

# 基于图像数据挖掘的有向图模型检索方法

龚 晖<sup>1</sup>, 夏开建<sup>2</sup>, 金兆岩<sup>3</sup>

(1. 常熟市第五人民医院, 江苏 常熟 215500; 2. 中国矿业大学 信息与控制工程学院, 江苏 徐州 221116;

3. 国防科技大学, 长沙 410073)

**摘要:** 为充分挖掘图像数据信息, 提出了一种有向图模型检索方法, 结合距离测度初次检索和有向图距离二次检索提高图像检索性能; 首先, 采用传统的纹理、边缘和颜色特征以及特征之间的欧氏距离测度来进行初次检索, 得到一个查询排序列表; 在此基础上, 结合距离测度与余弦测度设计图像之间的相关测度, 在不同的相关测度阈值下构建图像数据集的有向图模型集合; 最后, 计算有向图距离, 据此进行二次检索, 降低误检现象; 在 COREL 和 ImageCLEF 两个数据集上的图像检索实验结果表明, 该方法的平均精确度和平均召回率指标高。

**关键词:** 图像检索; 有向图模型; 欧氏距离; 图距离; 余弦测度

## A retrieval Method Based on Directed Graph Model for Image Data Mining

Gong Hui<sup>1</sup>, Xia Kaijian<sup>2</sup>, Jin Zhaoyan<sup>3</sup>

(1. Changshu No. 5 People's Hospital, Changshu 215500, China; 2. School of

Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China;

3. National University of Defense Technology, Changsha 410073, China)

**Abstract:** To fully dig the image data information, a retrieval method based on directed graph model is proposed, for improving performance of image retrieval by combining first retrieve with distance metric and second retrieve with directed graph distance. First, it executes first retrieve by using traditional features including texture, edge and color and the Euclidean distance among features, and obtains a query sort list; On this basis, it builds directed graph models with different correlation metric thresholds, according to correlation metric that is designed by combining with distance metric and cosine metric; finally, it computes directed graph distance and executes second retrieve, for reducing false retrieval phenomena. The results of image retrieval experiments on COREL and Image CLEF datasets show that, this method has average precision and average recall.

**Keywords:** image retrieval; directed graph model; Euclidean distance; graph distance; cosine metric

## 0 引言

多媒体和互联网技术的快速发展, 带来了海量的图像资源, 如何从海量图像数据中挖掘有用信息是一个具有挑战的研究课题。其中, 研究基于内容的图像检索 (Content-Based Image Retrieval, CBIR) 技术是挖掘有用图像信息的有效手段之一<sup>[1]</sup>。CBIR 的基本思想是提取图像中具有鉴别能力的特征, 然后通过相似性测度或者分类器来实现图像的检索<sup>[2]</sup>。基于纹理特征的检索是最常用的图像检索方法, 此类方法常提取图像的灰度共现矩阵特征、小波变换特征等纹理特征, 然后采用欧氏距离等相似性测度来实现图像的检索<sup>[3-5]</sup>。为了提高特征的鉴别能力, 提出了许多多特征融合的图像检索方法, 包括融合纹理、颜色、边缘等特征。还有许多新的特征描述子用于图像检索领域, 如线性判别分析特征、高层语义特征、词袋特征等。近些年, 随着深度学习方法的深入研究, 采用深度学习思想自动提取鉴别能力强的特征也是 CBIR 领域常用的特征提取方式之一, 包括提取图像的卷积神经网络特征 (CNN) 等<sup>[6-10]</sup>。除了特征提取之外, 特征的分类也是图像检索技术的重要环节之一。欧氏距离、卡方距离、相关系数、余弦测度等都是测量特征相似性的有效手段, 还有支持向量机、随机森林

等机器学习方法也常用来进行特征分类<sup>[11-13]</sup>。理论上讲, 特征分类方法本身并没有明显的优劣之分, 其分类性能主要与应用场景相关, 包括特征的表达力、数据集的构建以及应用中偏重的评价指标等。就图像检索而言, 目前在一些简单的测试数据集上已经达到了很高的检索指标。然而对于一些复杂的测试数据集, 误检现象还是过于频繁。为了降低误检现象, 本文源于二次检索的思路, 提出了一种基于有向图模型的图像检索方法。设计思想是在传统图像检索方法的基础上, 增加一级基于有向图距离测度的二次检索环节, 降低误检现象。其核心是依据图像之间的距离测度和相关特性构建有向图模型, 借助图像之间的相关性降低单纯依赖距离造成的误检现象, 提高图像检索性能。

