Semi-Automatic Road Extraction From VHR Images Based on Multiscale and Spectral Angle in Case of Earthquake

Idrissa Coulibaly, Student Member, IEEE, Nicolas Spiric, Richard Lepage and Michele St-Jacques

Abstract—Road extraction offers great potential for research initiatives because of the complexity due to its great topological variability. The use of remote sensing imagery to accomplish this mapping is an interesting option. Indeed, satellite images can be acquired shortly after the event, and cover a large area of territory. We hope to produce a mapping of the present facilities from very high resolution images shortly after a disaster. This availability of very high spatial resolution (VHSR) images brings added value to the study in urban areas and their mapping. Increasing the spatial resolution generates noise, which makes extraction difficult, especially in the event of an earthquake in an urban context. This problem increases false alarm rates and generally affects the performance of road extraction algorithms in detecting linear features used to locate and extract roads on such images. During major disasters, short deadlines demand an effective response in terms of updating the mapping of affected areas. Our aim is to improve the road extraction quality after adaptation of Lowe’s SIFT (scale-invariant features transform) descriptors jointly with spectral angle (SA) algorithms. An illustration is performed on three high-resolution images, respectively representing a rural, suburban and urban disaster area, captured by the Quickbird satellite. Our approach significantly reduces the false detection rate and shows an increase in overall quality of up to nearly 30% in some cases as compared to what obtain in the literature.

Keywords: Road extraction, Multiresolution analysis, Multi-spectral images, Major disaster, Spectral angle.

I. INTRODUCTION

Recent years have seen a spike in the number of natural disasters occurring around the world: the tsunami in Indian Ocean (2004), the earthquake in Haiti (2010) and the earthquake in Japan (2011) are recent examples. All relevant scenes are spread over a wide area and in all cases the damage was immense.

I. Coulibaly, N. Spiric and R. Lepage are with École de technologie supérieure, Montréal, Quebec, Canada, Department of Automated Production Engineering (e-mail: idrissa.coulibaly.1@ens.etsmtl.ca; nicolas.spiric.1@ens.etsmtl.ca, richard.lepage@etsmtl.ca).

Michele St-Jacques is with École de technologie supérieure, Montréal, Quebec, Canada, Department of Construction Engineering (e-mail: michele.st-jacques@etsmtl.ca).

One prerequisite for properly organizing assistance following such disasters is the availability of reliable and fast data pertaining to the nature of operational or damaged networks resulting from such events. Mapping these infrastructures and their respective statuses is a major priority for first responders and experts from international organizations immediately following a disaster. A mapping request is normally made upon activation of the International Charter on Major Disasters [1].

High-accuracy resolutions generate a cost, since the increased volume of information to be processed and the level of detail produce noise. In fact, high detail levels disrupt the road detection process as the presence of artifacts and visible objects such as cars, shrubs and buildings increase the level of noise. This is problematic as it increases the false alarm rate, and affects the performance of road detection and extraction algorithms. While the techniques used to solve this problem differ from one context to the next, they all nonetheless strive to increase their respective precisions in order to ensure better extraction. Multiresolution analysis (MRA) approaches operate with filters to detect low-resolution roads and verify the assumptions made on the preceding resolution. Smoothing is underpinned by reducing both intrinsic noise (from sensors) and extrinsic noise (artefacts creating occlusion zones that affect extraction). These detection methods usually produce a few false detections, but their results are often limited by the variability of major roads, which is easily identifiable by the human eye. There is a large body of literature on automatic road detection combined with multi-resolution approaches. Long and Zhao [2] address the automatic detection of roads on optical image using multiscale pretreatment steps. The urban environment presents more difficulties compared to rural areas, due to the high variability of gray shades on certain sections of the road. So, Mnih and Hinton [3] propose a new learning approach to large scale: (1) generation of synthetic labels road / non-road...
(2) many features and many marking predictions are driven by a neural network (3) a post-treatment improves the predictions of the neural network. Shi et al., [4] present a new approach for urban main-road centreline extraction from optical remotely sensed images integrating spectral-spatial information for the classification. It’s sufficiently flexible to simultaneously detect curved and rectilinear structures. Miao et al., [5] integrated shape and spectral information to extract candidate road segments from binary map obtained from the edge filtering segmentation. Precise extraction of the central axis of the road segments by the adaptive multivariate regression technique is used to avoid coarse road segments that keep the central axis smooth. Li et al. [6] use a road map before disaster to get a picture of road areas (buffer zones). Next, the authors use segmentation and classification for the detection of damage on the roads. To reduce the spectral confusion between roads and debris, they incorporate a texture measure.

The performance and expected outcomes are closely related to the techniques employed for the road model adopted and to the associated methodological approach. Some road extraction methods are largely based on correlation profiles or initialization. However, these methods are highly sensitive to noise and environmental changes or the nature of the road (intrinsic and extrinsic variability of the road). Other big extraction families are based on a multi-resolution approach (hierarchical network at different scales). Everything is focused on the intrinsic characteristics of the road (contrast neighborhood, parallel edges, road width, low curvature, etc.) to allow the introduction of restrictions guiding all detection, thus providing robustness. Herold et al. [7] explore relationships between remotely sensed parameters (i.e., spectral reflectance) and road condition parameters (Pavement Condition Index) to compare the reference library of a true asphalt with undamaged structures without fissures. In the same vein, Mei et al. [8] analyze different laboratory data, spectral signatures of asphalt and specific wavelength to improve the differentiation and correlation with multispectral and hyperspectral images in order to extract and classify roads.

