



基于结构一致性能量模型的异源光学和 SAR 图像变化检测

孙玉立, 雷琳*, 匡纲要

国防科技大学电子科学学院, 电子信息系统复杂电磁环境效应国家重点实验室, 长沙 410073

* 通信作者. E-mail: leilin98@nudt.edu.cn

收稿日期: 2022-10-17; 修回日期: 2023-01-25; 接受日期: 2023-04-11; 网络出版日期: 2023-10-13

国家自然科学基金面上项目 (批准号: 61971426) 资助

摘要 异源遥感图像变化检测是遥感领域越来越受到关注的课题. 由于成像机制的不同, 无法通过直接比较不同源的图像来检测变化. 为了克服这一难题, 本文利用与成像模态无关的结构一致性建立异源图像之间的联系. 首先, 通过构造 K 近邻图来表征图像结构, 并挖掘图像结构关系与节点变化状态之间的关联. 而后, 提出了基于图像结构一致性的能量模型来刻画这种关联, 并设计了两个能量函数分别度量节点发生变化和未发生变化时的能量损失. 最终, 通过求解该能量模型得到每个节点发生变化的概率, 并对其进行分割得到变化检测结果. 由于本文所提出的方法是基于图像结构的, 所以它对图像噪声、成像条件等因素不敏感, 同时由于该方法对节点状态直接建模, 能充分利用能量模型的灵活性, 所以它可适用于不同的任务场景. 在 5 个实际数据集上的实验结果验证了本文方法能够取得更高效、更准确的检测结果.

关键词 变化检测, 异源数据, 能量模型, 结构, 多模态

1 引言

变化检测是指通过比较在同一地点不同时刻获取的多幅遥感图像提取地物变化信息的技术^[1], 它广泛应用于环境监测、城市规划和灾害评估救援等方面^[2,3]. 根据获取的前后时刻图像是否来自同一传感器, 可以将变化检测分为同源 (单模态) 和异源 (多模态) 两个类别. 相较于前者, 异源变化检测具有以下两个方面的优势^[4,5]: 一是在突发事件的应急响应中, 当事件后的同源图像受成像条件限制不可用时 (如光照、天候等) 或受到卫星重访周期限制短期内无法获取时, 异源变化检测可以使用任意可用的事件前后图像来快速获取变化信息; 二是在动态监测过程中, 异源变化检测可以利用来自不同源的图像提升检测时间分辨率, 尤其是能够充分利用早期的遥感影像数据 (如来自停止服务的卫星).

引用格式: 孙玉立, 雷琳, 匡纲要. 基于结构一致性能量模型的异源光学和 SAR 图像变化检测. 中国科学: 信息科学, 2023, 53: 2016–2033, doi: 10.1360/SSI-2022-0408
Sun Y L, Lei L, Kuang G Y. A structure consistency-based energy model for heterogeneous optical and SAR images change detection (in Chinese). Sci Sin Inform, 2023, 53: 2016–2033, doi: 10.1360/SSI-2022-0408

同源变化检测可以通过直接比较两幅图像得到变化,而在异源变化检测中,前后时刻图像对同一地物的表征不同,无法通过直接比较异源图像提取变化信息.所以异源变化检测的首要任务是建立异源图像之间的关联,将“不可比”的图像转换到同一个域内使其“可比较”.基于这个目的,研究者们提出了不同的解决方法,可以粗略地分为传统方法和基于深度学习的方法.其中,传统方法利用统计分布模型^[6,7]、字典学习^[8]、分形投影^[9]、图信号处理^[10]、结构相似性度量^[11,12]等方式建立异源图像间的关联;基于深度学习的方法通常使用深度神经网络将两幅图像转换到同一潜在的特征空间内进行比较^[13,14]或者首先利用生成对抗网络将一幅图像转换到另一幅图像域内,而后在同一图像域内利用同源变化检测方法提取变化^[15,16].

图像的拓扑结构可以用来描述图像内部的连接关系,它对噪声的抵抗性较强,且具有一定的模态不变性,因而被广泛地应用于遥感图像智能解译中^[17~19].最近,研究者们也提出了一些基于图像结构的无监督异源变化检测方法.Luppino等^[20]将图像分成重叠的图像块,并计算两个亲和矩阵来表征不同源图像块内部的结构,而后直接比较亲和矩阵来衡量图像块内所有像素的变化水平,并以此挑选未变化的样本用于训练传统的图像回归函数^[20]或者辅助深度图像回归网络的训练^[15,16].Mignotte^[9]利用图像的自相似性使用分形编码表征事件前图像的结构,而后利用分形投影将事件前图像转换到事件后图像域内并计算差异图,最后利用基于马尔可夫(Markov)随机场的分割方法提取变化图.在我们之前的工作中,利用K近邻图刻画图像的非局部结构,而后通过图映射的方法将一幅图像的结构映射到另一个图像域内,从而比较图像结构之间的差异^[12,21].此外,还利用结构回归的方法,将事件后的图像分解为回归图像和变化图像,并约束回归图像的结构和事件前图像的结构一致,得到基于结构回归的图像分解模型^[22,23].Touati等^[24]考虑每个像素与其他所有像素之间的关系,并建立能量模型来刻画像素对关系和像素对标签之间的联系,而后使用多维尺度缩放求解差异图.随后,文献^[25]利用一个度量函数计算任意两个像素属于同一类标签(同时为变化或未变化)和属于不同类标签(一个为未变化,另一个为变化)的概率,并假设该度量分别服从指数分布和高斯(Gauss)分布,而后建立马尔可夫模型求解每个像素的标签.尽管这些基于像素/图像块对关系的结构比较方法很直观,但它们仍然存在3个缺点:一是这些模型以像素或图像块作为基本处理单元,因而在利用像素对或图像块对之间相似性关系时的复杂度为 $O(N^2)$,使得模型求解难度增大,限制了实际应用效果;二是这些模型求解大都比较复杂,如文献^[9]需要计算分形编码和分形投影,文献^[23]需要求解自相似性矩阵和图像分解两个优化模型,文献^[24]需要使用多维尺度缩放求解能量模型,文献^[25]需要使用迭代估计分布参数和使用模拟退火算法求解分割模型.尤其是当使用像素或图像块作为基本单元时,算法求解难度进一步加剧,变化检测效率大大降低;三是这些方法都忽略了变化样本(像素或图像块)对于结构比较的影响.如文献^[26]中所分析的,变化样本会直接影响与其相连接的样本,即通过边传递变化信息到其他样本上,造成对其他样本标签(变化/未变化)的误判.尽管之前的工作在尝试消除变化的影响,但它仍然需要一个复杂的、冗余的从粗到细的迭代过程^[26].这也促使本文探索更为简单、灵活和稳健的异源变化检测方法.

为了克服以上缺点,本文提出了基于结构一致性的能量模型(structure consistency based energy model, SCEM).它首先将异源图像分割成具有相同轮廓的超像素,并以超像素为节点构建K近邻图(k-nearest neighbor graph, KNNG)来刻画每幅图像的结构,而后利用结构一致性建立相邻节点之间特征相似性关系与节点标签之间的联系.同时受文献^[25]的启发,本文使用能量模型来刻画这种联系.据此,本文设计了两个能量函数分别度量节点发生变化和未发生变化时的能量损失,即利用基于结构一致性的能量损失来约束节点未发生变化的概率和利用基于稀疏变化先验的能量损失来约束节点发生变化的概率,从而得到可用于异源变化检测的能量模型.最终,使用优化方法对该模型进行求解,得到

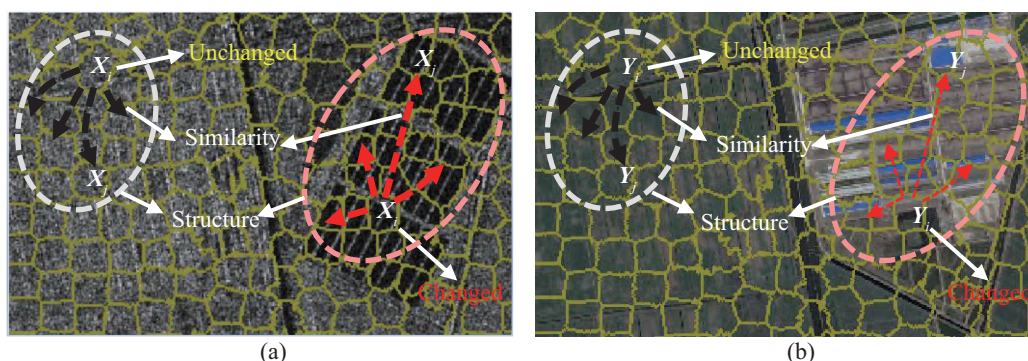


图 1 (网络版彩图) 图像结构示意图. (a) SAR 图像; (b) 光学图像. 图像的结构被定义为图像内部超像素之间的相似性关系, 图中使用连接线的粗细表示超像素之间的相似性大小. 在 SAR 和光学图像中, 未变化区域的结构是一致的, 而在变化区域的结构是不同的

Figure 1 (Color online) The structure defined in the proposed method. (a) SAR image; (b) optical image. The thickness of the connecting line indicates the degree of similarity between superpixels. The structure is defined as the superpixel similarity relationships within the whole image. In the SAR image and the optical image, the structure of the unchanged part is consistent, while the changed part is not

每个节点发生变化的概率, 并对其进行分割提取变化和未变化类别. 本文提出的方法挖掘了图像结构与节点状态之间的关联, 这种关联是普适的、稳健的, 所以该方法对图像噪声和成像条件等因素不敏感; 同时它利用能量模型对约束关系直接建模, 具有很强的灵活性和可拓展性; 此外, 该方法利用超像素作为节点, 减小了算法的复杂度, 很大程度上提高了变化检测的效率.

