Article

Fast Unsupervised Multi-Scale Characterization of Urban Landscapes Based on Earth Observation Data

Claire Teillet [1,*], Benjamin Pillot [1], Thibault Catry [1], Laurent Demagistri [1], Dominique Lyszczarz [2], Marc Lang [3], Pierre Couteron [4], Nicolas Barbier [4], Arsène Adou Kouassi [5,6], Quentin Gunther [7] and Nadine Dessay [1]

1   Espace-DEV, IRD, University of Montpellier, 34093 Montpellier, France; benjamin.pillot@ird.fr (B.P.);
    thibault.catry@ird.fr (T.C.); laurent.demagistri@ird.fr (L.D.); nadine.dessay@ird.fr (N.D.)
2   Entente Interdépartementale des Causses et des Cévennes, 48400 Florac, France;
    observatoire@causses-et-cevennes.fr
3   DYNAFOR, University of Toulouse, INRA, 31326 Castanet-Tolosan, France; marc.lang@toulouse-inp.fr
4   AMAP, University of Montpellier, CIRAD, CNRS, INRAE, IRD, 34000 Montpellier, France;
    pierre.couteron@ird.fr (P.C.); nicolas.barbier@ird.fr (N.B.)
5   Institute of Tropical Geography, University of Félix Houphouët-Boigny,
    01 BPV 34 Abidjan 01, Abidjan, Côte d'Ivoire; kouassiarsene.adou@ird.fr
6   MIVEGEC, University of Montpellier, IRD, CNRS, 34000 Montpellier, France
7   ENSG-Geomatique, Departement of Cartography, 77420 Champs-sur-Marne, France; q.gunther@orange.fr
*   Correspondence: claire.teillet@ird.fr

Abstract: Most remote sensing studies of urban areas focus on a single scale, using supervised methodologies and very few analyses focus on the "neighborhood" scale. The lack of multi-scale analysis, together with the scarcity of training and validation datasets in many countries lead us to propose a single fast unsupervised method for the characterization of urban areas. With the FOTOTEX algorithm, this paper introduces a texture-based method to characterize urban areas at three nested scales: macro-scale (urban footprint), meso-scale ("neighbourhoods") and micro-scale (objects). FOTOTEX combines a Fast Fourier Transform and a Principal Component Analysis to convert texture into frequency signal. Several parameters were tested over Sentinel-2 and Pleiades imagery on Bouake and Brasilia. Results showed that a single Sentinel-2 image better assesses the urban footprint than the global products. Pleiades images allowed discriminating neighbourhoods and urban objects using texture, which is correlated with metrics such as building density, built-up and vegetation proportions. The best configurations for each scale of analysis were determined and recommendations provided to users. The open FOTOTEX algorithm demonstrated a strong potential to characterize the three nested scales of urban areas, especially when training and validation data are scarce, and computing resources limited.

Keywords: remote sensing; multi-scale; unsupervised; urban landscapes; texture

1. Introduction

Urban landscapes are composed of objects of various sizes and materials arranged by humans in a complex manner [1]. Urban areas are defined as the footprint of impervious surfaces (buildings and roads), including pervious surfaces such as vegetation within the footprint [2,3].It is the built environment in which population is distributed. Urban areas are thus characterized by a multi-scale structure. Urban analysis using Earth Observation (EO) data is particularly scale dependent [4]. Current optical and SAR sensors offer a wide range of spatial, temporal and spectral resolutions. Medium and high resolution (10–500 m) images (MODIS, Landsat, Sentinel) can easily provide information on the urban footprint and delineate basic urban classes, while very high resolution (~1 m) data (Pleiades, SPOT 6/7, DigitalGlobe...) provide more precise information at a finer scale to delineate individual buildings and even estimate volumes [5]. The diversity of sensors

and data are associated with a wide range of pixel or object-based methods using spectral, spatial, textural, or contextual information to characterize urban areas [6]. Most studies reduce the characterization of urban areas to the application of a specific method at a specific scale, usually on a single city. Actually, the complexity of urban areas requires the use of a procedure allowing characterizing various urban areas at different scales using a single methodological approach, as suggested by Zhu et al. [7]. The results presented in this paper propose a first attempt to fill this gap, introducing the use of a single texture-based analysis for the multi-scale characterization of urban areas.

While many urban products are available, they are almost exclusively dedicated to the delineation of large-scale urban footprints (named macro-scale hereafter, with a resolution ranging from a few hundreds of meters to a few kilometers). The extraction of urban footprint is usually carried out by supervised methods involving a substantial amount of training data, as well as strong material resources and computing time especially with the current development of artificial intelligence [8–10]. For instance, the Global Human Settlement Layer (GHSL) [11–15], the World Settlement Footprint 2015 (WSF-2015) [16], the Global Land Cover [17] or the Human Built-Up And Settlement Extent Data [18] were produced from Landsat and Sentinel data at 20 to 30 m resolution. Only the Global Urban Footprint (GUF) from TerraSAR-X data at 12 m resolution was produced using an unsupervised procedure [19,20]. Lehner et al., studied urban footprint with unsupervised algorithm but only focused on Europe [21]. These products rely on spectral information but some, such as the GHSL, also integrate texture for the identification of urban footprint [22–25]. The detection of the urban footprint by texture analysis can be associated with the extraction of Haralick indices thanks to the Grey Level Co-occurrence Matrix (GLCM) or the use of other texture approaches [26–29]. More importantly, the quality of urban footprint products is usually well estimated in Europe [21,30], due to the availability of validation datasets, as illustrated by Reba and Seto [9] showing that 75% of studies focused on high or upper-middle-income countries. In southern countries, training and validation datasets are often missing or of low quality. This lack is not even bridged by OpenStreetMap, a common and essential source of training and validation datasets in supervised approaches, since only 11% of the contribution between 2014 and 2018 was located in Africa [31]. That is why in this study we favor an unsupervised texture-based approach.

When focussing on medium-scale description of urban areas (named meso-scale hereafter), there is no generic method that can be applied everywhere. This is partly due to the fact that there is no global definition of urban areas at the meso-scale. The meso-scale is usually referred to as neighbourhood scale. This term may have many different meanings or uses, but it can refer to similar units nested within larger sets [32]. One good example of meso-scale approach is the nomenclature of local climate zones (LCZ) in urban areas by Steward and Oke [33] which was designed to be applied globally at the at the meso-scale. However, many studies of urban heat islands do focus on urban analysis at the meso-scale [34–36] but are specific to one city without generic perspective. Only Stokes and Seto [37] developed a new approach to study "sub-metropolitan" urban units, designed to be "comparable across and within urban areas". Accordingly, using texture to characterize urban areas at the meso-scale has a strong potential for genericity since every city, although having a specific organization, has an inherent textural signal [38]. In this study, we define urban units as homogeneous subsets (with size of urban units ranging from ten meters to a few hundreds of meters) of the urban footprint.

This lack of generic method is even more important at the scale of urban objects (micro-scale hereafter, metric scale of buildings). Both for meso and micro-scale, this is mainly due to the limited availability of very high resolution imagery that is usually required for such scales of study. Some studies have highlighted the use of texture to characterize more finely urban areas [39–41].

