

多源遥感影像像素级融合分类与 决策级分类融合法的研究

贾永红¹ 李德仁²

(1 武汉大学遥感信息工程学院, 武汉市珞喻路 129 号, 430079)

(2 武汉大学测绘遥感信息工程国家重点实验室, 武汉市珞喻路 129 号, 430079)

摘要: 首先探讨了基于像素的多源遥感影像高频调制融合法, 根据成像系统特性和 Heisenberg 测不准原理, 设计的高斯滤波器对高分辨率影像滤波的方法是合理有效的。在研究 BP 神经网络的基础上, 采用动量法和学习率自适应调整的策略, 提高了 BP 神经网络学习算法收敛速度, 并增强了算法的可靠性。提出并实现了多源遥感影像像素级融合分类与决策级分类融合两种分类方法, 并进行了比较。采用 Landsat TM 3, 4, 5 和航空 SAR 影像进行试验, 结果表明两种分类方法是行之有效的, 均适用于多源遥感影像分类。

关键词: 高通滤波; 影像融合; BP 神经网络; 分类

中图法分类号: TP751; TP753; P231.5

在设计遥感成像装置过程中, 往往对空间分辨率和信噪比采取折中方式。同全色传感器相比, 多光谱到超光谱传感器, 波谱段变窄了, 为了保持影像信噪比必须采集更多的光能, 瞬时视场 IFOV 也必须扩大。因此, 获取的多光谱影像光谱分辨率高, 光谱信息丰富, 但其空间分辨率低; 而获取的全色影像具有较高的空间分辨率, 但缺乏光谱信息。显然, 将这两类影像进行优势互补, 以获得对环境或对象更精确、更可靠、更全面的估计和判断是非常重要的。多源遥感数据融合则是富集多种传感器信息的最有效途径之一。本文首先探讨了基于像素的多源遥感影像高频调制融合法, 并采用 BP 神经网络对其融合影像进行分类; 其次探讨基于 BP 神经网络的分类融合法, 并同前者的分类结果进行分析和比较。

1 多源遥感影像像素级融合分类

多源遥感影像像素级融合分类是指采用基于像素的融合法获得多源遥感影像的融合影像, 然后对融合影像进行分类。为此本文首先采用基于像素的高频调制融合法对空间配准的多源遥感影

像融合, 获得融合影像, 然后采用 BP 神经网络分类法对融合影像分类^[1]。

1.1 高频调制融合法

Schowengerdt^[2]首先提出采用一个较小的空间高通滤波器对高空间分辨率影像滤波, 滤波得到的结果保留了与空间信息有关的高频分量信息, 然后把高通滤波的高频分量依像素加到空间配准的各低分辨率多光谱影像上, 获得空间分辨率增强的多光谱影像。采用的公式如下:

$$F_k(i, j) = M_k(i, j) + HPF(i, j) \quad (1)$$

式中, $F_k(i, j)$ 表示第 k 波段像素 (i, j) 的融合值; $M_k(i, j)$ 表示低分辨率多光谱影像第 k 波段像素 (i, j) 的值; $HPF(i, j)$ 表示空间高通滤波器对高分辨率影像 $P(i, j)$ 滤波得到的高频影像像素 (i, j) 的值。

经过这种处理就把高分辨率影像的高频信息(细节)和多光谱影像的光谱信息融合了。这种方法虽然简单, 但会产生一种人为效应。

为了改善融合效果, 一般采用高频调制融合算法。它是将高分辨率影像 $P(i, j)$ 与空间配准的低分辨率多光谱影像 $M_k(i, j)$ 相乘, 并用高分辨率影像 $P(i, j)$ 经过低通滤波后得到的低频影

像 $LPF(i, j)$ 进行归一化处理, 得到增强的第 k 波段融合影像。其公式为:

$$F_k(i, j) = M_k(i, j) \cdot P(i, j) / LPF(i, j) \quad (2)$$

高分辨率影像 $P(i, j)$ 经过低通滤波后分解成 $LPF(i, j)$ 和 $HPF(i, j)$ 两部分。即

$$P(i, j) = LPF(i, j) + HPF(i, j) \quad (3)$$

将式(3)代入式(2)得:

$$F_k(i, j) = M_k(i, j) + M_k(i, j) \cdot HPF(i, j) / LPF(i, j) \quad (4)$$

式(4)就是高频调制融合法的计算式。高频调制融合法能提高低分辨率多光谱影像的空间分辨率, 同时能够有效保护原始多光谱影像的光谱信息。这就是本文采用高频调制融合法的依据。

