Partial differential equation-based object extraction from remote sensing imagery

LI Zhong-Bin, SHI Wen-Zhong

(Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China)

Abstract: Object extraction is an essential task in remote sensing and geographical sciences. Previous studies mainly focused on the accuracy of object extraction method while little attention has been paid to improving their computational efficiency. For this reason, a partial differential equation (PDE)-based framework for semi-automated extraction of multiple types of objects from remote sensing imagery was proposed. The mathematical relationships among the traditional PDE-based methods, i.e., level set method (LSM), nonlinear diffusion (NLD), and active contour (AC) were explored. It was found that both edge- and region-based PDEs are equally important for object extraction and they are generalized into a unified framework based on the derived relationships. For computational efficiency, the widely used curvature-based regularizing term is replaced by a scale space filtering. The effectiveness and efficiency of the proposed methods were corroborated by a range of promising experiments.

Key words: active contour, building extraction, level set method, object extraction, partial differential equation, nonlinear diffusion, road extraction

PACS: 07.05. Pj

Introduction

Man-made object extraction has always been an intensive research topic in the field of applied Earth observation\(^1\). It is one of the most commonly used data acquisition methods in remote sensing and geographical sciences\(^2\). For instance, it is beneficial to the timely update of GIS database\(^3\) and the decision-making of the urban planning. Generally, man-made objects in high spatial resolution remote sensing images appear as homogeneous regions with similar spectral signatures. However, automatic detection of man-made objects is still a challenging task. That is mainly because the complex...
scenes (e.g., bare land, cropland, vegetation, and shorelines) often share common features (e.g., geometric shape, radiometric intensity, or texture) with the desired objects\textsuperscript{45}. Thus, there is a tremendous need to develop more operational object extraction methods\textsuperscript{10}. This study primarily focuses on man-made object (i.e., roads and buildings) extraction from optical images using partial differential equation (PDE)-based methods. The aim of this paper is to propose practical semi-automated methods that can reduce user’s load considerably.

PDE-based methods\textsuperscript{7-12} have been widely used for object extraction from remote sensing imagery. Previous methods are mainly adapted from the region-based Chan-Vese (CV) model\textsuperscript{13}. Cao and Yang\textsuperscript{7} incorporate the fractal error metric and textural information into CV model for man-made object extraction from aerial images. Similar idea can be found in Ref. \textsuperscript{8}. In Ref. \textsuperscript{9}, a region-based level set method (LSM) adapted from CV model was used to extract roads, buildings, and airport runways from satellite images. Xu, et al.\textsuperscript{10} integrated intensity and texture information into CV model to extract salient objects from satellite images. Kim and Shan\textsuperscript{11} extracted building roofs from point cloud data using the multiphase CV model. Ardila, et al.\textsuperscript{12} used the region-based active contour to extract urban trees for change analysis. By contrast, edge-based methods have received much less attention for object extraction over the past decades. Laptev, et al.\textsuperscript{14} used scale space theory and an edge-based parametric snake for rural road network extraction from aerial images. More recently, an edge-based level set evolution has been used to extract building roofs from aerial images\textsuperscript{16}, in which the initial level curves are generated using a corner detector. However, it has difficulties in handling man-made objects that do not have corners.

\begin{equation}
\phi_i = F(\kappa) \mid \nabla \phi_i \end{equation}

where $\phi_i$ is the level set function. $\phi_i$ is the partial derivative of $\phi$ with respect to the temporal variable $t$. $\kappa = \frac{\text{div} (\nabla u / |\nabla u|)}{\nabla |\nabla u|}$ is the mean curvature of the zero-level set and $F(\kappa)$ is a function with respect to curvature $\kappa$. $F(\kappa)$ serves as not only the driving force in LSM (Eq. (1)), but also the regularization term to keep the moving zero-level set smooth.

