

Morphological path filtering at the region scale for efficient and robust road network extraction from satellite imagery[☆]

Luc Courtrai, Sébastien Lefèvre*

Univ. Bretagne-Sud, UMR 6074 IRISA, F-56000 Vannes, France



ARTICLE INFO

Article history:

Available online 24 May 2016

Keywords:

Mathematical morphology
Path filtering
Hit-or-miss transform
Road extraction
Remote sensing

ABSTRACT

Roads are important elements in geographic information systems and remote sensing applications. Their automatic extraction is challenging when only aerial or satellite images are used. Recently, some promising attempts have been made with (incomplete) path opening/closing, morphological filters able to deal with curvilinear structures. We propose here to apply morphological path filters not on pixels directly but rather on regions representing road segments, in order to improve both efficiency and robustness. The overall process is organized in two steps: first we map road segments by rectangular areas made of similar content, before we connect such segments into paths of segments or polylines using region-based path filtering. Robustness to occlusion is ensured through the adaptation of the incomplete path filtering strategy to the region scale, while better discrimination between road segments and other objects is achieved through an hit-or-miss transform that exploits background knowledge. Experiments conducted on several satellite images illustrate the interest of the proposed approach, and shows it outperforms pixelwise detection.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Road networks are important elements for urban planning or environmental monitoring. Despite being often modeled as geographic information systems (or GIS), their extraction from remote sensing data eases GIS updating (on a regular basis or after a disaster). It is even mandatory when no GIS is available. Manual extraction of the road network can be achieved on small areas with photo interpretation, but it is not possible anymore when satellite or aerial images of larger extent are considered. Automatic extraction of road networks from aerial or satellite images has thus been addressed since 40 years [1] and led to numerous works, see [2] for a review. Various frameworks have been used in this intent, e.g., Markov random fields [3,4], neural networks [5] and deep learning [6], mathematical morphology [7–10], graph modeling [11,12], spatially-adaptive classification [13], multiscale analysis [14], etc.

Among them, mathematical morphology has recently led to promising results in providing fast but accurate solutions, with the work from Valero et al. [15] that was relying on path opening and closing. These morphological filters aim to extract (or highlight)

curvilinear structures [16]. However, as most of the automatic techniques, such an approach is relying on pixelwise analysis and filtering, thus presenting two drawbacks. On the one side, the volume of information to be processed (pixels) prevents from efficient (fast) extraction and does not allow processing large remotely-sensed images. On the other side, it is not adapted to images with a very high (spatial) resolution where roads are described by a large set of pixels (e.g., a road of 7 m width is mapped by image segments of 10 to 100 pixels wide for spatial resolutions of 0.7 m to 0.07 m per pixel, corresponding respectively to VHR satellite and aerial images).

Inspired from the promises of path filtering operators, we propose here a novel morphological technique for road network extraction from remote sensing. Conversely to Valero et al. [15], we consider the region level instead of the (standard) pixel one when applying morphological filters. We thus propose a 2-step approach which first extracts rectangular regions corresponding to possible road segments, before connecting these regions through region-based morphological operators. Our contributions also consist in: adapting the (incomplete) path opening paradigm from Talbot and Appleton [16] to a region-based representation; and taking into account background knowledge through an HMT (hit-or-miss transform) procedure that has led to satisfying results in remote sensing [17]. Experiments conducted on several satellite images show our method outperforms the state-of-the-art.

* This paper has been recommended for acceptance by Chennai Guest Editor.

[☆] Corresponding author. Tel.: +33 2 97 01 72 66; fax: +33 2 97 01 72 79.

E-mail addresses: luc.courtrai@irisa.fr (L. Courtrai), sébastien.lefeuvre@irisa.fr (S. Lefèvre).

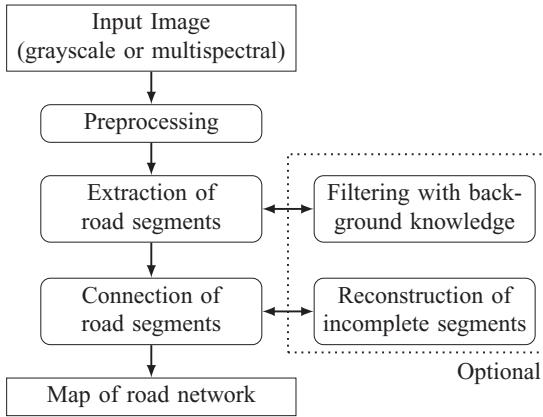


Fig. 1. Flowchart of the proposed method.

This paper extends our previous work [18] with several additional heuristics improving both robustness and efficiency of the initial method. Furthermore, we provide an in-depth analysis of the method parameters, as well as new results obtained on images acquired from various sensors, including some from publicly available datasets. Ability to deal with predefined road width ranges (both low and high bounds), to extract more curvilinear features, at a lower computational cost are the main advantages of our method w.r.t. the initial work of Valero et al. [15].

Our paper is organized as follows. Section 2 provides a comprehensive description of the method and its different steps, namely preprocessing, extraction and connection of road segments, as well as some additional improvements that support robustness or efficiency. Section 3.3 is devoted to experiments, that are conducted on several images coming from different satellite sensors. A quantitative comparison with representative methods from the state-of-the-art is included, as well as an in-depth analysis of the method parameters. Section 4 concludes this paper and discusses future research directions.

2. Method

The proposed method is made of two main steps: extraction and connection of road segments. Robustness to false positives and negatives is ensured through two additional steps: filtering with background knowledge and reconstruction of incomplete segments, respectively. To address various kinds of input data (i.e., grayscale, color or multispectral images), a preprocessing step is also included. Fig. 1 shows the overall flowchart.

For the sake of illustration, we provide in Fig. 2 two sample images, and for each of them the intermediate results (after extraction and connection of road segments) as well as the resulting road network map.

2.1. Preprocessing

The morphological operators involved in the overall process are applied on grayscale images, assuming that the road network is made of rather dark pixels in such images. This is usually the case with asphalt roads observed from panchromatic images. But roads can be made of other material (e.g., concrete or lightly colored rocks), and can be extracted from more informative data such as color or multispectral images.

Thus, the goal of the preprocessing step is to convert the input data $f : \mathcal{E} \rightarrow \mathcal{T}$, with $\mathcal{E} \subset \mathbb{Z}^2$ the image grid and $\mathcal{T} \subset \mathbb{R}^n$ the initial n-d value space, into a grayscale image $g : \mathcal{E} \rightarrow \mathcal{V}$ (with $\mathcal{V} \subset \mathbb{R}$, e.g., $\mathcal{V} = [0, 255]$) where road pixels are darker than the background. If no additional information is available, the multiple input channels