高频调制融合法的关键是设计滤波器。合理的方法是用低通滤波器去匹配多光谱影像的点扩散函数, 但精确测定多光谱影像的点扩散函数是很困难的。而且利用点扩散函数作为低通滤波器存在以下问题: ①瞬时视场不同; ②不同传感器所对应的点扩散函数不同; ③与两种传感器影像进行配准时所使用的重采样函数有关。

为此, 作者根据成像系统点扩散函数特性, 提出采用高斯滤波器对高分辨率影像滤波来获取低频成分和高频成分, 然后进行融合。因为一般成像系统的点扩散函数呈正态分布, 因此, 采用高斯正态分布函数作为滤波器对高分辨率影像滤波是合理的。若假定成像系统为线性不变系统, 那么二维高斯滤波器可转化为一维高斯滤波器形式表示。对于一维高斯滤波器而言, 只要给出方差 σ 参数, 高斯滤波器也就可以确定了。根据 Heisenberg 测不准原理, 采用高斯滤波器能保证滤波得到的高频成分产生的空间位置误差小, 因而融合的影像对后续分类、边缘检测和分割处理都是有效的。

针对不同空间分辨率之比的影像融合, 高斯滤波器大小取多光谱影像和高分辨率影像的空间分辨率之比的两倍^[4]。试验表明, σ 取多光谱影像和高分辨率影像的空间分辨率的比值是较合理的, 可得到满意的融合影像。

1.2 BP 神经网络学习算法

BP 算法就是 Back-Propagation(反向传播)算法的简称。BP 算法的学习过程包括正向传播和反向传播两个过程。在正向传播过程中, 输入信息从输入层经隐含层逐层处理, 并传向输出层, 每一层神经元的状态只影响下一层神经元的状态。如果输出层不能得到期望的输出, 则转入反

向传播, 将误差信号沿原来的路径返回, 通过修改各层神经元的权值, 使误差最小。

在含有一个隐含层的 BP 神经网络中, 假设输入层、中间层和输出层的单元数分别是 N, L 和 M ; x_0, x_1, \dots, x_{N-1} 是加到网络的连续值输入矢量; h_0, h_1, \dots, h_{L-1} 是中间层输出矢量, y_0, y_1, \dots, y_{M-1} 是网络的实际输出矢量, 并用 d_0, d_1, \dots, d_{M-1} 来表示训练组中各模式的目标输出矢量。输入单元 i 到隐单元 j 的权重是 v_{ij} , 而隐单元 j 到输出单元 k 的权重是 w_{jk} 。另外用 θ_k 和 ϕ_j 来分别表示输出单元和隐单元的阈值。于是, 中间层各单元的输出为:

$$h_j = f\left(\sum_{i=0}^{N-1} v_{ij} x_i - \phi_j\right) \quad (5)$$

而输出层各单元的输出为:

$$y_k = f\left(\sum_{j=0}^{L-1} w_{jk} h_j - \theta_k\right) \quad (6)$$

隐含层神经元的变换函数采用 log-sigmoid 型函数, 输出层神经元的变换函数采用线性函数。

应用 BP 神经网络分类的关键问题涉及网络结构设计、网络学习等^[3]。在 BP 神经网络结构确定后, 就可利用输入输出样本集对网络进行训练。但标准 BP 算法具有收敛速度慢、出现局部极值、难确定隐含层数和隐节点数等主要问题, 实际应用能力受到限制。对 BP 算法已有许多改进算法, 本文采用动量法和学习率自适应调整的策略, 以提高学习效率并增强算法的可靠性。三层 BP 网络的学习算法如下。

- 1) 选定训练组。
- 2) 将各权重 v_{ij}, w_{jk} 以及阈值 θ_k, ϕ_j 置成小的随机值。
- 3) 从训练组中取一个模式 x_0, x_1, \dots, x_{N-1} 加到网络, 并给定它的目标输出矢量 d_0, d_1, \dots, d_{M-1} 。
- 4) 根据式(5)计算出一个中间层输出矢量 h_0, h_1, \dots, h_{L-1} , 根据式(6)计算网络的实际输出矢量 y_0, y_1, \dots, y_{M-1} 。

