Mapping Vegetation-Covered Urban Surfaces Using Seeded Region Growing in Visible-NIR Air Photos

Jianhua Zhou, Yan Huang, and Bailang Yu, *Member, IEEE*

***Abstract*—Unreliability involved in the extraction of shaded vegetation-covered surfaces (VS) is a common problem in urban vegetation mapping. Serving as a solution to it, a novel method named Nonlinear Fitting-based Seeded Region Growing (NFSRG) is explored. With NFSRG, a series of classified results are organized by a seeded-region-growing process. In order to adapt to the variable separability between VS and background, the growing is limited in several weighted buffers defined by some nonlinear fitting relationships. When searching new VS members (member means both pixel and patch) within such a buffer, a gradually reduced weight makes the buffer width continually narrowed as the separability worsens. To avoid unexpected entrances of water and smooth shaded background members, a during-growing constraint, named *expansion rate,* is proposed. Accuracy assessments reveal that more than 96% of VS members can be accurately extracted by the proposed method.**

***Index Terms*—Classification, seeded region growing, shadow, urban, vegetation.**

I. Introduction

URBAN VEGETATION is an important component of urban ecosystems and performs significant environmental, recreational, and aesthetic functions [1]. Existing studies have revealed that the layout and abundance of urban vegetation significantly influence boundary-layer climates at local and micro-scales [2], [3]. Thus, timely and accurate information on the layout of urban vegetation is important for urban environment modeling, ecological benefits assessment, and horticulture planning, etc.

Several types of vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) and soil-adjusted vegetation index (SAVI), are commonly used to derive vegetation information [4]–[6] and to evaluate ecological benefits [7] typically with moderate-resolution remote sensing data. When the derivation is in downtown area at individual tree level, the extremely high heterogeneity of vegetation details and their surroundings may lead to significant uncertainties over the deriving accuracy [8]. Therefore, high spatial resolution (e.g., IKONOS, Quickbird, and aerial photography) [9]–[12] and high spectral resolution (e.g., hyperspectral scanner) [13] data have been utilized for urban VS mapping. Although these carefully improved data have verified their potential for urban applications, nonetheless, many challenges and problems remain to be addressed.

For example, a severe shadow effect, typically caused by the interaction between lower Sun angles and higher buildings in urban area, may significantly damage the accuracy of classification of shaded members [14], [15]. Several works related to the identification of shaded members have been reported [4], [16]–[18]. Some of them assigned the whole shaded surfaces into a single class, probably limited by the poor separability between different shaded classes which is often beyond the capabilities of applied classification methods [4], [16]. Although some researchers have paid attention to the detection and reconstruction of shaded scene, due to the difficulties in compensating for each band of weakened reflections in the scene, so far, only visually, instead of spectrally, reconstructed shadow-free imagery can be obtained [15], [19]–[21].

Fortunately, it is not a rare situation that a shaded member geographically connects to a bright one involved in the same class (e.g., the shaded and the bright VS members). This lays the conceptual foundation for **NFSRG**. The shadow effect to the accuracy of vegetation mapping may be promised to be resolved because the work mainly relies on the geographical connectivity defined by image coordinates, rather than the spectral similarity defined by domains of image features. The objective of this research is to improve the performance of extracting shaded VS members at individual tree level for urban vegetation mapping thus to avoid some almost inevitable mistakes involved in typical classification.

II. Study Site and Test Images

The study site is in the downtown area of Shanghai, located on the eastern coast of mainland China. Vegetation covered surfaces often show intensive planting and complex layout in the city and most of them are surrounded by buildings and other urban facilities therefore making a large proportion of them shaded.

More than 20 aerial false-color near infrared (NIR) images, referred to as "NIR images" in the following text, serve as the experimental data. The images were purchased from the geographic information service institution of the government. The sensors aboard on were of photogrammetric cameras. The original photographic scales were from 1:8000 to 1:15 000 and therefore the original spatial resolution at nadir was better than 2 m. Each photo was later scanned and geometrically

Manuscript received March 29, 2014; revised July 21, 2014; accepted September 30, 2014. This work was supported by the National Natural Science Foundation of China under Grant 41071275. *(Corresponding author: Jianhua Zhou.)*

The authors are with the Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai 200241, China (e-mail: jhzhou@geo.ecnu.edu.cn; hyansea@hotmail.com; blyu@geo.ecnu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2014.2362308

and orthographically corrected into digital image before selling. The image colors of red, green, and blue indicate NIR (760–850 nm), R (red, 630–690 nm), and G (green, 520–600 nm) bands, respectively, because the photographic film was sensitive to the reflection of NIR band. To verify the generality of the **NFSRG** method, the test images were randomly selected from nine image groups which were acquired in different years and also likely with different sensors and processing equipment.

In Shanghai, likewise with all rainy southern cities in China, NIR images are widely used for city surveying and mapping due to the difficulties in acquiring satellite images all seasons with satisfying cloud cover. Therefore, it is desirable for these cities to gain access to the technology of mapping VS from NIR images.

III. Methodology

A. Overview

NFSRG can be understood as a kind of hybrid classification in which a series of pixel-based classifications are organized as a seeded-region-growing process. From the view point of supervised classification, the iteratively captured VS members always serve as training samples for the next classification. From the perspective of seeded region growing, the members will serve as seeds for the next growing.

A whole **NFSRG** process involves three linked stages: extracting bright VS, extracting shaded VS within the right and the left buffers. Each stage consists of two steps: seed collection and region growing. Two key issues contribute to its good performance. 1) *The weighted buffer*: The growing is limited in several weighted buffers to fit for the variable separability between VS and background. The one, defined by an NDVI-S relationship where S is the saturation in the Hue-saturation-value color model, is for extracting bright VS members. The other two, defined by a pair of NDVI–NDSV relationships where NDSV is the normalized difference between saturation and brightness, are for extracting shaded VS members. 2) *The during-growing constraint*: To avoid unexpected entrances of water and smooth shaded background members, a during-growing constraint, named *expansion rate*, is imposed on the growing process for extracting shaded VS members. Fig. 1 shows the flowchart of NFSRG.

B. Image Sampling and Nonlinear Fitting

During the image sampling, a cursor-pointed pixel indicating a certain class serves as a sample (see Fig. 2). Its features are calculated in the same time and prepared for the nonlinear fitting. For each test image, there are no less than 50 samples (cursor-pointed pixels) for each of the classes.

The sampling data reveal that shaded VS samples usually possesses higher S and lower V (brightness) therefore larger value of S subtracting V than those of shaded background ones. Above-mentioned NDSV is derived by normalizing the value (1) [22]. NDSV may offer an opportunity to distinguish between shaded VS and shaded background

$$NDSV = (S - V)/(S + V). \tag{1}$$

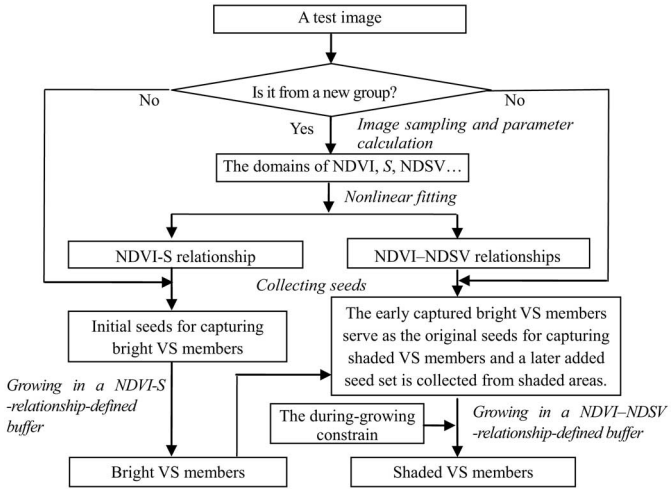

Fig. 1. Flowchart of **NFSRG**.