本文的主要贡献总结如下:

- 利用与成像模态无关的结构一致性建立起了异源图像之间的联系, 并提出了基于图像结构一致性的能量模型用于异源变化检测任务, 解决了异源遥感图像难以比较的问题.
- 利用能量模型刻画了图像结构关系和节点变化概率之间的关联, 并考虑到了变化节点对于其他节点的影响, 有效地减小了未知的变化样本对于模型的影响, 使得模型更加准确和稳健.
- 在 5 个数据集上与当前最优方法进行了对比实验, 结果验证了本文方法在准确性和时效性方面都优于其他方法. 相关代码已开源在网站¹⁾.

2 基于结构一致性的能量模型

定义事件前后时刻 (t_1 和 t_2) 获取的已配准异源图像分别为 $\mathbf{X} \in \mathbb{R}^{M \times N \times C_x}$ 和 $\mathbf{Y} \in \mathbb{R}^{M \times N \times C_y}$, 定义其像素分别为 $x(m, n, c)$ 和 $y(m, n, c)$, 变化检测的目的是求解一个二值变化图来标记每个像素是否变化. 显然, 直接比较异源图像的像素值是无意义的, 所以需要寻找图像 $\mathbf{X}$ 和 $\mathbf{Y}$ 之间的联系.

基于图像的自相似性, 图像中每个小的部分 $\mathbf{X}_i$ (图像块或超像素) 总能在图像中找到与其十分相似的部分 $\mathbf{X}_j$, 本文将这种图像块或者超像素之间的相似性关系定义为图像的结构. 据此, 结构一致性可以描述为: 对于 t_1 时刻图像中代表同一类地物的 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ (表现为 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 十分相似), 如果它们在事件中均未发生变化, 则在另一幅 t_2 时刻图像中相同区域的 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 也会十分相似, 即对于未变化的区域, 它们的相似性关系是一致的; 相反地, 当 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 在事件中发生了变化, 这种相似性关系在另一幅图像中将被破坏, 此时 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 是不相似的, 如图 1 所示. 本文利用异源图像的结构一致

1) <https://github.com/yulisun/SCEM>.

性构建能量模型, 主要包含以下 3 个步骤: (1) 构建 K 近邻图表征图像结构; (2) 建立能量模型并求解; (3) 提取变化图.

2.1 表征图像结构

如上文所述, 我们需要考虑成对的相似性关系来表征图像结构. 不同于以往文献使用像素或图像块作为基本处理单元^[12, 20, 21, 23~25], 本文采用超像素作为基本处理单元(节点), 具有两个方面的优势: 一方面它能够大大减小节点数目, 降低算法复杂度, 例如对于一幅 500×500 的图像, 使用像素为基本处理单元时节点数目为 2.5×10^4 ; 另一方面超像素能够充分利用上下文信息和保留地物的边缘轮廓.

为了得到共同的超像素分割图, 首先构造一个伪 RGB 图像, 其三通道分别为 t_1 和 t_2 时刻图像的灰度图通道和一个全零通道; 而后, 使用文献 [27] 提出的基于高斯混合模型 (Gaussian mixture model, GMM) 的方法生成超像素, 它在分割效率 (图像像素数目的线性复杂度) 和超像素边缘保持方面都十分出色. 当伪 RGB 图像被分割为 N_S 个超像素并得到分割图 Λ 后, 可以利用 Λ 将 t_1 和 t_2 时刻图像分割成 N_S 个具有相同边缘轮廓的超像素 $\mathbf{X}_i$ 和 $\mathbf{Y}_i$, $i = 1, \dots, N_S$, 定义为

$$\begin{aligned}\mathbf{X}_i &= \{x(m, n, c) \mid (m, n) \in \Lambda_i, c = 1, \dots, C_x\}, \\ \mathbf{Y}_i &= \{y(m, n, c) \mid (m, n) \in \Lambda_i, c = 1, \dots, C_y\}.\end{aligned}\quad (1)$$

此时, $\mathbf{X}_i$ 和 $\mathbf{Y}_i$ 代表相同的地理区域, 并且它们各自内部都是匀质的 (代表同一类地物). 为了简单起见, 本文提取每个波段的均值和中位值作为超像素特征, 得到特征矩阵 $\mathbf{X}'$ 和 $\mathbf{Y}'$. 当然这些特征不是唯一的, 其他特征也可以使用 (如方差、灰度共生矩阵等).

接下来, 以每个超像素为节点, 为每个图像构造一个 K 近邻图来表征其结构. 以 t_1 时刻图像 $\mathbf{X}$ 为例, 构造图 $G^{t1} = \{\mathcal{V}^{t1}, \mathcal{E}^{t1}, \mathbf{A}^{t1}\}$ 为

$$\begin{aligned}\mathcal{V}^{t1} &= \{\mathbf{X}_i \mid i = 1, 2, \dots, N_S\}, \\ \mathcal{E}^{t1} &= \{(\mathbf{X}_i, \mathbf{X}_j) \mid i = 1, 2, \dots, N_S, j \in \mathcal{N}_i^x\}, \\ A_{i,j}^{t1} &= 1, \forall (\mathbf{X}_i, \mathbf{X}_j) \in \mathcal{E}^{t1},\end{aligned}\quad (2)$$

其中, $\mathcal{N}_i^x$ 表示 $\mathbf{X}_i$ 的 K 近邻的位置集合, $\mathbf{A}^{t1}$ 表示邻接矩阵, 当 $(\mathbf{X}_i, \mathbf{X}_j) \notin \mathcal{E}^{t1}$ 时, $A_{i,j}^{t1} = 0$. 定义 $\mathbf{D}^x$ 为图像 $\mathbf{X}$ 的特征距离矩阵, 其元素 $d_{i,j}^x = \|\mathbf{X}'_i - \mathbf{X}'_j\|_2^2$ 表示超像素 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 之间的特征距离, $\mathcal{N}_i^x$ 可以通过对距离向量 $\mathbf{D}_i^x$ 进行排序并取出除 i 外的前 k 个最小值的位置得到.

对于 t_2 时刻图像 $\mathbf{Y}$, 同样地定义其特征距离 $d_{i,j}^y = \|\mathbf{Y}'_i - \mathbf{Y}'_j\|_2^2$ 和特征距离矩阵 $\mathbf{D}^y$, 并按照类似地步骤可以构造其 K 近邻图 $G^{t2} = \{\mathcal{V}^{t2}, \mathcal{E}^{t2}, \mathbf{A}^{t2}\}$ 为

$$\begin{aligned}\mathcal{V}^{t2} &= \{\mathbf{Y}_i \mid i = 1, 2, \dots, N_S\}, \\ \mathcal{E}^{t2} &= \{(\mathbf{Y}_i, \mathbf{Y}_j) \mid i = 1, 2, \dots, N_S, j \in \mathcal{N}_i^y\}, \\ A_{i,j}^{t2} &= 1, \forall (\mathbf{Y}_i, \mathbf{Y}_j) \in \mathcal{E}^{t2},\end{aligned}\quad (3)$$

其中, $\mathcal{N}_i^y$ 表示 $\mathbf{Y}_i$ 的 K 近邻的位置集合, $\mathbf{A}^{t2}$ 表示邻接矩阵.

本文构建 K 近邻图表征图像结构, 相比文献 [23, 27, 28] 中的全连接图具有两个方面的优势: (1) 文献 [23] 以图像块内所有像素为节点构建全连接图来刻画图像块内部的结构, 因而其无法检测出整体变化 (即整个图像块都发生变化), 而本文使用 K 近邻图连接每个超像素与图像内其他相似的超像素, 刻画的是非局部结构, 关注的是整个超像素节点的变化状态; (2) 文献 [27, 28] 以整个图像内

像素为节点并考虑所有像素对之间的关系, 其节点数目庞大, 导致算法复杂度很高. 此外, 在这些全连接图中, 它们采取直接比较的方式来度量图结构的差异, 类似于 $|d_{i,j}^x - d_{i,j}^y|$ 或 $|e^{-d_{i,j}^x} - e^{-d_{i,j}^y}|$. 我们认为, 在异源图像中这种跨图像比较不同地物之间的差异程度是不合适的. 例如, 当 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 代表不同的地物 (如草地和水体), $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 也代表不同的地物 (如草地和水体, 或者草地和建筑物), 此时直接比较 $d_{i,j}^x$ 和 $d_{i,j}^y$ 的值是没有意义的 (因为它们来自不同的图像域), 即 $|d_{i,j}^x - d_{i,j}^y|$ 值的大小无法评判第 i 和 j 个节点是否发生了变化. 所以, 本文使用 K 近邻图而未使用全连接图来表征图像结构, 它的目的是限制相连的超像素至少在某个图像中代表同一类地物, 从而使得结构比较具有意义.

2.2 构建能量模型

能量模型通过对输入变量的每个组合赋予能量捕获变量之间的相关性^[28]. 能量模型的优势在于它可以自由地以任意合理的方式构建能量约束, 从而使其具有很强的灵活性和表现力^[29~31]. 本文利用能量模型建立输入图像结构与节点标签之间的联系.

定义指标集 $\mathcal{I} = \{1, 2, \dots, N_S\}$, 待求解的变化向量为 $\mathbf{p} \in \mathbb{R}^{N_S}$, 其中第 i 个元素 $p_i \in [0, 1]$ 表示第 i 个超像素所代表区域发生变化的概率. 可以构造能量模型如下:

$$\mathbf{p}^* = \arg \min_{\mathbf{p} \in \mathbb{R}^{N_S}} E_H(\mathbf{p}; \mathbf{X}, \mathbf{Y}), \quad (4)$$

其中, 能量函数 E_H 用来衡量 $\mathbf{X}, \mathbf{Y}, \mathbf{p}$ 在不同取值组合下的合理程度.