- 5) 将输出矢量中的元素 y_k 与目标矢量中的元素 d_k 进行比较, 计算 M 个输出误差:

$$\hat{\alpha} = (d_k - y_k) y_k (1 - y_k) \quad (7)$$

对中间层的隐单元也计算 L 个误差:

$$\hat{\alpha}^* = h_j (1 - h_j) \sum_{k=0}^{M-1} \hat{\alpha} w_{jk} \quad (8)$$

- 6) 依次计算各权重的调整量:

$$\Delta w_{jk}(n) = \eta \hat{\alpha} h_j \quad (9)$$

$$\Delta v_{ij}(n) = \eta \hat{\alpha}^* x_i \quad (10)$$

η 是一个控制学习速度的正增益常数(学习率), 一般取值在 0~1.0 之间。

7) 调整权重:

$$w_{jk}(n+1) = w_{jk}(n) + \Delta w_{jk}(n) + \mu \Delta w_{jk}(n-1) \quad (11)$$

$$v_{ij}(n+1) = v_{ij}(n) + \Delta v_{ij}(n) + \mu \Delta v_{ij}(n-1) \quad (12)$$

式中, μ 为小于 1 的正常数(称动量参数), 必须在 0 和 1 之间, 一般取 $\mu = 0.8$ 。 $\Delta w_{jk}(n-1)$ 和 $\Delta v_{ij}(n-1)$ 是上次训练时计算出的调整量。式(11)和式(12)中的最后一项(动量项), 可以滤除权重空间中误差曲面的高频偏差(即误差曲面中高曲率的剧变), 使有效的权重间隔加大(即学习次数减少), 能加快收敛速度, 并使收敛过程中权重的变化较平滑。

在梯度下降法中, 算法的收敛速度与学习率 η 大小有关。学习率小, 收敛慢; 学习率大, 则有可能修正过头, 导致振荡甚至发散。这里采用自适应调整学习率的改进算法, 使式(9)和式(10)变为:

$$\Delta w_{jk}(n) = \eta(n) \hat{\delta} h_j \quad (13)$$

$$\Delta v_{ij}(n) = \eta(n) \hat{\delta}^* x_i \quad (14)$$

其中, $\eta(n) = 2^\lambda \eta(n-1)$ (15)

$$\lambda = \text{sign}[\Delta w(n) \Delta w(n-1)] \quad (16)$$

当连续两次迭代权修正值符号相同时, 表明下降太慢, λ 取 +1, 可使 η 步长加倍; 当连续两次迭代权修正值符号相反时, 表明下降过头, 这时 λ 取 -1, 使 η 步长减半。可见自适应调整学习率有利于缩短学习时间。

8) 返回 3), 继续迭代, 直到权重 v_{ij} 和 w_{jk} 达到稳定。

这时训练任务已完成, 固定当前权重后, 该网络便构成了一个模式分类器。当训练组以外的输入模式加到网络时, 计算出相应的输出矢量来判别该模式所属的类别。

这里权重稳定的含义是指在训练组中所有模式作用下, 权重调整量 Δw_{jk} 和 Δv_{ij} 均接近于 0。实际进行训练时, 常通过查看输出矢量与目标输出矢量间的接近程度来确定训练过程是否结束。只有训练组中所有模式加到网络上后, 输出矢量都和对应的目标输出矢量接近, 才认为训练已完成。

从上面的学习过程可以看到, 先计算输出层的误差 $\hat{\delta}$, 然后将它逐层往回传递, 来调整层与层间的权重。这种误差 BP 学习算法可以推广到

有若干个中间层的层状网络。

多层神经网络能够以任意精度逼近任意连续的非线性映射。对于给定的问题, 网络的层数取决于问题的要求, 同时直接影响其推广能力。根据实践经验, 本文选用三层 BP 神经网络结构, 其输入层节点数与图像的特征数相同, 输出层节点数与分类类别数相同, 隐含层节点数视问题的复杂度经试验确定。一般确定隐节点数的方法采用试探法: 先设定一隐节点数; 如果训练误差不能下降到所需范围就增加节点数。如果误差已经很小而分类效果依然很差, 说明隐节点数过多, 此时需适当减少隐节点数。

2 决策级分类融合

决策级分类融合就是把单源遥感影像分别进行分类, 然后将分类结果按照一定的融合规则综合, 得到最终的分类结果。基于改进的 BP 神经网络模糊分类融合是先采用改进的 BP 神经网络分别对各个遥感数据源分类, 然后按一定规则融合, 得到决策级分类融合的结果。具体方法如下。

