特征图,依据此特征图,使用 SOFM 网络对原始图像像素进行分类,并利用进化策略对各类像素求出最优的融合系数。实验结果表明该算法比拉普拉斯变换法和小波变换法具有更好的融合效果。

**关键词:** 图像融合; 小波分解; 多聚焦; 自组织特征映射网络; 进化策略

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2007)12-2800-05

## A New Pixel-Level Multi-focus Image Fusion Algorithm

Wu Yan<sup>①②</sup> Liu Chong-yang<sup>②</sup> Liao Gui-sheng<sup>①</sup>

<sup>①</sup>(National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China)

<sup>②</sup>(School of Electronics Engineering, Xidian University, Xi'an 710071, China)

**Abstract:** A new fusion method for fusing two spatially registered multi-focus images is proposed in this paper. It is based on multi-resolution wavelet decomposition, Self-Organizing Feature Map (SOFM) neural networks and Evolutionary Strategies (ES). First, a normalized feature image, which represents the local region clarity difference of two source images, is extracted by redundant wavelet transform, then the feature image is clustered by SOFM learning algorithm and every pixel pair in source images is classified into a certain class which indicates different clarity differences. Finally, to each pixel pair in different classes, different fusion factors are used to fuse it; these fusion factors are determined by evolution strategies to achieve the best fusion performance. Experimental results show that the proposed method outperforms the Laplace transform and wavelet transform methods.

**Key words:** Image fusion; Wavelet transform; Multi-focus; SOFM neural networks; Evolution strategies (ES)

### 1 引言

多传感器图像融合是指将多个传感器采集的关于同一目标或场景的图像数据根据某个算法进行适当的综合处理,产生一幅新的、满足某种需求的新图像<sup>[1,2]</sup>,而这一图像是单传感器无法得到的。从不同传感器获得的图像信息之间具有冗余性和互补性,经融合处理后的图像可以更全面、更精确地反映研究对象,因此,图像融合技术近年来已被广泛地应用于自动目标识别、计算机视觉、医学图像处理等领域。

本文主要研究多聚焦图像融合问题,即对同一景物,在成像条件相同情况下,使聚焦目标不同的多个图像通过图像融合得到一个目标都聚焦清晰的融合图像。最早人们提出了像素算术平均的方法,这种方法的缺点是融合图像的对比度很差,为了克服这一问题,人们又提出了基于金字塔的图像融合方法,其中包括拉普拉斯金字塔、梯度金字塔、比率低

通金字塔等<sup>[3-5]</sup>,80年代中期发展起来的小波变换技术<sup>[6-8]</sup>为图像融合提供了新的工具,尽管小波变换法的融合效果相对比较理想,但其分解重构实际是一个高通和低通滤波的过程,仍在一定程度上丢失了原始图像中的一些边缘信息,融合图像出现振铃效应,为此本文提出了一种新的图像加权融合方法,该方法利用冗余小波分解从聚焦区域不同的原始图像中提取反映图像清晰差异的特征,并依据此特征,使用自组织特征映射网络将原始图像中的像素进行分类,对属于各个类的像素再利用进化策略求出各自的最优融合系数,最后加权融合原始图像。仿真结果表明本文提出的融合算法比拉普拉斯变换法和小波变换法具有更好的效果。

### 2 基于 SOFM 网络和进化策略的图像融合算法

假设原始图像为  $A_1f$ , 则  $A_1f$  可分解到分辨率  $2^{-J}$  下,也就是说,  $A_1f$  可以用分辨率  $2^{-J}$  下的 1 个近似和  $3J$  个差值来表示:

$$(A_{2^{-J}}f, (D_{2^j}^1f)_{-J \leq j \leq -1}, (D_{2^j}^2f)_{-J \leq j \leq -1}, (D_{2^j}^3f)_{-J \leq j \leq -1}) \quad (1)$$