Some authors point out that even if no global dataset is available, future applications on remote sensing over urban areas have to be based on methodologies that can be applied globally on any urban areas [7,42,43].The methodological framework we propose aims

at describing urban areas at three scales to characterize urban footprint (macro-scale), urban units or neighbourhoods (meso-scale) and buildings (micro-scale). We focused on an unsupervised approach to try and overcome the current limitations of remotely-sensed urban characterizations described previously, especially the lack of training and validation datasets in southern countries. Among unsupervised methods, the Fourier-based Textural Ordination (FOTO) algorithm already gave in-depth results for describing vegetation patchworks based on high (HR) and very high resolution (VHR) satellite images [44–46]. As a texture-based method, it differentiates vegetation types, through the identification of spatial pattern wavelengths and orientations in the frequency domain due to object types and their spatial distributions. We believe that the same method can be applied to urban landscapes: those are typically defined as spatial arrangements of urban objects (micro-scale) and neighbourhoods (meso-scale) with different orientation, size and type [37,47–49]. Even more, we expect that vegetation and urban objects are characterized by distinct frequency content thus making the approach especially relevant in identifying urban footprints (macro-scale). In this paper, we developed the FOTOTEX algorithm to apply the FOTO method to the urban context, in order to identify and characterize urban landscapes at the three scales mentioned above. Our contributions, are the following: (1) we tailor the unsupervised FOTO method formerly used for the discrimination of vegetation patterns to identification of urban footprints (macro-scale) and characterization of urban patterns (meso-scale) and objects (micro-scale), (2) we use only one texture-based method to study urban areas at three nested scales, especially over the characterization of spatial urban unitsat the meso-scale which is rarely studied and represents a real interest for applications in many environmental studies (3) we implement new options and optimize the algorithm so that the whole HR or VHR image processing only takes minutes on a conventional computer, and finally (4) we show the added value of our approach for the South through both a detailed analysis of urban landscapes in southern cities from different continents, and a comparison with existing global products such as the GHSL and GUF.

2. Materials and Methods

2.1. Study Area and Satellite Imagery

We applied our method on two study areas (Figure 1): Brasilia (Brazil) and Bouake (Ivory Coast) and over the three scales (macro, micro, meso). Brasilia, the capital of Brazil, was built in 1960, and covers an area of 5802 km^2. The city is located on a plateau at an altitude of 1172 m, right in the heart of the Cerrado savanna and close to the watershed boundaries between the Amazon (north-west), Rio Tocantins (north-east), Rio Sao Francisco (north) and La Plata (south) river basins.

The city of Bouake, built in 1910, covers an area of 71.79 km^2 and is located in the center of Ivory Coast, on a flat relief, with an important constellation of villages around it. Bouake is located over the Bandama watershed, more specifically straddling two sub-basins: the Bandama Blanc to the north and east and the N'Zi to the south and east, the city is also crossed by the Kan River.

For the characterization at meso-scale, given the extent of the city of Brasilia, we performed the analysis over the administrative region of Sao Sebastiao that was created in 1993 by the Federal District Government of Brazil (Figure 1). Located in an area with hills cut by a fairly dense network of streams, Sao Sebastiao is a mainly rural territory with a peripheral urban area where the very rapid expansion is characterized by a phenomenon of urban sprawl at the expense of rural areas. Sao Sebastiao's urban infrastructure is characterized by a road network consisting mainly of unpaved roads, a limited sewage system and a lack of public health facilities in particular. We also performed a meso-scale analysis over the northwestern part of Bouake city (Figure 1). This area is in the process of being urbanized, without urban plan at the moment, and still remains very rural. The tracks are in very bad condition with a very bad accessibility during the rainy season. Electricity has started to be established. Then there is no water supply network, just some pumps essentially installed by an association.

Macro scale : Sentinel 2   Meso / micro scale : Pleiades
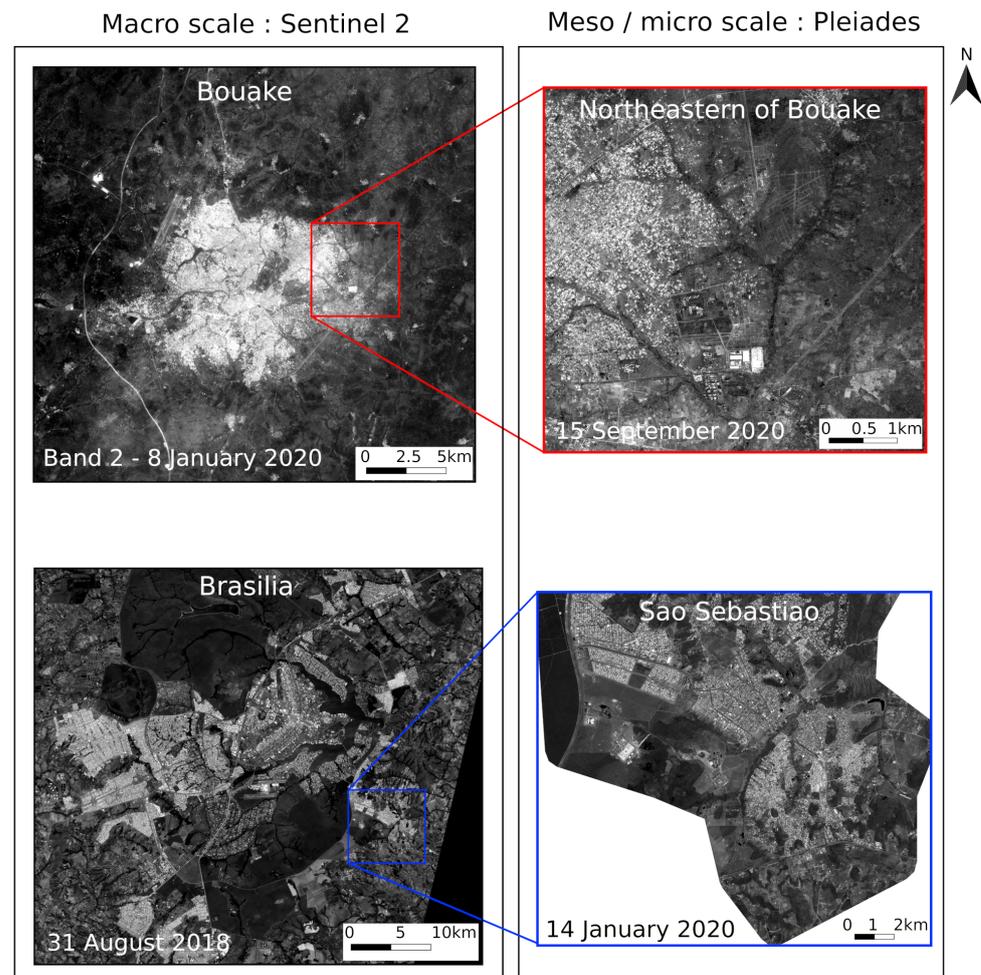
Figure 1. Study areas and satellite imagery used in this study. On the left, band 2 (blue) of Sentinel 2 images on Bouake (**top**) and Brasilia (**bottom**) are shown (Copernicus Sentinel data (2018-2020)/ESA). On the right Pleiades images (panchromatic channel) on the northeastern part of Bouake and Sao Sebastiao, a satellite City of Brasilia (Contains information © CNES 2020, Distribution Airbus DS, all rights reserved. No commercial use.).