2.2.1 结构与变化之间的关联

在得到表征图像结构的 K 近邻图 G^{t1} 和 G^{t2} 后, 我们考虑相连节点相似性关系与节点状态之间的联系.

(1) G^{t1} 上相连的节点. 对于任意节点 $i \in \mathcal{I}$, 首先考虑其在 G^{t1} 上相连的节点 j . 由于 $j \in \mathcal{N}_i^x$, 可以认为超像素 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 代表同一类地物 (如草地). 而对于 t_2 时刻图像中相对应的超像素 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 的关系, 区分两种不同的情形分别讨论, 如图 2 所示.

情形 #1. 如果 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 也代表同一类地物时, 则可以认为第 i 和 j 个节点的状态应该是一样的, 即均未发生变化 (如 #1a 中, $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 同属于草地) 或同时发生相同的变化 (如 #1b 中, $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 同属于水体).

情形 #2. 如果 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 代表不同类地物时, 包含: #2a, $\mathbf{Y}_i$ 和 $\mathbf{X}_i$ 代表同一类地物, $\mathbf{Y}_j$ 和 $\mathbf{X}_j$ 代表不同类地物 (如 $\mathbf{Y}_i$ 代表草地, $\mathbf{Y}_j$ 代表水体), 此时 i 节点未变化, j 节点发生变化; #2b, $\mathbf{Y}_i$ 和 $\mathbf{X}_i$ 代表不同类地物, $\mathbf{Y}_j$ 和 $\mathbf{X}_j$ 代表相同类地物 (如 $\mathbf{Y}_i$ 代表水体, $\mathbf{Y}_j$ 代表草地), 此时 i 节点发生变化, j 节点未发生变化; #2c, 如 $\mathbf{Y}_i$ 和 $\mathbf{X}_i$, $\mathbf{Y}_j$ 和 $\mathbf{X}_j$ 都代表不同类地物 (如 $\mathbf{Y}_i$ 代表水体, $\mathbf{Y}_j$ 代表建筑物), 此时第 i 和 j 个节点同时发生变化但变化类型不一样.

由以上两种情形可以得到相连节点间相似性关系与节点标签之间的联系. **结论 (a):** 对于任意节点 $i \in \mathcal{I}$ 和 $j \in \mathcal{N}_i^x$, 当第 i 和 j 个节点都未变化时, $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 的差异应该很小 (代表同一类地物); 当 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 的差异很大时 (代表不同类地物), i 和 j 至少有一个节点发生了变化.

(2) G^{t2} 上相连的节点. 同样地, 对于 G^{t2} 上任意相连的节点 $i \in \mathcal{I}$ 和 $j \in \mathcal{N}_i^x$, 也有类似的情形 #3 和情形 #4.

情形 #3. 如果 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 也代表同一类地物, 则 i 和 j 节点的状态一致, 即均未发生变化或同时发生相同的变化.

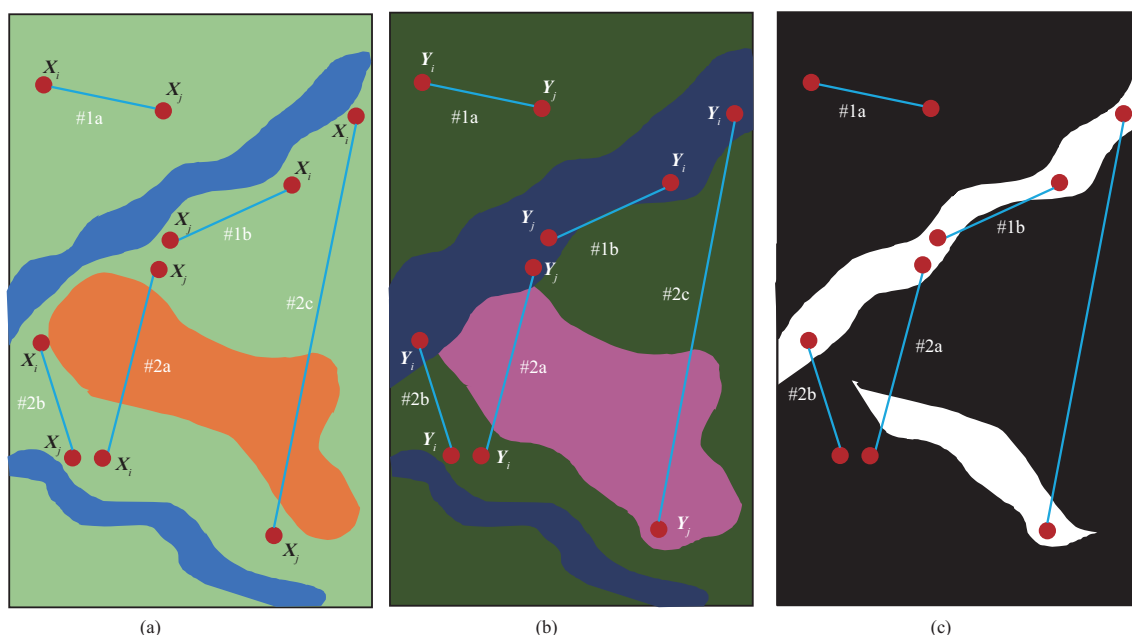


图 2 (网络版彩图) 结构与变化之间关联的示意图

Figure 2 (Color online) Illustration of the association between structure and change. (a) Pre-event image $\mathbf{X}$; (b) post-event image $\mathbf{Y}$; (c) round truth

情形 #4. 如果 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 代表不同类地物, 则 i 和 j 的节点状态可能为: 其中一个发生变化, 另一个未变化; 或者同时发生变化但变化类型不一样.

类似地, 可以得到结论 (b): 对于任意节点 $i \in \mathcal{I}$ 和 $j \in \mathcal{N}_i^y$, 当第 i 和 j 个节点都未变化时, $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 的差异应该很小 (代表同一类地物); 当 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 的差异很大时 (代表不同类地物), i 和 j 至少有一个节点发生了变化.

2.2.2 能量模型

根据情形 #1, #2 和结论 (a), 可以构造如下的能量函数:

$$E_1 = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{N}_i^y} \tilde{d}_{i,j}^y (1 - p_i) (1 - p_j), \quad (5)$$

其中, $\tilde{d}_{i,j}^y = d_{i,j}^y - \min_{j \in \mathcal{N}_i^y} \{d_{i,j}^y\}$ 用来衡量超像素 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 的特征差异, 在 E_1 中表示第 i 和 j 个节点均未变化时 ($p_i = p_j = 0$) 的能量. 可以看出, 当 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 的差异较大时 (此时 $\tilde{d}_{i,j}^y$ 较大), 若要使 E_1 取值较小则需要较大的 p_i 或 p_j , 即变化的 i 或 j 节点.

类似地, 根据情形 #3, #4 和结论 (b), 可以构造如下的能量函数:

$$E_2 = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{N}_i^y} \tilde{d}_{i,j}^x (1 - p_i) (1 - p_j), \quad (6)$$

其中, $\tilde{d}_{i,j}^x = d_{i,j}^x - \min_{j \in \mathcal{N}_i^x} \{d_{i,j}^x\}$ 用来衡量超像素 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 的特征差异, 在 E_2 中表示节点 i 和 j 均未变化时的能量. 从 E_1 和 E_2 中可以看出, 本文考虑的是节点对的变化状态而不是单独地计算某个节点变化概率, 即考虑到了每个节点对于其相连节点的影响.

定义矩阵 $\mathbf{B} \in \mathbb{R}^{N_S \times N_S}$, 其元素 $B_{i,j}$ 如下:

$$B_{i,j} = \tilde{d}_{i,j}^x \delta(j \in \mathcal{N}_i^y) + \tilde{d}_{i,j}^y \delta(j \in \mathcal{N}_i^x), \quad (7)$$

其中, $\delta(\cdot)$ 为判别函数: 当括号内条件成立时, 取值为 1; 相反地, 当括号内条件不成立时, 取值为 0. 从 $B_{i,j}$ 和 $\tilde{d}_{i,j}^x, \tilde{d}_{i,j}^y$ 的构造过程中可以看出, 当第 i 和 j 个节点都未发生变化时, $\tilde{d}_{i,j}^x$ 和 $\tilde{d}_{i,j}^y$ 值较小, 此时相应的 $B_{i,j}$ 值也会较小; 相反地, 当 $B_{i,j}$ 值较大时, 说明此时第 i 或 j 个节点大概率发生了变化.

合并能量函数 E_1 和 E_2 并利用式 (7), 可以得到基于结构一致性的能量函数

$$E_{sc} = (\mathbf{1} - \mathbf{p})^T \mathbf{B} (\mathbf{1} - \mathbf{p}), \quad (8)$$

其中, $\mathbf{1}$ 表示所有元素为 1 的 N_S 维列向量. 在 E_{sc} 中, 当 $B_{i,j}$ 越大, 表明图像相似性关系越不一致, 为使得能量值 E_{sc} 减小, 会倾向于得到更大的 p_i 或者 p_j , 即 i 或 j 节点发生变化的概率越大. 能量函数 E_{sc} 的意义在于: 当图像的相似性关系 (结构) 不一致时, 例如 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 相似, 但 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 不相似, 或者 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 相似, 但 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 不相似, 可以利用能量损失检测出变化.