在图像的小波分解中,  $A_{2^j}f$  是图像的低频部分,  $D_{2^j}^1f$  是图

2006-05-15 收到, 2006-09-30 改回

国家自然科学基金(60402038)和中国博士后科学基金(J63104020156)资助课题

像在垂直方向上的高频部分,  $D_2^2 f$  是图像在水平方向上的高频部分,  $D_2^3 f$  是图像在对角方向上的高频部分, 具体参见文献[9]。

## 2.1 特征提取

在多聚焦图像融合中, 首先将不同聚焦点的图像进行 3 层小波分解, 图像经 3 层小波分解后生成一组多方向、不同分辨率的子图。实验表明, 这些子图已包含图像的绝大多数高频信息, 它们直接与图像清晰度相关。为了使一系列小波分解子图像与原始图像大小相同, 在小波分解时没有进行下采样, 这样产生的小波系数将是冗余的, 在此基础上再计算反映原始图像清晰度差异的特征图, 此时提取的特征对图像融合是有帮助的。

假设两幅聚焦点不同的原始图像分别为  $f_A$ ,  $f_B$ , 经过 3 层无下采的小波分解之后, 各图像的高频信息叠加  $Df_A$ ,  $Df_B$  可定义为

$$Df(n, m) = \sum_{-3 \leq j \leq -1} \left[ |D_2^1 f(n, m)| + |D_2^2 f(n, m)| + |D_2^3 f(n, m)| \right] \quad (2)$$

在高频信息叠加层上, 分别计算出两幅图像对应的局部区域能量  $E_A$ ,  $E_B$ , 即

$$E(n, m) = \sum_{n' \in l, m' \in k} [Df(n+n', m+m')]^2 \quad (3)$$

式中  $E(n, m)$  表示以  $(n, m)$  为中心位置的局部区域能量;  $l, k$  定义了局部区域的大小(如  $3 \times 3, 5 \times 5, 7 \times 7$  等);  $n', m'$  的变化范围在  $l, k$  内。

本文用  $E_f g(n, m)$  表示原始图像清晰度差异的归一化能量特征图, 它定义为

$$E_f g(n, m) = E_A(n, m) / (E_A(n, m) + E_B(n, m)) \quad (4)$$

$E_f g(n, m)$  与原始图像有着相同的大小, 在空间上与原始图像  $f_A$ ,  $f_B$  相对应, 它客观地反映了原始图像清晰度的差异: 即在  $f_A$  比  $f_B$  清晰的区域中,  $f_A$  的细节特征更丰富, 小波高频系数能量较强,  $E_f g(n, m)$  的值接近于 1; 反之, 在  $f_B$  比  $f_A$  清晰的区域中,  $E_f g(n, m)$  接近于 0; 在  $f_A$ ,  $f_B$  清晰度相同的区域  $E_f g(n, m) = 0.5$ 。

## 2.2 像素分类

原始图像中  $f_A$  比  $f_B$  清晰,  $f_A$  比  $f_B$  模糊以及  $f_A$  与  $f_B$  清晰程度相近的 3 个区域, 分别对应特征图  $E_f g(n, m)$  中灰度值大, 灰度值小, 灰度值适中的 3 个区域, 而  $E_f g(n, m)$  的灰度直方图也相应地呈现出 3 个峰值(见图 1)。依照灰度值对  $E_f g(n, m)$  中的像素进行分类, 即可把原始图像的像素按清晰度差异分为  $f_A$  比  $f_B$  清晰,  $f_A$  与  $f_B$  清晰程度相近和  $f_A$  比  $f_B$  模糊 3 类。

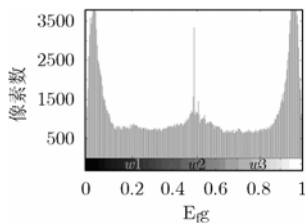


图 1  $E_f g(n, m)$  的直方图

本文采用一维自组织特征映射神经网络(SOFM)<sup>[10]</sup>对  $E_f g(n, m)$  分类。Kohonen提出的自组织特征映射网络是一种具有侧向联想能力的两层结构网络, 它是一种竞争式学习网络, 在学习中能无监督地进行自组织学习, 即在无教师监督的情况下, 通过对输入模式的自组织学习, 在竞争层将分类结构表示出来。这种网络分为输入层和输出层两层, 输入层神经元数目与输入向量维数相等, 输出层是竞争层, 其神经元采用一维的排列形式, 神经元数等于分类数。每个输入层神经元  $i$  与输出层神经元  $j$  间( $j = 1, \dots, c$ ,  $c$  是分类数)通过权值  $w_{ij}$  相连接, 每个输出神经元都有一个相应的邻域。训练结束后, 一维自组织网络的连接权值就是聚类的中心。由图 1 可见, 特征图像的灰度直方图有 3 个峰值, 即其灰度值集中在 3 个区域中, 据此, 网络输出节点数可选为 3, 即分类数  $c$  为 3。网络输入集  $\{x_1, x_2, \dots, x_n\}$  为  $E_f g(n, m)$  中各个像素的灰度值。网络采用 SOFM 算法确定权值, SOFM 算法可概括如下:

- (1) 初始化, 从输入节点到输出节点的连接权值  $w_j$ ,  $j = 1, \dots, c$  可取为一组较小的随机数, 输出节点数  $c$  为 3, 因为分类数  $c$  较小, 原始输出节点邻域距离确定为 1。
- (2) 提供一个输入模式。
- (3) 计算每个输入样本  $x_k$  与每个输出神经元  $j$  之间的欧氏距离  $\xi_j$ 。

$$\xi_j = \sum_{i=1}^n (x_i - w_{ij})^2$$

(5)

- (4) 选取具有最小距离  $\xi_j^* = \min_{j=1-c} \{\xi_j\}$  的输出神经元  $j^*$ 。

- (5) 按式(6)修改权值:

$$w_{ij}(k+1) = \begin{cases} w_{ij}(k) + \eta(t)[x(k) - w_{ij}(k)], & j^* \in N_q(k) \\ w_{ij}(k), & j^* \notin N_q(k) \end{cases} \quad (6)$$

$N_q(k)$  是  $j^*$  的邻域,  $0 < \eta(t) < 1$ , 随时间增长下降到零,  $t$  是迭代次数。

- (6) 如果增益项的值已减小至 0, 则结束; 否则转换至第(2)步。

## 2.3 融合规则

在图 1 中, 归一化特征图  $E_f g(n, m)$  经过 SOFM 网络分类后,  $E_f g(n, m)$  中对应源图像中清晰度差异不同的 3 类区域的中心分别为  $\omega_1, \omega_2, \omega_3$  ( $\omega_1, \omega_2, \omega_3$  为一维 SOFM 网络的权值), 采用不同的融合权值处理这 3 类区域中的像素, 可以使融合效果最佳。当原始图像  $f_A$  的权值被确定为  $\alpha_A(n, m)$  时, ( $\alpha_A(n, m) \in [0, 1]$ ), 原始图像  $f_B$  的权值即为  $\alpha_B(n, m) = 1 - \alpha_A(n, m)$ 。

$\omega_1$  左侧包含了  $E_f g(n, m)$  的第 1 个峰值(见图 1), 它对应  $f_B$  比  $f_A$  清晰程度大的区域, 在此区域中, 图像  $f_B$  主要描述了场景信息的细节, 融合图像应最大程度采用  $f_B$  中的像素, 所以令此区域中  $f_B$  的权值是 1, 则  $f_A$  的融合系数为 0; 类似

$\omega_3$  的右侧对应  $f_A$  比  $f_B$  清晰程度大的区域, 在此区域中,  $f_A$  的融合系数为 1,  $f_B$  的是 0。

对特征图  $E_{fg}(n, m)$  中分别与  $\omega_1, \omega_2$  或  $\omega_3$  相等的像素, 采用各自最优的融合系数  $\alpha_1, \alpha_2, \alpha_3$  进行融合, 这里  $\alpha_1, \alpha_2, \alpha_3$  不仅对与  $\omega_1, \omega_2$  或  $\omega_3$  相等的特征图  $E_{fg}(n, m)$  的像素有意义, 而且对在  $\omega_1$  和  $\omega_3$  之间变化的特征值像素具有更重要的作用。由上文分析可知, 特征值  $E_{fg}(n, m)$  在  $\omega_1$  和  $\omega_3$  之间变化的像素属于原始图像  $f_A$  和  $f_B$  中清晰度相近的区域, 对此我们使用模糊分类方法来确定这些像素的融合权值。模糊隶属度定义如下:

$$\mu_j = \begin{cases} 0, & d_i = 0, \quad (i \neq j, 1 \leq i, j \leq 3) \\ 1, & d_j = 0 \\ d_j^{-1} / \sum_{i=1}^3 d_i^{-1}, & \text{其他} \end{cases} \quad (7)$$