Analyzes were performed on both Sentinel 2 and Pleiades imagery. Sentinel 2 images were acquired on 31 August 2018 for Sao Sebastiao and on 8 January 2020 for Bouake. Sentinel 2 image consists of 13 multispectral bands of different spatial resolution. In this study we used bands from 2 to 8A and resampled all bands at a resolution of 10 m. Pleiades images were acquired on 14 January 2020 for Sao Sebastiao and on 15 September 2020 for Bouake. They were provided by the GEOSUD equipex (http://ids.equipex-geosud. fr/accessed on 20 September 2020). The Pleiades images consist of a panchromatic band (0.7 m resampled to 0.5 m spatial resolution) and 4 multi-spectral channels (the blue, green, red, near infrared bands at 2 m spatial resolution).

2.2. Method

Our methodology uses texture information extracted from EO data at different resolutions to characterize urban landscapes in southern countries at three different scales:

- the urban footprint (macro-scale)
- the urban units scale (meso-scale)
- the object scale such as buildings (micro-scale)

It is based on the "FOTO" algorithm (for FOurier-based Textural Ordination) [44–46]. FOTO was originally developed to characterize vegetation patterns on semi-arid ecosys-

tems with very high resolution imagery (aerial photography). Many references describing the procedure carried out with FOTO are already available [44–46,50,51]. We adapt and optimize the algorithm to study urban patterns and demonstrate the potential of FOTOTEX. This unsupervised approach allows studying urban landscapes with the same method and same tools at three nested scales (Figure 2). We first apply FOTOTEX on Sentinel 2 to delineate urban footprint, then FOTOTEX runs again on very high resolution images (Pleiades, unmanned aerial vehicle (UAV) imagery) to identify urban units and objects within this footprint. So from top to bottom (large scale to fine scale) the analysis at one scale "feeds" the next scale since the output at the macro-scale is used in the analysis at the meso-scale. Then when urban units are identified at the meso-scale, for each unit we study the arrangement of urban objects at the micro-scale. The influence of technical parameters in the FOTOTEX algorithm is evaluated over the three different scales.
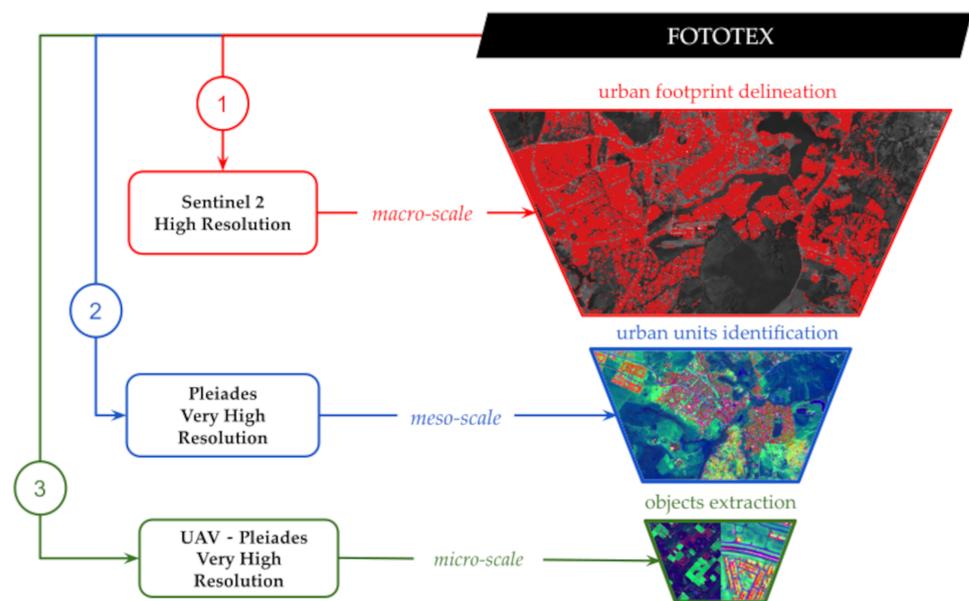

Figure 2. Potential of FOTOTEX for the multi-scale characterization of urban landscapes, with the scales of analysis and associated Earth observation data.

2.3. FOTOTEX Algorithm

Based on different versions of the original algorithm FOTO coded in R or Matlab, we developed the package FOTOTEX in python and added new features. To have an easy, robust and efficient tool, we decided to deploy the algorithm with Python thanks to the Python Package Index. The algorithm has been optimized in terms of process and calculation time: FOTOTEX is able to extract urban footprints and indicators from 50–100 millions pixels images within minutes on a conventional computer. We included in FOTOTEX the possibility for the users to activate or not several parameters such as the DC component or the normalization. In the original FOTO algorithm, the application was only dedicated to homogeneous areas of forest canopy, DC component and normalization were automatically on default mode.

Step 1: Image partitioning (Figure 3a)

The first step of the method is to partition the image. This partitioning corresponds to the definition of a parsing window's path rule on the image. The FOTOTEX algorithm allows two methods for slicing images over an image raster. The first is the block method (method = "block") that analyzes texture for each window in the image block by block. The second is the moving method (method = "moving_window") that analyzes texture for a given window sliding from west to east and north to south with a shift of one pixel.
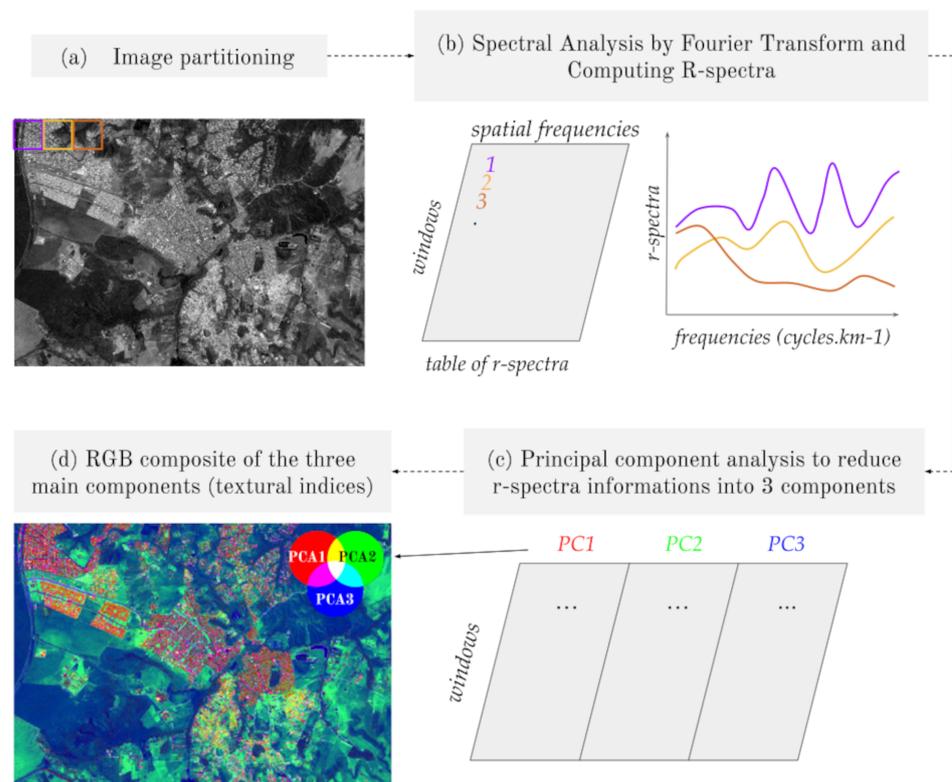
Figure 3. Simplified methodological framework of the FOTOTEX algorithm, divided in four steps (**a**) image partitioning, (**b**) spectral analysis by Fourier transform and R-spectra computation, (**c**) Principal Component Analysis and (**d**) RGB composite of the three main components (textural indices). These steps will be described in detail in the method section below.