\subsection{Nonlinear diffusion}

NLD proposed in Ref. \textsuperscript{18} is given as follows:

\begin{equation}
\psi_t = g\left( |G_\sigma \ast \nabla u| \right) \nabla u \left| \nabla u \right| \left( \frac{\nabla u}{\left| \nabla u \right|} \right),
\end{equation}

where $\psi_t$ is the partial derivative of $u$ with respect to the time variable $t$; $u$ is the image to be processed; $\nabla u = (u_x, u_y)$ is the gradient; $\left| \nabla u \right| = \sqrt{u_x^2 + u_y^2}$ is the magnitude of $\nabla u$; $G_\sigma$ is the Gaussian kernel with standard deviation $\sigma$; $\ast$ is the convolution operator; $\text{div}$ is the divergence operator; $\left( \frac{\nabla u}{\left| \nabla u \right|} \right)$ is the mean curvature as before; and $g(\cdot)$ is an edge detector with respect to the gradient such as $g\left( \left| \nabla u \right| \right) = 1/(1 + \left| \nabla u \right|^2)$. In low-level vision, NLD (Eq. (2)) was widely used for the noise removal. The term $\nabla u \left| \nabla u \right|$ diffuses along the boundaries but not across them under the control of the stopping term $g\left( |G_\sigma \ast \nabla u| \right)$\textsuperscript{21}.

\subsection{Edge-based active contour}

The seminal parametric snakes was first proposed in
Ref. [22]. However, it has difficulty in handling topological changes naturally. In contrast, geometrical models perform much better such as the edge-based active contour (EAC) proposed in Ref. [19]:

$$\phi_i = g(\nabla \phi) \left( \frac{\partial h}{\partial \phi} \frac{\nabla \phi}{|\nabla \phi|} + \nu \right),$$  \hspace{1cm} (3)

in which $\phi$ is the level set function as before. $g$ is the same as the one in NLD. Compared with NLD, EAC (Eq. (3)) has one more positive real constant $\nu$, which can keep the term $\partial h(\nabla \phi)/\partial \phi \cdot \nabla \phi$ always positive. It is clear that EAC (Eq. (3)) is essentially equivalent to NLD (Eq. (2)). The differences between them are twofold; 1) the former is originally developed for object extraction, whereas the latter is proposed for noise removal, and 2) the former often chooses signed distance function as level set function; whereas the latter often uses the original image as level set function.

1.1.4 Region-based active contour

In practical applications, EAC often suffers from limitations, e.g., 1) it is sensitive to noises and 2) it cannot extract the interior edges. To address these issues, region-based active contour (RAC) is proposed [10]:

$$\phi_i = \delta_{\nu} (\phi) \left[ \mu \frac{\partial h}{\partial \phi} \frac{\nabla \phi}{|\nabla \phi|} - \lambda_1 (I_0 - c_1)^2 + \lambda_2 (I_0 - c_2)^2 \right],$$  \hspace{1cm} (4)

in which $\delta_{\nu} (\cdot)$ is the Dirac function; $\mu \geq 0, \nu \geq 0$, and $\lambda_1, \lambda_2 > 0$ are free parameters; $I_0$ is the original image; and $c_1, c_2$ are intensity means inside and outside the zero-level set, respectively. Although RAC (Eq. (4)) is more advantageous than EAC, it also suffers from limitations. For instance, there are too many parameters in Eq. (4) that need to be tuned repeatedly before it can be employed in practical applications. However, it is often labor-intensive and time-consuming to obtain the optimal parameter values.

1.2 The proposed method

Based on the above analysis, some unequally mathematical relationships among LSM [17], NLD [18], EAC [10], and RAC [13] can be derived as follows:

1) NLD and EAC are theoretically equivalent when the level set function in Eq. (3) equals the image in Eq. (2), i.e., $\phi = u$.

2) PDE-based methods (i.e., NLD, EAC, and RAC) can be viewed as further developments of LSM, though they are derived from different mathematical models. In contrast to LSM, NLD and EAC take advantage of the edge detector $g(\cdot)$ as stopping term, whereas RAC employs intensity means as the stopping term.

3) As can be seen, when $g([G_x \ast |\nabla u|) = 1$, Eq. (2) becomes Eq. (1). When $g = 1$ and $\nu = 0$, Eq. (3) becomes Eq. (1). In addition, Eq. (4) becomes Eq. (1) when $\delta_{\nu} (\phi) \approx \nabla \phi$, $\mu = 1$, $\lambda_1 = \lambda_2 = \nu = 0$.

