

Automatic building extraction in VHR images using advanced morphological operators

Sébastien Lefèvre, Jonathan Weber, David Sheeren

▶ **To cite this version:**

Sébastien Lefèvre, Jonathan Weber, David Sheeren. Automatic building extraction in VHR images using advanced morphological operators. IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (URBAN), 2007, France. pp.INSPEC #9702691, 10.1109/URS.2007.371825 . hal-00516082

HAL Id: hal-00516082
https://hal.science/hal-00516082

Submitted on 8 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Automatic Building Extraction in VHR Images Using Advanced Morphological Operators

Sébastien Lefèvre, Jonathan Weber
LSIIT, University Louis Pasteur / CNRS
Pôle API, Bd Brant, BP 10413
67412 Illkirch cedex, France
Email: lefevre@lsiit.u-strasbg.fr

David Sheeren
DYNAFOR, INRA - INPT / ENSAT
Av de l'Agrobiopôle, BP 32607, Auzeville Tolosane
31326 Castanet Tolosan cedex, France
Email: david.sheeren@ensat.fr

***Abstract*— This paper presents a new method for buildings extraction in Very High Resolution (VHR) remotely sensed images based on binary mathematical morphology (MM) operators. The proposed approach involves several advanced morphological operators among which an adaptive hit-or-miss transform with varying sizes and shapes of the structuring element and a bidimensional granulometry intended to determine the optimal filtering parameters automatically. A clustering-based approach for image binarization is also introduced. This one avoids an empirical thresholding of input panchromatic images. Experiments made on a Quickbird VHR-image show the effectiveness of the method.**

Keywords: Mathematical Morphology, Hit Or Miss Transform, Bidimensional Granulometry, Fusion of Clusters, Quickbird VHR imaging, Building Extraction.

I. Introduction

Recent advances in the quality of satellite imagery open new prospects in the field of automatic detection of urban objects. The very high spatial resolution images offer the opportunity to recognize and individualize objects such as trees, buildings, roads, and so on. In this paper, we focus on automatic building extraction methods which are helpful to optimize the management of urban space by local politics [1]. Among these methods, mathematical morphology has already proved to be effective for many applications in remote sensing [2]–[5]. This is the approach we adopt here. The solution proposed differs from some previous works as it does not require any additional information or ancillary data to perform it (e.g. no digital elevation models [6]). The part of the user intervention is also limited (contrary to the approach of [7] for instance). Our method is based on a sequence of different morphological operators applied on binary images among which the Hit-or-Miss transform. We extend the solution developed in [8] by improving the binarization and the filtering steps. Here, a clustering-based approach is proposed to convert the input greylevel image into binary image. In addition, we compute a granulometry [9] to determine the parameters of the opertors in an automatic way. These extents enable to eliminate the limitations of our first solution and in particular: the images binarization through an empirical thresholding, the manual setting of the parameters of the algorithms, and the unability to process buildings with heterogeneous roofs.

The paper is organized as follows. In section 2, we present the building extraction method composed of three main steps: generation of binary images from panchromatic data, automatic morphological filtering, and building detection. In section 3, experimental results are given and discussed before to conlude the paper with final remarks in section 4.

II. Proposed method

The method we propose to extract building objects from VHR-images relies on the use of binary mathematical morphology operators which are based on set theory [9]. These operators are applied on the input image I with a predefined pattern called structuring element (SE). The two fundamental operators in MM are the erosion ($I \ominus S$) and the dilation ($I \oplus S$) respectively defined as:

$$I \ominus S = \{x \ : \ (S)_x \subseteq I\} \quad (1)$$

$$I \oplus S = \{x \ : \ (S')_x \cap I \neq \emptyset\} \quad (2)$$

with S' and S_x respectively denoting reflexion and translation by x of the set S. From these basic operators it is possible to define more complex operators as we will see throughout this paper.

The overall approach we propose is illustrated in figure 1. It is composed of three main steps. The first one consists in the input greylevel image binarization. Since, the method is applied on a panchromatic Quickbird image, it is necessary to convert it in binary data (compatible with binary MM operators). The second step is an automatic morphological filtering intended to eliminate some objects in the image and to determine the size of the structuring elements. The third step is the building extraction step itself based on the use of an adaptive Hit-or-Miss transform. Each of these steps is detailed below.