$d_j = \|E_{fg}(n, m) - \omega_j\|^2$  表示特征图  $E_{fg}(n, m)$  的像素到各类中心  $\omega_i$  ( $i = 1, 2, 3$ ) 的距离。据此, 对于原始图像  $f_A$ , 这些像素的融合权值为  $\alpha_A(n, m) = \sum_{j=1}^3 [\alpha_j \times \mu_j]$ , ( $\omega_1 < E_{fg}(n, m) < \omega_3$ )。值得注意的是, 对于  $d_j = 0$  的像素, 代入式(7)中的  $d_j^{-1} / \sum_{i=1}^3 d_i^{-1}$  得  $\mu_j = 1$ 。

综上所述, 图像  $f_A$  的融合系数  $\alpha_A(n, m)$  为

$$\alpha_A(n, m) = \begin{cases} 0, & E_{fg}(n, m) < \omega_1 \\ 1, & E_{fg}(n, m) > \omega_3 \\ \alpha_j, & E_{fg}(n, m) = \omega_j, \quad j = 1, 2, 3 \\ \sum_{j=1}^3 \left[ \alpha_j \times \left( d_j^{-1} / \sum_{i=1}^3 d_i^{-1} \right) \right], & \omega_1 < E_{fg}(n, m) < \omega_3 \end{cases} \quad (8)$$

当  $\alpha_A(n, m)$  和  $\alpha_B(n, m)$  确定后, 在  $(n, m)$  处融合图像  $f_M$  的值为

$$f_M(n, m) = \alpha_A(n, m) \times f_A(n, m) + \alpha_B(n, m) \times f_B(n, m) \quad (9)$$

## 2.4 $\alpha_1, \alpha_2, \alpha_3$ 的确定

**2.4.1 适应度函数** 如何确定 ( $\alpha_1, \alpha_2, \alpha_3 \in [0, 1]$ ) 以使融合图像的效果最佳? 本文采用进化策略<sup>[11]</sup>来求解。进化策略采用实数编码, 适用于实值函数的优化问题。

进化策略的搜索过程依赖于适应度函数。对于多聚焦图像的融合问题, 适应度函数应体现出融合图像同时最大限度地保持两幅不同聚焦点图像的高频信息的特点。本文采用了交叉熵测度进行评价, 图像  $X$  与  $Y$  的交叉熵为

$$CE_{X,Y} = \sum_{i=0}^{L-1} p_{X_i} \log \frac{p_{X_i}(X)}{p_{Y_i}(Y)} \quad (10)$$

其中  $L$  表示图像的总的灰度级数。  $p_i$  表示灰度值为  $i$  的像素数与图像总像素数之比。

由于本文中每组向量  $\mathbf{A} = (\alpha_1, \alpha_2, \alpha_3)$  选取对应一幅融

合图像, 因此可将融合图像的适应度函数作为向量  $\mathbf{A}$  的适应度函数。定义融合图像  $f_M$  的适应度函数为

$$\begin{cases} F(f_M) = \sqrt{(CE_{f_A, f_M}^2 + CE_{f_B, f_M}^2) / 2} \\ \text{s.t. } |E(f_M) - E(f)| \leq \varepsilon \end{cases} \quad (11)$$

其中  $E(\cdot)$  表示取均值, 且  $E(f) = E(f_A) = E(f_B)$ ;  $\varepsilon$  为允许的均值误差范围。

从适应度函数定义可看出, 适应度函数综合考虑了  $CE_{f_A, f_M}^2$  和  $CE_{f_B, f_M}^2$ , 也就是说适应度函数取决于融合图像与两幅原始图像间的综合差异, 即总体交叉熵, 同时又考虑了融合图像的均值与原始图像均值在允许的范围内接近相等, 另外, 多聚焦图像融合是在相同成像条件下的两幅图像进行融合, 也就是说, 两幅原始图像应具有相同的亮度和对比度。以上迭代过程向适应度函数减少方向进行, 适应度函数越小, 个体的适应度越高。

**2.4.2 进化策略** 进化策略的操作对象为问题的可行解组成的一个群体。定义群体中的个体为实数向量  $\mathbf{A} = (\alpha_1, \alpha_2, \alpha_3)$ , 每个个体的适应度高低可以用适应度函数来衡量, 问题求解的目标是找到一个适应度最高的个体  $\mathbf{A}^* = (\alpha_1^*, \alpha_2^*, \alpha_3^*)$ , 使该个体对应的融合图像  $f_M^*$  的适应度函数最小; 进化策略的基本运行过程如下:

(1)产生初始父代群体  $\{\mathbf{A}_i, i = 1, 2, \dots, \mu\}$ , 共  $\mu$  条个体, 每个个体的取值在  $[0, 1]$  范围内随机分布;

(2)变异: 产生子代个体  $\mathbf{A}_i^j = \mathbf{A}_i + N(0, \delta_j^2)$ ,  $i = 1, 2, \dots, \mu$ ,  $j = 1, 2, \dots, \lambda$ , 其中  $N(0, \delta_j^2)$  表示均值为 0, 方差为  $\delta_j^2$  的高斯噪声, 噪声的方差可以固定不变也可时变。父代中的每个个体经过变异产生  $\lambda$  条子代个体;

(3)选择: 采用确定选择方式, 从以上共  $\mu + \mu\lambda$  条个体中选出适应度最高的  $\mu$  条个体, 组成新的父代群体;

(4)重复(2), (3)两步操作, 直到适应度函数满足要求或算法运行代数达到最大允许运行代数。将最后一代群体中的适应度最高的个体作为求解结果。

## 3 实验结果与性能分析

本文对两组不同聚焦点的图像分别进行了融合实验。图像融合的性能评价除了采用总体交叉熵测度外, 还采用了空间频率<sup>[12]</sup>。空间频率用来测量图像整体活动性, 直接与图像清晰度相关, 定义为

$$SZ = \sqrt{RZ^2 + CZ^2} \quad (12)$$

其中  $RZ$  和  $CZ$  分别是行频率和列频率, 定义为

$$RZ = \sqrt{\frac{1}{NM} \sum_{n=1}^N \sum_{m=2}^M (Z(n, m) - Z(n, m-1))^2} \quad (13)$$

$$CZ = \sqrt{\frac{1}{NM} \sum_{m=1}^M \sum_{n=2}^N (Z(n, m) - Z(n-1, m))^2} \quad (14)$$

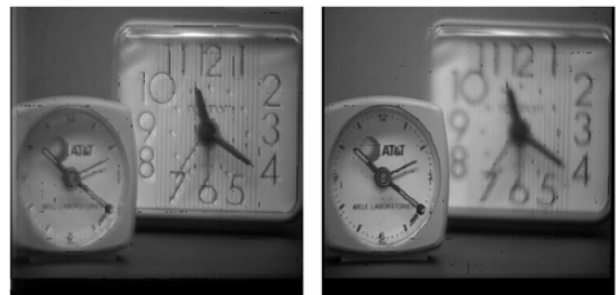
$Z(n, m)$  为图像在  $(n, m)$  处的灰度值。

表 1 小波变换法和本文方法所得结果的总体交叉熵和空间频率

	图像 Lab			图像 Clock		
	拉普拉斯变换法	小波变换法	本文方法	拉普拉斯变换法	小波变换法	本文方法
总体交叉熵	0.0669	0.6393	0.0144	0.0504	0.1853	0.0195
空间频率	29.0532	31.7464	39.4748	30.2575	31.0757	34.5398

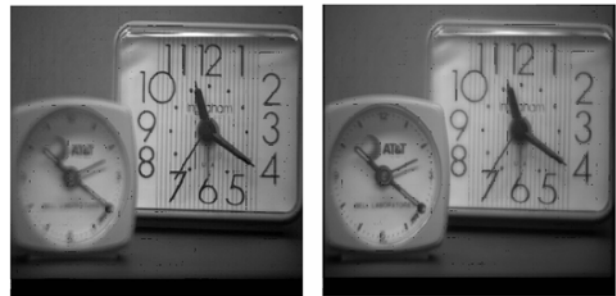
图 2 图像 Lab 的融合结果图

图 2(a)和 2(b)分别为聚焦在前面的钟表 Lab1 和聚焦在后面的学生 Lab2 的两幅不同聚焦点的图像，尺寸为  $480 \times 640$ ；图 3(a)和 3(b)分别为聚焦在前面的钟表 clock1 和聚焦在后面钟表 clock2 的两幅不同聚焦点的图像，尺寸为  $512 \times 512$ ，通过图像融合技术可得到一个目标都聚焦清晰的图像。本文分别采用了拉普拉斯变换法，小波变换法和本文提出的算法进行融合(采用 Daubechies 双正交样条小波 bior3.7)，结果如图 2(c)-2(e)和图 3(c)-3(e)所示，融合结果的总体交叉熵和空间频率如表 1 所示，可以看出与拉普拉斯塔形变换法和小波变换法的结果相比，本文方法的融合结果总体交叉熵较小，空间频率较大，即表明这种方法的融合结果携带的信息量较大，清晰度较高，与两幅原始图像的清晰区域差异较小，相对拉普拉斯塔形变换法和小波变换法，本文算法融合结果的细节更为清晰。