The relationships of NDVI-S and NDVI–NDSV can be expressed as the general forms of quadratic nonlinear correlation as in (2) and (3). Fig. 2 provides a case couple of the fitting curves for No. 4 image group. The curves for other eight groups are provided in Fig. 3

$$S = a1 \cdot NDVI^2 + b1 \cdot NDVI + c1 \tag{2}$$

$$NDSV = a2 \cdot NDVI^2 + b2 \cdot NDVI + c2 \tag{3}$$

where $a1$, $b1$, $c1$, $a2$, $b2$, and $c2$ are the undetermined coefficients.

It can be seen from Figs. 2 and 3 that the shaded VS samples can typically be split into left and right subsets by complying with $\mathrm{NDVI} = 0$. A pair of NDVI–NDSV relationships, the left and the right, are separately derived from the subsets and then used to define the left and the right buffers.

Table I lists the coefficients of the relationships for all nine image groups. Our experiments reveal that the coefficients did not change significantly between the images of the same group. Consequently, a set of relationships, typically one NDVI-S and two NDVI–NDSVs, for a certain group are applicable to all the images of that group.

Some phenomena can also be discovered from the table and figures. For example, the larger the contrast of DNs of an NIR image is, typically the larger the NDVI is. In this case, almost all the VS samples fall into the right subset (e.g., No. 6 and 7 groups in Fig. 3) and the case can be found out by the laws of $a21 = 0$ and $b21 = 0$.

C. Extracting VS Using Seeded Region Growing

The seeded region growing is under the constraint of "geographic-space connectivity," which means a pixel connecting to others involved in their image coordinates. The growing always begins from a more reliable end of a more reliable buffer. Within a buffer, the lower the NDVI is, often the more seriously confused the VS and background members become. However, the gradually narrowed buffer associated with a continually decreased weight and the constraint of

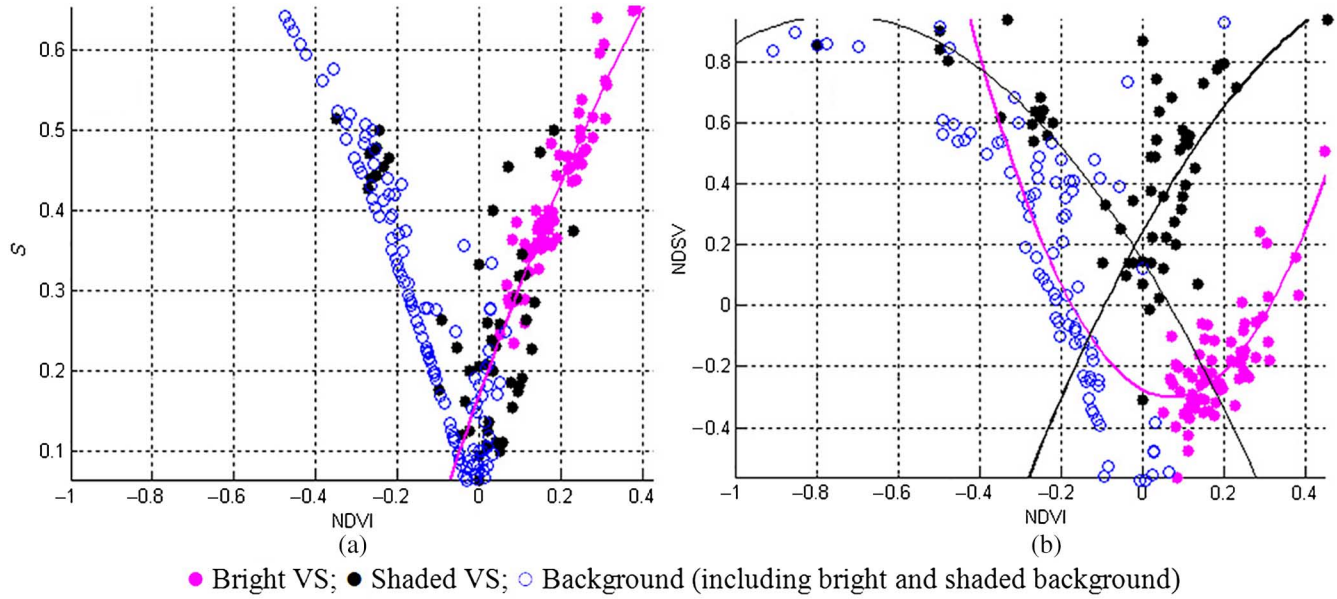

Fig. 2. Case couple of fitting curves. (a) Fitting in NDVI-S space. (b) Fitting in NDVI–NDSV space. It can be seen from (a) that the bright VS and the background samples are separable in an NDVI-S space. Therefore, an NDVI-S relationship (magenta curve) for extracting bright VS can be derived by a nonlinear fitting with the bright VS samples complying with (2). On the other hand, it is difficult to separate shaded VS from background samples in the same space. However, it is not a rare situation that they appear separable in an NDVI–NDSV space [see (b)] and typically can be split into left and right subsets. Thus, two NDVI–NDSV relationships (black curves) for extracting shaded VS can be derived complying with (3).

geographic-space connectivity as well as the during-growing constraint (see Section III-D) may ensure the full extraction of the VS members and also avoid the unexpected entrance of the background members.

As mentioned before, the extraction involves three linked stages: extracting bright VS, extracting shaded VS within the right and the left buffers. Next, they will be introduced separately.

1) Extraction of Bright VS: The initial seed set (*seedL*) consists of two subsets (4). The first [the main, e.g., the set in Fig. 8(a)] complies with NDVI $> VIL_{07}$, while the other [the extra, e.g., the set in Fig. 8(b)] is limited within a narrow region along an NDVI-S curve

$$\begin{aligned} seedL = \{seedL | NDVI &> VIL_{07} \cup S \\ &> (f_{VI-S}(VIL_{05}) \pm P_5 \cdot dSL) \cap S \\ &< (f_{VI-S}(VIL_{07}) \pm P_{20} \cdot dSL)\} \end{aligned} \tag{4}$$

where $maxVIL$ is the maximum NDVI of the bright VS samples; $f_{\mathrm{VI-S}}$ is an NDVI-S relationship derived from the bright VS samples; $VIL_{05} = \mathrm{c} \cdot maxVIL \cdot 0.5$, where c is an adjustable coefficient with an experimental default of 1.0; $VIL_{07} = \mathrm{c} \cdot maxVIL \cdot 0.7$; P is the weight of buffer, where P_5 and P_{20} mean $P = 5\%$ and $P = 20\%$, respectively; and *dSL* is the range of S of the bright VS samples. The region for collecting seeds is typically defined by certain samples. Those constants (e.g., 0.7%, 0.5%, 5%, and 20%) are experimental defaults and occasionally need to be tuned. The coefficient c is ready for some possible changes such as a variation in image type. Increasing c may reduce the number of VS members, and vice versa. The constants in (5a)–(9) are similar.