A. Generation of binary images

As we use binary morphological operators, panchromatic (or greylevel) input image I cannot be processed directly. In order to obtain a binary image B, the simplest solution is to choose arbitrarily a threshold value T (depending of the image), and classify all pixels as white or black according to whether the pixel values exceed or not this threshold:

$$B(x) = \begin{cases} 1 & \text{if } I(x) \geq T \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

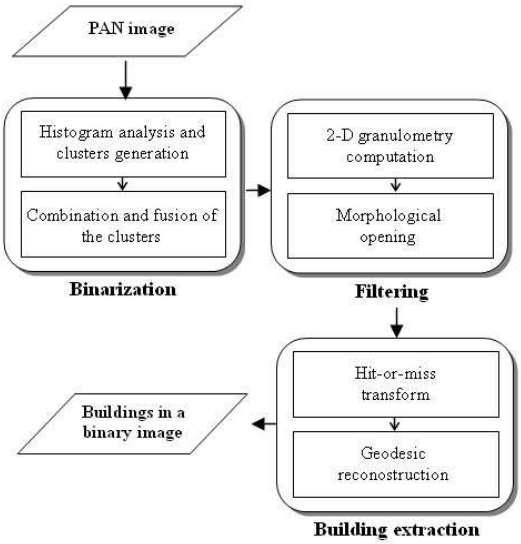


Fig. 1. The MM-based building extraction strategy.

This is a classical thresholding method but which presents well-known limitations. First, in many cases, finding one threshold to the entire image is very difficult. In our case, it assumes that buildings are either brigther or darker than the rest of the image which is not always the case. In addition, the threshold must be determined empirically. This threshold can vary from one image to another which supposes to repeat the analysis each time a new image is considered. This solution is not satisfactory.

In order to avoid the definition of the threshold empirically and to increase the genericity and automaticity of the solution, a clustering-based approach for binarization is proposed. This method is founded on the analysis of the histogram of the input image, which has been smoothed to avoid local optimums. More precisely, clusters are built iteratively by identifying the modes of the histogram and by selecting the highest local maxima with their neighbouring values. Each of these modes is assigned to a single cluster. Once a mode has been processed, all its values are set to 0 and the next mode is considered. At the end of the iterative process, when the most part of the pixels have been assigned to the clusters, the process stops and the remaining pixels are integrated to the closest clusters. This clustering method does not require to know a priori the number of classes to find in the histogram. The only parameter used in input is the stopping criterion, i.e. the percentage of the pixels that have to be assigned to the clusters.

Let us note $w(l) = i$ the classification function w which assign a class (or label) i to the value l. The image C resulting in the clustering in n classes can then be represented as:

$$C(x) = i \qquad \text{if } w(I(x)) = i \tag{4}$$

Thus, by applying this clustering method on the input image, a classified image C is obtained. This image can be also considered as a set of n binary images C_i (one by cluster), each one of them representing the binary membership of the pixels to a given cluster i:

$$C_i(x) = \begin{cases} 1 & \text{if } w(I(x)) = i \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

All of the images C_i can be therefore processed with binary morphological operators. However, this set of images is not sufficient. Indeed, some of the buildings can be characterized by heterogeneous roofs. These roofs can be composed of several parts having distinct spectral signatures. Consequently, they can be included into different clusters. For that reason, a combination of the initial binary images is also performed. For instance, if we consider the fusion of two clusters, the new image will be defined as follows:

$$C_{i,j}(x) = \max\{C_i(x), C_j(x)\} \tag{6}$$

The new set of binary images generated by fusion is added to the previous one and the morphological processing described below is applied on all of theses images.

B. Automatic morphological filtering

Before to extract the buildings from the generated images, an automatic morphological filtering is also performed. The aim of this filtering is to remove objects whose size is lower than the minimum size of a building in the raw image. These objects may be seen as noisy data capable of disturbing the extraction process. The filtering used is a morphological opening defined as a combination of erosion and dilatation:

$$\gamma_S(I) = (I \ominus S) \oplus S \tag{7}$$

where the size and the shape of the structuring element S are parameters of prime importance.