Today, information provided by the image is manually extracted by photo interpreters. Disaster management is complex. Automation or assistance to photo-interpreters in mapping the road condition is invaluable. As well, the color of the road provides a wealth of information which allows a distinction to be made between the road surface and vegetation. It’s better to use this information in order to discriminate some features in the image [9]. The main advantages of the SA algorithm is that it can be applied to multi-spectral images with any number of bands, and the choice of the reference pixel, which will be able to extract either tarred roads or dirt tracks.

Firstly, we adapted our MRA to the multispectral images as opposed to panchromatic images, using a Gaussian filter and the bilinear interpolation technique. Secondly, we went on to select a reference pixel and evaluate the angular distances for each pixel of the image using the spectral angle formula. A single-band image was generated. On this single-band distance image, a gradient filter in which each pixel is catheterized by its direction and intensity was applied. The road direction is obtained by applying the properties of the scalar product to the vectors. Thirdly, a representation level was selected for the versions of the generated images to enable detection of route segments, which were then projected toward the lower levels of the pyramid. These road segments were then combined on the multispectral image. The experiment ended with an evaluation of the performance of the road detection algorithm at different levels of the pyramid. We contribute in multiple ways to adapting the pyramidal approach to color images and to the modification of the interpolation technique. Our approach also helped to improve road traceability by reducing false alarms, and thus identifying the optimum road detection scale.

Our paper is organized as follows: Section 1 is the introduction. Section 2 presents methodology and the flowchart used (Fig. 1). Section 3 shows that the approach provides results whose performance is assessed according to a reference image and in comparison with existing methods. We conclude the paper with the summary in Section 4.

II. METHODOLOGY

A. Study area and data
All the results are obtained with Orfeo Toolbox (OTB) \(^1\). OTB is an open-source C++ library used for processing remote sensing images. The algorithm is very fast. An Intel Core 2, CPU 2.40 GHZ processor with 6 GB of memory took only 5 seconds to run a full resolution image (1000x1000). This calculation time enables great visualization and image analysis by photo interpreters. The urban and rural images present respectively 1000 x 1000, 4 spectral bands-R, G, B, IR with 0.6 m spatial resolution. The disaster areas (Haiti) is a VHSR Quickbird optical image (1000 x 1000, 4 spectral bands-R, G, B, IR) obtained on January 14, 2010 following the

\(^1\)The ORFEO toolbox software guide, URL: http://otb.cnes.fr, Version 4.0, 28 February 2014
Earthquake. The data consists of four multispectral bands with 2.4 m resolution and a panchromatic band with 0.6 m resolution. The multispectral and panchromatic images were fused to produce pan-sharpened multispectral images with a pixel size of 0.5 m. This applies for all the test images derived from the DigitalGlobal satellite, and were adopted as the standard for products from the CNES (National Center for Space Studies-France) constellation-pleiades.

These images used for testing allow the study an increasing levels of complexity in terms of road density and interconnections. With regard to rural roads, they were homogeneous and of constant width on the central axis of roads. The reference road map is hand-drawn [10] [4] and takes into account only the main axes of the road network. The second image shows a suburban area. This level exhibits higher complexity, where the presence of different widths and road textures is observable. Again, the reference includes only the main roads with similar width. The last image shows a dense urban area, which is part of downtown Port-au-Prince (Haiti), having areas defined by a road grid. The geographical location of the Earthquake area is in downtown of Port-au-Prince.

B. Methods

The aim of our research is to provide assistance to the photointerpreters when inputting the road network, as well as to optimize the extraction of the network. Our proposed road extraction method is modular. Our overall pattern is divided into three blocks (Fig. 1): (1) Generation of the scale on the horizontal axis space; (2) Road extraction and a selective projection recursive by a metric evaluation criterion; (3) Decimation of the image dimensions by a factor of 2 (vertical axis).

1) Generation of Scale-Space: The appearance of roads in a remote sensing image depends on many parameters such as spectral sensibility and the sensor resolution. Pyramidal approaches are mostly performed jointly with semi-automatic road extraction methods due to the difficulty in selecting a proper scale level for an object recognition application. The main problem encountered in these approaches is the scale invariance of image descriptors [11]. The various scales that are thus generated will be used to detect regions of interest in order to obtain points that are invariant to transformation functions. Adapting this approach (Algorithm 1) allows the multiscale decomposition of SIFT (scale invariant features transform) generated by the successive filtering of the algorithm of David Lowe [12].

The method uses a concept of representation with multiple trapezoidal pyramids based on the first step of the SIFT algorithm, which allows the processing of multi-spectral images at various resolution levels and scales. The multi-spectral image must be decomposed into individual bands for discrete filtering and then the whole image is recomposed.