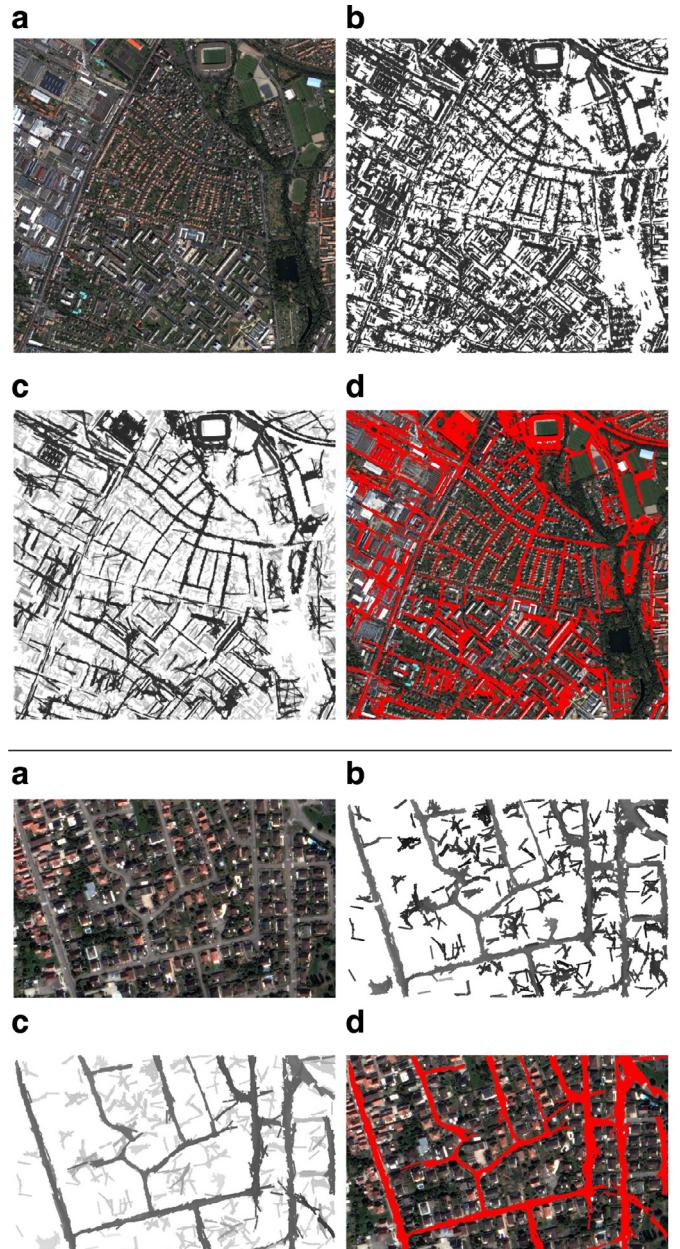


Fig. 2. Illustration of the main steps of our road extraction method: (a) input image extracted from a Quickbird image ©copyright Digitalglobe 2008 (top) and from a Pleiades image ©CNES 2012, Distribution Airbus DS (bottom), (b) extraction (in dark) of road segments and (c) connection through region-based path closing, and (d) final result after thresholding.

are merged to build a unique grayscale image, leading to an appropriate representation for asphalt roads. Otherwise, prior knowledge is used to derive a more accurate grayscale image to be further processed.

If the colors or spectral values of some n road materials are known [19] and noted $\{c_1, \dots, c_n\}$, we compute a distance map where each pixel p is assigned the distance from its spectral signature $f(p)$ to the closest road color c (e.g., asphalt and concrete colors), i.e., $g(p) = \min_i d(f(p), c_i)$, with d a predefined distance metric. Pixels that are made of the same material as roads are thus assigned very low values.

If training samples are available, we can also rely on a supervised classification procedure. We consider two classes (road c and non-road c') and compute membership probabilities m for each

pixel p , i.e., $m_c(p)$ and $m_{c'}(p)$. Pixels that are most likely to belong to roads (i.e., high m_c) are assigned low values, e.g., with the measure $g(p) = 1 - m_c(p) = m_{c'}(p)$.

In case no prior knowledge of road materials is available but multispectral images are given, vegetation can be discarded by computing the NDVI feature (normalized difference vegetation index, computed as $(NIR - R)/(NIR + R)$ with NIR and R denoting respectively the near infrared and visible red bands) and masking the high NDVI values indicating vegetation areas.

Let us note that other transformations might be involved in case of specific visible content of road materials, e.g., textured rock roads would require to compute texture features from which adequate gray values can be extracted. Such a strategy would allow limiting the large standard deviation of pixel values that may occur.

All these various strategies aim to deliver a grayscale image that will be the input for the next steps.

2.2. Extraction of road segments

As already recalled, existing methods for road extraction from remote sensing data usually operate at the pixel scale. However, recent sensors providing very high (spatial) resolution images (or VHR images) call for a higher level of analysis, similarly to the object-based image analysis paradigm [20]. Indeed, the pixel scale both requires a high volume of data to be processed (thus preventing scalability) and is not well adapted to the very high spatial resolution for which road width varies from 10 to 100 pixels. We address these two issues by introducing a segment extraction step, the goal of which is to change the analysis scale from pixels to regions, or in other words from the image support \mathcal{E} to the set of segments \mathcal{S} . Conversely to existing approaches relying on superpixels [3] or regions produced by a segmentation [21], here the regions are identified within the image as possible road segments. We rely here on the assumption that a road segment corresponds to a rectangular area of relative homogeneous content.

The extraction of road segments is then achieved through a probing step. It requires the definition of a rectangular template of predefined size (height and width), that is a priori defined based on the spatial resolution of the image to be processed (and is thus directly related to the visual appearance of roads in the image). The template is used as a sliding window when scanning the image. For each pixel, the template leads to the identification of its neighbors, all being compared with the pixel under analysis. From this comparison is computed the ratio of similar pixels within the template. If this ratio is higher than a predefined threshold, we consider that a possible road segment is present. Pixels are considered similar if their values are relatively close, i.e., within a predefined range. More formally, given a pixel location p , a new segment S will be extracted if

$$\left\{ \frac{\text{card}(p \sim q)}{\text{card}(q)} \geq T_{\text{sim}} \mid q \in \mathcal{N}(p) \right\},$$

with $\mathcal{N}(p)$ the neighborhood of p defined by the sliding window, T a given ratio threshold, and $p \sim q$ the similarity relation between a pixel p and its neighbor q . The similarity relation will be further discussed in Section 3.5. The grayscale value $g(p)$ is transferred to the segment S built from p through a new image of segments defined as $h : \mathcal{S} \rightarrow \mathcal{V} : S \mapsto h(S)$.

Despite leading to efficient further processing of road segments (instead of individual pixels), this probing step also comes with a significant computational cost. Indeed, it requires to analyze the full neighborhood of each single pixel. Such a neighborhood can be sampled, considering only a certain percentage of the pixels in the comparison process (thus using $\mathcal{N}' \subseteq \mathcal{N}$). As it will be demonstrated in the experiments, this allows greatly decreasing



Fig. 3. Illustration of overlap among detected road segments (shown with red borders) to be corrected with the maximal overlapping ratio threshold. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the computational cost without loss in detection accuracy. Let us observe that such an optimization would not have been possible if road segments were produced from superpixels or with a standard segmentation technique. Besides, geometrical properties of road segments are taken into account directly within the extraction step, with no need for post-processing beyond their connection.

While the proposed scheme allows for a great reduction in the number of elements to be subsequently processed, some additional heuristics can be involved to further increase the compression rate. Indeed, the overall computation cost is directly related to the number of road segments to be scanned by the path filtering operator. We thus include a constraint on the maximal overlapping ratio between extracted segments. In other words, a segment is kept as long as its overlapping surface with existing segments is lower than a given ratio, or in other words if it brings a given amount of new content in the detected image. Let us note that such a strategy is sensitive to ordering issues, first extracted segments being given a higher importance in the selection process. A careful analysis of this parameter will be given in Section 3.3. Fig. 3 illustrates the possible overlap between detected segments and the need for some overlap checking procedure as the one proposed here.

Besides, the shape of road segments can be dynamically adjusted to fit as best as possible actual roads. To do so, we consider variable widths for the roads segments. The neighborhood analysis described above assumes a predefined road width. If, for a given pixel, a segment is found, the neighborhood is iteratively extended (by step of 2 pixels) to consider larger road segment width. This allows considering roads with changing width (e.g., from 3 to 4 lanes and back).