同时, 对于异源变化检测问题, 还有一个变化稀疏性的先验知识: 在实际场景中, 发生变化的总是少部分地物, 大部分地物是未变化的. 这一稀疏先验也被广泛应用于其他变化检测方法^[22~25], 同时它也是无监督异源变化检测问题的基础. 理论上, 可以使用基于 ℓ_0 范数的 $\|\mathbf{p}\|_0$ 来作为稀疏惩罚项, 但考虑到 ℓ_0 的非凸性, 使用其凸松弛的 ℓ_1 范数惩罚项. 由此, 得到一个基于稀疏先验的能量函数

$$E_{sp} = \sum_{i \in \mathcal{I}} |p_i|. \quad (9)$$

合并上述基于结构一致的能量函数 E_{sc} 和基于稀疏先验的能量函数 E_{sp} , 可以得到最终能量模型

$$\begin{aligned} \min_{\mathbf{p}} \left\{ E_H := (\mathbf{1} - \mathbf{p})^T \mathbf{B} (\mathbf{1} - \mathbf{p}) + \lambda \mathbf{p}^T \mathbf{1} \right\}, \\ \text{s.t. } 0 \leq p_i \leq 1, i = 1, 2, \dots, N_S, \end{aligned} \quad (10)$$

其中, $\lambda > 0$ 为平衡参数. 从式 (10) 可以看出, 两个能量函数是相互对抗的, 其中 E_{sc} 倾向于得到变化的 $\mathbf{p}$, 即 $\mathbf{p} = \mathbf{1}$; 而 E_{sp} 倾向于得到未变化的 $\mathbf{p}$, 即 $\mathbf{p} = \mathbf{0}$. 另外, 可以看出模型 (10) 是直接建立在节点状态上的 (即对 $\mathbf{p}$ 直接求解), 形式十分简单, 具有明显的物理意义和可解释性. 同时, 由于目标函数形式简洁, 使用简单的梯度下降法即可近似求解.

2.2.3 提取变化

当求解模型 (10) 得到 $\mathbf{p}^*$ 后, 可以得到差异图: 使用 p_i^* 表示第 i 个超像素内每个像素的变化概率. 而后, 对差异图进行分割即可得到二值变化图, 可以采用阈值分割法 (如大津阈值^[32])、聚类法 (如 k 均值聚类^[33] 或者模糊 c 均值聚类^[34]), 或采用基于马尔可夫随机场 (Markov random field, MRF) 的方法. 本文直接使用文献 [22] 中提出的 MRF 分割方法, 它能够充分利用变化信息和上下文信息, 取得良好的分割效果.

本文提出异源变化检测方法的总体框架总结为算法 1. 在能量模型 (10) 的求解中, 它设置了迭代停止准则: 达到最大迭代次数 N_{iter} 或者相邻两次迭代的结果相对差值 $\epsilon^{(k)} = \frac{\|\mathbf{p}^{(k)} - \mathbf{p}^{(k-1)}\|_2}{\|\mathbf{p}^{(k)}\|_2}$ 小于给定的门限 ϵ^0 .

表 1 异源变化检测数据集介绍

Table 1 Description of the heterogeneous data sets

Dataset	Sensor	Size (pixels)	Date	Location	Event (& spatial resolution)
#1	Landsat-5/Google Earth	$300 \times 412 \times 1(3)$	Sept. 1995 ~ July 1996	Sardinia, Italy	Lake expansion (30 m)
#2	Pleiades/WorldView2	$2000 \times 2000 \times 3(3)$	May 2012 ~ July 2013	Toulouse, France	Construction (0.52 m)
#3	Radarsat-2/Google Earth	$593 \times 921 \times 1(3)$	June 2008 ~ Sept. 2012	Shuguang Village, China	Building construction (8 m)
#4	Radarsat-2/Google Earth	$343 \times 291 \times 1(3)$	June 2008 ~ Sept. 2010	Yellow River, China	Embankment change (8 m)
#5	QuickBird 2/TerraSAR-X	$4135 \times 2325 \times 3(1)$	July 2006 ~ July 2007	Gloucester, England	Flooding (0.65 m)

算法 1 Structure consistency based energy model

Input: The matrices of $\mathbf{X}$, $\mathbf{Y}$, parameters of N_S , λ , $\alpha > 0$.

Structure representation:

Segment images into N_S superpixels;

Extract the features to obtain $\mathbf{X}'$ and $\mathbf{Y}'$;

Construct the graphs of G^{t1} with (2) and G^{t2} with (3);

Energy model solving:

Initialize: calculate $\mathbf{B}$ and set $\mathbf{p}^{(0)} = (\mathbf{B}\mathbf{1} + \mathbf{B}^T\mathbf{1})/2$, $\mathbf{v}^{(0)} = \mathbf{0}$, $\beta = 0.5$;

Repeat:

1. Calculate the gradient $\nabla^{(k)} = (\mathbf{B} + \mathbf{B}^T)\mathbf{p}^{(k-1)} + \lambda\mathbf{1}$;

2. Update $\mathbf{v}^{(k)} \leftarrow \beta\mathbf{v} + (1 - \beta)\nabla^{(k)}$;

3. Update $\mathbf{p}^{(k)} \leftarrow \mathbf{p}^{(k-1)} - \alpha\mathbf{v}^{(k)}$;

4. Set $p_{\{i|p_i < 0\}}^{(k)} = 0$ and $p_{\{i|p_i > 0\}}^{(k)} = 1$;

Until the stopping criterion is met;

Change extraction:

Compute the difference image;

Segment the difference image into the binary change map.

3 实验和讨论

3.1 数据集和评价指标

(1) **数据集.** 为了验证本文提出的算法的性能, 在 5 个实际数据集上进行对比实验, 如表 1 和图 3(a)~(c) 所示. 这些数据集的前后时刻图像来自: (i) 不同的光学传感器, 如数据集 #1, #2; (ii) 不同类别的传感器, 如数据集 #3, #4, #5. 同时, 这些数据集涵盖了不同的分辨率 (30~0.52 m)、不同的图像尺寸 ($343 \times 291 \sim 4135 \times 2325$)、不同的变化事件 (如洪灾、耕地变化、建筑物构造), 它们能够充分评估不同算法的适应性和稳健性. 这些数据集已经完成了图像配准, 对于其他实际数据可以利用现有的异源图像匹配算法进行配准^[35,36].

(2) **评价指标.** 本文使用两类指标分别评估得到的差异图 and 变化图. (i) 使用受试者工作特征 (receiver operating characteristics, ROC) 曲线和精准率 - 召回率 (precision-recall, PR) 曲线评估差异图, 同时使用曲线下面积 AUR 和 AUP 作为定量指标分别评估 ROC 曲线和 PR 曲线; (ii) 使用不同的颜色在变化图上标记真阳 (true positive, TP)、真阴 (true negative, TN)、假阳 (false positive, FP)、假阴 (false negative, FN), 同时使用准确率 (overall accuracy, OA)、F1 得分和卡帕系数 (κ) 定量评估变化图, 其计算公式分别为 $OA = (TP + TN)/(TP + TN + FP + FN)$, $F1 = (2TP)/(2TP + FP + FN)$ 和 $\kappa = (OA - PRE)/(1 - PRE)$, 其中

$$PRE = \frac{(TP + FN)(TP + FP) + (TN + FP)(TN + FN)}{(TP + TN + FP + FN)^2}. \quad (11)$$

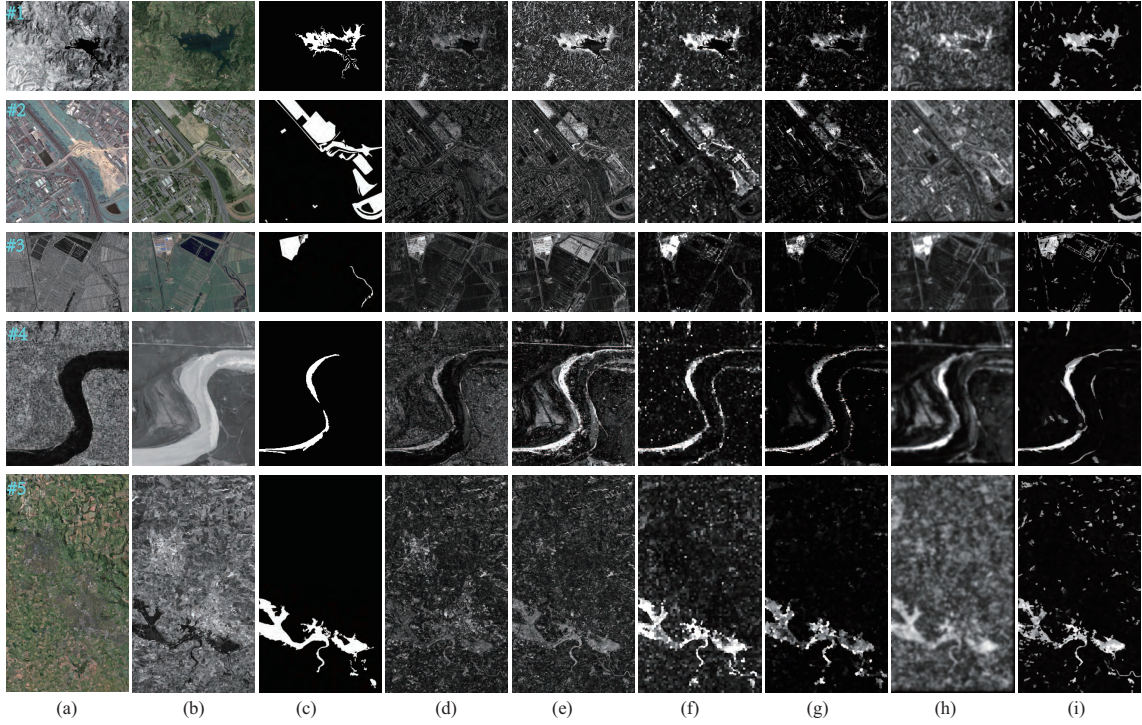


图 3 (网络版彩图) 不同方法所得到的差异图。第 1~5 行分别对应数据集 #1~#5。(a) 事件前图像; (b) 事件后图像; (c) 变化真值图。(d)~(i) 分别为不同方法得到的差异图: (d) AMD; (e) HPT; (f) IRG-McS; (g) SCASC; (h) FPMS; (i) SCEM

Figure 3 (Color online) Difference images of different methods. From top to bottom, they correspond to the datasets #1 to #5, respectively. (a) Pre-event image; (b) post-event image; (c) ground truth. From (d) to (i) are change maps obtained by (d) AMD, (e) HPT, (f) IRG-McS, (g) SCASC, (h) FPMS, and (i) SCEM

(3) 实验设置. 实验中, 将参数统一设置为: 超像素数目 $N_S = 5000$, 平衡参数 $\lambda = \frac{(1-p^{(0)})^T B(1-p^{(0)})}{0.25 \times N_S}$, 梯度下降步进 $\alpha = 0.01$, 最大迭代次数 $N_{\text{iter}} = 20$, 相对差值门限 $\epsilon^0 = 0.01$. 为了方便描述, 将所提出的基于结构一致性能量模型的方法简称为 SCEM.