4) All the aforementioned PDE-based methods can be implemented using finite difference scheme.

The unexpected relationships mentioned above make it easier to understand the essences of PDE-based methods. It is worth mentioning that the Dirac function $\delta_{\nu}$ in Eq. (4) can be replaced by $|\nabla \phi|^{1/2}$. In addition, according to the theory in Ref. [23], it is mathematically sound to replace the curvature term $div \left( \frac{\nabla \phi}{|\nabla \phi|} \right)$ commonly used in traditional PDE-based methods by the scale space filtering. It has been proved that the best filtering for constructing scale space is the Gaussian kernel $G_x^{[24]}$. Based on all these facts, it can be found that both edge- and region-based PDE-based approaches are equally important in practical applications, and thus, they are generalized into the following unified framework:

$$\phi_{i,j}^n = G_x \ast \text{sign}(\phi_{i,j}^n + \Delta t \ast F |\nabla \phi_{i,j}^n|),$$  \hspace{1cm} (5)

where $(i,j)$ denotes spatial position; $n$ means iteration number; $\text{sign}(\cdot)$ is the sign function; $\Delta t$ is the time step; and $F$ is the data term that is used to guide active contours toward desired object boundaries. In Eq. (5), $\phi_{i,j}^n$ is defined as follows:

$$\phi_{i,j}^n = \begin{cases} c_0, &\text{if } (i,j) \in R_0 \\ \phi_{i,j}^n - c_0, &\text{otherwise} \end{cases},$$  \hspace{1cm} (6)

in which $c_0 > 0$ is a constant. It is fixed at 1 in this study. $R_0$ is an arbitrarily given region (i.e., the initial zero-level set) in the image domain. Essentially, PDE (Eq. (5)) can be regarded as a further development of the seminal LSM, which is the most basic component of the popular methods NLD, EAC, and RAC. $F$ in Eq. (1) is a function with regard to the curvature of zero-level set, whereas $F$ in Eq. (5) is a generic data term that can be specifically written as follows:

$$F_e = g(G_x \ast |\nabla I|),$$  \hspace{1cm} (7)

$$F_b = -\lambda_1 (I - c_1)^2 + \lambda_2 (I - c_2)^2,$$  \hspace{1cm} (8)

where $F_e$ and $F_b$ are edge- and region-based data terms, respectively. $I$ is the image to be processed. In fact, data terms defined by Eqs. (7) and (8) are derived from EAC and RAC, respectively. Despite this, it leads to a considerable improvement of the original LSM. That is corroborated by the following experiments.

1.3 The implementation of the proposed method

PDE is an iterative process that exploits finite difference scheme to approximate its solution. Before its implementation, initial zero-level sets need to be given in the desired object regions interactively. In this paper, for numerical stability, scale space technique is used to regularize the evolving zero-level sets. The basic idea of this technique is to convolve the level set function with a scale space filtering from coarser to finer scales [25]. With the increase of the scale parameter $\sigma$, the level set function becomes coarser and smoother, and noises disappear in coarser scales. In this respect, it is very useful for edge-based PDE-based methods such as NLD and EAC as they are often sensitive to noise.

However, different from PDE methods NLD and EAC, in which the scale space filtering is mainly used for noise removal, PDE Eq. (5) mainly utilizes scale space filtering to convolve the iterated level set function, thus 1) removing small spurious objects and 2) making the level set function regular during evolution.

Overall, the advantages of the proposed PDE-based method over other existing methods in the following three aspects:

1) LSM does not take into account data term. This means that it cannot be used for object extraction directly.

In contrast, the proposed PDE Eq. (5) takes advantage of the image features (e.g., the edge detector in Eq. (7) and the intensity mean in Eq. (8) to drive the
zero-level set toward desired object boundaries.

2) To maintain the moving zero-level set smooth, the traditional approaches represented by Eqs. (1-4) exploit the curvature of the zero-level set; the proposed method (Eq. (5)), in contrast, uses the scale space filtering to convolve the zero-level set directly.

3) For numerical stability, traditionally methods can only use a very small time step, and thus, they often consume too much CPU time. By contrast, the proposed method can use a relatively larger time step due to the removal of the curvature. In this respect, it is computationally much more efficient.