Equation (5a) formulates an iterative process to perform seeded region growing within a weighted buffer defined by an NDVI-S relationship. Figs. 4 and 5 show a case form and the general form of the process, respectively

$$VL = \sum_{\substack{ND_i = VIL_{07} \\ P_i = 0.4 \\ i=0}}^{\substack{ND_i = \min VIL \\ P_i = 0 \\ i=n}} (f_{VI-S}(ND_i) \pm P_i \cdot dSL) \tag{5a}$$

$$VL_0 = seedL \tag{5b}$$

$$VL_{i-1} \subseteq VL_i = TURE \tag{5c}$$

where VL is an iteratively increasing set of bright VS members [e.g., the set in Fig. 8(c)]; ND_{i} is the ith NDVI with an increment of -0.01; P_{i} is the ith weight; and n is the number of iterations and $n = (VIL_{07} - minVIL)/0.01$. Condition (5b) means that $seedL$ serves as the growing start. Condition (5c) says that the iteratively captured VS members always serve as the newly added seeds for the next growing.

2) Extraction of Shaded VS: The shaded VS samples can be typically split into left and right subsets. If the number of samples in one of them is less than a given value (e.g., less than one-fourth of the total shaded VS samples) the subset will be dispensed with. Otherwise, both buffers derived from the two subsets will be considered and the growing will conduct within the right one at first and then the left.

An extra seed set for capturing shaded VS members is always required because some of the shaded may not connect geographically to the early captured members in *VL*. Equations (6) and (8) depict the rules to collect the extra seeds within the right

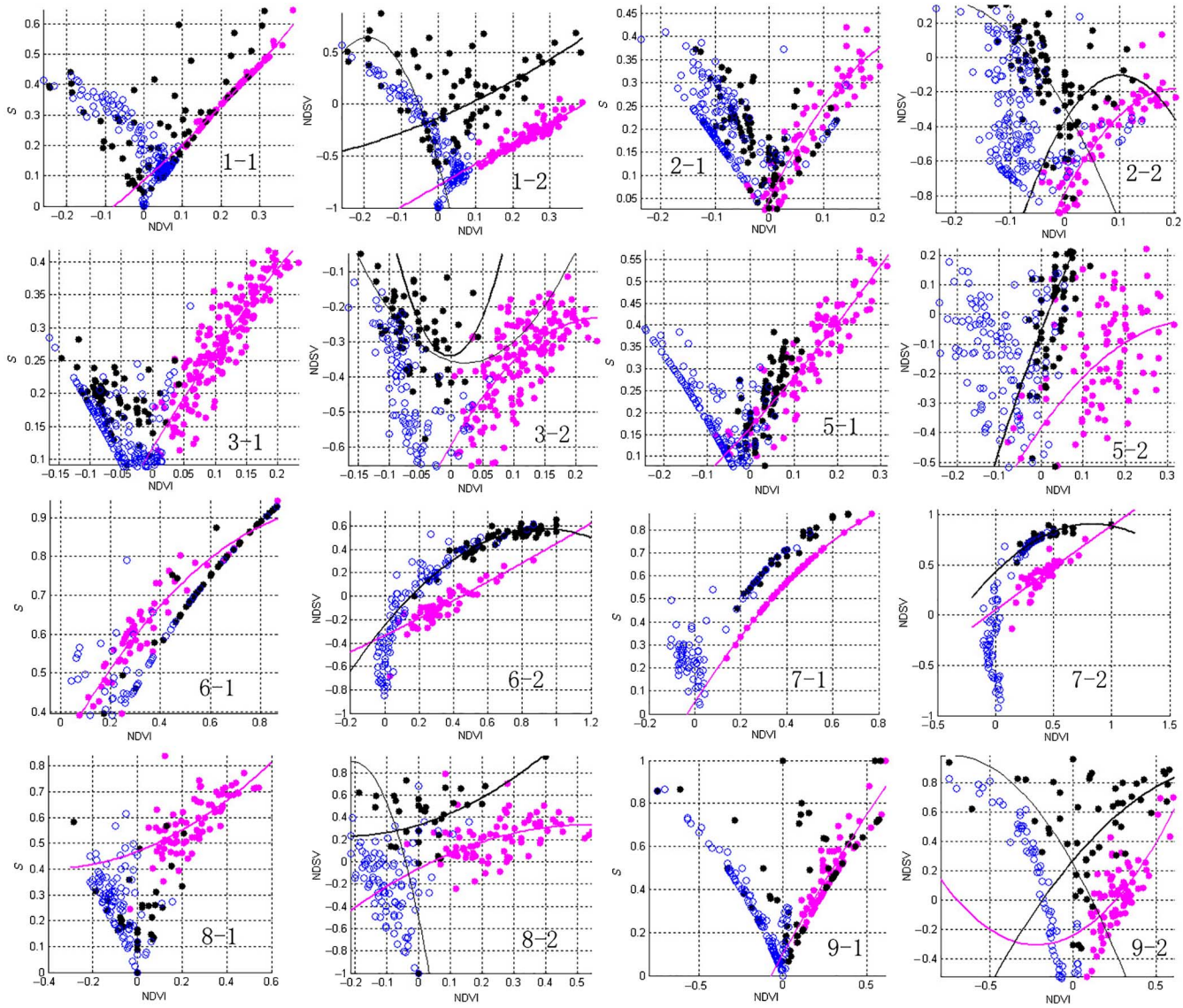

Fig. 3. Eight couples of fitting curves in NDVI-S and NDVI–NDSV spaces. There are the fitting curves for other eight image groups with the same legends as in Fig. 2. For each label, the first number is the serial number of image group and the other is the index of relationship, NDVI-S and NDVI–NDSV referring to as "1" and "2," respectively. For example, 3-2 means the NDVI–NDSV relationship/curve for No. 3 image group. The absent graphs which should be labeled as 4-1 and 4-2 are those given in Fig. 2.

and left buffers, (7a) and (9a) formulate the iteratively growing processes within the two buffers, and Figs. 6 and 7 show the general processes, respectively. Fig. 8 provides an example of extracting VS by **NFSRG** with intermediate outputs

$$\begin{aligned} seedD_1 = \{seedD_1 | NDSV \\ > (f_{VI-SV}(VID_{07}) \pm P_5 \cdot dSVD) \cap NDSV \\ < (f_{VI-SV}(\max VID) \pm P_{20} \cdot dSVD)\} \end{aligned} \tag{6}$$

where $seedD_1$ is an extra seed set collected from the right buffer [e.g., the set in Fig. 8(d)]; $f_{\mathrm{VI-SV}}$ is a right NDVI–NDSV relationship; $maxVID$ is the maximum NDVI of the shaded VS samples; $VID_{07} = c \cdot maxVID \cdot 0.7$; $dSVD$ is the range of NDSV of the shaded VS samples

$$VDR = \sum_{\substack{ND_i=\min VID \\ P_i=0 \\ i=0}}^{\substack{ND_i=\max VID \\ P_i=0.4 \\ i=n}} (f_{VI-SV}(ND_i) \pm P_i \cdot dSVD) \tag{7a}$$

$$VDR_0 = VL \cup seedD_1 \tag{7b}$$

$$VDR_{i-1} \subseteq VDR_i = TURE \tag{7c}$$

where VDR is an iteratively increasing set of shaded VS members extracted within a right buffer [e.g., the set in Fig. 8(e)];