Step 2: Spectral analysis by Fourier transform and R-spectra computing (Figure 3b)

A fast Fourier transform (FFT) is computed for each analysis window of the original image. FFT aims at expressing the variance of an image as a weighted sum of cosine and sine waveforms of varying travelling direction and spatial frequency [25,44,45]. It consists in converting the spatial signal (patterns in the image) into a frequency signal (number of times that a pattern is repeated within a defined area). The signal is thus represented by a function composed of a fundamental frequency (lowest frequency), the harmonics (the multiples) and coefficients of the Fourier series. These coefficients quantify the contribution of each frequency to the signal formation [52]. The frequency representation of a Fourier transform is a spectrum of amplitude corresponding to energy values for each frequency. In the subsequent step, the phase information is ignored, and the periodogram is computed for each window, which allows us to have the proportion of variance explained by each frequency pair (p,q), where p and q are spatial frequencies along the image column and row directions [45]. When computing the r-spectra, the DC component (or "zero frequency component") of the FFT which is the basis of the signal corresponding to the average reflectance in each window can be kept or not. This periodogram can be equivalently expressed in polar form with:

$$I(r, \theta) \tag{1}$$

values referring to a waveform with a spatial frequency and a direction.

Then we can process the calculation of an isotropic r-spectra by not taking into account a specific direction but by averaging the periodogram in all directions over the entire azimuth (from 0° to 360°) to yield an azimuthally averaged radial spectrum, called "r-spectrum" denoted by I(r).

Step 3: Texture ordination (Figure 3c,d)

A table compiling the r-spectra is produced where each row corresponds to the r-spectrum of a given window, and each column corresponds to a spatial frequency [46]. The table can be normalized, i.e., the values of the power spectrum over the image are divided by the variance in each given window in order to make the contribution of each frequency comparable. Then, a PCA is applied to the table in order to analyze the variability between spectra. The objective is to ordinate the analysis windows according to their frequency along a coarseness gradient [44]. The three first axes resulting from the PCA are then represented as a red-green-blue (RGB) composite image where colors correspond to the projection of the r-spectra of the various analysis windows upon the three axes of the PCA [45,53]. The spatial resolution of this composite image depends on which partitioning method is chosen. For the block method the output spatial resolution corresponds to pixel size multiplied by the window size used for the initial partitioning of the image. For the moving window, the step being one pixel, the output spatial resolution corresponds to the original pixel size.

2.4. Influence of Technical Parameters

The information corresponding to the three textural information for each window was combined in a RGB color composition in which the red channel represents the first texture axis (PC1), the green channel the second texture axis (PC2) and the blue channel the third texture axis (PC3). This representation allows us to have a spatial visualisation of the distribution of the frequencies that compose our initial image. Pixels values are the projection upon the 3 axes of the PCA of the analysis windows r-spectra. Different tests have been carried out to observe the influence of different parameters (Table 1) at the three different scales. All of the calculations were performed on a computer composed of a processor intel Xeon (6 processors HT, 2.8 Ghz) and 32 GB RAM.

Table 1. The various parameters and possible options to be implemented by the users in the FOTOTEX algorithm.

Parameters	Choice
Spectral band	Depends on the satellite image
Partitioning method (PM)	"Block" or "Moving window"
Pixel size (PX)	Any
Window size (in pixel) (WS)	Any positive odd number
Keep DC component (DC)	True or False
Normalize (N)	Normalize or non-normalized

The FOTOTEX algorithm can be used on different spectral bands. The choice of bands is important because the radiometric information influences the type of patterns distinguished by the algorithm. At the macro-scale, the impact of the use of the several bands available with Sentinel 2 from band 2 to band 8A is observed for the delimitation of urban footprint. For the meso and micro-scale, we decided to focus only on the panchromatic image of Pleiades which has the highest resolution (here 1 m) to work at this scale.

In order to observe the influence of the partitioning method over the characterization of urban areas, the block method and the moving window method are compared. The moving window method can be run in memory or use HDF5 file storage on the fly and run an increment PCA.

The modification of the window and/or pixel size influences the output resolution, and thus changes the number of sampled spatial frequencies (thus the diversity of the patterns). Regarding pixel size, at the macro-scale we use 10 m resolution sentinel 2 images, this resolution being commonly used for global products such as GHSL (10 m) or GUF (12 m). At the meso- and micro-scales, very high resolution images are required. We tested

Pleiades images at 1 and 3 m (resampled) resolution. We finally tested the algorithm on a 5 cm resolution UAV image for the detection of urban objects.

To better understand the impact of the windows size, we compare the smallest window size with a larger window size (31 is shown in this article). The smallest window size corresponding to the case whereby the number of spatial frequencies equals the number of PCA components (two with DC component = False and three with DC component = True). In this case, there is no need for dimensionality reduction, and so the output RGB corresponds to the r-spectra values. For the macro-scale and meso-scale, the smallest window size used is 5 and 9 respectively, which is the minimum window size that allows sampling a sufficient number of frequencies with FOTOTEX.

The DC component being the basis of the signal, corresponding to the average reflectance in each window, we compared its influence at the macro-scale and meso-scale.The normalization which makes the contribution of each frequency comparable and avoid an excessive influence of low frequency, we evaluate it's use to study urban areas at the macro-scale and meso-scale.