3.2 实验结果

(1) 差异图. 本文选择已有的 5 种异源变化检测方法作为对比: 基于结构比较的 AMD^[20] 和 IRG-McS^[26]、基于结构回归的 SCASC^[22] 和 FPMS^[9]、基于同质像素变换的 HPT (使用 40% 的未变化像素对作为监督数据)^[5]. 图 3 显示了所有方法在不同数据集下得到的差异图, 其 ROC 和 PR 曲线如图 4 所示, 表 2 中列举了相应的 AUR 和 AUP 值.

如图 3 所示, 大部分方法得到的差异图都能体现出一些变化信息, 即能够区分出大部分变化区域和未变化区域. 通过仔细比较, 可以看出 AMD 对于变化和未变化的区分能力相对较弱, 尤其是在数据集 #5 上 AMD 得到的差异图具有很大的混淆性, 难以区分变化区域, 这主要由于两个方面的原因: 一是 ADM 对每个图像块单独构图, 表征的是图像块内部的局部结构, 所以当图像块整体发生变化时, AMD 的区分性较弱; 二是 ADM 使用直接比较的方式对在不同图像上构造的图进行比较, 容易造成异源数据的混淆. 同时, 从图 3(f) 和 (h) 中还可以看出 SCASC 和本文方法 SCEM 得到的差异图都具有一定的稀疏性, 这是由于它们在模型中都利用了变化的稀疏先验知识. 此外, 由于这种稀疏性, 它们的

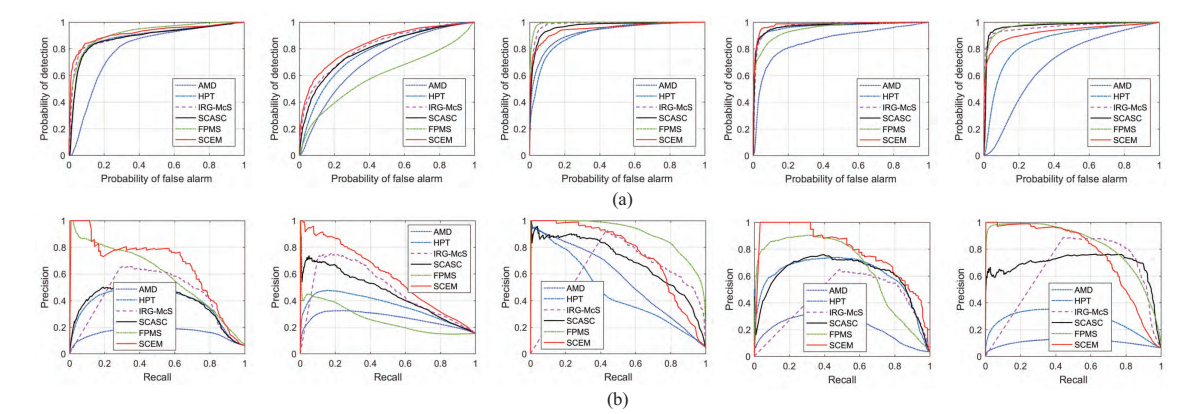


图 4 (网络版彩图) 不同方法得到差异图的 (a) ROC 和 (b) PR 曲线. 从左到右分别为数据集 #1~#5 上的结果

Figure 4 (Color online) (a) ROC and (b) PR curves of DIs generated by different methods. From left to right are the results on datasets #1 to #5, respectively

表 2 不同异源数据集下得到的差异图的定量比较^{a)}

Table 2 Quantitative measurement of difference images on the heterogeneous data sets^{a)}

Methods	Dataset #1		Dataset #2		Dataset #3		Dataset #4		Dataset #5		Average	
	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP
AMD [20]	0.795	0.155	0.708	0.264	0.922	0.564	0.859	0.216	0.691	0.103	0.795	0.260
HPT [5]	0.889	0.373	0.770	0.355	0.911	0.470	0.964	0.602	0.857	0.271	0.878	0.414
IRG-McS [26]	0.897	0.424	0.813	0.469	0.984	0.582	0.972	0.414	0.964	0.605	0.926	0.499
SCASC [22]	0.885	0.383	0.793	0.458	0.968	0.695	0.969	0.597	0.971	0.681	0.917	0.563
FPMS [9]	0.922	0.590	0.597	0.258	0.994	0.904	0.952	0.652	0.977	0.836	0.888	0.648
SCEM	0.919	0.659	0.853	0.598	0.954	0.759	0.982	0.780	0.929	0.763	0.926	0.712

a) The highest scores are highlighted in bold.

ROC 曲线性能受到了一定的影响. 对比图 3 和表 2 中的结果, 可以看出所提出的 SCEM 方法能够获取很高的 AUR 和 AUP 值, 这表明即使使用简单的阈值分割方法, SCEM 也能取得令人满意的变化检测结果. SCEM 的平均 AUP 值为 0.712, 比排名第二的算法 FPMS 高出 6.4%.

(2) 变化图. 除了上述差异图对比中的 4 种基于结构的异源变化检测方法, 本文还增加了另外 4 种方法用于变化图对比, 包括基于能量模型的 M3CD [25], 基于循环不变卷积模型的 CICM [4] 和基于深度神经网络的 SCCN [37], CGAN [38], X-Net [15] 和 ACE-Net [15]. 图 5 显示了 SCEM 和其他 8 种对比方法在所有数据集上的变化图, 在这些图上分别使用白色、黑色、红色和绿色标记 TP, TN, FP 和 FN. 表 3 列出了相应的定量评价结果 (OA, κ 和 F1).

从图 5 中可以看出, 部分方法在一些数据集上表现的不够稳健, 例如 AMD 在数据集 #4 和 #5 上有大量的虚警 (κ 分别为 0.115 和 0.171), X-Net 和 ACE-Net 在数据集 #5 上也表现不佳. 同时, FPMS 在数据集 #2 上有较大的漏检 (κ 仅为 0.215), M3CD 在数据集 #4 上有很高的虚警 (κ 仅为 0.158), 这主要是由于: FPMS 忽略了变化样本对于结构回归的影响, 导致回归后的图像不够准确, 因而检测性能不够稳健; M3CD 在衡量变化水平时使用了直接相减的方法对来自不同域的距离度量进行比较, 造成了异源数据的混淆, 导致变化度量无法有效区分变化和未变化类别. 相反地, 本文提出的基于能量模

表 3 不同异源数据集下得到的变化图的定量比较^{a)}
Table 3 Quantitative measurement of change maps on the heterogeneous data sets^{a)}

Methods	Dataset #1			Dataset #2			Dataset #3			Dataset #4			Dataset #5			Average		
	OA	κ	F1	OA	κ	F1	OA	κ	F1	OA	κ	F1	OA	κ	F1	OA	κ	F1
AMD [30]	0.799	0.255	0.328	0.724	0.259	0.411	0.950	0.572	0.597	0.686	0.115	0.170	0.782	0.171	0.254	0.788	0.274	0.352
HPT [5]	0.912	0.504	0.546	0.815	0.415	0.523	0.859	0.318	0.368	0.927	0.447	0.476	0.860	0.397	0.455	0.875	0.416	0.474
IRG-McS [26]	0.971	0.739	0.754	0.882	0.420	0.478	0.983	0.794	0.804	0.976	0.690	0.702	0.971	0.740	0.755	0.957	0.677	0.699
SCASC [22]	0.947	0.593	0.621	0.892	0.464	0.516	0.979	0.741	0.751	0.977	0.700	0.711	0.973	0.774	0.788	0.954	0.654	0.677
FPMS [9]	0.938	0.593	0.625	0.838	0.215	0.296	0.938	0.569	0.597	0.866	0.274	0.314	0.970	0.770	0.786	0.910	0.484	0.524
M3CD [25]	0.958	0.650	0.672	0.827	0.269	0.368	0.976	0.684	0.696	0.856	0.158	0.204	0.952	0.618	0.643	0.914	0.476	0.517
CICM [4]	0.943	0.451	0.481	0.867	0.270	0.321	0.974	0.745	0.759	0.789	0.024	0.080	0.896	0.371	0.423	0.894	0.372	0.413
SCCN [37]	0.919	0.522	0.562	0.818	0.240	0.342	0.903	0.315	0.359	0.789	0.183	0.232	0.907	0.474	0.521	0.867	0.347	0.403
CGAN [38]	0.965	0.724	0.742	0.863	0.338	0.402	0.943	0.402	0.432	0.971	0.584	0.599	0.911	0.364	0.412	0.931	0.482	0.517
X-Net [15]	0.958	0.674	0.696	0.864	0.389	0.461	0.954	0.586	0.609	0.959	0.490	0.510	0.901	0.447	0.497	0.927	0.517	0.555
ACE-Net [15]	0.967	0.723	0.740	0.868	0.463	0.538	0.964	0.651	0.670	0.962	0.526	0.546	0.878	0.341	0.401	0.928	0.541	0.579
SCEM	0.967	0.717	0.735	0.898	0.570	0.628	0.984	0.813	0.822	0.982	0.712	0.722	0.977	0.791	0.803	0.961	0.721	0.742

a) The highest scores are highlighted in bold.