2.3. Connection of road segments

Once road segments have been identified, they are used as the elementary units of the image for further processing. While the set of segments is formally defined as \mathcal{S} and associated gray values stored in the image h , in practice we store them in a dedicated grid structure, namely a matrix allowing to access both starting and ending segments at a given image location. Each cell contains two lists of segments, respectively starting and ending at the coordinates indicated by the cell indices. Since the number of segments is much smaller than the number of image pixels, the matrix is indeed sparse. Fig. 4 illustrate the data proposed structure, showing that one segment ends in (3, 2) while two segments start from this cell.

Inspired by the work from Valero et al. [15], we consider here path closing [16], a morphological filter able to extract curvilinear structures. Indeed, it is possible to distinguish between a road (i.e., a chain of segments with significant length) and a building (an unordered set of segments) while these objects can share the same spectral properties.

Original algorithms for path filtering from Talbot and Appleton [16] require some adaptation to process segments instead of pixels.

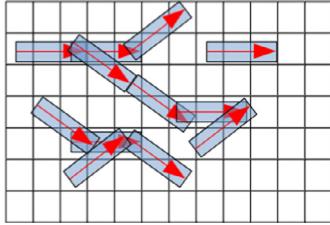


Fig. 4. Data structure used to store the set of segments.

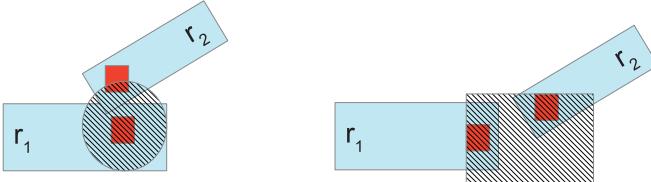


Fig. 5. The two possible strategies to connect successive road segments: using a disc of predefined radius (left) or any predefined shape (right). The latter one allows for greater flexibility and avoids misconnections.

Paths were originally built by linking neighboring pixels depending on their graylevel and following a given direction. Based on our grid data structure, we define here the neighborhood \mathcal{N} of a given segment S as the set of all segments overlapping S and having a grayscale lower or equal to the one of S , i.e. $\mathcal{N}(S) = \{S' \in \mathcal{S} | S \cap S' \neq \emptyset \text{ and } h(S') \leq h(S)\}$. To ensure robustness, we modify this definition by also considering segments relatively close to S (distance between segments lower than a threshold), and with a similar orientation (orientation difference lower than a threshold). Several strategies can be used to connect these non-overlapping segments (see Fig. 5): i) two segments are connected if the end of the first segment is within the neighborhood of the beginning of the second segment (Fig. 5 left); ii) the connection can also rely on a more advanced definition of the junction area, i.e., any predefined shape instead of a single disc of a given radius (Fig. 5 right). The second strategy is more flexible and allows for a more precise definition of the connection process. To illustrate, the second strategy is able to avoid misconnection of segments from the configuration given in the left part of Fig. 5.

We consider here a morphological path closing ϕ , whose aim is to identify long dark paths, while all other structures will be brightened (the dual operator, path opening γ , should be used if the road network is represented by bright pixels over a dark background). The overall algorithm is applied iteratively on graylevels, from the lowest (black) to the highest (white) levels. For each graylevel or value v , paths are built from segments with graylevel lower or equal to v . Each segment S is set to graylevel v if it belongs to a path whose length is higher than a threshold. Thus, it will be kept unchanged if it is possible to build a path of significant length from its graylevel. Otherwise, it will necessarily belong to a path of higher graylevel, and thus will have its graylevel set accordingly. The path closing operator adapted for segments is thus a transform $\phi : \mathcal{V}^{\mathcal{S}} \rightarrow \mathcal{V}^{\mathcal{S}}$ that operates on an image of segments (or mapping from the space of segments \mathcal{S} to the value space \mathcal{V}):

$$\phi(h(S)) = \min(v) | \exists \mathcal{P}, \text{length}(\mathcal{P}) \geq T_{\text{path}}, S \in \mathcal{P}, \\ \forall S' \in \mathcal{P}, h(S') \leq v. \quad (1)$$

Algorithm 1 provides a formal definition of the path closing operator when applied on segments. Depending on the scanning order (from image top to bottom (North to South), left to right (West to East), or conversely), the northern/southern segments are the segments with ending/startng coordinates in the neighborhood of the considered segment. The list L is thus obtained through the

Algorithm 1: Path closing on segments.

```

// Parameters
lmin: minimal path length for roads
// Preprocessing and extraction of segments
L: list of segments sorted by row N/S and column W/E
// Fields of segment structure
seg.active : active flag used in the algorithm
seg.value : segment value (graylevel) after detection
seg.length : segment length (in pixels)
seg.fvalue : final value of the segment (after closing)
seg.lup : maximal length for North path
seg.ldown : maximal length for South path
// Initialization
seg.active=false
seg.fvalue=null
seg.lup=0
seg.ldown=0
// Processing
forall the values V from black to white do
    // First Scan (list N/S and W/E, update lup)
    activate segments from list L when seg.value = V
    forall the all active segments S in list L do
        forall the attached northern segments aSeg do
            | S.lup ← max (S.lup, aSeg.lup + aSeg.length)
        end
        if S.lup is updated AND S.value = null then
            | if S.lup + S.down + S.length ≥ lmin then
                | | S.fvalue ← V
            end
        end
        if S.lup is updated OR S.value = V then
            forall the southern rectangles aSeg do
                if aSeg.value ≤ V then
                    | | aSeg.active ← true
                end
            end
        end
        S.active ← false
    end
    // Second Scan (with inverted list S/N and E/W)
    // Third and Fourth Scans, updating seg.ldown and reactivating northern
    // segments instead of southern ones

```

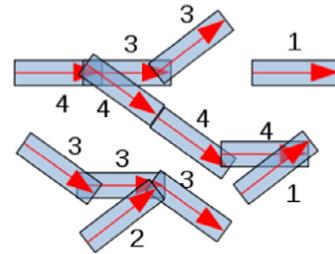


Fig. 6. Illustration of path length computation. For each segment, the length of the longest path it belongs to is shown.

analysis of the grid data structure (Fig. 4). For the sake of illustration, we provide in Fig. 6 the length of the paths built from segments stored in the data structure shown in Fig. 4.

The road segments contained in the filtered image have their graylevels brightened depending on the size of the longest path they belong to. It is then straightforward to produce the final road network map with a simple thresholding.

2.4. Filtering with background knowledge

Accuracy of the road network map directly relies on the quality of road segments extracted in the first step. Among these segments, a significant subset corresponds to large and homogeneous areas (e.g., parking lots, courts, fields, etc.) and leads to false positives. We propose to take into account background knowledge for a