In our previous works, the filtering parameters were determined manually. For instance, a structuring element corresponding to a square of 15×15 pixels was retained in [8]. Here, we involve an automatic process to determine the optimal parameters. It consists in applying a bidimensional granulometry on the binary image. A granulometry (also known as a morphological profile or differential morphological profile) is a kind of morphological histogram [9]. It is computed using a sequence $\Gamma_{S,n}$ of morphological openings with structuring element S of increasing size k:

$$\Gamma_{S,n}(I) = (v(\gamma_{kS}(I)))_{k\in[1,n]} \tag{8}$$

with kS representing S dilated k times, v the surface function (i.e. returning the number of pixels equal to 1). Thus, a granulometry is not a histogram which reflects the distribution of the spectral values. It rather considers the distribution of the shapes and sizes of the objects existing in the image. This operator has been already used with success in remote sensing [10]–[12].

Most of the time, a modified version of the morphological profile is used: the function is both normalised (i.e. all values are divided by $v(I)$) and derived (i.e. all values $v(\gamma_{kS}(I))$ are replaced by $v(\gamma_{(k+1)S}(I)) - v(\gamma_{kS}(I))$). We use this version in our approach. In addition, the granulometry we compute is bidimensional as both height and width of the structuring

element can vary independently. It can be expressed as follows:

$$\Gamma'_{S,(m,n)}(I) = \left(v(\gamma_{(k,l)S}(I))\right)_{(k,l)\in[1,m]\times[1,n]} \quad (9)$$

where $(k,l)S$ means S dilated k times in the horizontal dimension (i.e. by a horizontal line of 3 pixels length) and l times in the vertical dimension (i.e. by a vertical line of 3 pixels length). In other words, k and l help to define the size (height and width) of the SE S. For each binary image, the highest peak in the bidimensional granulometry curve enables us to identify the size under which the objects will be removed. The structuring element of the morphological opening is defined according to these values. If the size of the structuring element is too low, we consider the related binary image as irrelevant and not containing any building. Otherwise, the filtered image is processed with the next step of the method.

C. Building detection

The goal of the two previous steps was to prepare the detection of the buildings by generating a set of filtered binary images from a greylevel image. Now, the building detection itself will be considered. We propose to use the Hit Or Miss Transform (HMT) which consists in a double erosion of the image I and its complement I^c (i.e. the background) with two disjoint structuring elements E and F. This transform is particularly useful for template matching and is defined as:

$$I \circledast (E,F) = (I \ominus E) \cap (I^c \ominus F) \quad (10)$$

$$= \{x \ : \ ((E)_x \subseteq I) \wedge ((F)_x \subseteq I^c)\} \quad (11)$$

where a pixel x is keep as long as it ensures a successful match of both the ES E with I and the ES F with I^c, both ES being centered into x.

Since we try to detect square or rectangular buildings of various sizes, we adapt the HMT to be able to take into account some structuring elements E and F with varying sizes and shapes. Our adaptive HMT is defined as:

$$I \circledast_{\mathcal{K},\mathcal{L}} (E,F) = \bigcup_{\substack{k\in\mathcal{K} \\ l\in\mathcal{L}}} (I \ominus E_{\alpha k,\alpha l}) \cap (I^c \ominus F_{k,l}) \quad (12)$$

Thus, the result of this adaptative HMT is defined as the union of all the results of the transform applied with a given pair of structuring elements. The two variable structuring elements $E_{a,b}$ and $F_{c,d}$ are respectively defined as a rectangle of size $a \times b$ and a frame (contour of a rectangle) of size $c \times d$, with the constraints $c > a$ and $d > b$. The sets $\mathcal{K}$ and $\mathcal{L}$ contain respectively all the possible heights and widths of the SE, and α is a coefficient used to determine the uncertain area between E and F. In other words, it helps to mark the area between pixels which surely belong to buildings and pixels which surely belong to background (or not to buildings).

At the end of this operation, if the parameters of the HMT have been correctly defined, only the buildings are retained with their respective position. However, the shape of these buidings do not correspond any more to the initial shape. Indeed, the HMT is based on erosions which reduce the size of the objects. Thus, a postprocessing is necessary to rebuild the shape of the detected buildings. An additional morphological operator is used for this task which corresponds to a geodesic reconstruction:

$$I \triangle_B M = (M \oplus_I B)^\infty \quad (13)$$

using two images, an input image I and a marker image M, and applying until convergence a conditional dilation with SE B defined as:

$$M \oplus_T B = (M \oplus B) \cap T \quad (14)$$

The geodesic reconstruction eliminate all the objects which do not appear in the result of the Hit Or Miss Transform (defined as marker M) from the result of the filtered binary images (defined as input I).