3. Results

3.1. Macro-Scale: Urban Footprint

3.1.1. Influence of Spectral Bands as Input

The SENTINEL 2 data allows for testing the different spectral bands to determine if one is more efficient than the other in delineating the urban macro-scale footprint. Tests showed that texture information from blue (B2), green (B3), red (B4), and red-edge (B5) bands are more discriminating to identify urban footprint than bands B6 to B8A. Using B2 to B5 as input allowed us to discriminate the urban footprint thanks to a RGB composite of the 3 first principal components (Figure 4).
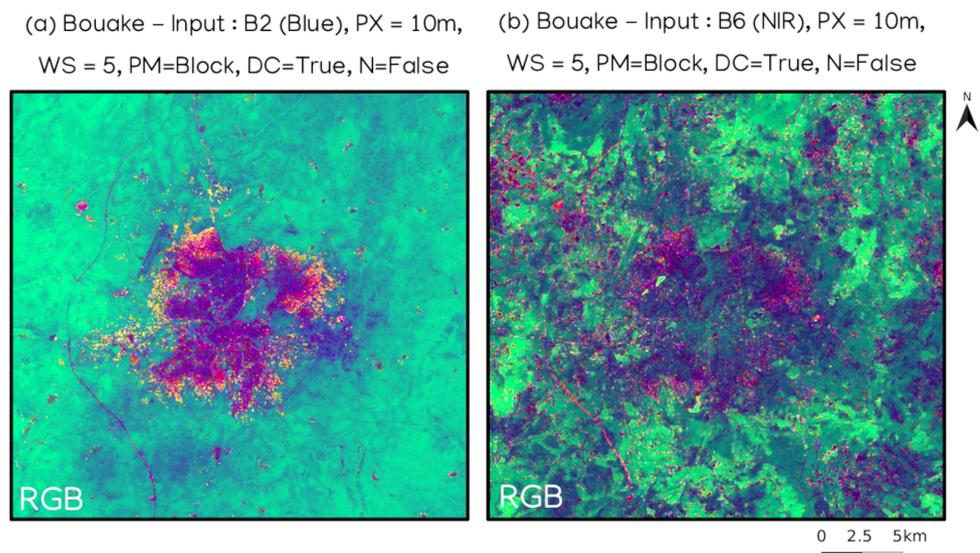
Figure 4. Influence of spectral bands used as input in the FOTOTEX algorithm for the characterization of urban areas at the macro-scale using Sentinel 2 images on Bouake. (**a**) B2 (blue) sentinel 2 band, and (**b**) B6 (NIR) Sentinel 2 band. (PX = pixel size, WS = window size, PM = partitioning method, DC = DC component, N = normalize).

3.1.2. Influence of Partitioning Method

The choice of the partitioning method, block or moving window, has no major impact on the identification of the urban footprint. Both methods allow for an accurate delineation of the urban footprint. However, they differ greatly in terms of computation time, especially for large images. For Bouake imagery (2801 × 2751 pixels, spatial resolution 10 m) with the smallest window size possible of 5, the block method took less than 25 s to compute the
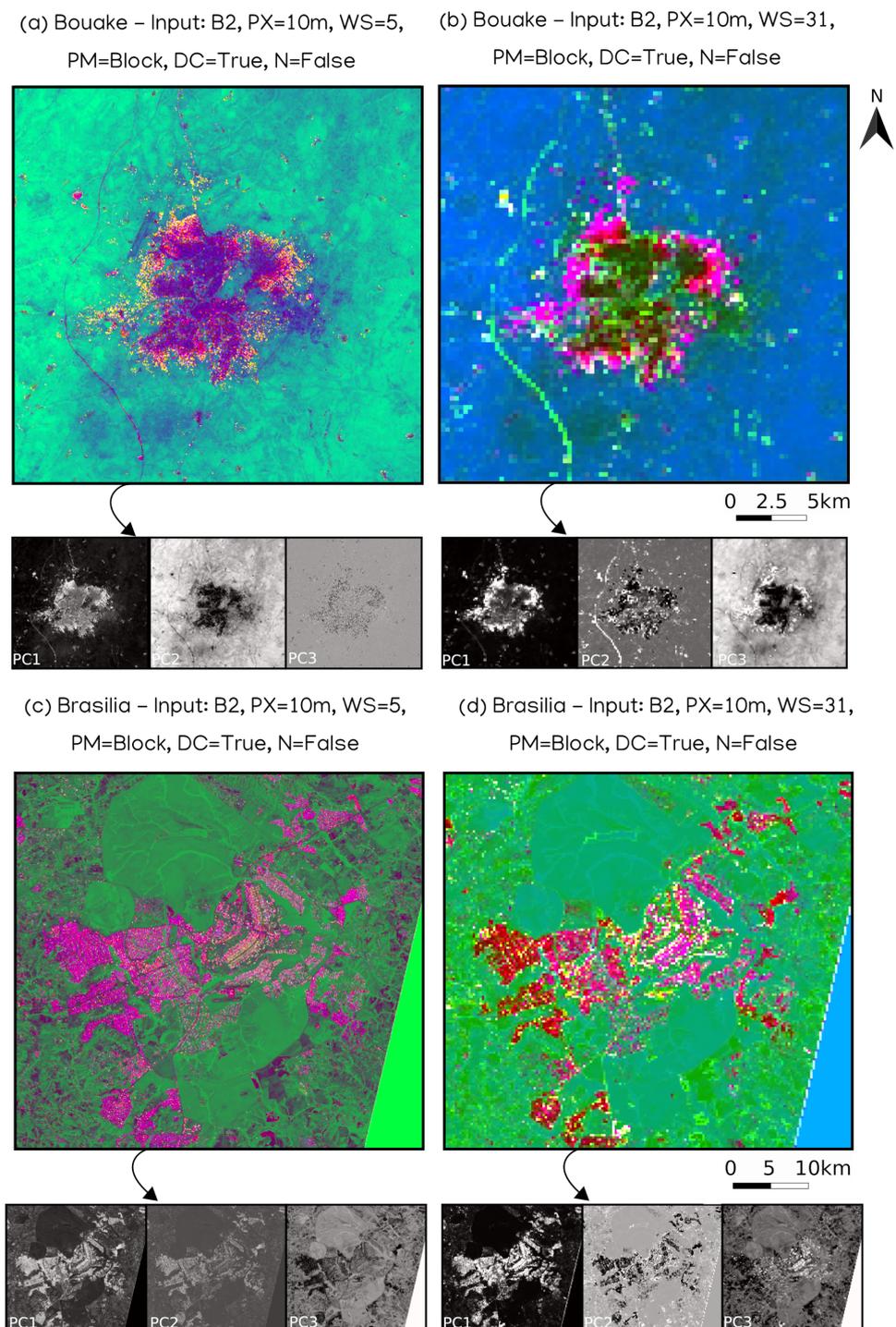
r-spectra and process the PCA while it took a little over 1 min in moving window (Table 2). For Brasilia imagery which is larger than the Bouake imagery (9306 × 6192 pixels, spatial resolution 10 m), the calculation time for the block method is around 2 min and a little over 9 min in moving window. Although computing time is more important using a moving window, this method has the advantage of providing texture maps with finer details. Default setting of moving window mode is also very demanding in computer power, saturating the computer memory. One way of avoiding this, is to work with HDF5 files, running an incremental PCA. Memory is then preserved but computing time degraded (from 9 to 15 min).

Table 2. Computing time using the FOTOTEX algorithm on Sentinel 2 images over Brasilia (scene of 9306 × 6192 pixels) and Bouake (scene of 2801 × 2751 pixels). Block and Moving window (with or without HDF5 + incremental PCA) methods are compared in this table.

	Bouake (Scene of 2801 × 2751 Pixels)	**Brasilia (Scene of 9306 × 6192 Pixels)**
Block Method	25 s	2 min
Moving window method	1 min	9 min
Moving window method (HDF5 and incremental PCA)	3 min	15 min

3.1.3. Influence of Window Size

To better understand the impact of the window size parameter, we compare the smallest window size possible (5 pixels) and a larger window size of 31 pixels using the block method. The output resolution is 50 m for the window size of 5 pixels and of 310 m for a window size of 31 pixels (Figure 5). The delineation of the urban footprint is much coarser in the second case because of the higher output resolution. However, when we observe the axes that compose the RGB, the three axes are useful to distinguish the urban footprint (Figure 5b) whereas for a smaller window size, axis 3 is not discriminating at all.