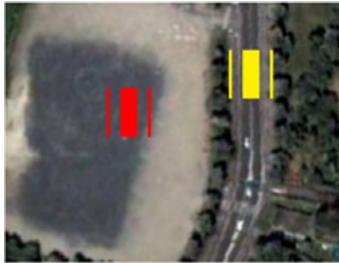


Fig. 7. HMT ability to deal with background knowledge and distinguish between actual road segments (yellow) and outliers (red). Foreground and background are respectively probed with the central and side rectangles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

more accurate identification of road segments. We thus follow the principle of the Hit-or-Miss Transform (HMT) that led to the successful design of template matching solutions for a wide range of problems, including in remote sensing [17]. Assuming that roads can be distinguished from other areas by their width, we modify the first step to take into account both a foreground rectangular template and a composite background template. The latter is made of two parallel rectangular areas located on each side of the foreground template. Distance between these templates and the foreground one is set depending on the expected road width and its possible variation (e.g., varying number of lanes). The pixels not contained in the foreground nor the background templates (i.e., in-between area) corresponds to an uncertainty zone, and are not involved in the template matching process. The probing now operates on two criteria: all pixels covered by the foreground template shall have a similar value to the central pixel, i.e., foreground color, while all pixels covered by the background template shall have a different value to the central pixel, i.e., background color. Let us remark that the ratios of similar pixels for the foreground and background template have to differ, since the uniformity imposed on road segments is not valid anymore when dealing with background (that can be made of any material). The proposed filtering with background knowledge step leads to the extraction of a segment S from pixel location p if both following conditions hold:

$$\left\{ \begin{array}{l} \frac{\text{card}(p \sim q)}{\text{card}(q)} \geq T_{\text{sim-fg}} \mid q \in \mathcal{N}_{\text{fg}}(p), \\ \frac{\text{card}(p \not\sim q)}{\text{card}(q)} \geq T_{\text{sim-bg}} \mid q \in \mathcal{N}_{\text{bg}}(p) \end{array} \right\},$$

with \mathcal{N}_{fg} and $T_{\text{sim-fg}}$ (resp. \mathcal{N}_{bg} and $T_{\text{sim-bg}}$) the neighborhood defined by foreground (resp. background) template and its related ratio threshold.

Fig. 7 shows this step allows distinguishing between actual road segments and regions that compose large pieces of the same material (e.g., asphalt, concrete) but are not roads.

2.5. Reconstruction of incomplete segments

While the previous step aims to remove false positives, it is also necessary to remove false negatives. Indeed, some roads might not be completely detected due to a partial occlusion (e.g., by trees, shadow, cars, pedestrian crossing, etc.) of road segments, leading to their heterogeneous content and a subsequent missed detection. This in turn results in roads of insufficient length, discarded by the path closing process.

Similarly to Talbot and Appleton [16], we also rely on incomplete path filtering to deal with missing elements. Let us recall that here, elements are segments instead of pixels. A scanning process is started for each road extremity. It tries to fill with a small set of successive segments the possible gap between the current

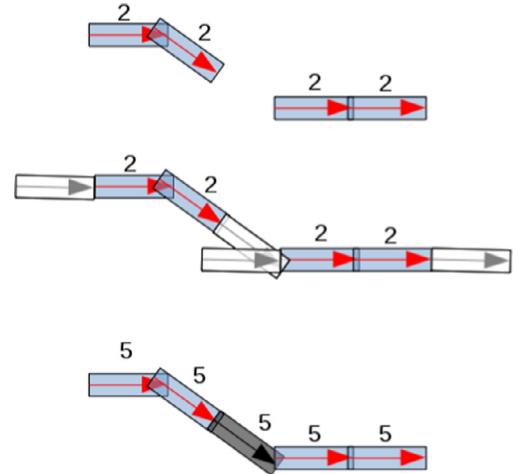


Fig. 8. Illustration of path length computation (from top to bottom): a set of segments with their path length, the candidate segments to complete existing paths (only segments located at extremities of existing paths are considered), and finally the selected segments (connecting two paths) with path length recomputed.



Fig. 9. Illustration of the ability of incomplete path filtering to reconnect disconnected road segments: initial segments are given in red, added segments are in blue, thus resulting in length updating in the path filtering process that is finally able to return new segments (in green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

road and another one. To ensure coherence between the added segments and the original ones, and to avoid the inclusion of outliers (e.g., a building located between two roads), the connecting segments shall share some common properties. Here we rely on the same criteria than in the path closing step, but using some lowered constraints (thresholds). This reconstruction process is included in the general path closing step, thus enabling to connect road segments even in presence of missing segments.

For the sake of illustration, we provide in Fig. 8 an example of path completion, where the candidate segments located at path extremities are used to connect two paths, leading to a further path length updating step.

Fig. 9 shows this step allows reconnecting segments that were initially unconnected (in red), relying on both filling segments (in blue) and new segments (in green) extracted following the update of the path length. While present in the set of segments extracted after the first step of the method, such new segments were initially discarded since they were not belonging to a path of sufficient length. The proposed incomplete path filtering procedure prevents discarding such segments.

3. Experiments

Our method has been assessed on various satellite images. We provide here some comparative results with the state-of-the-art, as

well as some analysis and discussion of the different parameters involved in the different steps of our method.

3.1. Datasets and evaluation measures

To assess our contribution, we have performed a series of experiments on various satellite images. For the sake of comparison, we will consider here Quickbird images and show results obtained both on a grayscale image of 420×300 pixels used by Valero et al. [15], and on a multispectral image of 2832×2772 pixels from Strasbourg, France. The latter has been converted to grayscale and is associated with some validation data provided by French Geographic Institute IGN. We also assess our method on some Ikonos images from an almost 10 years old dataset [22]. As observed in results reported in recent works (e.g., Shi et al. [13]), this dataset can still be considered as a reference, and the last years have not seen significant advances of the state-of-the-art. Comparative results as well as quantitative measurements will be provided in Section 3.3.

Evaluation is performed using standard measures. We consider here the criteria proposed by Wiedemann et al. [23] and still commonly adopted e.g., by Bae et al. [12]: completeness (or recall, sensitivity), correctness (or precision), as well as a quality score inspired from the accuracy measure. The latter is defined using TP , FP , FN (respectively denoting true positives, false positives, and false negatives) by the ratio $TP/(TP + FP + FN)$. It is very similar to the standard accuracy measure, the only difference being the lack of TN (true negatives) in the denominator, in order to accommodate with the unbalanced classes (roads vs. non-roads). As indicated by Mayer et al. [22], practical usefulness of road network extraction methods requires completeness and correctness scores above 0.6 and 0.75, respectively. Let us recall that they are respectively computed as ratios $TP/(TP + FN)$ and $TP/(TP + FP)$. Furthermore, we also use the F1-score (harmonic mean of completeness and correctness, i.e., $(2TP)/(2TP + FP + FN)$) when considering the IKONOS dataset, to ease fair comparison with the state-of-the-art results achieved on this dataset.

3.2. Parameter settings

The method proposed in this paper relies on a number of parameters that have been set empirically from image properties (e.g., spatial resolution, contrast). In the experiments that will be reported in the comparative evaluation (Strasbourg image), we have used the following settings. Road segments have width of 5 pixels, and height of 45 pixels. Segments are detected in all orientations with an angular step of 5° . Maximal orientation change and maximal distance between segments are respectively set to 30° and 5 pixels. Pixels are considered similar if their graylevels are equal $\pm 5\%$. The rate of similar pixels within a segment is set to 100% (all pixels shall be similar). Background knowledge requires the definition of an uncertainty zone in the HMT process. It has been set here equals to road width (5 pixels). The ratio of similar pixels within the background template is set to 20% since less assumptions can be made on the uniformity of the background (i.e., non road) areas. Reconstruction fills gaps of length once or twice the length of segments (i.e., 45 or 90 pixels). A path is considered as a road if it contains at least 5 successive segments (ca. 250 pixels). Finally, pixels from the filtered image are kept in the road network map if their graylevel is lower to 128. Since we consider here asphalt roads, the final map contains the dark half of the filtered image, while the other objects are made of pixels brightened by the region-based morphological filtering. But as described in Section 2.1, the preprocessing allows dealing with other contexts and images.

Since parameters have been set here empirically, we provided in Section 3.5 some analysis of their impact on the end-results.