- Decomposition: with a multi-spectral image (having three or four spectral bands), and a single processing, the intensity of the values of each band must be scaled and processed separately and individually;
- Recursive filter [Module 1]: in this representation of the pyramid, the scale parameter is initialized. Here, the size of the Gaussian kernel is very important. A large size will have a significant effect on the contrast, which allows the borders of a road to be discriminated from its immediate environment. The coarse resolution does not contain the details (in the spatial sense) as we show the image at full resolution. Conversely with a small Gaussian kernel size, we avoid the scattering effect of the low-pass filter of objects of interest. However, the goal of our approach is to detect and extract the road image in a noisy environment. We must find a good compromise in terms of the kernel size to use, depending on the characteristics of a given image. To overcome the difficult choice respecting the size of the Gaussian kernel, we proceed by variable size.

The processing is performed on three dimensions, namely, the Cartesian coordinates \((x, y)\) and the scale factor \(\sigma\). Scale images are generated by the application of a Gaussian filter whose standard deviation is the scale factor. The images are distinguished by factors of scales separated by a fixed coefficient \((k)\), which is calculated to discretize the space of scales following a geometrical progression \((\sigma, k\sigma, k^2\sigma, \ldots, 2\sigma)\).
This pyramid is generated by emphasizing the physical elements present in the original image by excluding less relevant information (noise) of a level (octave). Furthermore, to maintain the invariance of scales at different octaves, the same filter kernel is used in each of the pyramid decomposition levels. The filtering process is applied recursively until the last image of the same octave level is generated. Then, every two pixels in the image are sampled allowing the Nyquist-Shannon theorem to be respected, and enough data to be retained in order to restore the original information. The sampling is carried out by bilinear interpolation to assign a weighted intensity by the four neighboring pixels of the lower level of the pyramid.

The base of our pyramid is derived from the original image, and the different scale levels (octaves) are obtained by dividing the resolution of the image of the preceding level by 2, which effectively doubles the scale factor (2).

Among other things, this basic technique allows the reduction of visual distortions that can appear after the image is resized. The use of Gaussian kernel is to considerably improve the computation time and also the best way to approximate an impulse response (kernel) by having few values is to apply the Moivre-Laplace theorem (Eq.1).

\[
\binom{n}{k} p^k q^{n-k} = \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}} \quad (1)
\]

where \( p \) and \( q \) are probability parameters. \( p \) is the event probability and \( q \) its complement;

\[
\binom{n}{k} = \frac{n!}{(n-k)!k!} = \binom{n}{n-k} \quad (2)
\]

The first expression of the equation represents the binomial coefficients, which indicate the probability that event \( k \) occurs in \( n \) tests. These factors provide the best approximation of the Gaussian filter needed to respect the constraints of finite dimensions and discrete values. The kernel size is then specified according to the first normalized coefficients of the binomial expansion; for example, the coefficients for a binomial filter of width \( n = 3 \) is \( \frac{1}{2}(1,2,1) \).

- **Recomposition:** the three bands are filtered individually and merged to reform a multi-spectral picture.

### Algorithm 1 Generating images of the pyramid

1. Input: Multispectral images to \( M_i \) bands
2. Output: Series of multi-spectral images filtered
3. Initialization : \( I_n = I_0 \)
4. Adjust intensity and extraction of \( M_i \) bands
5. Gaussian filtering on each individual bands on the two-dimensional \((X,Y) : J_n = F_0(I_n)\) recursive filtering loop : \( \sigma_i = k\sigma_{i-1} \)
6. Recomposition of multi-spectral image:
   - Add \( I_n \) images to the radial list
7. Subsample level \((n)\) by a factor 2 in each dimension:
   - \( I_{n+1} = E(J_n) \)

Weighted sum of 4 pixels of the higher level with
the binomial coefficients: \( b_{n,k} = \frac{n^l}{(n-k)!k!} \)
8. Next high level : \( I_n = I_{n-1} \) and return to step 5.

Figure 2. Gaussian pyramid: 4 gradients and 3 octaves

2) Road Network Extraction and Evaluation: The initial algorithm [13] is for the alignment of the road segments by a line detector with gradient constraints. It is applied to panchromatic images to solve the problem of discontinuity between the road segments. Its method is not adapted to multispectral images. Christophe and Ingлада [9] improved the algorithm [13] for multispectral images by using the SA technique. The SA algorithm allows the transformation of the intensities of each of the image bands scalar values. This transformation considers all spectral information to generate a monochrome image compatible with linear structure detection applications based on the calculation of the Gaussian gradient filter. The second block consists in the application of the road extraction algorithm on each reconstituted image. The results of this detection are projected toward the original image for comparison purposes. The second block consists of three sequential modules: (1) application of the algorithm for road
extraction according to the spectral angle, (2) selective projection, (3) quantitative assessment with metrics to assess the degree of detectability and the false alarm rate.