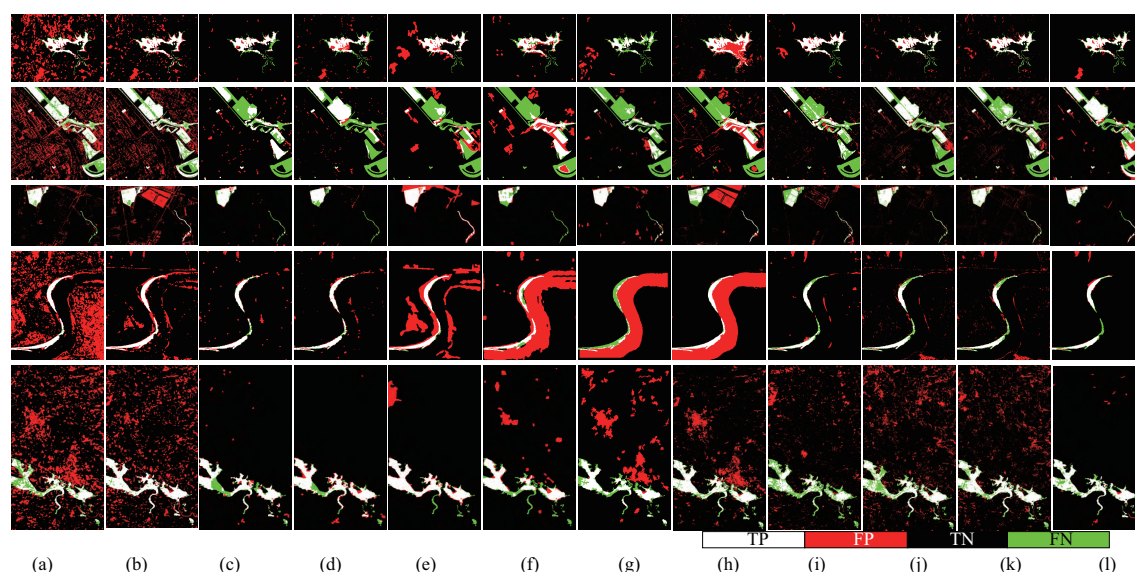


图 5 (网络版彩图) 不同方法所得到的变化图. 第 1~5 行分别对应数据集 #1~#5. (a)~(l) 分别为不同方法得到的变化图: (a) AMD; (b) HPT; (c) IRG-McS; (d) SCASC; (e) FPMS; (f) M3CD; (g) CICM; (h) SCCN; (i) CGAN; (j) X-Net; (k) ACE-Net; (l) SCEM. 在变化图中, 分别使用白色、红色、黑色和绿色标记 TP, FP, TN 和 FN

Figure 5 (Color online) Change maps of different methods. From top to bottom, they correspond to the datasets #1 to #5, respectively. From (a) to (l) are change maps obtained by (a) AMD, (b) HPT, (c) IRG-McS, (d) SCASC, (e) FPMS, (f) M3CD, (g) CICM, (h) SCCN, (i) CGAN, (j) X-Net, (k) ACE-Net, and (l) SCEM. In the change map, white: TP; red: FP; black: TN; cyan: FN

型的 SCEM 方法在这些数据集上均能取得较为满意的变化图, 在数据集 #2~#5 上的检测结果都为最优. 此外, 从图 5 和表 3 还可以发现 SCEM (也包括其他对比方法) 在数据集 #2 上的检测结果比在其他数据集上的要差, 这主要是由以下两个方面造成的: 一是数据集 #2 的图像分辨率更高 (0.52 m), 图像的细节和纹理更加丰富, 本文所使用的特征 (均值和中位值) 在表示超像素内容时效果受限; 二是数据集 #2 中包含了更多的地物类别, 包括道路、房屋、树木、球场、工地等, 而且这些地物所占比例不均. 基于以上两点可知数据集 #2 中的图像结构更为复杂, 所构造的 K 近邻图无法充分表征图像的结构, 导致算法在该数据集上的性能弱于其他数据集. 表 3 中 SCEM 的平均 OA, κ 和 F1 值分别为 0.961, 0.721, 0.742, 均高于其他 8 种对比方法. 例如, SCEM 的平均 F1 值比排名第二的 IRG-McS 高出 4.3%. 这也再一次证明了本文提出的基于结构一致性能量模型的优越性.

3.3 讨论

3.3.1 模型扩展

由于能量模型的灵活性, 可以对提出的结构一致性能量模型形式 (10) 进行一系列的拓展. 我们知道空间上相邻的节点之间也是存在一定约束的 (它们往往代表同一类地物), 所以它们的变化状态往往也是相同的, 即具有空间光滑性. 因而, 可以在模型目标函数 $E_H(10)$ 中加入基于空间光滑的能量函数 E_{ss} , 定义为

$$E_{ss} = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{N}_i^s} W_{i,j} \|p_i - p_j\|_2^2, \quad (12)$$

其中, $j \in \mathcal{N}_i^s$ 表示节点 i 与 j 互为空间上相邻的邻居节点, $\mathbf{W}$ 为权重矩阵, 其元素为 $W_{i,j} = W_{j,i} \geq 0$

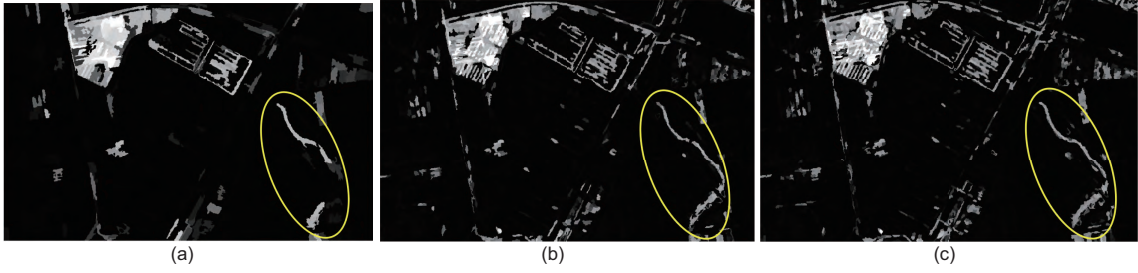

 图 6 (网络版彩图) 使用不同 N_S 时 SCEM 在数据集 #3 上得到的差异图

Figure 6 (Color online) Difference images calculated by SCEM on the Dataset #3 with different N_S . (a) $N_S = 1000$; (b) $N_S = 5000$; (c) $N_S = 9000$

为基于空间距离的加权系数, 并且有如果 $j \in \mathcal{N}_i^s$, 则 $W_{i,j} > 0$; 否则, $W_{i,j} = 0$. 定义图的度矩阵 $D_w \in \mathbb{R}^{N_S \times N_S}$ 是以 $\sum_{j=1}^{N_S} W_{i,j}$ 为第 i 个对角元素的对角矩阵, 则图的拉普拉斯 (Laplace) 矩阵为 $L = D_w - W$. 此时, 可以将 E_{ss} 化简为

$$E_{ss} = 2p^T L p. \quad (13)$$

将 E_{ss} 加入本文模型 (10) 中, 即可得到包含空间光滑约束的能量模型. 类似地, 也可将其他可用的、与变化相关的先验知识纳入本文提出的能量模型中, 如文献 [10] 中所用的图光滑约束, 进而提升模型的表达能力以提高变化检测性能.

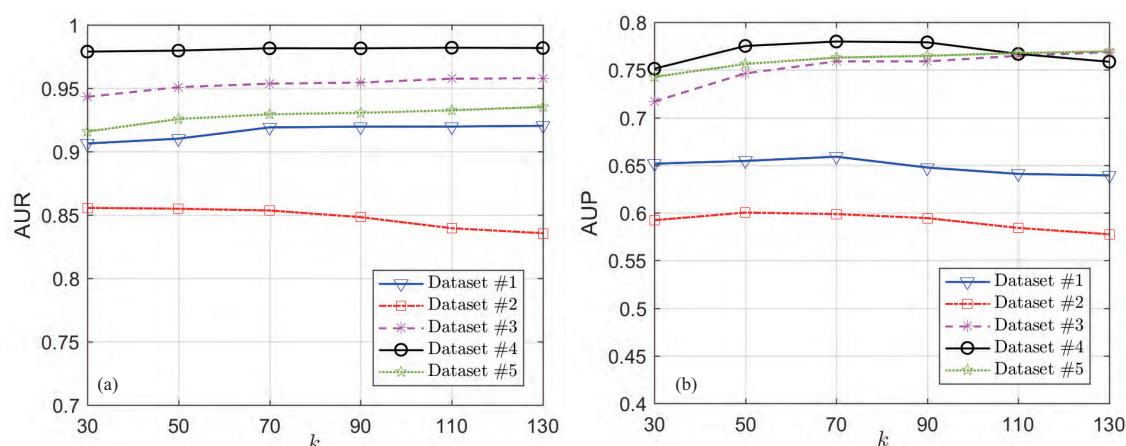
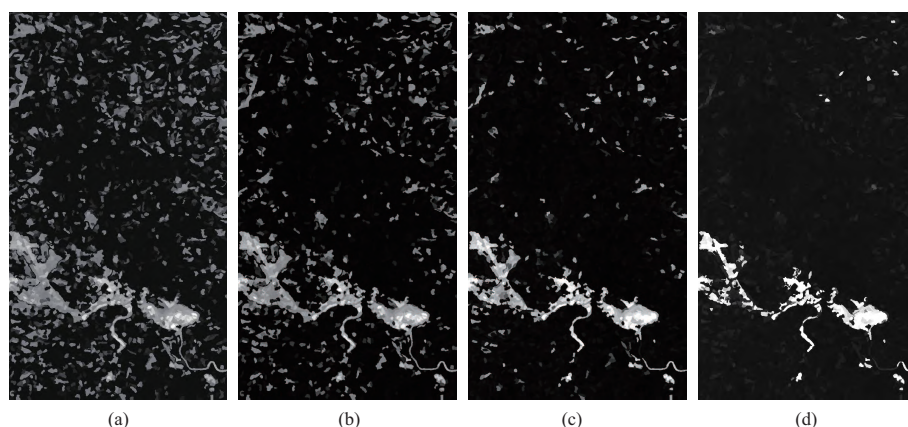
3.3.2 参数分析

本文算法中主要参数为超像素数目 N_S 、邻居数目 k 和平衡参数 λ , 下面分别讨论.