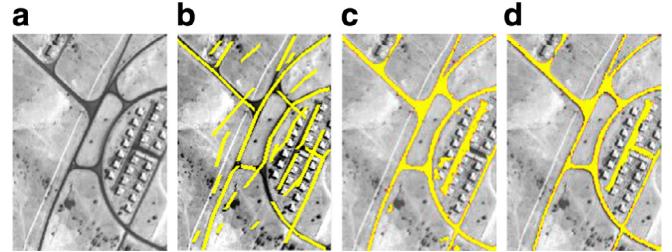


Fig. 10. Road network extraction, easy dataset: original image from [15] (a) and results from OTB (b), Valero et al. (c), and our method (d).

Table 1

Comparison of road network extractors, easy dataset.

| Method | Completeness | Correctness | Quality |
|---------------|--------------|-------------|-------------|
| OTB | 0.63 | 0.70 | 0.50 |
| Valero et al. | 0.95 | 0.78 | 0.76 |
| Proposed | 0.93 | 0.85 | 0.81 |

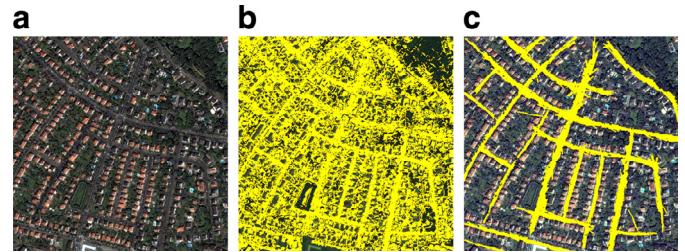


Fig. 11. Road network extraction, challenging dataset: RGB composition (a) of the original image ©Digitalglobe 2008, and results from Valero et al. (b) and our method (c).

Their automatization is left for future work. Indeed, the parameters greatly depend on the image properties but also on the kind of road network considered, since the latter may vary considerably from one country to the other, and even within a given country (urban versus rural areas).

3.3. Accuracy evaluation

We compare our method with two representative techniques from the state-of-the-art: the road extraction algorithm available in OTB (CNES ORFEOTool Box, <http://www.orfeo-toolbox.org/otb>) and the method from Valero et al. [15] that motivated our work. The latter has been here initialized manually with the parameters leading to the best results.

We first consider the easiest dataset already suggested in [15]. We can observe in Fig. 10 that both our method and Valero et al. [15] provide similar results, better than with OTB which does not rely on paths. A quantitative analysis is provided in Table 1, confirming the conclusions driven from visual analysis.

A second and more challenging experiment has been performed with the Strasbourg image. We can see in Fig. 11 (a subset of the original image for better visualization) that roads have a very similar color to the background. While such a difficult scenario cannot be tackled by Valero et al. [15] (nor by the OTB method which is unable to analyze such a complex image), our method (with all steps from Fig. 1) is able to extract the most important elements of the road network. We clearly observe the relevance of mapping road segments by rectangular areas, making possible to distinguish between road segments and neighboring pixels of similar color or graylevel.

Once roads have been extracted, the availability of some reference data provided by IGN allows us performing a quantitative

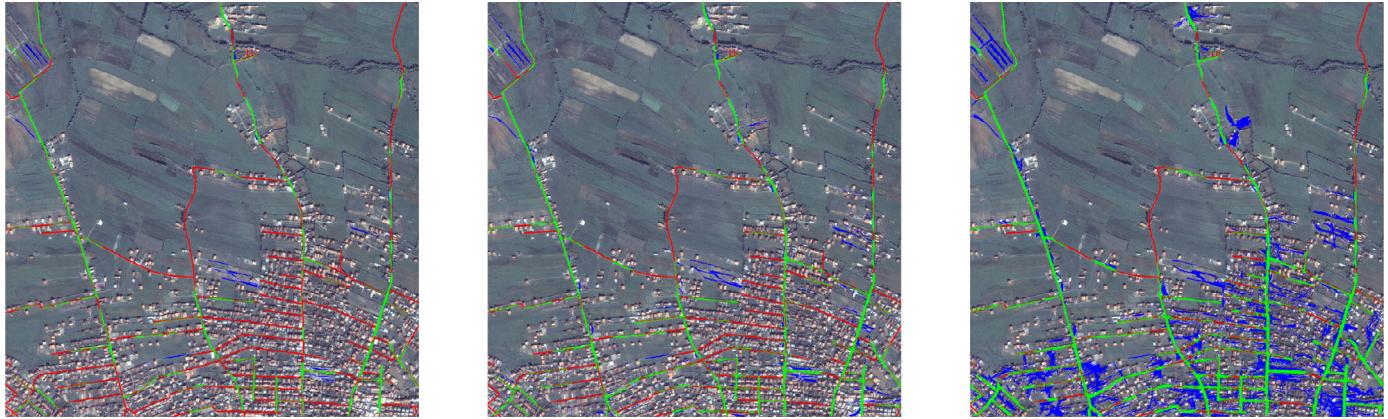


Fig. 12. Illustration of road network extraction with correctness preferred (left), completeness preferred (right), or compromise between both objectives (center). True positives, false negatives and false positives are respectively displayed in green, red, and blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Comparison of road network extractors, challenging dataset.

| Method | Completeness | Correctness | Quality |
|----------------------------|--------------|-------------|-------------|
| Baseline | 0.68 | 0.80 | 0.58 |
| Baseline + Filtering | 0.70 | 0.82 | 0.61 |
| Reconstruction | 0.69 | 0.81 | 0.60 |
| Reconstruction + Filtering | 0.71 | 0.83 | 0.62 |

Table 3
Quantitative evaluation on the IKONOS dataset, various settings favoring completeness, correctness, or both.

| Settings | Completeness | Correctness | F1 score |
|-----------------|--------------|-------------|-------------|
| Low FN | 0.71 | 0.42 | 0.53 |
| Low FP | 0.42 | 0.81 | 0.56 |
| Low FN and FP | 0.60 | 0.69 | 0.64 |
| Shi et al. [13] | 0.34 | 0.63 | 0.44 |

evaluation. Only a part of the Strasbourg image is available with some ground truth (Fig. 11(a), 1000 × 1000 pixels) and leads to quantitative evaluation. Table 2 gives the results obtained with our baseline method, as well as alternative steps to improve robustness to false positives/negatives. We can quantitatively assess the relevance of these options.

Finally, we also report accuracy evaluation on the IKONOS dataset used in the comparative evaluation performed by Mayer et al. [22]. Results obtained on the Ikonos1-Sub1 image are given in Fig. 12, using the same color code as in the original paper, and with different settings favoring completeness, correctness, or both. Indeed, parameter settings can be achieved to increase completeness or correctness, as shown in Table 3, leading to similar overall accuracy scores (measured here with the F1 Score). Finding a compromise between completeness and correctness (thus trying to lower both false positives and false negatives) allows reaching a higher level of overall quality, as shown in Table 3. We can see on the last line of the table that the proposed settings is not far to meet the requirements for practical usefulness (completeness of 0.6, correctness of 0.75) defined by Mayer et al. [22] on this challenging dataset. Let us recall that F1 scores reported in the initial study were varying between 0.37 and 0.57, and that even recent methods such as [13] hardly tackle this dataset.

3.4. Performance evaluation

Beyond a major improvement in detection accuracy, the proposed region-based morphological scheme also allows a significant

Table 4
Comparison of processing times (in seconds).

| Method | Reconstruction | CPU time |
|---------------|----------------|----------|
| Valero et al. | No | 98 |
| Valero et al. | 1 pixel | 310 |
| Valero et al. | 2 pixels | 720 |
| Proposed | No | 84 |
| Proposed | 1 segment | 117 |
| Proposed | 2 segments | 119 |

decrease of computation time. We compare here our method with the initial work of Valero et al., for which we have realized a Java-based implementation based on the C++ code provided by the authors (so ensuring a fair level of optimization). As far as our method is concerned, it comes with a moderate level of optimization (single-threaded, with all optimization tricks described in the previous sections). Using a similar coding language and operating environment allows for more objective comparison. Furthermore, both methods are characterized with a similar computational complexity, since both rely on the principles of the path opening operator. There is however a major difference, that lies in the kind of elements to be processed (segments instead of pixels).