TABLE I
COEFFICIENTS OF NDVI-S AND NDVI–NDSV RELATIONSHIPS

No.	$a1$	$b1$	$c1$	$a21$	$b21$	$c21$	$a22$	$b22$	$c22$
1	0.471419	1.134739	0.084456	–33.8295	–12.8968	–0.5819623	1.121596	1.540996	–0.13063
2	–3.88165	2.401574	0.049572	–11.8596	–5.43128	–0.2884827	–25.4788	5.174452	–0.36628
3	–1.05015	1.574821	0.117327	10.18969	–0.46004	–0.3540134	40.62821	0.225477	–0.33942
4	–0.56923	1.436223	0.168372	–1.4273	–2.138	0.14948219	–1.66695	2.388114	0.24005
5	0.44189	1.103763	0.166552	–23.3732	0.258543	–0.089493	–13.0497	3.573503	0.02049
6	–0.49968	1.112252	0.307598	0	0	–4.90E-05	–0.87734	1.654847	–0.20183
7	–0.65272	1.561395	0.050504	0	0	0.51998597	–0.58949	1.012249	0.466479
8	0.407037	0.333345	0.468264	–33.6438	–13.7331	–0.5052613	1.83227	0.869191	0.32972
9	–0.20832	1.400181	0.1031	–1.34945	–2.00099	0.24416544	–0.68335	1.342842	0.270617

No. is the serial number of image group and the numbers are the same as those in Table III. The later three coefficients are further divided into $a21$, $b21$, $c21$, $a22$, $b22$ and $c22$, of which the first three are for the left and the others for the right relationships.

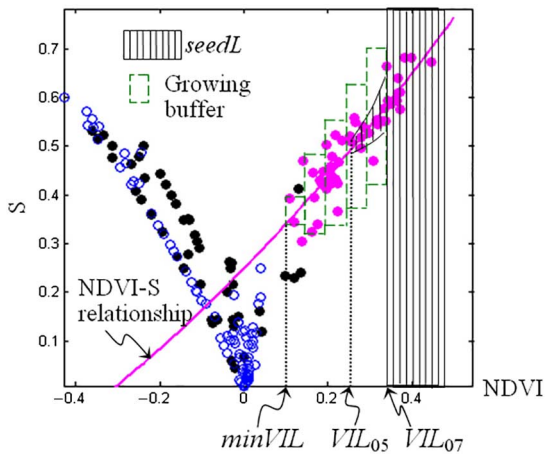

Fig. 4. Case process of extracting bright VS members. The samples were acquired from an image of Group 4 by image sampling. Parameters VIL_{05} and VIL_{07} demarcate the relatively reliable region for bright VS seeds.

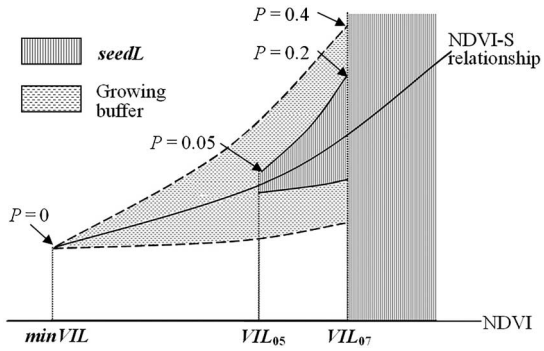

Fig. 5. General process of extracting bright VS members. This figure shows a general form from Fig. 4 for all the cases labeled with the relationship index "1" in Fig. 3. The main part of $seedL$ complies with NDVI $> VIL_{07}$ while the other (the extra) is limited in a narrow region along the NDVI-S curve. These constants in it are the experimental defaults for weight P and the same to those in Figs. 6 and 7.

$n = (maxVID - minVID)/0.01$. The meanings of ND_i and P_i, are the same as those in (5a). Equation (7b) depicts the growing start. The meaning of (7c) is similar to that of (5c)

$$seedD_2 = \{seedD_2 | NDSV > (fL_{VI-SV}(VI) + P_{20} \cdot dSVD) \cap NDSV < (fL_{VI-SV}(VID_{15}) + P_5 \cdot dSVD)\} \quad (8)$$

where $seedD_2$ is another extra seed set for the case of only left buffer available; VID is the NDVI at the intersection of the left and the right NDVI–NDSV curves; $fL_{\mathrm{VI-SV}}$ is a left NDVI–NDSV relationship; *minVID* is the minimum NDVI of the shaded VS samples; $VID_{15} = minVID \cdot 1.5$

$$VDF = \sum_{\substack{ND_i=\min VID \\ P_i=0 \\ i=0}}^{\substack{ND_i=\max VID \\ P_i=0.4 \\ i=n}} (fL_{VI-SV}(ND_i) \pm P_i \cdot dSVD) \quad (9a)$$

$$VDF_0 = VDR \cup seedD_2 \quad (9b)$$

$$VDF_{i-1} \subseteq VDF_i = TURE \quad (9c)$$

where *VDF* is an iteratively increasing set of shaded VS members extracted within a left buffer. Equation (9b) depicts the growing start. The meaning of (9c) is similar to that of (5c).

D. During-Growing Constraints

The main purpose of utilizing seeded region growing is to take advantage of the so-called "during-growing analysis," i.e., taking the previous results as a reference to decide whether a candidate is qualified or not. Some presetting constraints have remarkable influences over the decision. The *expansion rate* is an example of such constraints. Others associated with various applications may also be designed by referencing this.

With **NFSRG**, the growing is limited within a series of intervals (e.g., those green dashed rectangles in Fig. 4). The number of the qualified members during one time iteration can be determined by the domains of an interval. However, pixels on a whole tree crown unlikely fall within a single interval due to their uneven reflectance, i.e., more time iterations being required for all of them to enter. Other background members likely enter in the same time due to the difficulty in distinguishing between the former and the latter. Fortunately, the majority of unexpected entrance can be controlled by some

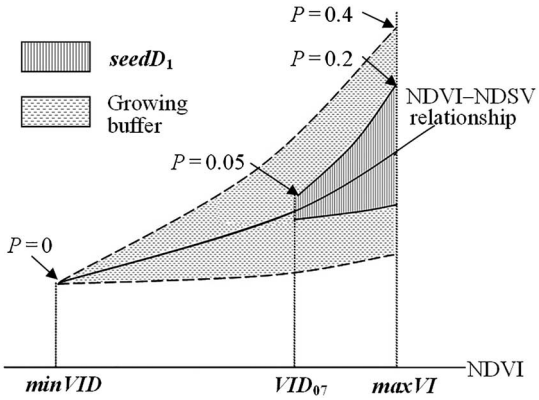

Fig. 6. General process of extracting shaded VS members in right buffer. An extra seed set referring to as $seedD_1$ is necessary for extracting shaded VS members which connect to no bright VS one. The collection is limited in a narrow region along the NDVI–NDSV curve within which shaded VS and background members appear separable to some degree.