In the next section, a range of experiments is demonstrated to verify the effectiveness of the proposed method.

2 Experiments

2.1 Dataset and experiment setup

In this experiment, six remote sensing images acquired from different sensors were used to test the proposed methods. As shown in Figs. 2 and 3, they are named R_1, R_2, B_1, B_2, S_1, and S_2 (R, B, and S denote road, building, and stadium), respectively. They all contain three bands (i.e., R, G, and B), except for R_2 that just has the panchromatic band. As shown in the figures, road networks in R_1 and R_2 have opposite intensities. Building in B_1 is very noisy. Stadiums in S_1 and S_2 have different illuminations and there are intensity variations inside the stadium in R_2. In addition, as all the desired objects are characterized by different shapes, they can be used to validate the capability of the proposed method to extract objects without the need of the geometrical a priori information. The detailed information of these images is given in Table 1. To corroborate the advantages of the proposed method, it was compared with EAC and RAC in all the experiments. All the experimental results were compared with ground truths qualitatively and quantitatively. Ground truths are generated by manual delineation. For quantitative evaluation, three indices are used [16], i.e., Completeness = TP_u/TP_u + FN_u, Completeness = TP_u/TP_u + FN_u, and Completeness = TP_u/TP_u + FN_u, in which TP_u is the total pixel number of the extracted object that is matched with the ground truth, TP_u is the total pixel number of the ground truth, TP_u is the total pixel number of the extracted object, and TP_u is the total pixel number of the ground truth that is unmatched with the extracted object. The proposed algorithms were run under MATLAB R2013a 64 b in Windows 7 OS with Intel(R) Core(TM) i7-3770 CPU @ 3.40 GHz, 16 GB RAM. The source code will be publicly available at http://www.lsgi.polyu.edu.hk/academic_staff/John_Shi/index.htm.

Table 1 Dataset description (SR means spatial resolution)

<table>
<thead>
<tr>
<th>Images</th>
<th>Sensor</th>
<th>Size (pixel x pixel)</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_1 (0.1)</td>
<td>Aerial</td>
<td>1722 x 2321</td>
<td>Indiana, USA</td>
</tr>
<tr>
<td>R_2 (0.3)</td>
<td>Worldview-1</td>
<td>179 x 239</td>
<td>N/A</td>
</tr>
<tr>
<td>B_1 (0.5)</td>
<td>Pleiades-1</td>
<td>725 x 285</td>
<td>Washington, USA</td>
</tr>
<tr>
<td>B_2 (0.5)</td>
<td>Pleiades-1</td>
<td>329 x 334</td>
<td>Melbourne, Australia</td>
</tr>
<tr>
<td>S_1 (0.5)</td>
<td>Pleiades-1</td>
<td>532 x 519</td>
<td>Fortaleza, Brazil</td>
</tr>
<tr>
<td>S_2 (0.5)</td>
<td>Pleiades-1</td>
<td>463 x 529</td>
<td>Salvador, Brazil</td>
</tr>
</tbody>
</table>

Table 2 Parameter values used in each test methods

<table>
<thead>
<tr>
<th>Images</th>
<th>EAC (3)</th>
<th>RAC (4)</th>
<th>PDE (7)</th>
<th>PDE (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_1</td>
<td>a = 0.2, r = 3</td>
<td>µ = 0.5, σ = 0.5, λ = 500, a_1 = 4, a_2 = -4, a_3 = 1</td>
<td>σ = 1, σ = 2</td>
<td>σ = 1, σ = 2</td>
</tr>
<tr>
<td>R_2</td>
<td>a = 0.2, r = 3</td>
<td>µ = 0.5, σ = 0.5, λ = 500, a_1 = 4, a_2 = -4, a_3 = 1</td>
<td>σ = 1, σ = 2</td>
<td>σ = 1, σ = 2</td>
</tr>
<tr>
<td>B_1</td>
<td>a = 0.2, r = 3</td>
<td>µ = 0.5, σ = 0.5, λ = 500, a_1 = 4, a_2 = -4, a_3 = 1</td>
<td>σ = 1, σ = 2</td>
<td>σ = 1, σ = 2</td>
</tr>
<tr>
<td>B_2</td>
<td>a = 0.2, r = 3</td>
<td>µ = 0.5, σ = 0.5, λ = 500, a_1 = 4, a_2 = -4, a_3 = 1</td>
<td>σ = 1, σ = 2</td>
<td>σ = 1, σ = 2</td>
</tr>
<tr>
<td>S_1</td>
<td>a = 0.2, r = 3</td>
<td>µ = 0.5, σ = 0.5, λ = 500, a_1 = 4, a_2 = -4, a_3 = 1</td>
<td>σ = 1, σ = 2</td>
<td>σ = 1, σ = 2</td>
</tr>
<tr>
<td>S_2</td>
<td>a = 0.2, r = 3</td>
<td>µ = 0.5, σ = 0.5, λ = 500, a_1 = 4, a_2 = -4, a_3 = 1</td>
<td>σ = 1, σ = 2</td>
<td>σ = 1, σ = 2</td>
</tr>
</tbody>
</table>