Measures obtained with the full Strasbourg image are reported in Table 4 (using a standard workstation, with a i5 2.5 GHz CPU and 8 GB RAM). Let us observe that our method requires 22 seconds to extract road segments, which lowers the efficiency gain if no reconstruction is performed. However, processing regions instead of pixels leads to a major improvement in efficiency as soon as a reconstruction step is involved. Moreover, the complexity increases only slightly when considering gaps of higher length (1 or 2 segments, i.e., ca. 50 or 100 pixels). Let us note that these results shall be considered carefully since optimization was only moderate.

Most of the papers introducing new road detection methods do not provide indications of computation time. Nevertheless, our method seems very competitive w.r.t. the state-of-the-art, e.g., more than one hour is required by Ouled Sghaier and Lepage [14] to process a 2048 × 2048 image.

As far as the memory cost is concerned, we have measured the memory footprint with a profiling tool (namely jVisualVM). The Java user heap was 1.86 GB for Valero et al., while only 0.95 GB with our algorithm, for processing the Strasbourg image (31 MB). Both methods were performing an incomplete path opening. Let us recall that these values were obtained without specific optimization, and the structure we are using (a sparse matrix to store two

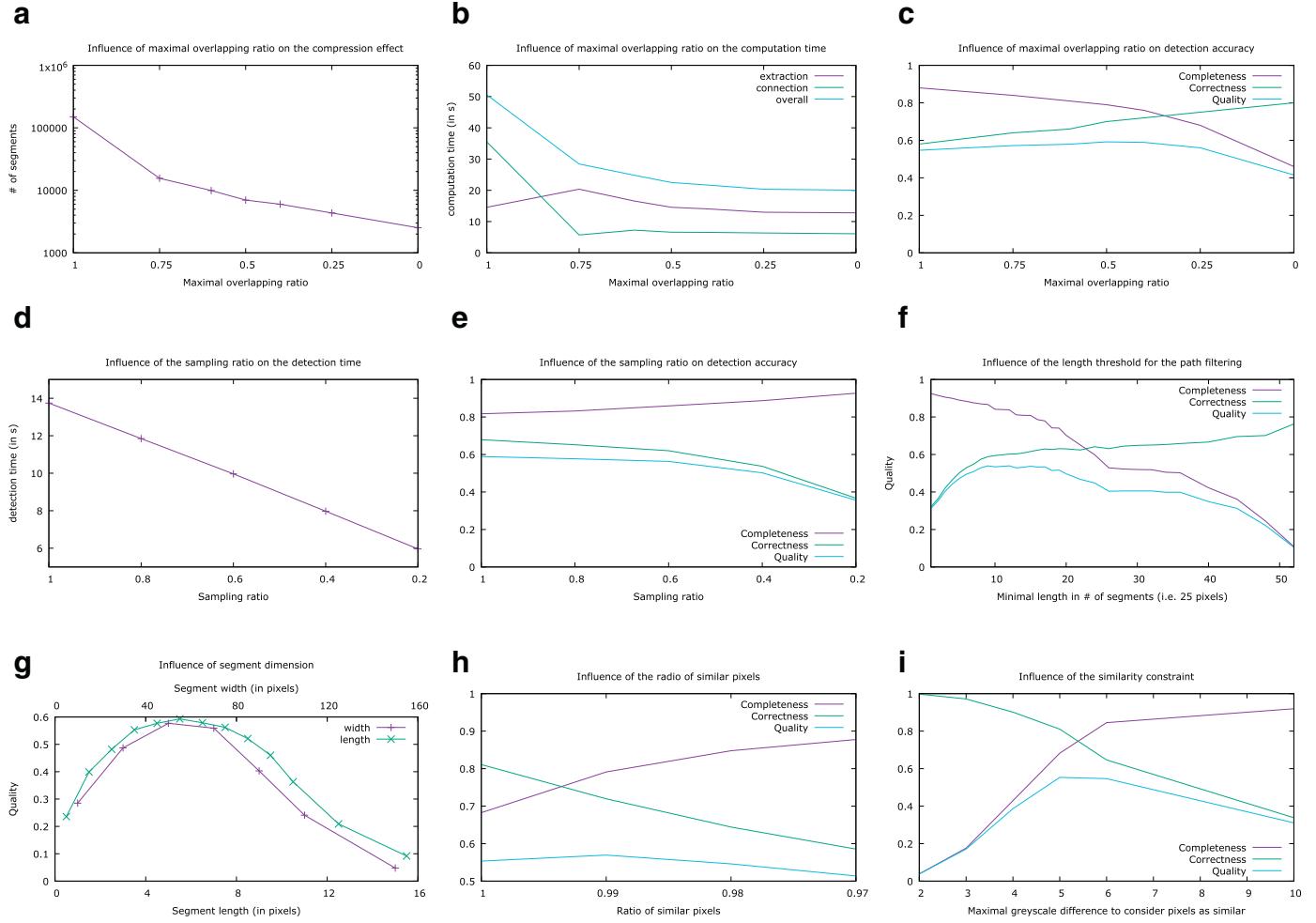


Fig. 13. Influence of parameters settings: maximal overlapping ratio effect on the number of extracted segments (a), the overall complexity (b) and quality of the full detection process (c); sampling ratio effect on the complexity of the detection step (d) and the quality of the full detection process (e); observed quality measures for different values of the road segment connection parameter (f) and of the road segment extraction parameters: segment dimensions (g), ratio of similar pixels (h), and grayscale difference to define similarity (i).

lists of segment extremities) contributes to a gain in CPU cost w.r.t. a linked list structure, but at an additional memory cost. Nevertheless, we can observe a significant gain in terms of memory cost, that is easily explained by the change of representation: while the original work from Valero et al. was operating on pixels, the proposed method is processing road segments, i.e., objects that are far less numerous than image pixels.

3.5. Discussion on heuristics and parameters

Several heuristics related to the extraction of segments were introduced in Section 2.2. We report here some experimental results obtained when conducting further analysis to understand how they behave in the overall process.

We first study how the overlapping ratio threshold influences the number of extracted segments, and subsequently the overall computation time as well as the accuracy of the end results. We consider various ratios from 1 (no overlapping constraint, overlapping allowed until 100%) to 0 (strict overlapping constraint, no overlapping allowed). The results are given in first line of Fig. 13 and show the great effect of this constrained selection procedure on the number of extracted segments that composed the space on which the path filtering process operates. We can observe that reducing the number of segments leads to a very significant gain in performance, while keeping satisfying accuracy levels (e.g.,

keeping a segment only if it brings 75% of new content, i.e., maximal overlapping ratio lower than 25%, reduces by a factor of 35 the number of segments, by 5 the path filtering time and by 2.5 the overall time, for a 20 points loss in completeness and no loss in overall quality. Conversely, a slight gain in quality can be observed. It is probably due to the fact that the unconstrained scenario extracts some erroneous rectangles located on road borders (partly inside, partly outside the actual road). Such a situation is alleviated with the overlapping constraint that force extracted rectangles to be aligned (on the road direction). But this gain is lowered when the overlapping ratio becomes too low, since there is then not enough remaining road segments for the filtering step. Considering a maximal overlapping ratio of 50% allows reaching a quality level of 0.59 (i.e., gain of 4.5 points), for only a half of the initial CPU cost.