一般来说, 超像素的数目 N_S 应根据数据集的空间分辨率和异源变化检测任务的时效要求来确定. 较大的 N_S 可以使得分割的超像素面积更小, 因而可以提高变化检测粒度, 但同时由于 N_S 的增大, 算法的计算复杂度增加导致算法时效性变弱 (如后文 3.3.3 小节所示); 相反地, 减小 N_S 可以提升计算效率, 但如果 N_S 取值过小会使分割的超像素面积过大, 此时有可能导致部分超像素内包含不同类别的地物, 即会破坏超像素的内部匀质性, 会使得算法性能大大降低. 图 6 显示了 SCEM 在数据集 #3 中使用不同 N_S (1000, 5000, 9000) 时产生的差异图, 并在差异图上使用黄色椭圆标记了一些区域以便于对比. 从图中可以看出, 当 N_S 越小时, 产生的超像素越大, 导致差异图中的块效应越明显, 导致一些细小的变化容易被忽略. 为了综合考虑算法的检测准确率和时效性, 在实验中将其简单的设置为 $N_S = 5000$. 当然, 也可按文献 [39] 中的方法选择合适的分割尺度.

在 K 近邻图中, 邻居数目 k 具有重要的作用. 图 7 展示了 k 取值从 30 到 130 (以 20 为等间隔) 时, 模型在不同数据集上得到差异图的 AUR 和 AUP 值. 可以看出, 模型对于 k 的取值不是很敏感. 当然, 过大和过小的 k 都是不合适的: 过小的 k 会导致 K 近邻图 G^{t1} 和 G^{t2} 不够稳健, 并且会使矩阵 B 中非零元素数目很少, 使得基于结构一致性的能量约束 E_{sc} 无法发挥作用; 相反地, 过大的 k 会使得计算复杂度增加, 同时会导致相似性混淆, 即 K 近邻图中会连接不相似的图像块, 使得能量约束 E_{sc} 不够准确. 因此, 根据经验将其设置为 $k = \sqrt{N_S}$, 这也和基于 KNN 的密度估计问题 [40, 41] 和分类问题 [42] 中 k 的设置是一致的.

对于平衡参数 λ , 它用于调节能量函数中基于稀疏先验的 E_{sp} 的比重, 即用来控制变化图的稀疏程度. 从模型 (10) 中可以看出, 当 $\lambda \rightarrow 0$ 时, 模型的解 $p^* = \mathbf{1}$, 即所有区域都是变化的; 当 $\lambda \rightarrow \infty$ 时, 模型的解 $p^* = \mathbf{0}$, 即所有节点都是未变化的. 因而, 应根据变化区域的占比来调节 λ . 而实际中的变化区

图 7 (网络版彩图) SCEM 对参数 k 的敏感性分析Figure 7 (Color online) Sensitivity analysis of parameter k in SCEM: (a) AUR value; (b) AUP value图 8 使用不同 λ^* 时 SCEM 在数据集 #5 上得到的差异图Figure 8 Difference images calculated by SCEM on the Dataset #5 with different λ^* . (a) $\lambda^* = 1$; (b) $\lambda^* = 2$; (c) $\lambda^* = 4$; (d) $\lambda^* = 8$

域的大小是无法事先准确估计的, 为了增强算法对 λ 的稳健性, 将 λ 设置为 $\lambda = \lambda^* \times \frac{(1-p^{(0)})^T B(1-p^{(0)})}{N_S}$ 并通过调节 λ^* 来取代 λ , 它可以减小数据集和其他参数 N_S , k 等因素对平衡参数的影响. 图 8 展示了 λ^* 取值不同时 ($\lambda^* = 1, 2, 4, 8$) 的能量模型得到的差异图, 可以看出 λ^* 取值越大, 差异图越稀疏. 同时, 也可看出虽然 λ^* 取值不同, 但模型输出的差异图都能在一定程度上区别变化和未变化部分, 如图 8(a)~(d) 所示. 图 9 展示了 λ^* 取值从 2^{-2} 到 2^5 时 (以 2 为等比), 模型在不同数据集得到差异图的 AUR 和 AUP 取值情况. 从图 9 中可以看出, 算法对于 λ^* 的取值变化相对稳健, 并且设置 $\lambda^* = 4$ 在所有数据集下均能取得一个相对较好的结果.

3.3.3 复杂度分析

本文所提算法主要包括两个步骤: 构造 K 近邻图和求解能量模型, 如算法 1 中的框架流程所示. 关于算法的空间复杂度, 存储图像特征距离矩阵和 B 矩阵各需要 $\mathcal{O}(N_S^2)$, 求解能量模型过程中存储向量 ∇ , p 和 v 各需要 $\mathcal{O}(N_S)$, 则算法的空间复杂度为 $\mathcal{O}(N_S^2)$. 关于算法的时间复杂度, 在构造 K

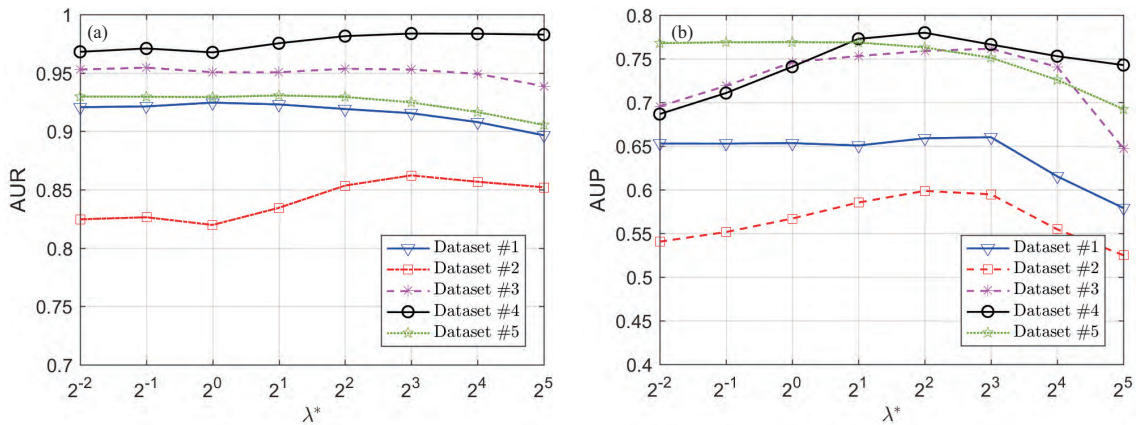


图 9 (网络版彩图) SCEM 对参数 λ^* 的敏感性分析

Figure 9 (Color online) Sensitivity analysis of parameter λ^* in SCEM. (a) AUR value; (b) AUP value

表 4 不同算法在数据集 #2 上的运行时间

Table 4 The computational time of different methods on Dataset #2

Methods	AMD [20]	HPT [5]	IRG-McS [26]	SCASC [22]	FPMS [9]	M3CD [25]	CICM [4]	SCEM
Time (s)	109.6	78.1	20.5	26.7	92.4	2611.6	113.0	5.8

近邻中使用高斯混合模型分割超像素的时间复杂度与图像像素数目成线性关系, 即 $\mathcal{O}(MN)$ [27], 计算前后时刻图像的特征距离矩阵需要 $\mathcal{O}((C_x + C_y)N_S^2)$, 对特征距离进行排序以得到每个超像素的 K 近邻共需要 $\mathcal{O}(N_S^2 \log N_S)$; 在使用梯度下降法求解能量模型中, 每次迭代计算梯度向量需要 $\mathcal{O}(N_S^2)$. 由此可见, 本文算法的时间复杂度较低, 且主要与超像素数目 N_S 而不是图像像素数目相关. 表 4 中列举了对比的传统算法在数据集 #2 上的运行时间, 其中 FPMS, M3CD 和 CICM 使用 C++ 编译并运行在 Intel Xeon Silver 4110 CPU 的 Linux 环境下, AMD, HPT, IRG-McS, SCASC 和 SCEM 均使用 MATLAB 2016a 软件并运行在 Intel Core i7-8700K CPU 的 Windows 环境下. 从表 4 中可以看出本文所提出的基于超像素的变化检测方法所需时间最少, 算法效率很高, 具有很强的实际应用价值.

4 结论

本文提出了一种新颖的基于结构一致性能量模型的异源遥感图像变化检测方法. 该方法构造 K 近邻图表征图像结构, 而后利用结构一致性得到图像结构差异与节点变化状态之间的关联, 并利用能量模型刻画这种联系, 同时设计了两种能量函数分别度量节点发生变化和未变化时的能量损失. 该能量模型是直接建立在节点状态上的, 对其求解即可直接得到节点发生变化的概率. 本文提出的方法充分利用了图像结构的稳健性和能量模型的灵活性, 并以超像素作为基本单元, 计算效率很高. 在不同的实际数据集下的对比实验验证了本文提出方法的优越性. 本文中, 使用固定的 K 近邻图刻画图像结构, 所以在一些复杂场景中具有一定的局限性. 未来, 我们将考虑利用自适应图或者图神经网络提取图像结构信息, 以进一步提升变化检测性能.