2.2 Parameter tuning

All the parameters used in the test methods are listed in Table 2. The parameter values in EAC and RAC are determined via trial and error. The free parameters A_1 and A_2 in RAC have significant impacts on final results. Sometimes they are not equal to each other as such in the experiment of R_1 (see Fig. 1). Throughout the experiments, the template size of the Gaussian kernel is fixed as 9 x 9. Generally, the use of a relatively larger time step Δt can expedite the iteration of PDE-based methods. However, it may lead to unstable numerical results. To obtain stable results using traditional methods such as NLD, EAC, and RAC, the time step should be sufficiently small due to the curvature term \( \text{div}(\nabla u/\sqrt{\nabla u}^2) \) or \( \text{div}(\nabla \phi/\nabla \phi) \). Thus, a relatively small time step is often utilized. By contrast, due to the removal of the curvature term in the proposed PDE, the time step can be chosen as up to 40 for both the edge- and region-based PDEs (Eqs. (7) and (8)). Nevertheless, it is fixed at 18 for stable results in all the experiments. In addition, it is worth noting that edge-based model defined with Eq. (7) use the Gaussian kernel twice: once for the noise removal of the original image and once for the regularization of the zero-level set. Thus, there are two scale parameters σ_1 and σ_2 need to be tuned for Eq. (7). However, experiments show that the optimal pa-
2.3 Qualitative evaluation

As presented in Figs. 2-3, the proposed PDEs are capable of extracting multiple types of man-made objects from optical images. All the initial zero-level sets in the experiments are manually provided inside the object regions.

Figure 2 shows how the proposed methods can extract bright and dark roads from images $R_1$ and $R_2$, respectively. Despite the over-detection in some areas, they can extract the road networks completely. Different from the model in Ref. [15], which extracts the neighboring sideways (see Fig. 1), the proposed methods can extract the desired road networks accurately. As can be seen, over-detection occurs in the results of EAC and RAC in $R_1$. In addition, EAC cannot obtain the complete dark road network in $R_2$.

Figure 3 shows how they can extract the buildings and stadiums without the need of a priori geometric information. As shown in the figures, the proposed methods are capable of extracting the rectangular building with heavy noises from image $B_1$, and noises are automatically filtered out by the scale-space filtering, as presented in Fig. 3(a). The specific scale parameters $\sigma_1$ and $\sigma_2$ are given in Table 2. Also, they successfully extract the building immersed in the complex backgrounds from image $B_2$, examining Fig. 3(b). In addition, they are able to extract the elliptic stadiums from images $S_1$ and $S_2$ accurately, as presented in Fig. 3(c) and (d). The stadium in $S_1$ is bright and homogeneous, whereas the one in $S_2$ is relatively dark and heterogeneous. Despite this, the proposed PDEs can extract them accurately. By contrast, EAC performs relatively poor. It cannot extract the building in $B_1$ completely, and it only extracts part of the stadium in $S_2$ due to the heavy intensity variations.

![Fig. 2 Results for road extraction with different methods. (a) Aerial image $R_1$ with the bright road. (b) Satellite image $R_2$ with the dark road. From left to right, original images with green initial zero-level sets, results of EAC(3), RAC(4), edge-based and region-based PDE, and the ground truths, respectively.