(a) Clock1 ( 聚焦在右边 )

(b) Clock2 ( 聚焦在左边 )



(c) 拉普拉斯变换的融合结果

(d) 小波变换的融合结果



(a) Lab1( 聚焦在闹钟 )

(b) Lab2( 聚焦在学生 )



(c) 拉普拉斯变换的融合结果

(d) 小波变换的融合结果



(e) 本文方法的融合结果

图 3 图像 clock 的融合结果图

#### 4 结束语

本文提出了一种基于小波分解，SOFM 网络和进化策略相结合的多聚焦图像融合方法。它充分地利用了移不变小波的多尺度，多方向冗余信息提取反映图像清晰度差异的归一化能量特征图，并根据此特征图的直方图确定自组织特征映射网络的分类数，有效建立了分类融合的思想，其中分类中心融合加权值的确定采用了进化策略，将图像融合问题转化为一个优化问题，避免了小波反变换所带来的外来影响，成



(e) 本文方法的融合结果

功地实现了多聚焦图像融合。实验结果表明,该方法最大限度地保留了原图图像中的边缘信息,它比拉普拉斯变换法和小波变换法具有更好的融合效果。

### 参 考 文 献

- [1] Genderen J L and Van. Pohl C. Image fusion: Issues, techniques and applications, Intelligent Image Fusion. Proceedings EARSeL Workshop, Strasbourg, France, 11 Sept, 1994: 18-26.
  - [2] Hall D L. Mathematical Techniques in Multisensor Data Fusion. Boston, Artech House, 1992: 20-59.
  - [3] Burt P J and Adelson E H. The Laplacian pyramid as a compact image code [J]. *IEEE Trans. on Commun.*, 1983, 31(4): 532-540.
  - [4] Burt P T and Loaczynski R J. Enhanced image capture through fusion. IEEE Proceedings of the 4th international Conference On Computer Vision, Berlin, Germany, 1993: 173-182.
  - [5] Toet A. Hierarchical image fusion. *Machine Vision and Application*, 1990, 3(2): 1-11.
  - [6] Li H, Manjunath B S, and Mitra S K. Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 1995, 57(3): 235-245.
  - [7] David A Y. Image merging and data fusion by means of the discrete two-dimensional wavelet transform. *J. Opt. Soc. Am. A*, 1995, 12(9): 1834-1841.
  - [8] Zhang Z and Blum R S. A categorization of multiscale decomposition-based image fusion schemes with a performance study for a digital camera application. *Proc. IEEE*, 1999, 87(8): 1315-1326.
  - [9] Mallat S G. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. on Pattern Analysis and Machina Intelligence*, 1989, 11(7): 674-693.
  - [10] Kohonen T. Self-organized formation of topologically correct feature maps[J]. *Biological Cybernetics*, 1982, 43(1): 59-69.
  - [11] Fogel D B and Fogel L J. An introduction to simulated evolutionary optimization[J]. *IEEE Trans. on NN*, 1994, 5(1): 3-14.
  - [12] Li Shutao, Kwok J T, and Wang Yaonan. Multi-focus image fusion using artificial neural networks. *Pattern Recognition Letters*, 2002, 23(6): 985-997.
- 吴 艳: 女, 1965年生, 教授, 博士, 主要从事多传感器信息融合、图像融合、图像处理与理解模式识别等方面的研究工作。
- 刘重阳: 男, 1981年生, 硕士, 从事多传感器信息融合、图像处理与理解、模式识别等方面的研究工作。
- 廖桂生: 男, 1963年生, 教授, 博士生导师, 主要从事雷达信号处理、雷达探测系统的阵列信号处理、SAR 雷达成像及后处理等方面的研究工作。