(a) Bouake – Input: B2, PX=10m, WS=5, PM=Block, DC=True, N=False

(b) Bouake – Input: B2, PX=10m, WS=31, PM=Block, DC=True, N=False

(c) Brasilia – Input: B2, PX=10m, WS=5, PM=Block, DC=True, N=False

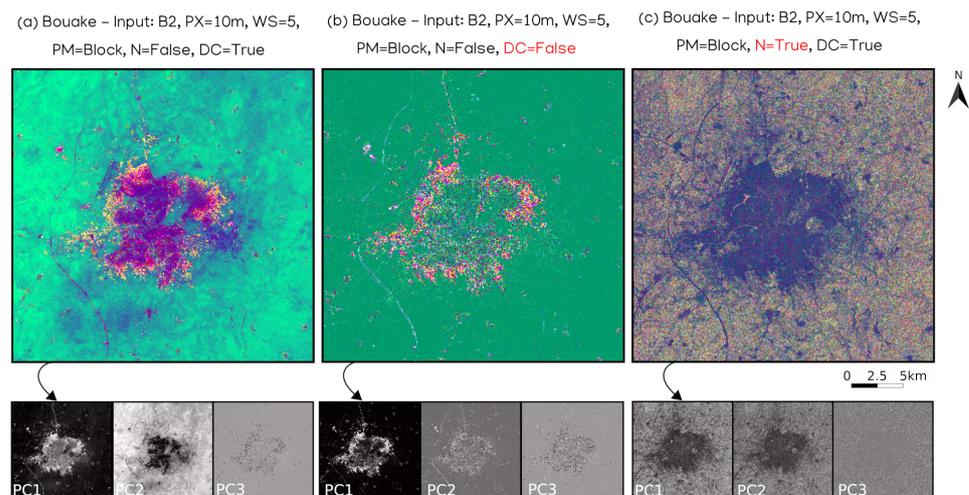(d) Brasilia – Input: B2, PX=10m, WS=31, PM=Block, DC=True, N=False

Figure 5. Influence of the window size parameter in the FOTOTEX algorithm for the characterization of urban areas at the macro-scale using Sentinel 2 images on Bouake and Brasilia. (**a**) B2 (blue) sentinel 2 band on Bouake with a window size of 5 pixels, (**b**) B2 (blue) sentinel 2 band on Bouake with a window size of 31 pixels, (**c**) B2 (blue) sentinel 2 band on Brasilia with a window size of 5 pixels and (**d**) B2 (blue) sentinel 2 band on Brasilia with a window size of 31 pixels . (PX = pixel size, WS = window size, PM = partitioning method, DC = DC component, N = normalize).

3.1.4. Influence of Keeping the DC Component

The use of the DC component is important for discriminating the urban footprint. First of all, we can clearly see that on the vegetation area, it seems very homogeneous without the DC component (Figure 6b). On the contrary, the use of the DC component allows

discriminating more contrast in this vegetation area. It also allows the identification of the urban limits (Figure 6a). However, in Figure 6b, the result without the DC component shows a more scattered urban area and does not allow the correct delineation of the whole urban area of Bouake. When we take a look at the different axes composing the RGB, we can observe that the axis 1 and 2 over Figure 6 make the urban footprint more visible than the third. None of the axes completely and homogeneously delimit the urban footprint (Figure 6b). Our tests on Brasilia showed that FOTOTEX performs similarly in the detection of urban footprint with or without DC component. However, on Bouake, the algorithm is more efficient with the DC component.

Figure 6. Influence of the dc component and the normalization parameters in the FOTOTEX algorithm for the characterization of urban areas at the macro-scale using Sentinel 2 images on Bouake. (**a**) B2 (blue) sentinel 2 band on Bouake with DC component (DC = True) and without normalization (N = False), (**b**) B2 (blue) sentinel 2 band on Bouake without DC component (DC = False) and without normalization (N = False) and (**c**) B2 (blue) sentinel 2 band on Bouake with DC component (DC = True) and normalization (N = True). The text in red highlights the modified parameter compared with image (**a**). (PX = pixel size, WS = window size, PM = partitioning method, DC = DC component, N = normalize).

3.1.5. Influence of the Normalization

Since textures have different frequencies between urban areas and vegetation, normalization will tend to smoothen the differences and thus reduce the algorithm's ability to distinguish the urban footprint. With normalization, the result is noisy and does not allow to delineate the urban footprint properly (Figure 6c). In both cases, PCA 3 is not relevant for the identification of urban footprint (Figure 6a,c). Urban footprint is overestimated on PCA 1 and 2, when the r-spectra are normalized (Figure 6c).

3.1.6. Extraction of the Urban Footprint over Optimal Configuration

The urban footprint was extracted using a threshold on the axis 1 of the PCA (PC1 of FOTOTEX in the RGB composite) in order to identify the pixels considered as urban. The quality of the detection performed by FOTOTEX is assessed by comparing the urban footprint with the GHSL product and GUF product over Brasilia and Bouake. For the partitioning method, the extraction results show an insignificant difference between the urban footprint extracted on an image using the block method and on an image using the moving window method. The only difference is related to the output resolution which will be finer with the moving window. In order to compare the different window sizes, the GHSL and the GUF was resampled according to the FOTOTEX results (Figure 7). For a window size of 5 (50 m output resolution), the urban footprint of Bouake is clearly underestimated with the GHSL and the GUF (Table 3). FOTOTEX estimates the urban footprint of Bouake around