Finally, since we have initially generated a set of binary images (either associated to a single cluster or to a combination of several clusters), the last step of the method consists in merging the results obtained for each image which have been processed independently. A binary union is performed: a pixel is retained to form a building if this pixel has been declared as such in one of the images obtained after the geodesic reconstruction.

III. Experiments and discussion

In order to assess the effectiveness of our method, we applied it on several pieces of panchromatic Quickbird VHR-images from Strasbourg, France. This image has a spatial resolution of 0.7 meter. The quality of the results obtained is variable and depends directly of the heterogeneity of the buildings roofs. The figure 2 illustrates the possible complexity of the observed roofs.

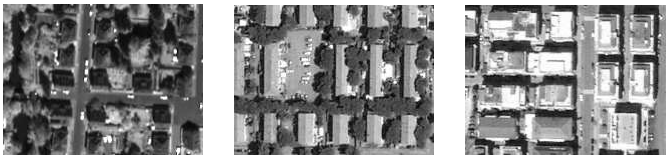


Fig. 2. VHR Images containing buildings with roofs of variable visual complexity.

The method returns very accurate results for images containing buildings with simple roofs (made of a single part). If building roofs are composed of two symmetrical parts with different spectral signatures due to sunlight illumination, the accuracy of the results is still ensured. However, the accuracy strongly decreases with roofs of high heterogeneity. An illustration of some buildings detected is given in figure 3.

Accuracy has been assessed through a confusion matrix (table III). We computed a global precision rate of 88 % with a Kappa value of 63 %.

Several remarks can be formulated about the results obtained. Firstly, as we seen in figure 3, we can effectively detect buildings of various sizes during the HMT, by applying structuring elements of increasing length and width. Secondly, the buildings with shapes that do not strictly correspond to

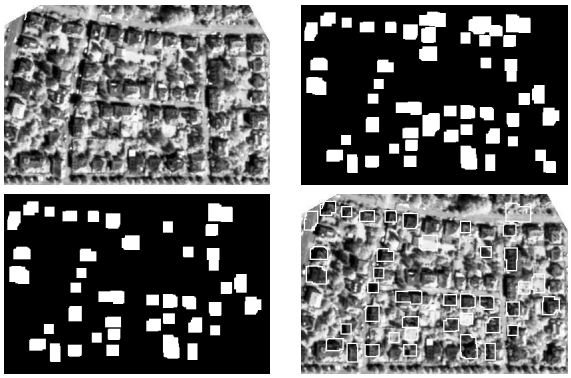


Fig. 3. Input image (top left), filtered image (top right), detected objects (bottom left), and final result (bottom right) on one of the test images.

TABLE I

PIXEL-BASED EVALUATION THROUGH A CONFUSION MATRIX COMPUTED ON THE ENTIRE SET OF PROCESSED IMAGES

	building	background
building	17673	4577
background	10138	90012

squares or rectangles can also be detected (e.g. buildings having small recess or projection construction, or having hidden parts). This is made possible thanks to the uncertain area introduced in the HMT. This area is particularly important since it garantees a tolerance on the shapes of analyzed objects. It lets the process ignore some pixels between the buildings and background areas (figure 4). Finally, the HMT enables to eliminate some elements in the binary images that are not buildings (for instance, some shades). Thus, the number of objects existing in the HMT image is lower than the one existing in the filtered image. As regards to the shape of objects used during the geodesic reconstruction, it corresponds to the one of the elements present in the filtered (marker) image.

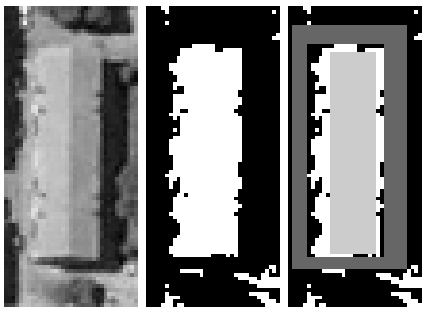


Fig. 4. Relevance of the HMT for unperfect building shapes: original image (left), binary image (middle), and application of the HMT (right). The two SE used in the HMT appear in light grey (for the foreground) and dark grey (for the background), the uncertain area is located in between.