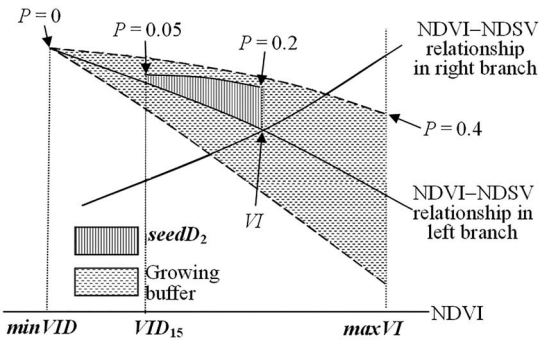

Fig. 7. General process of extracting shaded VS members in left buffer. In the case of only the left buffer available (e.g., the case labeled with 3-2 in Fig. 3), another extra seed set referring to as $seedD_2$ is required. The faint differences in NDSV between shaded VS and shaded background may offer a promise to collect such seeds.

during-growing constrains. For example, if the candidates are the members of "smooth dark background," such as water and shaded roads, the number of the entered during one time iteration likely increases suddenly. Such accidents can be detected by the expansion rate defined as

$$ER_i = (EA_i - EA_{i-1})/EA_{i-1} \tag{10}$$

where ER and EA are the expansion rate and expansion area, respectively; and i is the loop variable.

Fig. 9 shows EA and ER during a case growing process. The best separation between the shaded VS and the smooth dark background can be achieved by adjusting the threshold of ER. Such during-growing constrain, in many cases, contributes to a better separability between different shaded members with different roughness.

E. Clumping VS Members Together

The originally captured VS members by **NFSRG** are often very scattered. It is necessary to clump them together for objects. In general, as long as the VS members geographically connect with each other to some degree they can be clumped together using morphological operations. The process involves: bonding discrete members into objects by morphological closing, then smoothing the object boundaries by morphological opening, and removing the objects smaller than a given size by area filtering. Fig. 10 gives two examples after the clumping.

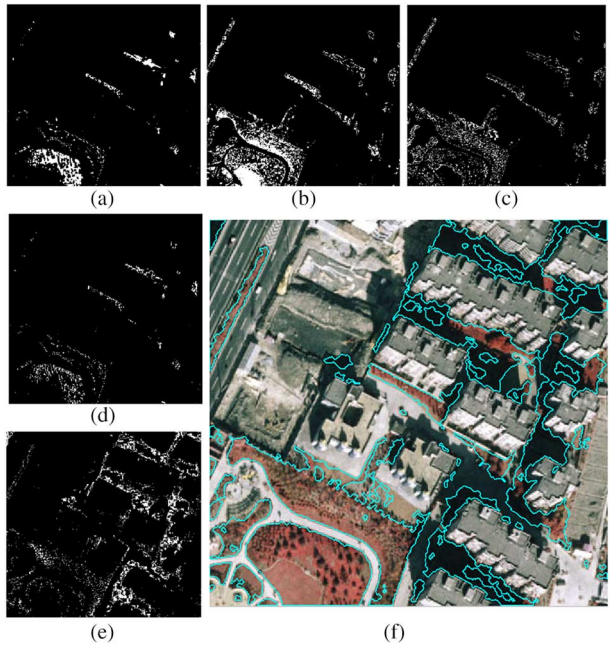

Fig. 8. Results and process of extracting VS using **NFSRG.** (a) The main part of $seedL$ complying with NDVI $> VIL_{07}$ (4); (b) The other part of $seedL$ (4); (c) VL [(5a)–(5c)]; (d) $seedD_1$ (6); (e) VDR [(7a)–(7c)]; (f) The final results of VS (sketched by cyan lines).

F. Refining VS Objects

A refining process sometimes needs to be taken into account to make an extra improvement in the accuracy of VS objects. This is conducted by human–computer interaction and includes to find and remove the false-present objects by the hitting algorithm and to generate the false-absent objects by the point expansion algorithm.

1) Hitting Algorithm: The original VS object set in the form of binary image refers to as VS_0. The so-called hitting algorithm comes from the hitting set theory, but much simpler. The algorithm is applied to detect a cursor-pointed member from VS_0 in a mouse-click event. It is relatively easier to detect the false-present objects because they are parts of the true members of VS_0. With the hitting algorithm, a false-present object is pointed by cursor and then removed from VS_0.

2) Point Expansion Algorithm: However, to determine a false-absent object is not easy. Its absence results from that the feature domains defined by its pixels likely to deviate from those defined by most other pixels in VS_0. Otherwise, it would not miss. Based also on **NFSRG**, a novel algorithm, named *point expansion*, has been explored to capture such pixels. Controlled also by the expansion rate, the desired new objects can be "generated."

A point expansion process includes these steps: (1) taking a cursor-pointed pixel and all the pixels in its size-given

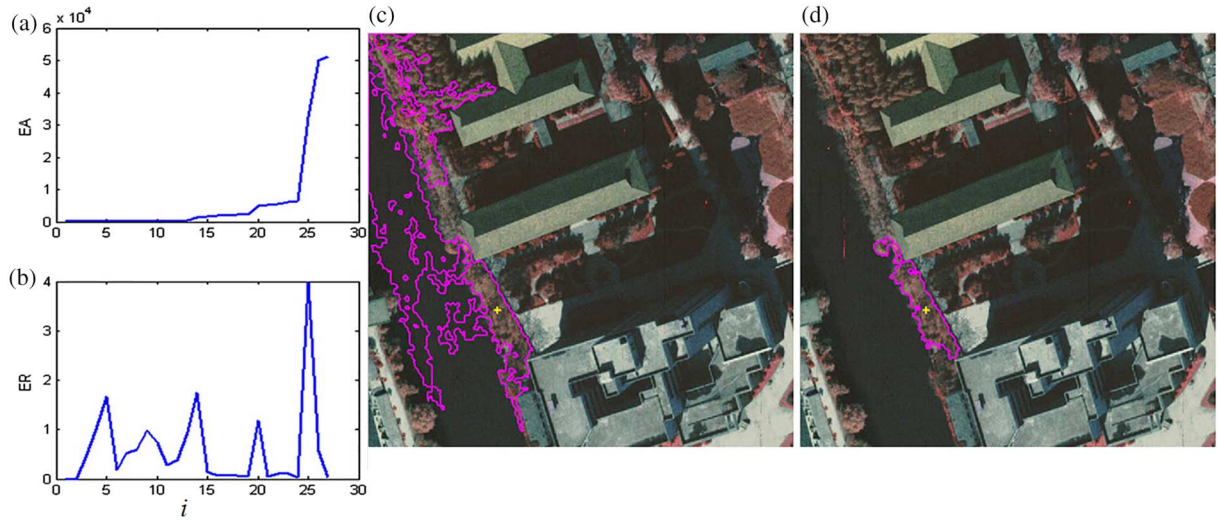

Fig. 9. Extracting VS with and without the expansion rate constrain. (a) and (b) *EA* and *ER* change with i. (c) and (d) without and with the expansion rate. With the continual increase of i, *ER*s are relatively stable at most steps whereas $EA_{\rm i}$ often steadily increase with small amplitudes. However, both $EA_{\rm i}$ and $ER_{\rm i}$ increase suddenly as i equates to 25 due to the entrance of a larger body of water. After this, *ER* is back again to the earlier stability. (c) and (d) are the magenta lines sketch the growing results, whereas the yellow crosses show the sampling points. The process of capturing VS members in (d) is controlled under the condition of $ER < 3$.

