

Development of an algorithm to analyze cartographic features extraction methods in digital images

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 IOP Conf. Ser.: Earth Environ. Sci. 17 012219

(<http://iopscience.iop.org/1755-1315/17/1/012219>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 186.217.236.119

This content was downloaded on 25/02/2015 at 16:53

Please note that [terms and conditions apply](#).

DEVELOPMENT OF AN ALGORITHM TO ANALYZE CARTOGRAPHIC FEATURES EXTRACTION METHODS IN DIGITAL IMAGES

Guilherme Pina Cardim and Erivaldo Antônio da Silva

Univ. Estadual Paulista – UNESP, Presidente Prudente, Brazil

Email: gpcardim@gmail.com; silva.erivaldo@gmail.com

Abstract. The algorithm creates a buffer area around the cartographic features of interest in one of the images and compare it with the other one. During the comparison, the algorithm calculates the number of equals and different points and uses it to calculate the statistical values of the analysis. One calculated statistical value is the correctness, which shows the user the percentage of points that were correctly extracted. Another one is the completeness that shows the percentage of points that really belong to the interest feature. And the third value shows the idea of quality obtained by the extraction method, since that in order to calculate the quality the algorithm uses the correctness and completeness previously calculated. All the performed tests using this algorithm were possible to use the statistical values calculated to represent quantitatively the quality obtained by the extraction method executed. So, it is possible to say that the developed algorithm can be used to analyze extraction methods of cartographic features of interest, since that the results obtained were promising.

1. Introduction

First of all it is necessary to mention the importance of Cartography to the World. Cartography performs many studies measuring Earth spaces, calculating flooded areas and analyzing remote sensing images, for example. Talking specifically about the remote sensing area, it is known the interest of Cartography on these images since the 1970's, when the first commercial satellite was launched to the Earth's orbit [1]. From this moment on many satellites were launched to Earth's orbit aiming to obtain images of our planet.

Thus, Cartographers realized the advantages of the remote sensing images, such as its low cost and, nowadays, good spatial resolution. So they started studies using the images, mainly with the objective of extracting, or detecting, the features that interest and that are present in the image of study. These studies aim to obtain just the target, or feature of interest present in the image. Thus, many algorithms to extract the features were developed over the years.

Having obtained the results of an extraction algorithm, it is extremely important to know the quality obtained on the extraction method. However, literature has just some numbers presented [2] or adaptations of the same numbers implemented by algorithms [3], [4]. However, the adaptations implemented do not use the buffer area proposed in the metrics [2].

As mentioned, the quality of analysis of the cartographic features extraction methods is very important and necessary, but there are just a few works that intend to help the Cartographer to find a quality value to the extraction method proposed. Thus, this paper proposes an automatic way to calculate statistical values about the result of an extraction method. The automatic way was possible



through an algorithm implementation that aims to calculate the numbers presented in the literature using the buffer area differently of the adaptations found.

The paper is organized in five sections. The first one is the subject introduction. The second one shows some definitions that are necessary to understand the applied methodology, which is presented on the third section. Section number four shows the results obtained using an example for this paper. The fifth section presents the paper conclusion.

2. Theoretical Recital

This section presents some definitions necessary to understand the paper methodology, such as the statistical metrics calculated by the algorithm implemented that aims to evaluate the result of an extraction method.

2.1. Statistical Metrics

The evaluation of any automatic system is essential and to evaluate an extraction method it is always necessary a reference image [5]. The reference image is necessary as a perfect model of the result desired by the extraction method. All statistical values are calculated considering that the reference image is the result from a more accurate extraction.

Having the reference image and the resulting image of the extraction method, a comparison between both images can be made. The comparison is made through the creation of a buffer area in the reference image and overlapping the extracted image. Thereby the algorithm can count the number of pixels that are the same in both images and how many of them are different, in other words, how many pixels can be considered errors. After that, the same comparison is made, but this time creating the buffer area in the extracted image overlapping the reference image. Both comparisons described can be simplified by the scheme presented on Figure 1 (Adapted from [2]).

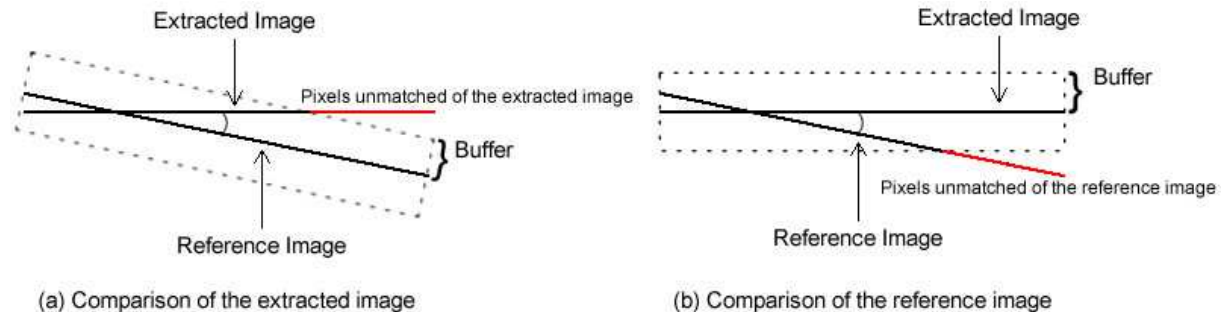


Figure 1. Comparison between the extracted and the reference images

After the comparison, the statistical values are calculated using the total number of pixels that matched or unmatched during both comparison. The values can be used by the user to evaluate the quality of the extracted method executed and they are presented in the next sections.