**Road Network Extraction:** in [9], the authors present an interesting road extraction technique, which uses spectral information to distinguish roads from other land uses. The SA principle, which takes into account the color in an image, is indicative of capital and beneficial information for the discrimination of land (roads, vegetation, buildings, soil, water, etc.) on an image. The SA is a common, powerful spectral recognition technique, in which an unknown spectral characteristic is compared to a reference spectral characteristic. This information, which is based mainly on color, is generally composed of three or four spectral bands. The unknown and the reference are treated as vectors, with their dimensionality equal to the number of bands. The representation of the bands in a multidimensional space is used to measure the angular deviation between the measured color vector of the road \( p \) and a reference vector \( r \) sampled by the analyst, and representing a road sample (Fig. 3). This algorithm is simple to implement, and is based on search for similarity between the reference pixel and the other pixels of the multi-spectral image. One of the advantages of using of the spectral angle is that many algorithms use a scalar representation. However, one way to effectively use the spectral information in scalar data is to carry out a conversion using the SA technique [9]. The algorithm runs with an image of the spectral angle (distance according to the spectral angle) and this algorithm does not depend on the number of spectral bands (Fig. 3). In addition, choosing the reference pixels based on the color allows the detection of tarred roads and road tracks. For each image, only one reference pixel is specified. On a visible road segment, we identify a pixel which the spectral band information is given by the Cartesian coordinates \((x, y)\). For a multi-spectral image with \( M \) spectral bands, the SA between the reference pixel \( r \) and the current pixel \( p \) is defined by:

\[
SA = \cos^{-1}\left(\frac{\sum_{b=1}^{M} r(b) p(b)}{\sqrt{\sum_{b=1}^{M} r(b)^2} \sqrt{\sum_{b=1}^{M} p(b)^2}}\right) \tag{3}
\]

The assignment of a pixel to a road pixel is based mainly on the SA measurement, i.e. the angle between the vector and the reference spectrum of each image vector. The smaller the angle, the greater the similarity between the evaluated pixel and the reference. This measure is less sensitive to changes in brightness and does not affect the direction of the vector, but does its amplitude. By simply clicking a pixel-area road (a highly asphalted area) for example, we choose a pixel road. It must be recognized that a bad choice of reference pixels can lead to false matches or false alarms. The resulting image contains all roads in a darker color. We apply to the image of this spectral angle a line detection algorithm that uses a gradient constraint in order to determine the pixels delineating the road. A removal of non-maximum scalar values and the highest scalar value is obtained between \( a_1 \) and \( b_1 \) (Fig. 4). This approach is described in [13]. The

**Algorithm 2 Road extraction algorithm**

1: Choice of spectral information (reference pixel)
2: Evaluation of the angular distances for each image pixel
3: Building of distance image with SA to compute a greyscale image from the multi-spectral image
4: Application of the Gaussian gradient filter for obtaining the direction and intensity of each pixel:
\[
\sigma = \alpha (1.2/\text{resolution} + 1)
\]
5: Removal of pixel which is not maximum on the direction perpendicular to the road direction
6: Detection of linear structures
7: Removal of irregular linear structures
8: Separation of high angles segments outside \([-\frac{\pi}{8}; \frac{\pi}{8}]\)
9: Extension between two segments in the presence of occlusion

![Figure 3](image3.png)  
**Figure 3.** [Module 2(a)]: Visualization of the spectral angle for one pixel of a three-band image between a spectral vector (current) \( p \) and a reference vector, \( r \).

![Figure 4](image4.png)  
**Figure 4.** [Module 2(b)], Gradient representation and roads: (a) gradient direction toward the roads; (b) computation of the direction and scalar values for each pixel [14]
Refinement results and elimination of irregular structures are realized using an alignment technique, softening and bends connecting segments extending between two segments by the presence of occlusion (Algorithm 2). **Projection Looping:** A simple visual evaluation indicates a gradual decrease in the number of segments of roads detected in each octave, and for each scale level. The principle of selective projections matches the information from one higher octave to improve the quality of the result. For this, an operation is performed to loop back the information extracted from a higher level to a lower level road, up to the original image. The image projection of an upper level to a lower level is performed by a bilinear interpolation (image resizing). Then, a mask is generated from the image of the lowest level for mapping using binary values. This will keep the road pixels and remove those that are not. This operation is performed repeatedly. The selection of the optimal scale for the projection of the image at the lowest level can be done via a visual evaluation or based on different criteria. Rather than choosing the image at the same scale position at each iteration, we proceed by a selection based on an objective criterion, namely, the selective projection based on a recovery rate (R). Depending on our goals, we want a higher level of likelihood of the road segments to have a best match. The goal is to project the image best correlation with the reference trace [10] [4] which is only logical given that the amount of data in the upper level images is reduced. A quantitative assessment is performed to measure the degree of precision (P), recovery (R) and quality (Q) of the results obtained.