Another optimization consists in considering only a certain percentage of the pixels in the neighborhood when checking if a segment has to be selected or filtered out. Fig. 13(d) and (e) show that this sampling procedure leads to a significant decrease in computation time for the detection step (d), with no prohibitive loss in accuracy up to a certain sampling ratio (e). Indeed, the method is able to provide results of similar quality when considering only 50% of the pixels.

We have also studied how the dynamic width adjustment influences the quality of the end results. However, on the Strasbourg

image considered so far, the observed differences in quality score were not considered as significant (lower than 1%). Similarly, using an advanced neighborhood configuration to connect successive segments has not led to major improvement, since the gain in correctness has been counterbalanced by a loss in completeness, resulting in an overall gain in quality of ca. 1%. We are convinced that conducting additional experiments on other datasets will be helpful in assessing the relevance of these various heuristics.

The overall process relies both on extraction and connection of road segments, whose parameters' influence are studied in the remaining plots of Fig. 13. The first step is parametrized by the dimensions and heterogeneity of the road segments to be extracted. Fig. 13(g) shows how dimensions influence the overall result. When dimensions are close to 1, the region-based path filtering will act similarly as the standard one (i.e., pixelwise) and keep small areas located along the road. Conversely, high values for segment length and width will prevent selecting curved and narrowed roads, respectively. Appropriate dimensions have to be set depending on the visual appearance of roads in the processed image. We can observe that working with segments of elongation ratio around 1:10 (i.e., 50×5 for length \times width) as basic units for the morphological path filtering operator used in the second stage leads to the best results. Road segments are defined by their dimension but also by their content. More precisely, a road segment is extracted if it contains a significant ratio of similar pixels in a rectangular neighborhood (see Section 2.2). Thus, Fig. 13(h) and (i) show respectively the effect of the ratio parameter and similarity constraint. We can observe that accepting 1% of outliers leads to better results, while higher values tend to lower the overall quality. Furthermore, a tolerance in terms of graylevels within the segment is required to achieve satisfying results. Defining two pixels as similar if the difference between their graylevels is not lower than 5 or 6% of the graylevel range provides the best results. Let us note that a very low difference (e.g., 0, all pixels have to be strictly equal; or 1, all pixels should share the value of the centered pixel $\pm 1\%$) prevent detecting any road segment.

The second step, namely connection of road segments, consists in a morphological path filter. It relies on an additional parameter defining the minimal path length or the number of successive segments that a path should contain to remain after the filtering step (similarly to the path length used with the standard path filtering operator). As shown on the Fig. 13(f), this filtering is mandatory to eliminate false positives (resulting in gain of 15 points in correctness), and best results were obtained with a threshold value of 9. Let us note that a length of 9 segments is equal to 225 pixels (each segment being 25 pixels long in the settings used for this figure), and the processed image is here 1200×800 pixels. Higher threshold values result in discarding small roads, leading to a loss in detection quality.

3.6. Requirements for better spatial resolution

Results provided so far were promising and show how our method outperforms the state-of-the-art. Nevertheless, we have also observed its limitations in some specific cases, e.g., when road segments were very narrow (while our method is assuming a minimal width). As illustrated in Fig. 14, some road segments are then not extracted if only graylevels are used. Indeed, on the image given in the aforementioned figure, the method achieved low recognition rates (0.34 for completeness, 0.44 for correctness), below expected standards defined by Mayer et al. [22]. This demonstrates the need for relying on additional or updated data when detection is particularly challenging. Let us note that detection is eased with new sensors such as Pleiades that come with a higher spatial resolution (50cm/pixel for the Pleiades color image vs 2.4 m/pixel for the Quickbird multispectral image). To illustrate,

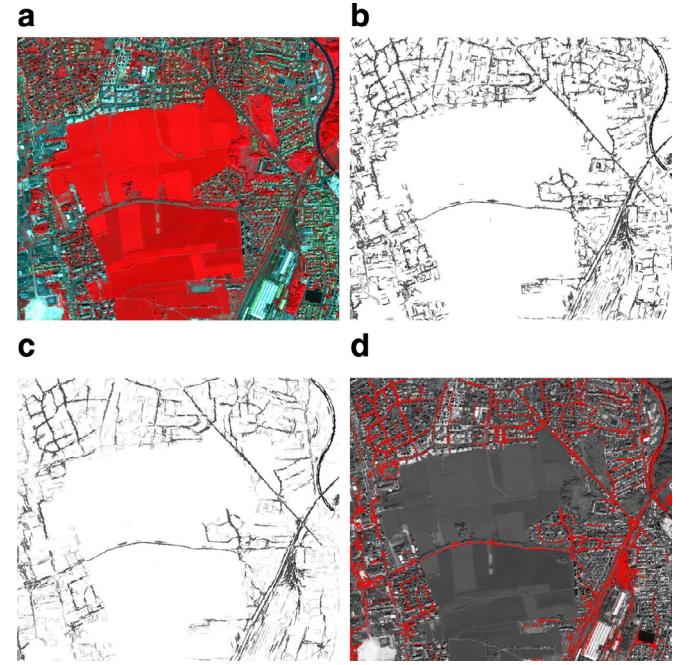


Fig. 14. Road network extraction with a very challenging dataset with narrow road segments: (a) RGB composition of the original Quickbird image – Strasbourg North area ©Digitalglobe 2008, outputs of the road segment (b) detection and (c) connection steps, and (d) final map superimposed on the panchromatic processed image.

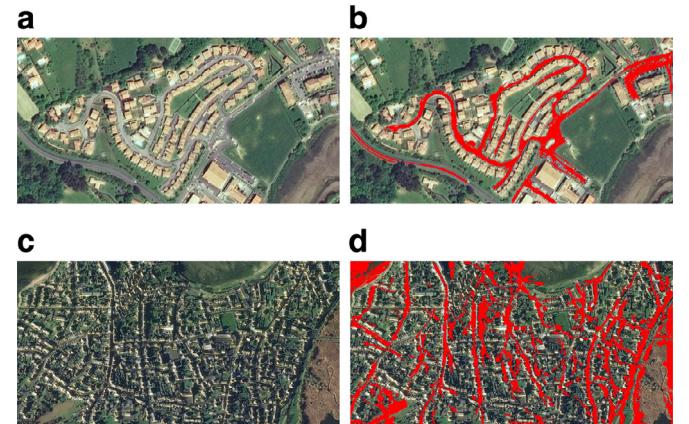


Fig. 15. Road network extraction on sample Pleiades images: (a) original image of St-Jean-de-Luz ©CNES 2013, Distribution Airbus DS and (b) extracted road network; (c) original image of Lorient ©CNES 2012, Distribution Airbus DS and (d) extracted road network.

Fig. 2 (bottom) used in Section 2 shows how road segments are extracted from a grayscale image of Strasbourg, and with the same setup as Fig. 14. Recognition rates climb up to 0.79 for completeness, 0.69 for correctness, leading to a global quality score of 0.58.

Experiments with Pleiades images have been conducted with two other geographical areas of France, namely St-Jean-de-Luz and Lorient, and results are given in Fig. 15. While in the first case the road network is accurately extracted from residential areas, the second one is much more challenging. Roads do not have a constant spectral value especially due to shadows, and can be sometimes barely visible. The proposed method is able to extract only a (significant) part of the road network. It can then be used as a very efficient deterministic initialization of more complex probabilistic strategies.