106.7 km^2 while the GHSL and the GUF estimates the surface area at around 34.7 km^2 and 38.5 km^2 respectively. With a window size of 31 (310 m output resolution), the urban footprint is even more overestimated with FOTOTEX (157.2 km^2) than with the GHSL (34.9 km^2) and than the GUF (39.5 km)^2. For Brasilia, a larger window size (31) also implies an overestimation of the urban footprint (663.3 km^2 instead of 453.8 km^2 for window size 5 since mixed pixels are classified as "urban".
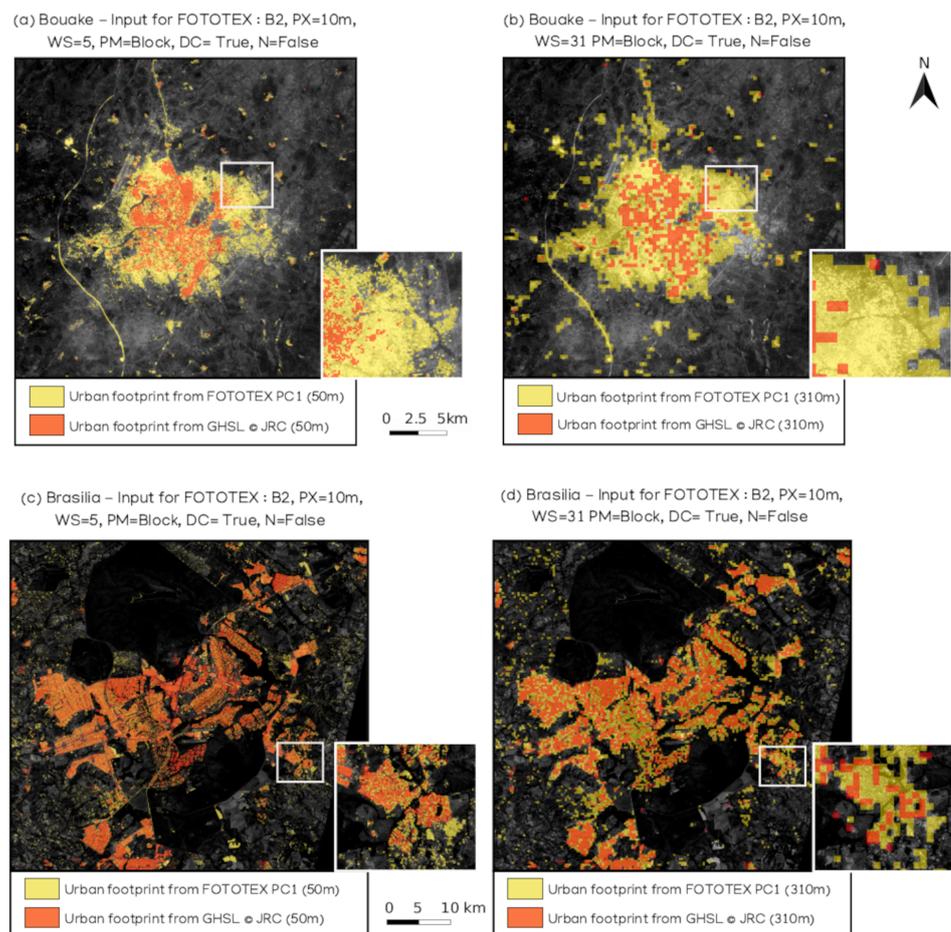
Figure 7. Comparison of urban footprints extracted from Sentinel 2 images on Bouake and Brasilia with FOTOTEX with the Global Human Settlement Layer (GHSL) product from JRC (© European Union) over different window sizes, (**a**) urban footprints on Bouake window size = 5 pixels, (**b**) urban footprints on Bouake window size = 31 pixels, (**c**) urban footprints on Brasilia window size = 5 pixels, (**d**) urban footprints on Brasilia window size = 31 pixels, Sentinel 2 Background © ESA (PX = pixel size, WS = window size, PM = partitioning method, DC = DC component, N = normalize).

Table 3. Surface of the urban footprints of Brasilia and Bouake from FOTOTEX using Sentinel 2 images under various configurations, and comparison with the Global Human Settlement Layer (GHSL) product (from JRC) and the Global Urban Footprint (GUF) product (from German Aerospace Center DLR) on both areas (PX =pixel size, WS = window size, PM = partitioning method, MW = Moving Window, DC = DC component, N = normalize).

	FOTOTEX			GHSL			GUF		
Input Information	MW WS = 5 (10 m)	Block WS = 5 (50 m)	Block WS = 5 (310 m)	10 m	50 m	310 m	10 m	50 m	310 m
Surface of the urban foorptint of Brasilia in km^2	457.2	453.8	663.3	338.9	338.5	342.5	385.2	385.6	391.8
Surface of the urban foorptint of Bouake in km^2	106.4	106.7	157.2	34.7	34.7	34.9	38.5	38.5	39.5

3.2. Meso-Scale: Urban Units

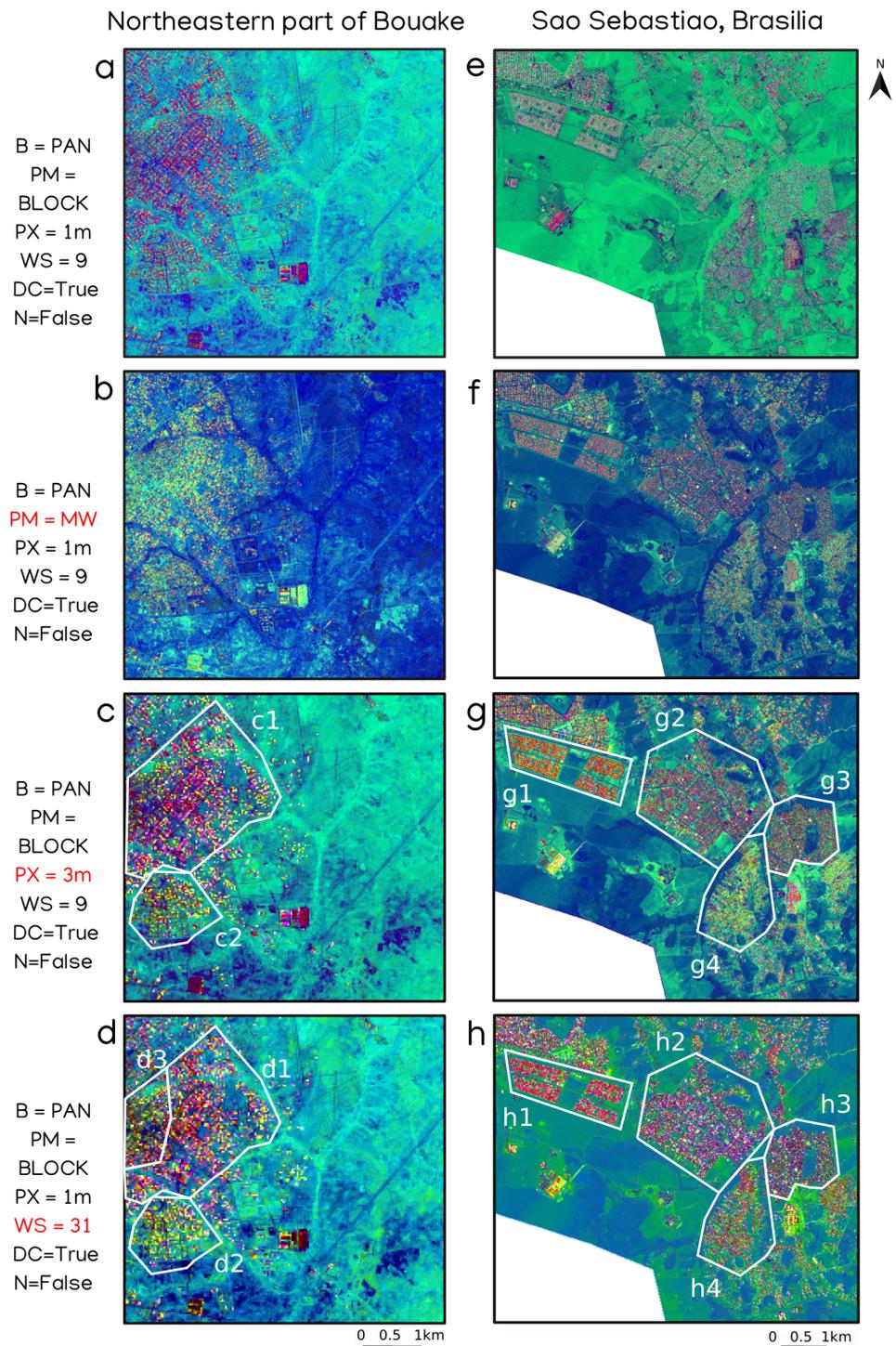3.2.1. Influence of Partitioning Method

First for analyzing urban areas at a meso-scale, we chose to compare the partitioning method between "block" and "moving window". The choice of partitioning method has no major impact to discriminate different neighbourhoods at the meso-scale on Bouake (Figure 8a,b) and on Sao Sebastiao (Figure 8e,f), despite different output resolutions (9 m in block vs 1 m in moving window). For the northeastern part of Bouake imagery (5666 × 8858 pixels) with a window size of 9 pixels, the block method took approximately 12 s to process the r-spectra and the PCA instead of 11 min with the moving window method.