2.1.1. Correctness. The first value calculated by the algorithm is the *correctness* value, which is the percentage of the extracted image pixels that matched with the reference image. The *correctness* is a value between [0:1] whereas the value 1 is the desired value. The equation 1 shows how the *correctness* value is obtained by the algorithm.

$$correctness = \frac{\text{matched pixels of extracted image comparison}}{\text{total number of pixels of the extracted image}} \quad (1)$$

2.1.2. Completeness. The *completeness* value is the percentage of the reference image pixels that matched the extracted image. The *completeness* is also a value between [0:1] whereas the value 1 is the desired value. The equation 2 shows how the *completeness* value is obtained by the algorithm.

$$completeness = \frac{\text{matched pixels of reference image comparison}}{\text{total number of pixels of the reference image}} \quad (2)$$

2.1.3. Quality. The last statistical value calculated by the algorithm shows the idea of *quality* obtained on the extraction method performance. The *quality* value does not have more information of the *correctness* and *completeness* values, since the *quality* is obtained combining the first two statistical values presented. The *quality* is also a value between [0:1] whereas the value 1 is the desired value and the equation 3 shows how the *quality* value is obtained by the algorithm.

$$quality = \frac{completeness * correctness}{completeness - completeness * correctness + correctness} \quad (3)$$

3. Methodology

The methodology applied by the algorithm consists of obtaining the statistical values presented in the previous section. So, the user has to inform the algorithm of the two images necessary, the reference and the extracted images.

After obtaining both images, the algorithm developed creates the buffer area around the features that interest and that are presented in both images. The buffer area is created following the direction of the interest feature but with one pixel larger than the reference or extracted image. One buffer area is created for each image as previously described.

The next step consists of performing the comparison of the extracted image and the reference image with the buffer area. During this comparison, the algorithm counts the total of pixels that are part of the feature in both images or only in the extracted image, which are denominated matched or unmatched pixels of the extracted image comparison.

Likewise, another comparison is performed using the reference image and the extracted image with the buffer area. So, the algorithm can count the total of pixels that are part of the feature in both images or only in the reference image, which are denominated of matched or unmatched pixels of reference image comparison.

During both comparisons the algorithm creates an image containing the result of the comparison. The result image is composed by black, white and red pixels. The black ones are the background of the image. The white ones are the matched pixels in the comparisons and the red ones are the pixels that are considered errors, in other words, the red pixels are the unmatched pixels in the comparisons.

After performing the comparisons and creating the result images, the algorithm calculates the statistical values presented by the equations 1, 2 and 3. These values allow the user to analyze the extraction method statistically.

4. Results

In order to test the methodology used by the developed algorithm, a remote sensing image was used. The image is from a river named Tietê that crosses the city of São Paulo, in Brazil. The image used to test the algorithm is presented on Figure 2.



Figure 2. Original Image from Tietê River - Brazil

After having the original image, an automatic extraction method was used to extract only the features that interest and that are present in the image. In this case, the interest feature is the river present in the original image. The extraction result is presented on Figure 3.

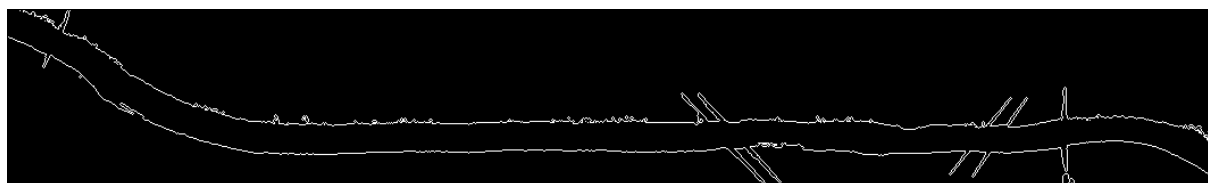


Figure 3. Extracted image

After getting the extracted image, a reference image was obtained manually, using an image edition software, to be considered the correct or ideal result of an extraction. The reference image obtained is present on Figure 4.