![Fig. 3 Results of the test methods for building extraction from satellite images. (a) Building $B_1$ with heavy noises. (b) Building $B_2$ with complicated backgrounds. (c) Bright homogeneous stadium $S_1$. (d) Dark heterogeneous stadium $S_2$. From left to right, original images with green initial zero-level sets, results of EAC(3), RAC(4), edge-based and region-based PDE, and the ground truths, respectively.]

![Table 4 Running times of the test methods (unit: second)]

<table>
<thead>
<tr>
<th>Images</th>
<th>$R_1$</th>
<th>$R_2$</th>
<th>$B_1$</th>
<th>$B_2$</th>
<th>$S_1$</th>
<th>$S_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAC(3)</td>
<td>7991.6</td>
<td>98.8</td>
<td>80.0</td>
<td>23.0</td>
<td>375.5</td>
<td>517.2</td>
</tr>
<tr>
<td>RAC(4)</td>
<td>3596.5</td>
<td>25.7</td>
<td>10.3</td>
<td>4.1</td>
<td>52.4</td>
<td>46.5</td>
</tr>
<tr>
<td>PDE(7)</td>
<td>230.5</td>
<td>0.7</td>
<td>0.7</td>
<td>0.3</td>
<td>2.2</td>
<td>3.5</td>
</tr>
<tr>
<td>PDE(8)</td>
<td>293.1</td>
<td>1.0</td>
<td>0.6</td>
<td>0.3</td>
<td>3.1</td>
<td>2.7</td>
</tr>
</tbody>
</table>

2.4 Quantitative evaluation

The quantitative evaluation results of the proposed PDEs for man-made object extraction are given in Table III. Overall, they have comparable quality to the traditional methods (i.e., EAC and RAC) in all the experiments. In comparison with the edge-based methods (i.e., EAC and PDE (7)), region-based methods (i.e., RAC and PDE (8)) have better completeness. That is mainly because the original images need to be smoothed by scale-space filtering in edge-based methods, and thus, edges become blurred and their locations are shifted in some sense. However, this denoising step is unnecessary in region-based methods. Sometimes the correctness of region-based methods is not as great as edge-based methods, as indicated by the bold text in Table 3. That is because region-based methods often detect neighboring undesired objects that are spectrally similar to the desired ones. From the perspective of quality, the performance of the proposed edge-based PDE method is not as great as other methods in $B_2$ and $S_1$. However, it outperforms EAC in other experiments. In addition, the proposed region-based PDE method clearly excels RAC.
in experiments of R_1. However, they have similar performance in other experiments. In terms of running time, the proposed methods outperform EAC and RAC significantly, as shown in Table 4. In all the experiments, the proposed edge-based method is 35, 141, 114, 77, 171, and 148 times faster than EAC; the proposed region-based method is 12, 26, 17, 14, 17, and 17 times faster than RAC. Thus, the proposed methods are computationally much more efficient than the traditional methods.

3 Discussion and conclusion

This paper has presented a PDE-based unified framework for object extraction from remote sensing imagery. Essentially, the framework is derived from the unexpected relationships among the traditional level set method (LSM), nonlinear diffusion (NLD), and active contour (AC). The curvature term widely used in traditional PDE-based methods is replaced by the scale space filtering in the new framework, which makes it possible to use a large time step in the numerical scheme. Meanwhile, this scale space filtering can keep the moving zero-level sets regular while filtering out noises. The proposed framework is finally implemented by using finite difference scheme. Experiments indicate that it is computationally much more efficient than previous methods while obtaining comparable performance.

In future research, it would be necessary to automate the proposed PDE-based methods. Also, it would be interesting to improve traditional PDE-based methods such as NLD and LSM by taking advantage of texture and spatial context information. Finally, the attention can be paid to extracting natural objects such as tree crown, cropland, and forest from remote sensing imagery.

Acknowledgment

The authors would like to thank Prof. Gem Unsalan, Yeditepe University for providing the GeoEye-1 panchromatic image R_2. Also, they would like to thank the handling editor and the two anonymous reviewers for their valuable and detailed comments, which improved the quality of the manuscript substantially.

References