**Evaluation Process:** An evaluation of the performance of the algorithms used to appreciate the level of detectability of roads may be relevant for a better understanding of the weaknesses of the algorithms. The evaluation technique involves comparing a detected image and a reference image, with 2 and 1 pixels of the road width, respectively. To facilitate this evaluation operation, the reference image road segments width was defined along the roads central axis on a constant width of a single pixel. We also know that the widths to quantify images are larger than that of the reference (road central axis). To facilitate this evaluation operation which requires two images (detecting / reference image), the reference image segment was defined along the central axis of roads on a constant width of a single pixel. The width of the reference segments ought to be increased using a structuring element having similar dimensions as the segments of the detected image, which in turn prevents significant overestimations of the reference segments. We opted for a minimum width of 2 pixels wide (L = 2 pixels) for the image to be quantified. Significant spectral homogeneity is observed along the center line of the road. With the mathematical morphology technique, we applied a structuring element, and then proceeded to the dilatation of the road segments of the detected image. Then we created a binary mask of the two respective images to standardize their greyscale intensities. Thus, we got the road pixels and no-road pixels. Fig. 5 illustrates an example of the technique used for the quantitative evaluation of a detection image. The evaluation process is based partly on the methodology presented in [15]. The approach is to compare the detection results compared to a reference trace by following the traffic lights paradigm, namely: green plots indicate a good match, the red strokes are associated with incorrect correspondence and orange plots require verification for possible omissions. These quantities of generated pixels are used to calculate the precision rate (P), the recovery rate (R) and the quality rate (Q). To quantify the results, we define the parameters used by [16]. Following the concept of confusion matrix, the evaluation matrices for precision assessment of the road surface detection accuracy assessment can be defined at the pixel level as follow: (1) Recovery rate = TP /(TP + FN), (2) Precision rate = FP /(TP + FP), and (3) Quality = TP /(TP +
FP + FN). In the above equation, TP (true positive) is the number of road surface pixels correctly identified, FN (false negative) is the number of road surface pixels identified as other objects, and FP (false positive) is the number of non-road pixels identified as road surfaces.

3) Decimation of the image: We do subsampling of the image by factor 2 following the vertical axis.

III. EXPERIMENTAL RESULTS AND INTERPRETATION

The results presented were obtained using the MRA analysis method, combined with the algorithm spectral angle. These results are compared with approaches that have not been processed by the MRA approach and other studies by another pyramid method (Beamlet transform). The tests show the advantages of using such an approach in both a noiseless and a noisy environment. These results can be used to help photo interpreters to produce a map of operational road communications during a major crisis.

One of the objectives of this study is thus to determine whether the MRA approach can significantly increase the detectability rate. Unlike other study sites, the Haiti image underwent a pre-processing (contrast enhancement) to increase the signal / noise ratio in order to improve the precision of extraction methods and ultimately the quality of the final results. An additional filtering had previously been applied to areas of interest in order to reduce disparities between road segments that interfered with the extraction process. For a road extraction algorithm, a single reference pixel was specified for each image studied (rural, suburban and Haiti), with the latter corresponding to the spectral information pertaining to a road segment visible on the image. The reference pixel coordinates were obtained using an OTB graphic interface software application. Among the parameters used to specify the road extraction algorithm, only two could influence the resulting images: the reference pixel and the distance threshold parameter. These two parameters are used to connect two road segments (road with ends that are close according to a Euclidian distance criterion. Binding values were adjusted according to the resolution of the original image. Most of the default settings of the algorithm are used for all tests in order to limit their influence on the results. The following table (Table I) shows the parameters of the extraction algorithm and their default value: For the results assessment, only main roads were considered. For the pyramid process, the values of the Gaussian kernels used respectively for rural, suburban and Haiti images are: \( \sqrt{2} \), \( \frac{1}{2} \), and \( \frac{\sqrt{2}}{2} \). Four resolution levels separated by a distance of one octave (O) are used, with four scales (S) used to describe each scene.

This choice of pyramidal decomposition was determined by the degree of observable information loss: after a number of levels, the structures were practically no longer detectable and the detection results are acceptable. Moreover, the generation on four levels of the pyramid provides enough samples to represent the evolution of the results using a graph. The evaluation process involves comparing the detection results to a reference trace obtained manually by following the paradigm of traffic lights: green plots indicate a correct match, red traces are associated with an incorrect correspondence with the reference, and orange plots require verification for possible omissions. The goal is to assess the degree of detectability of our developed approach in a noisy environment. This section presents only the best results of MRA, which we illustrate in Table II and III.

A. Qualitative and Quantitative Evaluation

General observations: Fig. 6 presents an overview of the results generated by the Gaussian pyramid on the 4 octaves, each with 4 scales. There is a gradual loss of information and a general reduction in noise as we go up the pyramid. In the first two octaves, we can observe a more complete detection of the segments at the field edges. Furthermore, a false positive is encountered in a field due to a sudden change in contrast. Continuous graphs are accurate in some areas of the image. A greater rate of false detections and isolated segments is observed for the suburban and rural images due to non-complete reference road, as well as the presence of occlusions (presence of shade). The curvature in the center of the rural image is partially detected. The opposite is observed in the disaster image, where plots are continuous, with a high level of confidence in certain parts of the image. In addition, the connections are more accurate and uniform in the latter. From the third octave, all test images show a significant noise reduction, but at