4. Conclusion

Extraction of roads from aerial or satellite images allows maintaining an up-to-date map of road networks. While explored for decades, automatic road extraction remains challenging. Very high spatial resolution sensors provide more details that should ease recognition techniques, but also bring a higher data volume and lead to visual appearance of roads made of elongated structures with a width higher than a single pixel. In this paper, we address these issues by introducing new region-based morphological operators. We first map an input image into a set of road segments which become the elementary image units. Such elements are then processed with an (incomplete) path closing algorithm, the aim of which is to emphasize segments belonging to the road network.

Experiments conducted on various satellite images demonstrate how our method outperforms existing solution using path closing at the pixel scale [15], both in terms of accuracy and efficiency. Results obtained with very high spatial resolution images (such as the ones provided by the Pleiades sensor) show that our method achieves automatic road extraction with a satisfying level of accuracy. Nevertheless, our method is not yet able to address aerial datasets such as the one provided by the ISPRS Test Project on Urban Classification and 3D Building Reconstruction [24]. We have indeed observed that shadows are playing a crucial role in this dataset, and are making roads appearing as disconnected paths (and thus are calling for a specific shadow processing to address such artefacts).

Among future research directions, we believe the color/spectral information could be addressed in a better way, through color/multispectral morphological operators [25] instead of the conversion to graylevels used as preprocessing. Beyond color, texture information can also be taken into account. Furthermore, the proposed approach being only deterministic, it could be combined with probabilistic modeling or refinement of road segments. Finally, improved path filtering algorithms have been recently proposed [26,27] and could also benefit to our method (if adapted to the region scale).

Acknowledgments

The authors would like to thank A. Puissant from LIVE UMR CNRS 7362 (University of Strasbourg) for providing the geometrically corrected images of Strasbourg (Quickbird and Pleiades sensors), and H. Mayer from Bundeswehr University Munich for providing the IKONOS dataset [22]. IKONOS images are ©Bundeswehr Geoinformation Office (AGeoBw), Euskirchen, Germany. This work has been supported by the French Agence Nationale de la Recherche (ANR) under reference ANR-13-JS02-0005-01 (Asterix project).

References

- [1] R. Bajcsy, M. Tavakoli, Computer recognition of roads from satellite pictures, *IEEE Trans. Syst. Man Cybernet.* 6 (9) (1976) 623–637.
- [2] J. Mena, State of the art on automatic road extraction for {GIS} update: a novel classification, *Pattern Recognit. Lett.* 24 (16) (2003) 3037–3058.
- [3] J. Wegner, J. Montoya-Zegarra, K. Schindler, Road networks as collections of minimum cost paths, *ISPRS J. Photogramm. Remote Sens.* 108 (2015) 128–137.
- [4] O. Besbes, A. Benazza-Benyahia, Joint road network extraction from a set of high resolution satellite images, in: Proceedings of the European Signal Processing Conference (EUSIPCO), 2014, pp. 2190–2194.
- [5] V. Mnih, G. Hinton, Learning to detect roads in high-resolution aerial images, in: Proceedings of the European Conference on Computer Vision, in: Lecture Notes in Computer Science, vol. 6316, 2010, pp. 210–223.
- [6] J. Wang, J. Song, M. Chen, Z. Yang, Road network extraction: a neural-dynamic framework based on deep learning and a finite state machine, *Int. J. Remote Sens.* 36 (12) (2015) 3144–3169.
- [7] T. Géraud, J. Mouret, Fast road network extraction in satellite images using mathematical morphology and Markov random fields, *EURASIP J. Appl. Signal Process.* 16 (2004) 2503–2514.
- [8] C. Zhu, W. Shi, M. Pesaresi, L. Liu, X. Chen, B. King, The recognition of road network from high-resolution satellite remotely sensed data using image morphological characteristics, *Int. J. Remote Sens.* 26 (24) (2005) 5493–5508.
- [9] R. Gaetano, J. Zerubia, G. Scarpa, G. Poggi, Morphological road segmentation in urban areas from high resolution satellite images, in: Proceedings of the International Conference on Digital Signal Processing, 2011, pp. 1–8.
- [10] C. Sujatha, D. Selvathi, Connected component-based technique for automatic extraction of road centerline in high resolution satellite images, *EURASIP J. Image Video Process.* 8 (2015) 1–16.
- [11] C. Unsalan, B. Sirmacek, Road network detection using probabilistic and graph theoretical methods, *IEEE Trans. Geosci. Remote Sens.* 50 (11) (2012) 4441–4453.
- [12] Y. Bae, W. Lee, Y. Choi, Y. Jeon, J. Ra, Automatic road extraction from remote sensing images based on a normalized second derivative map, *IEEE Geosci. Remote Sens. Lett.* 12 (9) (2015) 1858–1862.
- [13] W. Shi, Z. Miao, J. Debayle, An integrated method for urban main-road centerline extraction from optical remotely sensed imagery, *IEEE Trans. Geosci. Remote Sens.* 52 (6) (2014) 3359–3372.
- [14] M. Ouled Sghaier, R. Lepage, Road extraction from very high resolution remote sensing optical images based on texture analysis and beamlet transform, *Sel. Top. Appl. Earth Obs. Remote Sens. IEEE J.* 9 (5) (2015) 1946–1958.
- [15] S. Valero, J. Chanussot, J. Benediktsson, H. Talbot, B. Waske, Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images, *Pattern Recognit. Lett.* 31 (10) (2010) 1120–1127.
- [16] H. Talbot, B. Appleton, Efficient complete and incomplete path openings and closings, *Image Vis. Comput.* 25 (4) (2007) 416–425.
- [17] S. Lefèvre, E. Aptoula, B. Perret, J. Weber, Morphological template matching in color images, in: E. Celbi, B. Smolka (Eds.), *Advances in Low-Level Color Image Processing*, Lecture Notes in Computational Vision and Biomechanics, vol. 11, Springer, 2014, pp. 241–277, doi:10.1007/978-94-007-7584-8_8.
- [18] L. Courtrai, S. Lefèvre, Road network extraction from remote sensing using region-based mathematical morphology, in: Proceedings of the IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS), 2014.
- [19] M. Herold, D. Roberts, Mapping asphalt road conditions with hyperspectral remote sensing, in: Proceedings of the International Symposium Remote Sensing of Urban Areas (URS), ISPRS, 2005.
- [20] T. Blaschke, Object based image analysis for remote sensing, *ISPRS J. Photogramm. Remote Sens.* 65 (1) (2010) 2–10.
- [21] H. Long, Z. Zhao, Urban road extraction from high-resolution optical satellite images, *Int. J. Remote Sens.* 26 (22) (2005) 4907–4921.
- [22] H. Mayer, S. Hinz, U. Bacher, E. Baltsavias, A test of automatic road extraction approaches, in: Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2006, pp. 209–214.
- [23] C. Wiedemann, C. Heipke, H. Mayer, O. Jamet, Empirical evaluation of automatically extracted road axes, in: Proceedings of the IEEE Computer Society Workshop on Empirical Evaluation Techniques in Computer Vision, 1998, pp. 172–187.
- [24] F. Rottensteiner, G. Sohn, J. Jung, M. Gerke, C. Baillard, S. Benitez, U. Breitkopf, The ISPRS benchmark in urban object classification and 3d building reconstruction, *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. I-3* (2012) 293–298.
- [25] E. Aptoula, S. Lefèvre, A comparative study on multivariate mathematical morphology, *Pattern Recognit.* 40 (11) (2007) 2914–2929.
- [26] F. Cokelaer, H. Talbot, J. Chanussot, Efficient robust d-dimensional path operators, *IEEE J. Sel. Top. Signal Process.* 6 (7) (2012) 830–839.
- [27] V. Morard, P. Dokladal, E. Decencière, Parsimonious path openings and closings, *IEEE Trans. Image Process.* 23 (4) (2014) 1543–1555.