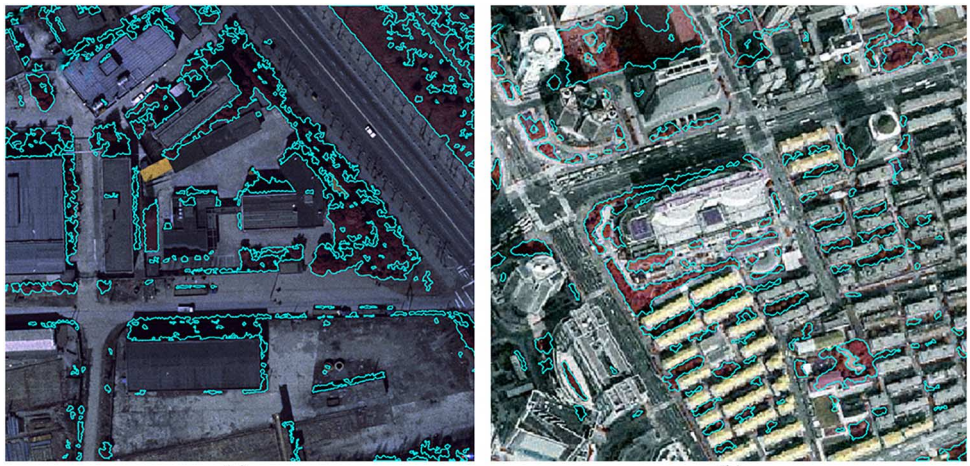

Fig. 10. Two examples of extracting VS objects (sketched by cyan lines) using **NFSRG** from (a) No. 7 and (b) No. 9 groups, respectively.

neighborhood as samples, (2) deriving an NDVI–NDSV relationship from these samples by nonlinear fitting, and (3) taking these samples as the initial seeds to conduct seeded region growing within the buffer defined by the relationship to capture those always under-extracted VS members. The process can be formulated as

$$V = \sum_{\substack{ND_i=\max VI \\ P_i=0.5 \\ i=0}}^{\substack{i=n \\ P_i=0.2 \\ ND_i=\min VI}} f_{VI-SV}(ND_i) \pm P_i \cdot dV \quad (11a)$$

$$V_0 = seed \quad (11b)$$

$$V_{i-1} \subseteq V = TURE \quad (11c)$$

$$ER < T_3 = TURE \quad (11d)$$

where V is an iteratively increasing set of the always under-extracted VS members during one time point expansion; V_0 and V_i are the original and the i-th increasing sets, respectively; $f_{\rm VI-SV}$ is the NDVI–NDSV relationship derived from the samples; *seed* is all the samples; $minVI$, $maxVI$, $minSV$, and $maxSV$ are the end values of NDVI–NDSV region decided by the samples [see Fig. 11(a)]; $dV = maxSV - minSV$.

Fig. 11(a) shows a certain procedure of deriving an NDVI–NDSV relationship. Fig. 11(b) provides an example of the refining process.

In addition, the hitting and the point expansion algorithms can also serve as a tool for the late accuracy assessment because the differences between before and after the refining can be referred to as the errors involved in the automatically extracted VS objects.

IV. Discussion

Next, the principle and characteristics of **NFSRG** will be discussed and the accuracy assessment will be given.

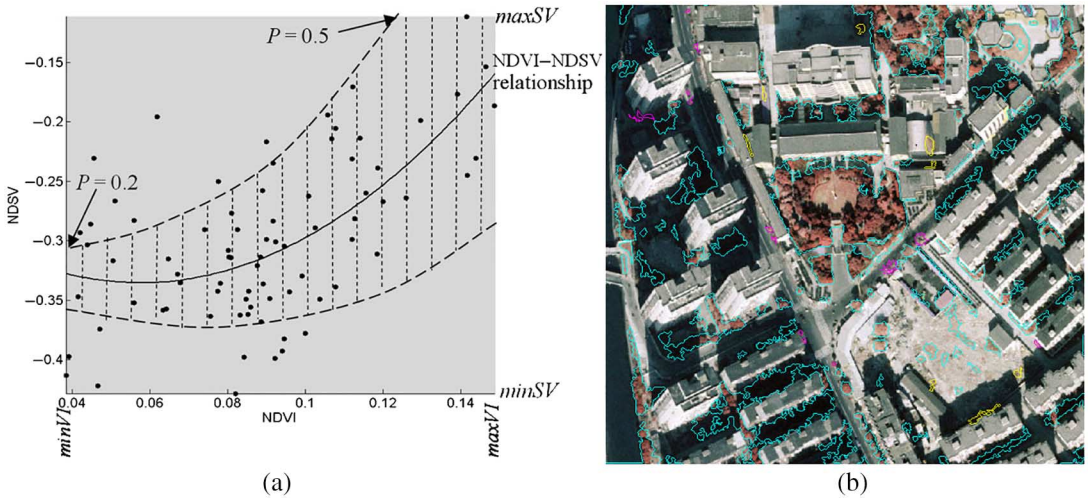

Fig. 11. Refining VS objects. (a) A case fitting process at a mouse click event. Gradually deduced P may continually narrow the buffer (in this case P from 0.5 to 0.2) as the separability between VS and background worsens. Black filled circles represent samples (a cursor-pointed pixel and its neighboring pixels). Gray colored represents the sample-DN-defined region within which the haloiline region is the buffer where point expansion algorithm works. (b) A case refining process. Those generated false-absent objects (magenta line sketched) reveal that the point expansion algorithm works well. Blue line represents original VS objects. Yellow line represents false-present objects removed by the hitting algorithm. Magenta line represents false-present objects generated by the point expansion algorithm.

A. Principle and Characteristics of NFSRG

NFSRG is a kind of serial classification organized as a seeded-region-growing process. The classifications conduct typically in a two-dimensional feature space, NDVI-S or NDVI–NDSV, for urban vegetation mapping. By imposing a constraint on the growing process, some almost inevitable mistakes involved in a typical "single" classification can be probably avoided. This is the actual motivation for exploring NFSRG.

Fig. 12 and Table II provide three examples of comparing two kinds of VS objects obtained by the single classification and **NFSRG**, respectively. The decision tree serves as the classifier model for the former. To ensure the complexity of feature space comparable between the two methods, the input vector for the former consists of three components (NDVI, S, and V) also derived from the DNs involved in both the RGB and the Hue-saturation-value color models.

It can be seen from Fig. 12 and Table II that the result accuracy of extracting VS by **NFSRG** is much better than that by single classification. The distinct difference in their capability to capture VS objects may also be proved by the up to 88.88% of differences in area of the VS objects.