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>Acceptance threshold of detected roads</td>
<td>1</td>
</tr>
<tr>
<td>Amplitude</td>
<td>Sensitivity of the vectorization stage</td>
<td>( 10^{-4} )</td>
</tr>
<tr>
<td>Tolerance</td>
<td>Sensitivity of the simplification stage</td>
<td>1</td>
</tr>
<tr>
<td>Angle</td>
<td>Limit of the segment curvature angle</td>
<td>( \frac{\pi}{2} )</td>
</tr>
<tr>
<td>Average distance 1</td>
<td>Distance criterion for false #1 positive elimination operation</td>
<td>1</td>
</tr>
<tr>
<td>Average distance 2</td>
<td>Distance criterion for false #2 positive elimination operation</td>
<td>10</td>
</tr>
<tr>
<td>Distance</td>
<td>Euclidean distance limit for linking two segments</td>
<td>25</td>
</tr>
</tbody>
</table>
Figure 6. Results of the Gaussian pyramid on the three study areas: \((r_1)-(r_3) =\) Hand-drawn reference road map of rural, suburban and Haiti images; \((a_0)-(a_3) =\) Road extracted from rural image with 4 levels (octave) and the best scale \([a_2]\); \((b_0)-(b_3) =\) Road extracted from suburban image with 4 levels (octave) and the best scale \([b_1]\); \((c_0)-(c_3) =\) Road extracted from Haiti image with 4 levels (octave) and the best scale \([c_1]\).

a cost of loss of contour precision. At the fourth octave, a significant loss of information is observed as compared to the previous octave. The pyramid technique allows the extraction of more relevant information because it has reached the limit of useful resolution.

**First test (rural image):** especially the high curvature

Figure 7. Progress of metrics for rural image

is detected. In the northeast of the image there is an omission and a large number of false alarms observed in the residential areas as some roads are not taken into account by the reference image. It is observed that most of the main roads associated with our reference were detected. In the center of the rural image, especially the high curvature is detected. In the north-east of the
image, there is an omission, and a large number of false alarms is observed in residential areas because some roads are not taken into account by the reference image. First observation, the first two octaves manifest a high reduction in the number of false positives, a quantity which tends toward the number of true positives. In fact, it is at the octave and scale where the highest quality rate was obtained that the quantity of true positives begins to diminish. In this specific case, we note that a compromise is struck between the rate of precision and recovery at the first step of the second octave (illustrated by the orange band in Figure 7). While the precision and quality increase from the first octave and then fall back to the first scale of the second octave. This result in the loss of some road segments and a reduction in the trust level. A simple glans shows a reduction of the quantity of false positives (represented by red pixels) when these results are compared with those of the unprocessed image. This noise reduction translates into a roughly 10% increase in quality, thus implying a 16% increase in precision at the cost of a 5% loss in recovery.

Second test (suburban image): in the context of the heterogeneity with varying widths of roads and the presence of structures in borders are factors which complicate the detection process. Notably, occlusions produced by shades of trees on the main road segments cause many discontinuities, as does the incomplete layout of the reference in some areas. The detection result obtained without the application of the technique AMR is shown in the figure 10. In this case, it can be observed that the false detections correspond to approximately 78% of the segments extracted. The variation of textures and widths of roads and the presence of artifacts along the road are factors that hinder the detection process. These occlusions produced by tree shades or the presence of cars create discontinuities in the road alignment segments. This is true with false detection less localized in some place, as the case for rural images. As the reference road based on main roads only, appears incomplete in some places, as compared to roads in the original image. As in the previous case, the number of false positives is significantly reduced from the first octave. Little change was observed at different scales. A relative decrease in false alarms was observed in residential areas. A significant reduction explains a high recovery rate for the urban image (Fig.8). As was the case for the first series of images, the number of false positives is significantly reduced starting from the first octave. However, few variations were observed on the different scales of the latter. The best compromise is found on the second level of the the first octave (illustrated by the orange band in Fig.8). The results of the second octave are only different through a slight increase in precision, with the average quality deviation at only 0.75%.

Third test (haiti image): the result is differentiated by the high quantity of false detections generated by the extraction algorithm, and this, despite the rise in interest zones. As compared to the original image, a high recovery rate (96%) and better overall precision (21%) was observed from the reference grid. The road segments that were correctly detected are not clearly distinguished from the others and many structures were interpreted as possible roads. Very similar spectral information between the different entities hinder overall extraction: false detections appear in virtually all parts of the image except in a few grids of greenery. It was observed that the quality rate was achieved starting with the third scale in the second octave. A significant demarcation appears starting from the following scale, where the recovery rate drops by 24% for an 8% gain in precision, this not being enough to reach a higher quality. The best quality is obtained when the TP trend is also descending, which in this case, occurs starting from the third scale of the second octave. A trace of the corresponding result is presented in the figure 9. Here, a reduction in the number of false detections can be visually observed, helping us to easily distinguish the segments detected. Quantitatively, we observe an increase in precision levels and quality of about 30% for a 5% loss in recovery. The best compromise rate is achieved for the second level of the first octave. An interesting occurrence in this example is that several items corresponding to false items have similar characteristics: Relatively short segments are sometimes grouped into aggregates. This suggests the possibility of additional thresholding to eliminate these small groups.

B. Interpretation

The results in Fig. 10 and Fig. 6 of the three study sites are mixed for various reasons. Indeed, the three contexts have different characteristics, including the density of structures, coating and variability of roads, and the great similarity between roads and their immediate environment to the presence of artifacts. The study of the differences observed can be used to improve the quality of results obtained by combining the MRA method with the SA algorithm. The representation of a road element on a particular frame depends on its spectral resolution. The size and shape of the structures of interest must be ample enough to ensure that the structures are not completely eliminated during decimation. To recap, the algorithm evaluates the spectral angle from the eight neighbors pixels of a central one (an operation that extracts linear
structures as fine as a pixel). The effects of the size of the filter kernel must be adapted to the different shapes of structures present in the image. Because there is a considerable loss of information between the different levels of the pyramid. The search for an optimal size or range of acceptable kernel values to use should be a path to explore. A great similarity can be observed between the various radiometric bands of each of the test images. The spreading degree distributions of the different bands can help increase the contrast of each band before the application of our approach, especially in a noisy environment.