假设 C_s 是分类精度较高的影像上某一像素的类别值, C_m 是分类精度较低的影像上相应像素的类别值, C_f 是分类融合结果, $\max(\cdot)$ 代表当前像素 3×3 邻域内相同类别值最多的像素的类别值, 采用的融合规则为:

$$\textcircled{1} \quad C_s = C_m, \text{ 则 } C_f = C_s = C_m;$$

$$\textcircled{2} \quad C_s \neq C_m, C_s = \max(C_m), \text{ 则 } C_f = C_s;$$

$$\textcircled{3} \quad C_s \neq C_m, C_s \neq \max(C_m), C_m = \max(C_s),$$

则 $C_f = C_m$;

$$\textcircled{4} \quad C_s \neq C_m, C_s \neq \max(C_m), C_m \neq \max(C_s), \max(C_m) = \max(C_s), \text{ 则 } C_f = \max(C_s);$$

$\textcircled{5} \quad C_s \neq C_m, C_s \neq \max(C_m), C_m \neq \max(C_s), \max(C_m) \neq \max(C_s)$, 则按模糊隶属度大小确定当前像素类别。首先, 利用 BP 神经网络对各数据源的分类结果, 统计数据源 j 具有相同类别 i 的像素各灰度级出现的频率并归一化, 作为数据源 j 像素属于类别 i 的隶属度 p_i^j 。同理, 确定每一像素属于其他类别的隶属度, 然后将所有数据源当前像素属于同一类别的隶属度按下式融合:

$$p_i = 0.5 + \left\{ \sum_{j=1}^n w_j (p_i^j - 0.5)^\alpha \right\}^{\frac{1}{\alpha}} \quad (17)$$

式中, p_i 表示类型 i 的可能性测度或称为可信度; w_j 为每个信源的相对重要程度, 且满足 $\sum_{j=1}^n w_j = 1$; α 为奇数, 对于航空 SAR 和 TM 影像而言, 其

分类精度分别可达 60% 和 80%, 因此, $w_1 = 0.6 / (0.6 + 0.8) \approx 0.43$, $w_2 \approx 0.57$, $\alpha = 3$ 。比较该像素各类别融合隶属度大小, 根据最大隶属度原则, 确定当前像素最终类别。

3 试验结果分析

本文的试验数据是采用覆盖广东三水地区空间配准的 Landsat TM3, 4, 5 三波段影像和航空 SAR 影像, 大小为 512×512 像素。该试验区为平坦地区, 且大致可分为水体、林地、旱地、裸露地 4 类覆盖物。根据前述, 采用 Matlab 语言编制了高频调制融合算法、改进的 BP 神经网络分类和分类融合程序。

3.1 高频调制融合影像评价

采用 Matlab 语言编制的高频调制融合算法对空间配准的 Landsat TM3, 4, 5 三波段影像和航空 SAR 影像进行融合, 得到的融合影像同原 TM 影像从视觉上相比, 融合的影像不仅在空间分辨率和清晰度上比原 TM 影像增强了, 且融合影像的彩色合成与原 TM 影像的彩色合成非常相似。由计算得出, 不仅融合影像的平均梯度、熵比对应 TM 影像的平均梯度、熵大, 而且与 TM 对应波段的相关系数高达 75%, 这定量表明融合影像清晰度高, 信息量更丰富, 保光谱特性好。由此可见, 定性和定量结果都表明高频调制融合法不仅能改善影像空间分辨率和清晰度, 而且能很好地保持原多光谱影像的光谱特征, 因而融合的影像有利于改善判读、分类和制图精度。

3.2 融合分类与分类融合的分类精度

首先分别从高频调制融合影像、Landsat TM3, 4, 5 和 SAR 影像上选取水体、林地、旱地、裸露地的训练样区, 每一类选取 4 个 $15 \text{ 像素} \times 20 \text{ 像素}$ 大小的样本, 将高频调制融合影像、TM 影像和 SAR 影像获取的每类样本像素灰度值作为输入值, 对应样区的地类编号作为输出值, 一起构成各自的学习样本对, 对神经网络进行学习训练。然后利用训练好的神经网络对融合影像、TM 影像和 SAR 影像进行分类, 最后采用融合规则将 TM 影像和 SAR 影像的分类结果融合, 得到决策级分类融合的结果。

为了定量比较二者的分类效果, 笔者计算了两种方法的平均分类正确率。平均分类正确率是从分类图中随机选取若干样区, 统计样区每一类正确分类像素数与实际该类像素数的百分比, 将各类正确分类率取平均得到的。计算结果见