## 1 本文方法

本文提出的图像检索方法的核心是构建图像数据集的有向图模型, 依据有向图模型描述图像之间的相关特性以及特征之间的距离关系, 通过距离测度的初次检索和有向图距离的二次判断, 来降低图像检索可能出现的误检现象。下面首先介绍有向图模型的元素组成以及构建方法, 然后再介绍基于有向图模型的图像检索方法。

### 1.1 有向图模型

令  $S = \{I_i \mid i=0, 1, \dots, n\}$  表示一个图像集合, 其中,  $n$  表示集合中图像中的数量。

令  $\varphi$  表示一个图像映射函数, 用于提取图像特征, 图像  $I_i$

收稿日期: 2018-02-24; 修回日期: 2018-03-01。

作者简介: 龚 晖 (1988-), 男, 江苏常熟人, 大学, 工程师, 主要从事计算机应用技术方向的研究。

对应的特征向量  $v_i$  可以表示为:

$$v_i = \varphi(I_i) \quad (1)$$

令  $D$  表示一个距离测度函数, 用于计算两个特征向量之间的距离。特征向量  $v_i$  和  $v_j$  之间的距离可以表示为:

$$d_{ij} = D(v_i, v_j) \quad (2)$$

本文将图像映射函数  $\varphi$  和距离测度函数  $D$  组成一个二元组  $\Psi = [\varphi, D]$ , 作为图像的描述子。

在图像查询时, 对于查询图像  $I_q$  而言, 可以计算该图像与数据库中各个图像之间的距离测度, 然后按照距离由小到大的顺序给出查询排序列表, 表示为:

$$\tau_q = (I_{n_1}, I_{n_2}, \dots, I_{n_s}) \quad (3)$$

其中:  $n_s$  表示查询到的图像总数, 满足条件  $n_s = n$ , 于是有  $\tau_q \subset S$ 。

本文采用有向图模型描述图像之间的关联特性, 记为:

$$G = (V, E) \quad (4)$$

其中, 顶点集合  $V$  由图像集合  $S$  代替, 也即  $V = S$ 。这样, 每一幅图像都是有向图模型中的一个顶点。边集合  $E$  用于定义图像之间的相关性。令  $c(q, n_j)$  表示图像  $I_q$  和  $I_{n_j}$  之间的相关测度, 两幅图像的相关测度得分越大, 说明两幅图像之间的关联特性越强。当两幅图像之间的相关测度得分超过给定阈值  $t_c$ , 且测试图像  $I_{n_j}$  处于查询图像  $I_q$  的查询排序列表  $\tau_q$  的前  $n_s$  个位置时, 本文认为图像  $I_q$  到  $I_{n_j}$  的方向上存在一条边, 可以表示为:

$$E = \{(I_q, I_{n_j}) \mid \tau_q(n_j) \leq n_s, c(q, n_j) \geq t_c\} \quad (5)$$

其中:  $\tau_q(n_j)$  表示排序列表  $\tau_q$  中图像  $I_{n_j}$  的排列位置。很明显, 相关测度阈值  $t_c$  越小, 得到的有向图模型越稠密。反之, 得到的有向图模型越稀疏。

综上所述, 用于图像检索的有向图模型构建涉及三个关键环节, 包括图像映射函数  $\varphi$ 、距离测度  $D$  和相关测度  $c$ 。详细介绍如下。

1) 图像映射函数: 图像映射函数用于提取图像的特征。对于图像检索而言, 常用的特征有纹理特征、颜色特征、边缘特征等。选择具有显著性和稳健性的特征对于图像检索而言意义重大。然而, 对于不同的图像类别, 其显著性和稳健性特征可能存在较大差异。因此, 需要根据实验数据的不同有针对性的选择图像映射函数, 提取图像特征。在图像特征提取方面目前已经有很多成熟的方法, 这不是本文的研究重点。在本文的仿真实验中, 根据本文所选的两个数据集的不同, 本文结合文献 [10] 所述方法以及本文的实验结果, 选择不同的纹理、边缘和颜色特征, 具体在实验部分详细介绍。