参考文献

- 1 Singh A. Review article digital change detection techniques using remotely-sensed data. *Int J Remote Sens*, 1989, 10:

- 989–1003
- 2 Lv Z Y, Liu T F, Benediktsson J A, et al. Land cover change detection techniques: very-high-resolution optical images: a review. *IEEE Geosci Remote Sens Mag*, 2022, 10: 44–63
 - 3 Wen D W, Huang X, Bovolo F, et al. Change detection from very-high-spatial-resolution optical remote sensing images: methods, applications, and future directions. *IEEE Geosci Remote Sens Mag*, 2021, 9: 68–101
 - 4 Touati R. Détection de changement en imagerie satellitaire multimodale. Dissertation for Ph.D. Degree. Montréal: Université de Montréal, 2019
 - 5 Liu Z, Li G, Mercier G, et al. Change detection in heterogenous remote sensing images via homogeneous pixel transformation. *IEEE Trans Image Process*, 2017, 27: 1822–1834
 - 6 Mercier G, Moser G, Serpico S B. Conditional copulas for change detection in heterogeneous remote sensing images. *IEEE Trans Geosci Remote Sens*, 2008, 46: 1428–1441
 - 7 Prendes J, Chabert M, Pascal F, et al. A new multivariate statistical model for change detection in images acquired by homogeneous and heterogeneous sensors. *IEEE Trans Image Process*, 2014, 24: 799–812
 - 8 Gong M G, Zhang P Z, Su L Z, et al. Coupled dictionary learning for change detection from multisource data. *IEEE Trans Geosci Remote Sens*, 2016, 54: 7077–7091
 - 9 Mignotte M. A fractal projection and Markovian segmentation-based approach for multimodal change detection. *IEEE Trans Geosci Remote Sens*, 2020, 58: 8046–8058
 - 10 Jimenez-Sierra D A, Quintero-Olaya D A, Alvear-Munoz J C, et al. Graph learning based on signal smoothness representation for homogeneous and heterogeneous change detection. *IEEE Trans Geosci Remote Sens*, 2022, 60: 1–16
 - 11 Lei L, Sun Y L, Kuang G. Adaptive local structure consistency-based heterogeneous remote sensing change detection. *IEEE Geosci Remote Sens Lett*, 2022, 19: 1–5
 - 12 Sun Y L, Lei L, Li X, et al. Structure consistency-based graph for unsupervised change detection with homogeneous and heterogeneous remote sensing images. *IEEE Trans Geosci Remote Sens*, 2022, 60: 1–21
 - 13 Liu J, Zhang W H, Liu F, et al. A probabilistic model based on bipartite convolutional neural network for unsupervised change detection. *IEEE Trans Geosci Remote Sens*, 2022, 60: 1–14
 - 14 Lv Z Y, Huang H T, Gao L P, et al. Simple multiscale UNet for change detection with heterogeneous remote sensing images. *IEEE Geosci Remote Sens Lett*, 2022, 19: 1–5
 - 15 Luppino L T, Kampffmeyer M, Bianchi F M, et al. Deep image translation with an affinity-based change prior for unsupervised multimodal change detection. *IEEE Trans Geosci Remote Sens*, 2022, 60: 1–22
 - 16 Luppino L T, Hansen M A, Kampffmeyer M, et al. Code-aligned autoencoders for unsupervised change detection in multimodal remote sensing images. *IEEE Trans Neural Netw Learn Syst*, 2022, : 1–13
 - 17 Zhang Y X, Li W, Zhang M M, et al. Topological structure and semantic information transfer network for cross-scene hyperspectral image classification. *IEEE Trans Neural Netw Learn Syst*, 2023, 34: 2817–2830
 - 18 Zhang Y X, Li W, Tao R, et al. Cross-scene hyperspectral image classification with discriminative cooperative alignment. *IEEE Trans Geosci Remote Sens*, 2021, 59: 9646–9660
 - 19 Zhang Y X, Li W, Zhang M M, et al. Graph information aggregation cross-domain few-shot learning for hyperspectral image classification. *IEEE Trans Neural Netw Learn Syst*, 2022, doi: 10.1109/TNNLS.2022.3185795
 - 20 Luppino L T, Bianchi F M, Moser G, et al. Unsupervised image regression for heterogeneous change detection. *IEEE Trans Geosci Remote Sens*, 2019, 57: 9960–9975
 - 21 Sun Y L, Lei L, Li X, et al. Nonlocal patch similarity based heterogeneous remote sensing change detection. *Pattern Recogn*, 2021, 109: 107598
 - 22 Sun Y L, Lei L, Guan D D, et al. Sparse-constrained adaptive structure consistency-based unsupervised image regression for heterogeneous remote-sensing change detection. *IEEE Trans Geosci Remote Sens*, 2022, 60: 1–14
 - 23 Sun Y L, Lei L, Li X, et al. Patch similarity graph matrix-based unsupervised remote sensing change detection with homogeneous and heterogeneous sensors. *IEEE Trans Geosci Remote Sens*, 2021, 59: 4841–4861
 - 24 Touati R, Mignotte M. An energy-based model encoding nonlocal pairwise pixel interactions for multisensor change detection. *IEEE Trans Geosci Remote Sens*, 2018, 56: 1046–1058
 - 25 Touati R, Mignotte M, Dahmane M. Multimodal change detection in remote sensing images using an unsupervised pixel pairwise-based Markov random field model. *IEEE Trans Image Process*, 2020, 29: 757–767
 - 26 Sun Y L, Lei L, Guan D D, et al. Iterative robust graph for unsupervised change detection of heterogeneous remote

- sensing images. *IEEE Trans Image Process*, 2021, 30: 6277–6291
- 27 Ban Z H, Liu J G, Cao L. Superpixel segmentation using Gaussian mixture model. *IEEE Trans Image Process*, 2018, 27: 4105–4117
- 28 LeCun Y, Chopra S, Hadsell R, et al. A tutorial on energy-based learning. In: *Predicting Structured Data*. Cambridge: MIT Press, 2006
- 29 Isack H, Boykov Y. Energy-based geometric multi-model fitting. *Int J Comput Vis*, 2012, 97: 123–147
- 30 Zhao J, Mathieu M, LeCun Y. Energy-based generative adversarial networks. In: *Proceedings of International Conference on Learning Representations*, Toulon, 2017
- 31 Wang Y, Li B, Che T, et al. Energy-based open-world uncertainty modeling for confidence calibration. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Montreal, 2021. 9302–9311
- 32 Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern*, 1979, 9: 62–66
- 33 Hartigan J A, Wong M A. Algorithm AS 136: a K-means clustering algorithm. *Appl Stat*, 1979, 28: 100
- 34 Bezdek J C, Ehrlich R, Full W. FCM: the fuzzy c-means clustering algorithm. *Comput Geosci*, 1984, 10: 191–203
- 35 Zhang H, Lei L, Ni W P, et al. Optical and SAR image dense registration using a robust deep optical flow framework. *IEEE J Sel Top Appl Earth Obs Remote Sens*, 2023, 16: 1269–1294
- 36 Xiang D, Xie Y Z, Cheng J D, et al. Optical and SAR image registration based on feature decoupling network. *IEEE Trans Geosci Remote Sens*, 2022, 60: 1–13
- 37 Liu J, Gong M G, Qin K, et al. A deep convolutional coupling network for change detection based on heterogeneous optical and radar images. *IEEE Trans Neural Netw Learn Syst*, 2018, 29: 545–559
- 38 Niu X D, Gong M G, Zhan T, et al. A conditional adversarial network for change detection in heterogeneous images. *IEEE Geosci Remote Sens Lett*, 2019, 16: 45–49
- 39 Wu J Z, Li B, Ni W P, et al. Optimal segmentation scale selection for object-based change detection in remote sensing images using kullback-leibler divergence. *IEEE Geosci Remote Sens Lett*, 2020, 17: 1124–1128
- 40 Loftsgaarden D O, Quesenberry C P. A nonparametric estimate of a multivariate density function. *Ann Math Stat*, 1965, 36: 1049–1051
- 41 Fukunaga K, Hostetler L. Optimization of k nearest neighbor density estimates. *IEEE Trans Inform Theor*, 1973, 19: 320–326
- 42 Mitra P, Murthy C A, Pal S K. Unsupervised feature selection using feature similarity. *IEEE Trans Pattern Anal Machine Intell*, 2002, 24: 301–312

A structure consistency-based energy model for heterogeneous optical and SAR images change detection

Yuli SUN, Lin LEI* & Gangyao KUANG

State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System, College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China

* Corresponding author. E-mail: leilin98@nudt.edu.cn

Abstract Heterogeneous change detection is a topic of increasing interest in remote sensing. Due to the different imaging mechanisms, it is impossible to compare the multitemporal images from different sources directly. To address this challenge, this paper uses structure consistency to establish the connection between heterogeneous images, which is imaging modality-invariant. First, we construct a K-nearest neighbor graph to characterize the image structure and explore the relationship between the image structure and the node state. Then, we propose a structure consistency-based energy model to characterize these relationships and construct two functions to measure the energy loss both when the change occurs and when it does not occur. Finally, we solve the energy model to obtain the change probability of each node and segment them to obtain the change detection results. Since the proposed method is based on the image structure, it is insensitive to image noise and imaging conditions. Meanwhile, the method models the node states directly and can fully exploit the flexibility of the energy model, enabling it to be applied to different change detection scenarios. The experimental results on five real datasets verify that the proposed method can achieve more efficient and accurate results.

Keywords change detection, heterogeneous data, energy based model, structure, multimodal