By the way, **NFSRG** is under the fundamental conceptions of classic seeded region growing (SRG) but with following modifications. 1) As using SRG, it is required to keep the connectivity in geographic space and the similarity in feature space between new entered members and original seeds. In the use of NFSRG, however, only the former is required. This enables the characterization of generating shaded VS members from bright VS seeds. 2) To avoid possible over-growing, SRG requires feature homogeneity in an iteratively renewed region whereas NFSRG only requires to limit the expansion rate to a lower level. The latter is favorable for the extraction of heterogeneous objects (e.g., tree crowns). 3) NFSRG has far better tolerance of seed noises than SRG dose. The noises usually do not significantly influence the final accuracy due to the usages of the weighted buffers and the during-growing constrain in a NFSRG process.

B. Accuracy Assessment

The assessment is to evaluate the differences in the layouts of VS objects between the automatically extracted by **NFSRG** and the visually interpreted, therefore, the so-called "error rate" as a measure being employed. The incorrect vegetation objects, including the over-extracted (false-present) and the under-extracted (false-absent) ones, cannot become aware before they being extracted. In most case, such objects/pixels may not be collected in the sampling phase as the reserved checking samples for accuracy assessment since they usually do not have the typical appearances of vegetation covers. Therefore, they are practically collected as accurately as possible by human–computer interaction in the refining phase (see Section III-F) with the hitting and the point expansion algorithms to promise an objective assessment. Nine test images are randomly selected from all nine groups for the assessment.

It can be seen from Table III that the mean error rate, including the over-extracted and the under-extracted, is 3.05%; in other words, more than 96% of VS objects can be automatically and correctly extracted.

V. Conclusion

NFSRG, a series of classified results are organized by a seeded-region-growing process. The iteratively captured VS members are always serve as newly added training samples for the next classification or as newly added seeds for the next growing. By imposing the expansion rate on the phase of capturing the shaded VS, some almost inevitable mistakes involved in typical single classification can be avoided. The accuracy assessments have revealed that more than 96% of VS objects can be automatically and accurately extracted by **NFSRG**.

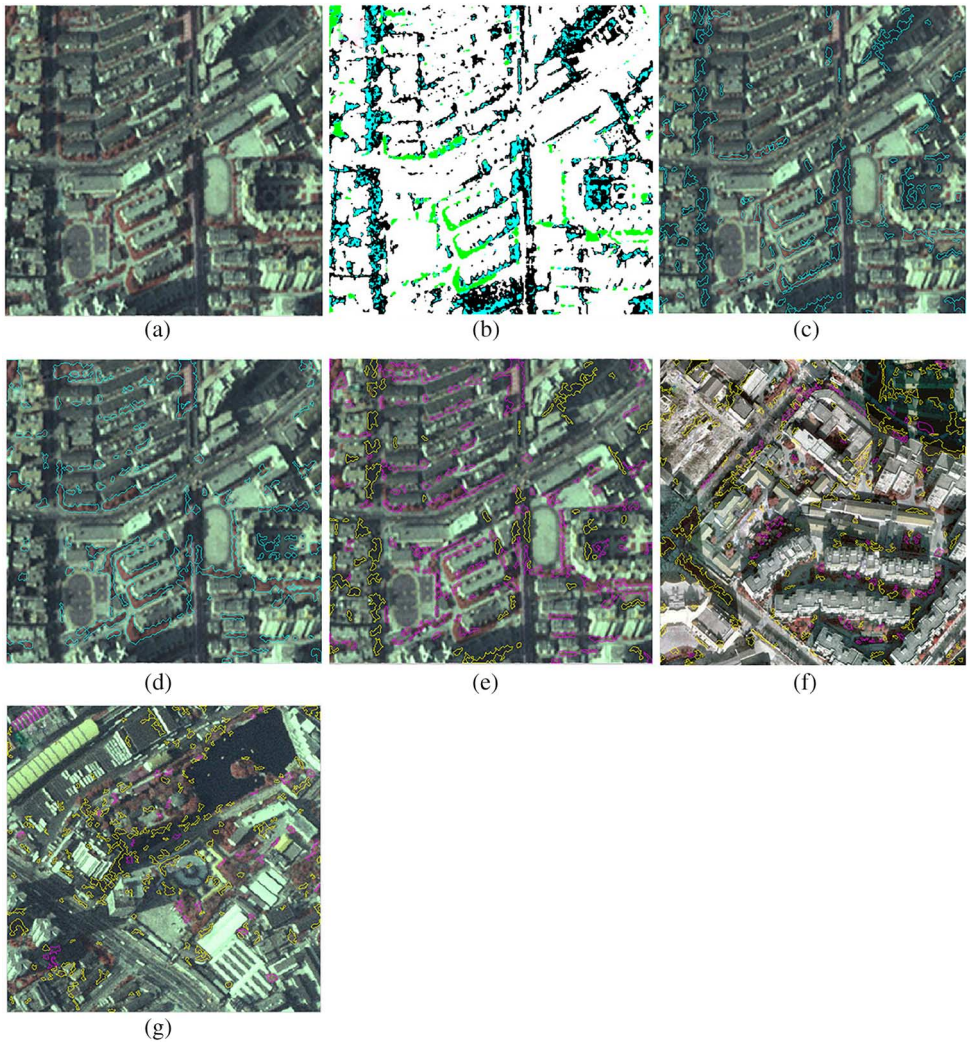

Fig. 12. Three examples of comparing the differences between using the single classification and **NFSRG.** (a)–(e) are for Example 1. Of them, (a) is the original image from No. 2 group, (b) shows the assigned members by the single classification, (c) gives the VS objects clumped from (b) by morphology; (d) provides the VS objects extracted by NFSRG, and (e) gives a comparison between the objects shown in (c) and (d). (f) and (g) are the similar comparisons for other two instances from No. 4 and No. 2 groups. By comparing these sketched objects with their background images in (e)–(g), it can be seen that the conformity between the NFSRG-captured VS-objects and the visually observed ones is much better than that by replacing NFSRG with single classification. The total error rate ($E_{\max}$) in Table II reveals the distinct differences in the results obtained by the two methods. In residential area with high-density buildings, $E_{\max}$ may raise further more [see (e)] due to a more serious shadow effect. The members in (b): green—bright VS; blue—shaded VS; white—bright background; and black—shaded background. The objects in others: blue—VS objects; yellow—over-extracted (should be removed) VS objects; and majenta—under-extracted (should be generated) VS objects.

TABLE II
EVALUATION FOR THE COMPARISONS SHOWN IN FIG. 12

Fig. No.	Year	A_0	A_{over}	E_{over}(%)	A_{under}	E_{under}(%)	E_{max}(%)
12(e)	2003	29 519	11 559	39.16	18429	62.43	101.59
12(f)	2000	40 115	28 685	71.51	6989	17.42	88.93
12(g)	2003	48 704	29 520	60.61	7554	15.51	76.12
Mean				56.94		31.79	88.88

A_0 is the number of automatically extracted VS pixels by **NFSRG**; A_{over} and A_{under}, the numbers of VS pixels over-extracted and under-extracted by the single classification respectively; $E_{\text{over}} = (A_0 - A_{\text{over}}) \cdot 100/A_0$. $E_{\text{under}} = (A_0 - A_{\text{under}}) \cdot 100/A_0$. $E_{\max} = E_{over} + E_{\text{under}}$.