2) 距离测度: 距离测度用于测量两个特征向量之间的相似程度, 常用的距离测度有欧氏距离、街区距离、明氏距离、马氏距离、卡方距离等。本文选择欧氏距离作为距离测度, 原因是该距离测度兼顾了运算效率和度量精度, 是目前应用最广的距离测度。于是有

$$D(v_i, v_j) = \|v_i - v_j\| \quad (6)$$

3) 相关测度: 相关测度是本文在构建有向图模型的过程中, 结合距离测度与夹角余弦测度设计的, 对于有向图模型的构建影响很大。图像检索领域通常采用距离测度来定义两幅图像之间的相关性, 对于图像  $I_q$  和  $I_j$  而言, 两幅图像越相似, 通常其距离越小。换言之, 相邻图像之间的距离是相关的。因此, 本文结合  $k$  近邻和距离测度构建图像的相关测度。详细描

述如下。

令  $N_k(q)$  表示一个包含了给定查询图像  $I_q$  的  $k$  个最近邻距离的图像集合。令  $N_k(q, j)$  表示一个包含了图像  $I_q$  和  $I_j$  的  $k$  个最近邻距离的图像集合, 也即有:

$$N_k(q, j) = N_k(q) \cup N_k(j) \quad (7)$$

对于图像集合  $N_k(q, j)$  中的第  $i$  个图像  $I_i$ , 其与图像  $I_q$  的距离可以表示为:

$$x_i = D(\varphi(I_q), \varphi(I_i)) \quad (8)$$

同样地, 图像  $I_i$  与图像  $I_j$  的距离可以表示为:

$$y_i = D(\varphi(I_j), \varphi(I_i)) \quad (9)$$

本文采用夹角余弦测度来计算图像  $I_q$  和与图像  $I_j$  之间的相关测度得分, 可以表示为:

$$c(q, j) = \frac{X^T Y}{\sqrt{(X^T X)(Y^T Y)}} \quad (10)$$

其中:

$$X = [x_1, x_2, \dots, x_{2k}] \quad (11)$$

$$Y = [y_1, y_2, \dots, y_{2k}] \quad (12)$$

相关测度得分越大, 说明对应图像之间的相关性越强。

## 1.2 图像检索方法

本文结合有效图模型来进行图像检索, 设计思想是: 先用传统的距离测度进行初次检索, 缩小图像检索范围。然后再依据有效图距离进行二次判断, 降低误检现象。具体实现方法是, 首先, 对于数据库中的每一幅图像  $I_i$ , 提取特征向量  $v_i = \varphi(I_i)$ , 并计算任意两幅图像对应的两特征向量  $v_i$  和  $v_j$  之间的距离  $d_{ij} = D(v_i, v_j)$ , 存储在特征数据库中。然后, 对于查询图像, 构建有向图模型, 计算有向图距离, 依据有向图距离选择图像检索结果, 具体实现步骤描述如下。

Step1, 对于查询图像  $I_q$ , 提取特征向量  $v_q = \varphi(I_q)$ , 并计算其与数据库中任一图像  $I_i$  对应的两特征向量之间的距离  $d_{qi} = D(v_q, v_i)$ 。

Step2, 按照距离值从小到大的顺序进行排序, 得到排在前  $n_s$  位的查询排序列表  $\tau_q = (I_{n_1}, I_{n_2}, \dots, I_{n_s})$ 。

Step3, 给定初始相关测度阈值  $t_c = t_s$ , 计算图像  $I_q$  与查询排序列表  $\tau_q$  中任意一幅图像  $I_j$  之间的相关测度  $c(q, j)$ , 构建有向图模型  $G_1 = (V, E_1)$ 。其中, 顶点集合  $V$  由数据库中所有图像再加上查询图像构成。

Step4, 相关测度阈值增加  $\Delta t$  (即  $t_c = t_c + \Delta t$ ), 重复 Step3, 构建有向图模型  $G_i = (V, E_i)$ 。其中, 序号  $i$  表示构建的第  $i$  个有向图模型。该过程重复执行, 直到  $t_c \geq 1$ 。