表 1。

表 1 融合分类与分类融合的分类正确率

Tab. 1 The Classification Accuracy of Fused Images and Classified Result Fusion

分类方法	各类正确分类率	平均分类精度
融合分类	水体 90%	87.8%
	旱地 86%	
	林地 87%	
	裸地 88%	
分类融合	水体 86%	87%
	旱地 88%	
	林地 88%	
	裸地 86%	

由此可见, 多源遥感影像像素级融合分类与决策级分类融合的分类精度很接近。当多源遥感影像用于分类时, 可直接采用基于改进的 BP 神经网络决策级分类融合法; 而当需要制作影像图和分类时, 则可先采用像素级高频调制融合法获得融合影像, 然后用改进的 BP 神经网络对融合影像分类。

参 考 文 献

- 贾永红. 人工神经网络在多源遥感影像分类中的应用. 测绘通报, 2000(7): 7~8
- Schowengerdt R A. Reconstruction of Multi-spatial Multi-spectral Image Data Using Spatial Frequency Content. PE & RS, 1980, 46(10): 1325~1334
- 楼顺天, 施阳编. 基于 Matlab 的系统分析与设计——神经网络. 西安: 西安电子科技大学出版社, 1998
- Chavez P S Jr, Sides S C, Anderson J A. Comparison of Three Different Methods to Merge Multi-resolution and Multi-spectral Data: TM & SPOT Panchromatic. PE & RS, 1991, 57(3): 295~303
- 何明一. 神经计算. 原理. 语言. 设计. 应用. 西安: 西安电子科技大学出版社, 1992
- Benediktsson J A, Swain P H, Ersoy O K. Neural Network Approaches Versus Statistical Methods in Classification of Multisource Remote Sensing Data. IEEE Trans. on Geoscience and Remote Sensing, 1990, 28(4): 540~551
- Dreyey P. Classification of Land Cover Using Optimized Neural Nets on SPOT Data. PE & RS, 1993, 59(5): 617~621
- Fisher P F, Pathirana S. The Evaluation of Fuzzy Membership of Land Cover Classes in the Suburban Zone. Remote Sensing of Environment, 1990, 34: 121~132
- Foody G M, Cox D P. Sub-pixel Land Cover Composition Estimation Using a Linear Mixture Model and Fuzzy Membership Functions. International Journal of Remote

Sensing, 1994, 15(3): 619~631

研究。代表成果:《计算机图像处理与分析》;多源遥感影像数据融合方法及其应用研究等。发表论文20多篇。

作者简介:贾永红,副教授。主要从事数字图像处理与遥感应用

E-mail: yhjia@wtusm.edu.cn

An Approach of Classification Based on Pixel Level and Decision Level Fusion of Multi-source Images in Remote Sensing

JIA Yonghong¹ LI Deren²

(1) School of Remote Sensing Information Engineering Wuhan University, 129 Luoyu Road, Wuhan, China 430079)

(2) National Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing,
Wuhan University, 129 Luoyu Road, Wuhan, China 430079)

Abstract: With the availability of multi-sensor, multi-temporal, multi-resolution and multi-spectral image data from operational Earth observation satellites, the image fusion has become a valuable tool in remote sensing image evaluation. It is a relatively new and rapidly developing research field in remote sensing. In this paper, a pixel-level fusion algorithm of multi-source images in remote sensing based on high frequency modulation is studied. According to the characters of imaging system and principle of Heisenberg, a Gaussian filter is designed and used in the algorithm, which is proved to be effective. A back-propagation feed forward artificial neural network using momentum and adjusting learning rate by self-adaptation is studied. The speed and reliability of BP neural network are improved. A pixel-level fusion procedure and a decision-level fusion procedure for classification of multi-source remotely sensed images are proposed. A multi-source image set including Landsat TM3, 4, 5 and SAR has been used in classification. Compared with their classification accuracy obtained by the two procedures, the results show that the two procedures applied in classification of multi-source remotely sensed images are effective.

Key words: high frequency modulation; image fusion; BP neural network; classification

About the author: JIA Yonghong, associate professor. His research interests include digital image processing, remote sensing and evaluation of image quality. His typical achievements are image processing and analysis by computer, the fusion methods of multi-source images and its application, etc. His published papers are more than 20.

E-mail: yhjia@wtusm.edu.cn