Some results of the clustering step are illustrated in figure 5. We can see that the method generates a set of binary images which may not contain the entire buildings. The fusion of the clusters invovled in the method enables to overcome this problem. The top left image in figure 6 which corresponds to the fusion of the clusters 2 and 3 contains this time all the elements composing the buildings. It is particularly relevant in this case.



Fig. 5. A building with a roof made of several parts (top left) and the five clusters returned by the histogram-based clustering step.

Figure 7 illustrates the relevance of the proposed 2-D granulometry to automatically determine the optimal SE to be used in the morphological filtering process (based on an opening operation). As we can observe, the proposed SE size (here 19×21 pixels) helps to greatly reduce the noisy areas which do not correspond to buildings.

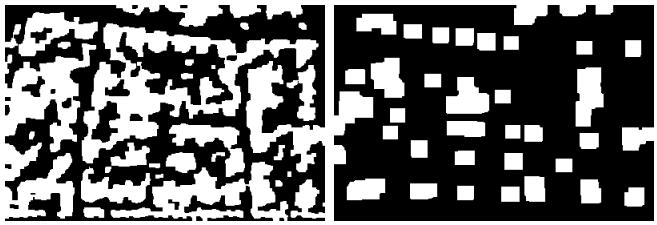


Fig. 7. Relevance of the 2-D granulometry to determine the optimal parameters for morphological filtering : unfiltered (left) and filtered (right) binary images.

Two different views of the 2-D granulometry obtained for this image are given in figure 8 in order to better understand the content of this morphological measure. As we can observe, the peak in the granulometry corresponds the optimal SE size.

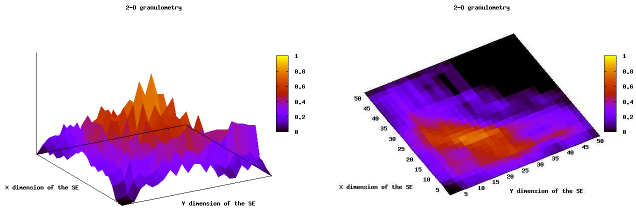


Fig. 8. Two different 3-D views of the 2-D granulometry.

The main limitations of the proposed method are related to the quality of the binary image processed. Indeed, close buildings could be agregated during the morphological smoothing step, thus resulting in large elements with relative complex shapes. These non-rectangular shapes are then misdetected by the HMT operator. These agregate elements should be processed with relevant SE. We have also observed that agregate may be built not only from buildings but also from other objects such as trees, which may have a similar spectral