Step5, 计算有向图距离。

Step6, 按照有向图距离值从小到大的顺序进行二次排序, 得到最新的查询排序列表  $\tilde{\tau}_q$ 。

Step7, 依据设置的查询余量  $T$ , 输出查询排序列表  $\tilde{\tau}_q$  中的前  $T$  个图像作为最终的查询结果。

本文将有向图距离作为图像二次检索的依据。考虑到图像越相似, 其相关性越强。因此, 本文依据不同相关测度阈值下有向图模型中边的数量来计算有向图距离, 具体实现步骤描述如下。

首先, 定义一个一维的图像对连接权重向量, 记为:

$$W = [\omega_{n_1}, \omega_{n_2}, \dots, \omega_{n_s}] \quad (13)$$

其中:  $\omega_{n_i}$  表示查询图像  $I_q$  与图像  $I_{n_i}$  之间的连接权重。本文只对查询排序列表中的图像计算其与查询图像之间的连接

权重, 因为列表之外的图像与查询图像之间的距离太大, 不可能是相似图像。在初始化阶段, 将该权重向量的各个元素的初始值置为 0。

然后, 遍历有向图模型集合  $\{G_i = (V, E_i) \mid i=0, 1, \dots, m\}$  中的每一个有向图模型, 其中,  $m$  表示前面构建的有向图模型的总数。对于任意一个有向图模型  $G_i = (V, E_i)$ , 遍历其中的每一条边, 依据边是否存在以及对应的相关测度值对权重向量中的各个元素执行累加操作, 具体可以表示为:

$$\omega_{n_j} = \begin{cases} \omega_{n_j}, & (I_q, I_{n_j}) \notin E_i \\ \omega_{n_j} + \frac{1}{c(q, n_j)}, & (I_q, I_{n_j}) \in E_i \end{cases} \quad (14)$$

这样, 待遍历完毕所有有向图模型中的所有边之后, 可以得到最终的权重向量。

最后, 计算查询图像与查询排序列表  $\tau_q = (I_{n_1}, I_{n_2}, \dots, I_{n_s})$  中各个图像之间的有向图距离。其中, 查询图像  $I_q$  与图像  $I_{n_j}$  之间的有向图距离可以表示为:

$$d_{n_j} = \frac{1}{\omega_{n_j} + 0.0001} \quad (15)$$

其中: 上式的分母项所加的小数是为了避免分母为零。

## 2 实验与分析

下面进行图像检索实验, 将本文方法与图像检索领域目前性能较好的方法进行对比分析, 定量评价本文方法的图像检索性能。首先, 介绍本文使用的图像检索数据集; 然后, 介绍图像检索领域常用的定量评价指标; 接着, 结合实验结果选择本文方法适用的参数和特征; 最后, 对比分析不同方法的图像检索结果, 评价本文方法的图像检索性能。

### 2.1 实验数据集

图像检索领域的公开测试数据集较多, 本文与文献 [10] 一样, 选择 COREL 和 ImageCLEF 两个数据集, 简要介绍如下。

1) COREL 数据集: 该数据集包含 50 个不同的类别的图像集合, 每一个类别有 100 幅图像, 共 5000 幅图像。每一个类别代表了不同的语义, 如花、猫、狗等。图像存储为 24 位彩色图像。部分图像样本示例如图 1 所示。



图 1 COREL 数据集样本示例

2) ImageCLEF 数据集: 该数据集为医学图像数据集, 包含 20 个不同类别的图像集合, 每一个类别至少有 100 幅图像, 共 6157 幅图像。每一个类别代表了不同的语义, 如手部、胸部、脚部等。图像存储为 256 色灰度图像。部分图像样本示例如图 2 所示。

### 2.2 性能评价指标

直观地讲, 图像检索算法所检索到的相关图像越多, 不相



图 2 ImageCLEF 数据集样本示例

关图像越少, 则图像检索算法的性能越好。基于此, 图像检索领域常用的性能评价指标有两个: 一是平均精确度 (average precision, AP) 指标, 二是平均召回率 (average recall, AR) 指标, 分别定义为:

$$AP = \frac{\text{检索到的相关图像数量}}{\text{所有检索到的图像数量}} \quad (16)$$

$$AR = \frac{\text{检索到的相关图像数量}}{\text{数据集中相关图像数量}} \quad (17)$$

平均精确度指标越高, 说明检索结果中干扰越少。平均召回率越高, 说明检索结果越有效。两个指标的值越高, 对应的图像检索算法的性能越好。

### 2.3 特征与参数选择

#### 2.3.1 特征选择

对于 COREL 数据集, 由于图像是彩色图像, 可以利用颜色信息辅助图像检索, 参考文献 [10] 的实验结果, 本文选择颜色、边缘和纹理三类特征, 对于每一幅图像, 颜色特征向量为 6 维, 包含图像 H、S 和 V 三个颜色通道的均值和方差。边缘特征的提取方法是, 先将彩色图像转换为灰度图像, 再采用 Canny 算子求取图像边缘, 接着计算边缘方向直方图, 其中, 边缘方向直方图按  $20^\circ$  一个方向进行量化, 最后得到一个 18 维的边缘特征向量。纹理特征的求取方法是, 在灰度图像上, 执行 5 个尺度和 8 个方向的 Gabor 小波变换, 得到 40 个子图像, 然后计算每一幅子图像的均值、方差, 得到一个 80 维的纹理特征向量。最终, 将颜色特征向量、边缘特征向量和纹理特征向量串联在一起, 每一幅图像可以得到一个 104 维的特征向量。

对于 ImageCLEF 数据集, 由于图像是灰度图像, 本文仅选择边缘和纹理两类特征, 边缘和纹理特征的提取方法如前所述, 这样, 每一幅图像可以得到一个 98 维的特征向量。

#### 2.3.2 参数选择

本文涉及的参数主要有 5 个, 包括初次查询图像数量  $n_s$ 、最近邻图像集合数量  $k$ 、查询余量  $T$ 、相关测度阈值的初始值  $t_s$  和增量  $\Delta t$ 。其中, 为了便于对比不同方法的性能, 所有对比方法的查询余量必须一致, 都设为  $T=20$ 。相关测度阈值的初始值  $t_s$  和增量  $\Delta t$  取经验值, 具体为  $t_s=0.3$ 、 $\Delta t=0.01$ 。其他参数依据实验结果来设置最优的值。

图 3 给出了参数  $n_s$  取不同值时的平均精确度和平均召回率曲线 (数据集选用的是 COREL 数据集), 其中, 参数  $k$  取 25。为了避免在初次查询时遗漏查询目标, 初次查询图像数量应当远大于最终查询的图像数量, 因此, 实验时参数  $n_s$  的初始值设置为查询余量的 5 倍, 也即  $n_s=100$ , 然后每间隔 20 进行一次实验, 从图 3 所示的实验结果可以发现, 随着  $n_s$  的增

大, 平均精确度和平均召回率都有所增加, 当  $n_s$  大于 160 之后, 平均精确度和平均召回率都趋于稳定。考虑到  $n_s$  越大, 图像检索效率越低, 因此, 本文取  $n_s = 160$ 。

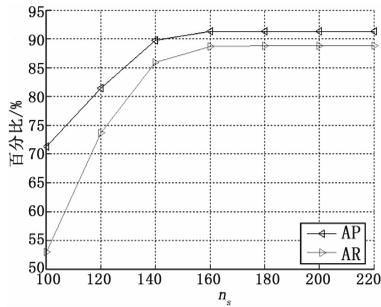


图 3 参数  $n_s$  取值不同时的实验结果

图 4 给出了参数  $k$  取不同值时的平均精确度和平均召回率曲线 (数据集仍选用 COREL 数据集), 其中, 参数  $n_s = 160$ 。实验发现, 当参数  $k$  取 25 时, 平均精确度和平均召回率都达到峰值。因此, 本文取  $k = 25$ 。

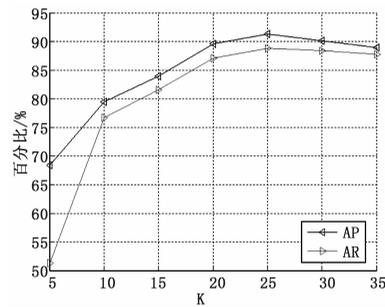


图 4 参数  $k$  取值不同时的实验结果

### 2.4 图像检索性能评价

下面在 COREL 和 ImageCLEF 两个数据集下进行图像检索实验, 其中, 本文方法所使用的参数如前所述, 对比方法所使用的参数出自对应文献。图 5 和图 6 分别给出了 COREL 和 ImageCLEF 两个数据集下图像检索的平均精确度和平均召回率对比结果。