After conducting a literature assessment of current practices in urban vegetation mapping with high-resolution remote sensing data, we found the NFSRG method can provide significant accuracy improvement.

As a hybrid classification method, **NFSRG** also has some limitations. First, it works only in a low-dimensional feature space. More other features, for extracting other members instead of vegetation, can be allowed, but a low dimensionality

TABLE III
ACCURACY ASSESSMENTS FOR ALL THE NINE IMAGE GROUPS

No.	*Year*	*entirePIX*	*underN*	*underRE* (%)	*overN*	*overRE* (%)	*maxRE* (%)
1	1995	139 934	3502	2.5219	4573	3.2932	5.8151
2	2003	72 310	1584	2.1639	692	0.9453	3.1092
3	2005	68 228	428	0.6246	133	0.1941	0.8187
4	2000	33 226	1574	4.5840	463	1.3484	5.9324
5	1993	132 717	1529	1.1470	939	0.7044	1.8514
6	1988	139 466	815	0.5829	470	0.3362	0.9191
7	2001	219 634	293	0.1369	5972	2.7912	2.9282
8	1997	223 169	9063	3.9157	781	0.3374	4.2532
9	2006	32 654	254	0.7799	341	1.0471	1.8270
Mean				1.8285		1.2219	3.0505

Year is the photography time of an image group; *entirePIX*, the number of the automatically extracted VS pixels; *underN*, the numbers of under-extracted VS pixels generated by the point expansion algorithm; *overN*, the numbers of over-extracted VS pixels detected by the hitting algorithm; $underRE = (underN/entireN) \cdot 100\%$; $overRE = (overN/entireN) \cdot 100\%$; $entireN = entirePIX + underN - overN$; $maxRE = underRE + overRE$.

is still required, otherwise too complex relationships may damage its universality. In addition, more buffers can be derived by sectional fitting but the geographic connection between members of relevant classes also needs to be promised (e.g., the connection between the bright VS and the shaded VS members), otherwise, the NFSRG method is not applicable.

ACKNOWLEDGMENT

The authors would like to express our gratitude to Dr. B. Finlayson, University of Melbourne, Australia, and M. Ji, East China Normal University, China for their reviews and helpful suggestions.

REFERENCES

[1] A. B. Morancho, "A hedonic valuation of urban green areas," *Landscape Urban Plann.*, vol. 66, no. 1, pp. 35–41, Dec. 2003.
[2] J. A. Voogt and T. R. Oke, "Complete urban surface temperatures," *J. Appl. Meteorol.*, vol. 36, no. 9, pp. 1117–1132, Sep. 1997.
[3] H. Akbari, "Shade trees reduce building energy use and CO2 emissions from power plants," *Environ. Pollut.*, vol. 116, no. S1, pp. S119–S126, Mar. 2002.
[4] C. Small, "Estimation of urban vegetation abundance by spectral mixture analysis," *Int. J. Remote Sens.*, vol. 22, no. 7, pp. 1305–1334, Jul. 2001.
[5] J. C. Price, "Estimating vegetation amount from visible and near infrared reflectances," *Remote Sens. Environ.*, vol. 41, no. 1, pp. 29–34, Jul. 1992.
[6] A. R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote Sens. Environ.*, vol. 25, no. 3, pp. 295–309, Mar. 1988.
[7] S. Myeong, D. J. Nowak, and M. J. Duggin, "A temporal analysis of urban forest carbon storage using remote sensing," *Remote Sens. Environ.*, vol. 101, no. 2, pp. 277–282, Mar. 2006.
[8] A. Buyantuyev and J. Wu, "Urbanization alters spatiotemporal patterns of ecosystem primary production: A case study of the Phoenix metropolitan region, USA," *J. Arid Environ.*, vol. 73, no. 4–5, pp. 512–520, Apr./May 2009.
[9] N. Goodwin, N. C. Coops, and C. Stone, "Assessing plantation canopy condition from airborne imagery using spectral mixture analysis and fractional abundances," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 7, no. 1, pp. 11–28, Jan. 2005.
[10] M. L. Cadenasso, S. T. A. Pickett, and K. Schwarz, "Spatial heterogeneity in urban ecosystems: Reconceptualizing land cover and a framework for classification," *Front. Ecol. Environ.*, vol. 5, no. 2, pp. 80–88, Mar. 2007.
[11] N. Thomas, C. Hendrix, and R. G. Congalton, "A comparison of urban mapping methods using high-resolution digital imagery," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 9, pp. 963–972, Sep. 2003.
[12] J. P. Ardila, W. Bijker, V. A. Tolpekin, and A. Stein, "Context-sensitive extraction of tree crown objects in urban areas using VHR satellite images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 15, pp. 57–69, Apr. 2012.
[13] M. Bochow, K. Segl, and H. Kaufman, "Comparison of multi- and hyperspectral remote sensing data for use in comprehensive urban biotope mapping," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Boston, MA, USA, 2008, pp. V5–V8.
[14] H. G. Sohn and K. H. Yun, "Shadow-effect correction in aerial color imagery," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 5, pp. 611–618, May 2008.
[15] J. H. Zhou, Y. F. Zhou, X. H. Guo, and Z. Ren, "Methods of extracting distribution information of plants at urban darken areas and repairing their brightness," *J. East China Normal Univ. (Natl. Sci. Ed.)*, vol. 2011, no. 6, pp. 1–9, Nov. 2011.
[16] R. L. Pu, P. Gong, R. Michishita, and T. Sasagawa, "Spectral mixture analysis for mapping abundance of urban surface components from the Terra/ASTER data," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 939–954, Jun. 2007.
[17] T. R. Tooke, N. C. Coops, N. R. Goodwin, and J. A. Voogt, "Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications," *Remote Sens. Environ.*, vol. 113, no. 2, pp. 398–407, Feb. 2009.
[18] S. van der Linden and P. Hostert, "The influence of urban structures on impervious surface maps from airborne hyperspectral data," *Remote Sens. Environ.*, vol. 113, no. 11, pp. 2298–2305, Nov. 2009.
[19] A. Prati, I. Mikic, M. M. Trivedi, and R. Cucchiara, "Detecting moving shadows: Algorithms and evaluations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 7, pp. 918–923, Jul. 2003.
[20] V. J. D. Tsai, "A comparative study on shadow compensation of color aerial images in invariant color models," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1661–1671, Jun. 2006.
[21] X. J. Gao, Y. C. Wan, S. Y. Zheng, and J. Li, "Automatic shadow detection and compensation of aerial remote sensing images," *Geomat. Inf. Sci. Wuhan Univ.*, vol. 37, no. 11, pp. 1299–1302, Nov. 2012.
[22] J. H. Zhou, Y. F. Zhou, and W. S. Mu, "Mathematic descriptors for identifying plant species: A case study on urban landscape vegetation," *J. Remote Sens.*, vol. 15, no. 3, pp. 524–538, Aug. 2011.

Jianhua, photograph and biography not available at the time of publication.

Yan Huang, photograph and biography not available at the time of publication.

Bailang Yu, (M'09) photograph and biography not available at the time of publication.