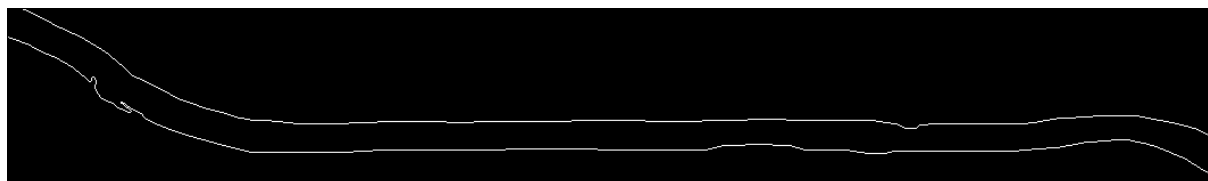


Figure 4. Reference Image

Following the methodology, on the next step the algorithm creates a buffer area around both images and after that, the comparison is performed. Figure 5 shows the resulting image from the comparison of the extracted image, while Figure 6 shows the result image from the comparison of the reference image. Both images were created by the algorithm during the comparisons and they show the matched pixels in white and pixels that are errors in red.

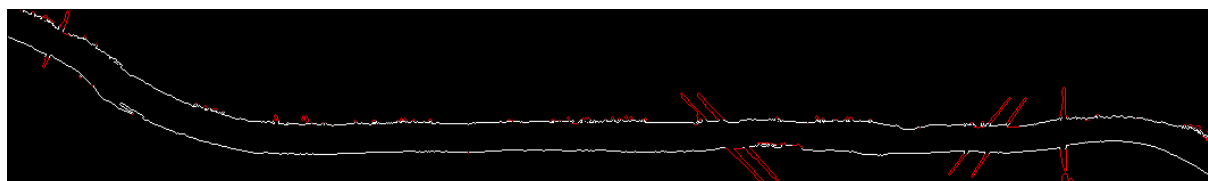


Figure 5. Comparison of the Extracted Image

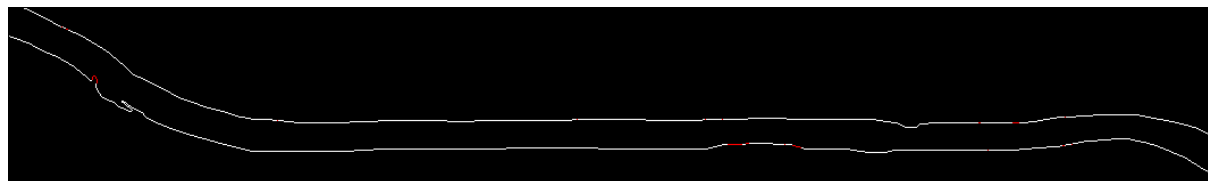


Figure 6. Comparison of the Reference Image

Figure 5 and Figure 6 show the errors of the extraction method performed, so the user can evaluate the extraction method by the positions of the found errors and what structure of the original image

could be responsible for the errors of the extraction method. After the images, the algorithm also calculates the statistical values that are related to the errors of the extraction method. The statistical values calculated in the example are shown by the Table 1.

Table 1. Statistical Values

	Values
Correctness	0.707858
Completeness	0.973501
Quality	0.694942

5. Conclusions

Trying to find an automatic way to calculate the statistical values related to the quality of an extraction method, an algorithm was developed to calculate the statistical values and create result images during the comparisons. So the algorithm was able to resolve the limitations of other works, since that this paper shows the results obtained by an algorithm that calculates the statistical values automatically. Furthermore, the algorithm is not restrict so it can be used by anyone. Moreover, the algorithm can calculate the statistical values regardless the type of feature that interests, being it a river, a road or any other desired spot.

The developed algorithm contributes by calculating the statistical values and creating the result images efficiently and quickly. Since it uses binary images to perform the comparisons, reducing the number of possibilities.

The results obtained by the developed algorithm confirm the importance of evaluating an automatic extraction method. Moreover, the algorithm shows efficiency in calculating the statistical values about the quality of the performed extraction method.

All the tests performed in this project were made with remote sensing image, which were the objective of the work. However, it is intended to make studies about the possibility to perform the same methodology and algorithm in other types of images and features.

Acknowledgements

The authors would like to thank FAPESP on the financial support given to the project by providing a research grant and enabling the purchase of the images used on the paper.

References

- [1] Hung M C and Ridd M K, Development of a supervised classifier for subpixel remote sensing 2001 *Geoscience and Remote Sensing Symposium - IGARSS '01. IEEE 2001 International* **4** 1892-4
- [2] Wiedemann C, External evaluation of road networks 2003 *ISPRS Archives* **XXXIV** 93-8
- [3] Da Silva E A and Cardim G P, Applying digital image processing to evaluate an extraction method of cartographic features in digital images 2012 *Journal of Earth Science and Engineering* **2** 241-6
- [4] Cardim G P and Da Silva E A, Análise da qualidade de processos automáticos de extração de feições cartográficas 2011 *Omnia Exatas* **4** 7-18
- [5] Wiedemann C, Heipke C, Mayer H and Hinz S, Automatic extraction and evaluation of road networks from MOMS-2P imagery 1998 *International Archives of Photogrammetry and Remote Sensing* 95-100