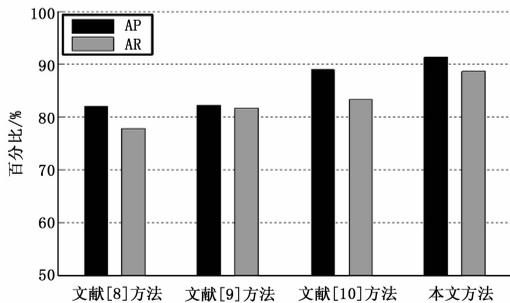


图 5 COREL 数据集实验结果

结合两幅实验结果图像可以发现, 在 COREL 数据集下, 本文方法平均精确度和平均召回率两个指标高于其他三种对比方法 2.4% 和 5.5% 以上。在 ImageCLEF 数据集下, 本文方法平均精确度和平均召回率两个指标高于其他三种对比方法 2.1% 和 4.2% 以上。尤其是与文献 [10] 所述方法相比, 两者所使用的图像特征基本相同, 但本文方法通过有向图距离的

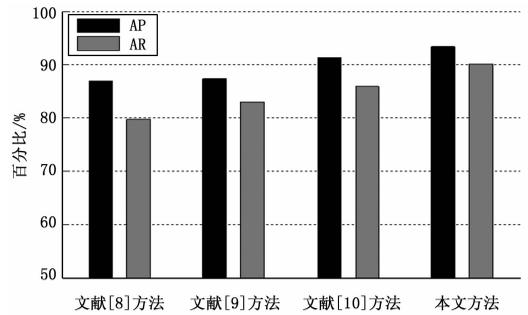


图 6 ImageCLEF 数据集实验结果

二次搜索大幅降低了误检现象, 平均精确度和平均召回率两个指标都明显高于文献 [10] 所述方法。综上所述, 本文方法的图像检索性能优于其他三种对比方法, 是一种有效的图像检索方法。

### 3 结束语

为了提高图像检索性能, 本文提出了一种基于有向图模型的图像检索方法。该方法依据图像之间的特征距离和相关特性, 设计图像数据集的有向图模型。在图像检索时, 先采用传统的欧氏距离测度初步检测符合图像检索条件的查询图像集合。在此基础上, 采用有向图距离测度进行二次检索, 对欧氏距离排序的查询图像列表进行二次排序, 降低误检现象。实验结果表明, 本文方法在图像检索实验中得到的平均精确度和平均召回率高, 是一种有效的图像检索方法。目前该技术主要可以应用在信息识别等相关领域。

### 参考文献:

- [1] Thenkalvi B, Murugavalli S. Review on CBIR trends and techniques to upgrade image retrieval [J]. International Review on Computers & Software, 2014, 9 (7): 1227-1240.
- [2] Gong Y, Lazebnik S, Gordo A, et al. Iterative quantization: a procrustean approach to learning Binary codes for large-scale image retrieval [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35 (12): 2916-2929.
- [3] 许元飞. 基于纹理的图像检索算法研究 [J]. 西安科技大学学报, 2013, 33 (4): 470-474.
- [4] 李明灵, 吴国平, 耿秀秀, 等. 基于 LBP-GLCM 纹理特征提取的服装图像检索 [J]. 电视技术, 2015, 39 (12): 99-103.
- [5] 葛 芸, 江顺亮, 叶茂茂, 等. 视觉词袋和 Gabor 纹理融合的遥感图像检索 [J]. 光电工程, 2016 (2): 76-81.
- [6] Liu G H, Yang J Y. Content-based image retrieval using color difference histogram [J]. Pattern Recognition, 2013, 46 (1): 188-198.
- [7] Hu R, Collomosse J. A performance evaluation of gradient field HOG descriptor for sketch based image retrieval [J]. Computer Vision & Image Understanding, 2013, 117 (7): 790-806.
- [8] Wang X, Wang Z. A novel method for image retrieval based on structure elements' descriptor [J]. Journal of Visual Communication & Image Representation, 2013, 24 (1): 63-74.
- [9] Yang Y, Newsam S. Geographic Image Retrieval Using Local Invariant Features [J]. IEEE Transactions on Geoscience & Remote Sensing, 2013, 51 (2): 818-832.