C. Comparisons With Existing Approaches

In this section, the proposed approach is compared with two existing road extraction methods from the literature. These two methods are introduced by Christophe and Inglada [9], Sghaier and Lepage [17]. Like the proposed approach, these two were programmed and run in OTB. The comparison with the approach is carried

Table II

<table>
<thead>
<tr>
<th>QuickBird Sensors</th>
<th>Coating</th>
<th>Noise</th>
<th>O</th>
<th>S</th>
<th>(\sigma_{initial})</th>
<th>R (%)</th>
<th>Q (%)</th>
<th>P (%)</th>
<th>(\Delta R) (%)</th>
<th>(\Delta Q) (%)</th>
<th>(\Delta P) (%)</th>
<th>TP (pix)</th>
<th>FP (pix)</th>
<th>FN (pix)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>Dirt path</td>
<td>Low</td>
<td>2</td>
<td>1</td>
<td>0.7071</td>
<td>69.42</td>
<td>27.97</td>
<td>31.90</td>
<td>-4.74</td>
<td>-32.10</td>
<td>-37.87</td>
<td>47.75</td>
<td>10648</td>
<td>11653</td>
</tr>
<tr>
<td>Suburban</td>
<td>Asphalt</td>
<td>Medium</td>
<td>1</td>
<td>2</td>
<td>0.5</td>
<td>43.01</td>
<td>16.90</td>
<td>21.78</td>
<td>-4.95</td>
<td>-6.93</td>
<td>-5.09</td>
<td>30.93</td>
<td>7072</td>
<td>15789</td>
</tr>
<tr>
<td>Haiti-Damaged</td>
<td>Asphalt</td>
<td>High</td>
<td>1</td>
<td>2</td>
<td>0.7071</td>
<td>96.19</td>
<td>20.76</td>
<td>20.94</td>
<td>-4.65</td>
<td>-4.15</td>
<td>-1.37</td>
<td>51.58</td>
<td>351448</td>
<td>329873</td>
</tr>
</tbody>
</table>

Table III

<table>
<thead>
<tr>
<th>QuickBird</th>
<th>Coating</th>
<th>Noise</th>
<th>O</th>
<th>S</th>
<th>(\sigma_{initial})</th>
<th>R (%)</th>
<th>(\Delta R) (%)</th>
<th>(\Delta Q) (%)</th>
<th>(\Delta P) (%)</th>
<th>TP (pix)</th>
<th>FP (pix)</th>
<th>FN (pix)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>Dirt path</td>
<td>Low</td>
<td>2</td>
<td>1</td>
<td>0.7071</td>
<td>64.57</td>
<td>-4.47</td>
<td>37.75</td>
<td>-4.15</td>
<td>10648</td>
<td>11509</td>
<td>3816</td>
</tr>
<tr>
<td>Suburban</td>
<td>Asphalt</td>
<td>Medium</td>
<td>1</td>
<td>2</td>
<td>0.5</td>
<td>38.06</td>
<td>-4.95</td>
<td>30.93</td>
<td>-5.09</td>
<td>7072</td>
<td>15789</td>
<td>329873</td>
</tr>
<tr>
<td>Haiti</td>
<td>Asphalt</td>
<td>High</td>
<td>1</td>
<td>2</td>
<td>0.7071</td>
<td>91.53</td>
<td>-4.65</td>
<td>51.58</td>
<td>-1.37</td>
<td>351448</td>
<td>32513</td>
<td>32513</td>
</tr>
</tbody>
</table>

Figure 10. Qualitative evaluation and comparison of approaches SA and (SA + MRA) on Quickbird images: (a1)-(a3) Input images of rural, suburban and Haiti; (b1)-(b3) Hand-drawn reference road map of rural, suburban and Haiti; (c1)-(c3) Results of Emmanuel and Inglada’s approach [9] on rural, suburban and Haiti images; (d1)-(d3) Results of the proposed approach on rural, suburban and Haiti images; (e1)-(e3) Road network and original image superposition of rural, suburban and Haiti. Good matching is shown in green (TP), incorrect matching in red (FP), and matching requiring verification is shown in orange (FN).
Table IV

<table>
<thead>
<tr>
<th>Approaches</th>
<th>Christophe</th>
<th>Sghaier</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>QuickBird Sensors</td>
<td>R (%)</td>
<td>Q (%)</td>
<td>P (%)</td>
</tr>
<tr>
<td>Study Area 1-Rural</td>
<td>69.42</td>
<td>27.97</td>
<td>31.90</td>
</tr>
<tr>
<td>Study Area 2- Suburban</td>
<td>43.01</td>
<td>16.90</td>
<td>21.78</td>
</tr>
<tr>
<td>Study Area 3-Haiti-Damaged</td>
<td>96.19</td>
<td>20.76</td>
<td>20.94</td>
</tr>
</tbody>
</table>