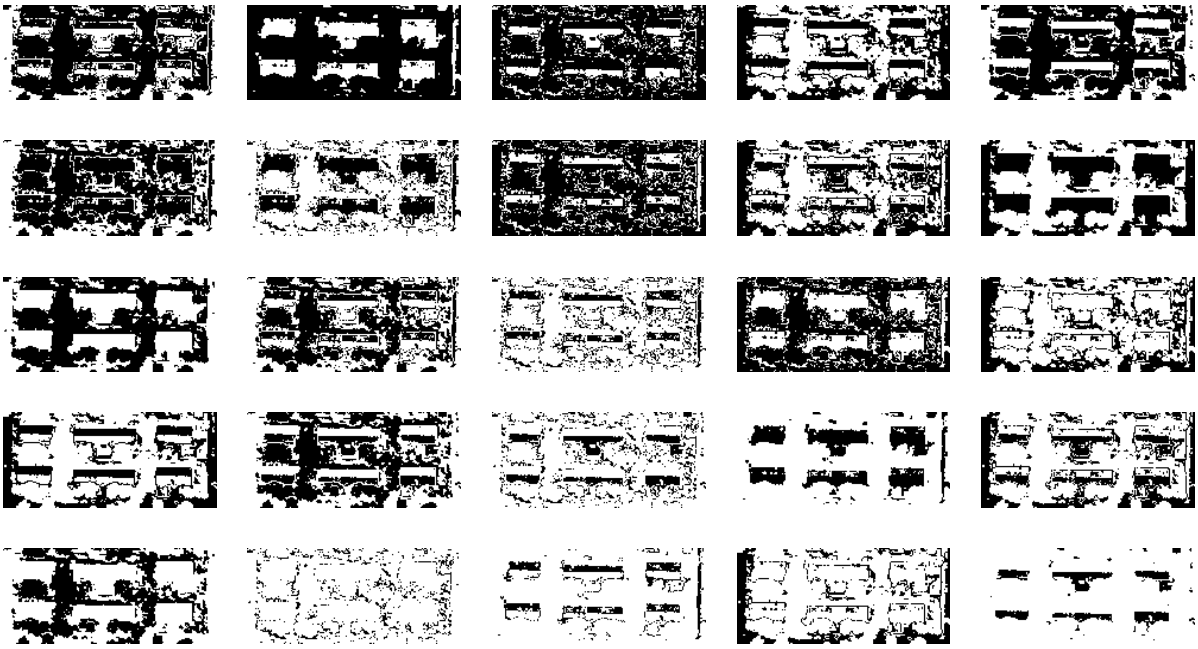


Fig. 6. The different results of the fusion of clusters (of increasing cardinality) for image given in figure 5.

signature in the panchromatic image. Thus, other SE should be considered in order to deal with agregated elements and improve the quality of the proposed approach. Moreover, some errors could be avoided if multispectral images were implied. In this case, confusions between buildings and trees would be easily solved by computing the NDVI index.

IV. Conclusion

Mathematical morphology offers some image processing tools which can be successfully used to solve urban remote sensing issues such as building detection in VHR images. In this paper, we proposed a morphological approach that deal with this problem. Our method is adapated to panchromatic images and do not require any ancillary data to be performed.

We extended our previous works [8] by introducing a bidimensional granulometry in the filtering step. This morphological profile helps to define automatically the structuring elements used in the adaptative hit or miss transform. In addition, the clustering method proposed to convert the input greylevel image into binary images avoids to determine the binarization threshold empirically. The fusion of the clusters also enables to take buildings with complex (composite) roofs into account.

Future works will focus on the application of greylevel or multivalued operators [13] on input or filtered image data. We are also considering an implementation of the solution on a grid-based architecture in order to reduce the computation time required by the morphological operations (bidimensional granulometry and adaptive hit or miss transform). Finally, the method should be applied on larger dataset in order to assess its performance in a more accurate way.

References

[1] S. Lhomme, C. Weber, D. He, and D. Morin, "Building extraction from vhrs images," in *ISPRS Congress*, Istanbul, Turkey, 2004.

[2] I. Destival, "Mathematical morphology applied to remote sensing," *Acta Astronautica*, vol. 13, no. 6/7, pp. 371–385, 1986.

[3] F. Laporterie, G. Flouzat, and O. Amram, "Mathematical morphology multi-level analysis of trees patterns in savannas," in *IEEE International Geosciences And Remote Sensing Symposium*, 2001, pp. 1496–1498.

[4] P. Soille and M. Pesaresi, "Advances in mathematical morphology applied to geoscience and remote sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 9, pp. 2042–2055, September 2002.

[5] S. Derivaux, S. Lefèvre, C. Wemmert, and J. Korczak, "Watershed segmentation of remotely sensed images based on a supervised fuzzy pixel classification," in *IEEE International Geosciences And Remote Sensing Symposium*, Denver, USA, July 2006.

[6] U. Weidner and W. Forstner, "Towards automatic building reconstruction from high resolution digital elevation models," *ISPRS Journal*, vol. 50, no. 4, pp. 38–49, 1995.

[7] C. Matti-Gallice and C. Collet, "Contribution of remote sensing to the definition of an indicator for the analysis of periurban dynamics," in *IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (URBAN)*, Berlin ,Germany, May 2003, pp. 182–185.

[8] J. Weber, S. Lefèvre, and D. Sheeren, "Building extraction in vhrs images with mathematical morphology," in *International Conference on Spatial Analysis and Geomatics*, Strasbourg, France, September 2006.

[9] J. Serra, *Image Analysis and Mathematical Morphology*. Academic Press, 1982.

[10] M. Pesaresi and J. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 2, pp. 309–320, February 2001.

[11] J. Benediktsson, M. Pesaresi, and K. Arnason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 9, pp. 1940–1949, September 2003.

[12] X. Jin and C. Davis, "Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information," *Eurasip Journal of Applied Signal Processing*, vol. 14, pp. 2196–2206, 2005.

[13] E. Aptoula and S. Lefèvre, "A comparative study on multivariate mathematical morphology," *Pattern Recognition*, January 2007, in revision.