Figure 11. Comparison with existing semi-automatic approaches: (a1)-(a3) Input images of rural and two damaged areas of Haiti; (b1)-(b3) Results produced by Sghaier and Lepage [17]; (c1)-(c3) Results shows by proposed approach. True positive are shown in green.

out on three types of images: (1) rural area with a dirt path coating and low noise, (2) suburban area moderately noisy and covered in asphalt, (3) damaged urban areas (Haiti) with an asphalt coating and very noisy. Fig. 10 and Fig. 11 give the results of comparisons of different road extraction approaches. To evaluate these approaches quantitatively, the recovery rate, the precision rate and the quality rate of each approach are computed. The respective results are presented in Table IV, where it can be seen that although both algorithms (proposed approach and Sghaier algorithm [17]) provide essentially the same performance in terms of the recovery and quality criteria, Sghaier algorithm obtained good results in term of recovery rate with two sub-images of the Haiti original image. Compared to our approach, there are no large differences with evaluation rates. The quality rate is low because of excessive noise, material change, shadow and presence of similarity between the road and its environment, it is difficult to set a suitable value of kernel filter. This leads to low quality in our approach for this complex Haitian image. Our proposed approach has performed more in the rural and suburban images with the precision rate. The comparison with the Christophe and Inglada approach [9] is on three different study sites presenting a gradual complexity in terms of noise and coating. As can be seen from Table IV with precision criteria, our approach seems to be the best. With the proposed approach, we got in the case of Haiti, gains of over 30%, and a modest increase for the rest of the images, despite the presence of false alarms and small segments of isolated roads in of the image locations.

IV. CONCLUSIONS AND OUTLOOK

The application of our approach increases the recovery rate to about 90% with a 30% gain in road detection and more than 20% in terms of quality for the Haiti image. Our multi-scale joint spectral angle representation unit for the processing and detection of roads on satellite images is semi-automatic. It is fast, and takes less than 5 seconds for a full resolution image. Color or spectral information is used as the sole input parameter with the SA for road detection. The approach is able to consider only one spectral information, what does not enable to take into account its variabilities. However, with our approach, we are able to distinguish the variabilities geometrically at the end of the road. The algorithm then uses the gradient for direction and vectors to refine road segments. The resulting image is used as an input parameter for the MRA. The MRA generates simplified versions (with reduced resolutions) of the original image by applying a Gaussian filter for each level of the pyramid. Our approach deletes the noise present iteratively and improves the degree of road detectability in the images.

The experimental results are very encouraging and show that MRA can significantly aid in the process of segmentation of objects under certain conditions. The application of our approach increases the recovery rate to about 90% with a 30% gain in road detection and more than 20% in terms of quality for the Haiti image. The results are not perfect, but could be improved by
exploring techniques for tracking algorithms, such as active contours, Kalman filter, etc. We could consider adding other types of information (texture or contextual), which could bring about added value and significantly reduce the rate of false alarms.

REFERENCES


Idrisss Coulibaly received the postgraduate degree (D.E.A.) in physical modeling, remote sensing (2001), M.Sc.A. degree in electronics and informatics engineering from Félix Houphouët Boigny University (UFBH), Abidjan, Cote d’Ivoire (2000) and M.ing. degree in automated manufacturing engineering (2013) from École de Technologie Supérieure (ETS), Université du Québec, Montréal, Canada. He is currently pursuing the Ph.D. degree in images processing and machine learning in Laboratory imagery, vision and artificial intelligence (LIVIA) from ÉTS. He is the IEEE-ETS graduate students branch president and member of remote sensing association of Québec, (AQT) Canada. From 1997-2000, he was consultant engineering in cartography and remote sensing at SODEFOR (state society development, management of heritage protection forest, Cote d’Ivoire). His current researches are centered around images satellites processing and GIS, especially in pattern recognition, machine learning and change detection.

Nicolas Spiric M.ing. received the Engineer degree in computer engineering from Sherbrooke University-Canada in 2007 and the Master degree in automated manufacturing engineering in 2011 from École de Technologie Supérieure, Université du Québec, Canada. He is currently working as an independent software engineer in the field of data analysis, images processing and component development.

Richard Lepage (M’97) received the degree in electrical engineering, the M.S. degree in signal processing and the Ph.D. degree in artificial neural networks applied to computer vision, all from Laval University, Quebec, P.Q., Canada, in 1974, 1982, and 1993 respectively. He is currently a Professor in the Department of Automated Production, École de Technologie Supérieure, Université du Québec, Montréal. From 1976 to 1989, he was with the Defence Research Center, Québec, Canada. His research interests include neural networks, computer vision, 3-D vision, and the design of artificial intelligence techniques in remote sensing.

Michele St-Jacques, ing. Department of Construction Engineering, École de Technologie Supérieure, Université du Québec, Montréal, Canada. Professor since 1998, she is specialized in the field of road engineering. As engineer and physicist, her research interests include Traffic calming; Roadway and Geometric design; Traffic analyses; Road Safety; Signalization; Road Materials; Bituminous mixes (hot, warm, coloured, noiseless). Before pursuing her academic career, she worked for over 15 years for one of the top engineering design firms in the world SNC-Lavalin.