3.2.2. Influence of the Pixel Size

The capability of FOTOTEX to discriminate urban units from input images at 1 m and 3 m resolution was compared. At 1 m resolution, there is no discrimination of urban units either on Bouake or Sao Sebastiao, the entire area showing the same textural signature on both sites (Figure 8a,e). At 3 m resolution, Bouake and Sao Sebastiao are characterized by distinct urban units with specific textural signature (Figure 8c,g). Two urban units stand out over northeastern Bouake with a pixel size of 3 m (Figure 8c). The first one b1 is distinguished by pink colors corresponding to a type of texture (Figure 8(c1)) and the second one is distinguished with yellow colors corresponding to another type of texture (Figure 8(c2)). It's even more visible on Sao Sebastiao, where 4 urban units stand out (Figure 8(g1)–(g4)). Here, when the resolution of the input image decreases from 1 m to 3 m, there is a better discrimination between the urban units over the Northeastern part of Bouake and Sao Sebastiao.

3.2.3. Influence of Window Size

Figure 8 shows that for a 1 m input resolution, a larger window size (9 vs. 31 pixels) allows for a better discrimination of urban units both on Bouake and Sao Sebastiao (Figure 8a,d,e,h). The influence of pixel and window size are very similar in the identification of urban units at the meso-scale. Indeed, the comparison between a 3 m input resolution and a window size of 31, reveals the same patterns of urban units (Figure 8c,d,g,h) on both areas. For Bouake, a larger window size even allows the identification of a third urban unit on Figure 8(d3).

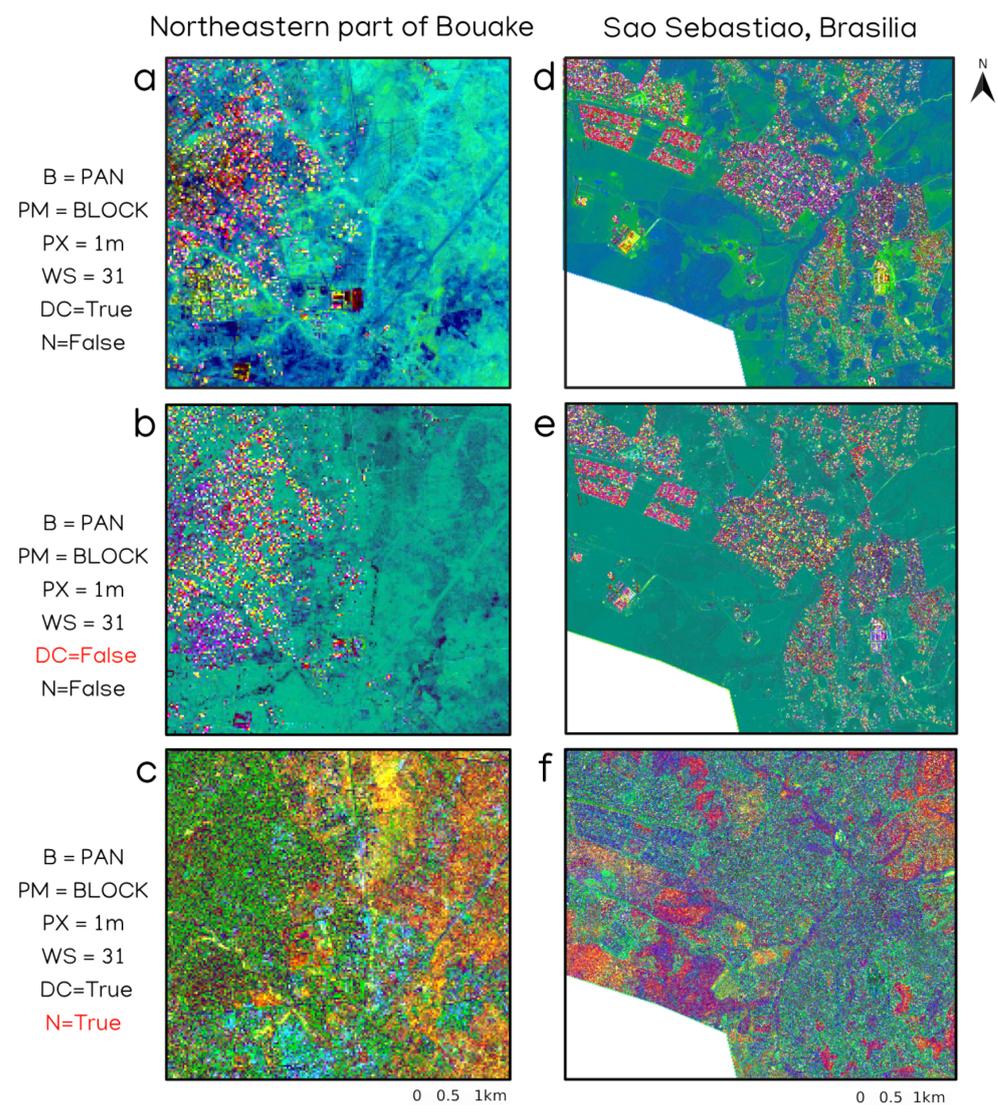Northeastern part of Bouake   Sao Sebastiao, Brasilia

Figure 8. Influence of various parameters (partitioning method, pixel size and window size) in FOTOTEX in the characterization of urban units at the meso-scale from Pleiades images on the northeastern part of Bouake and Sao Sebastiao, Brasilia. (**a,e**) use of block method, (**b,f**) use of moving window method, (**c,g**) use of a 3 m input pixel size over, (**d,h**) use of a window size of 31 pixels. The text in red highlights the modified parameter compared with image (**a**). (PX = pixel size, WS = window size, PM = partitioning method, MW = Moving Window, DC = DC component, N = normalize).

3.2.4. Influence of DC Component and Normalization

Figure 9a,b and d,e reveals that the use of the dc component is not strongly related to the identification of urban units in Bouake and Sao Sebastiao, where we can identify similar neighbourhoods with or without DC component. However, Figure 9a,d show that an effect of the DC component is to enhance the contrast between vegetated and urbanized areas, making the identification of urban units easier.

When the normalization is used (with or without DC component), the results show a noisy textural signal and do not allow to delineate the urban units properly (Figure 9c,f). Urban areas are clearly overestimated both on the northeastern part of Bouake (in green, Figure 9c) and on Sao Sebastiao (in blue-green, Figure 9f).

Figure 9. Influence of DC component and normalization in FOTOTEX in the characterization of urban units at the meso-scale from Pleiades images on the northeastern part of Bouake and Sao Sebastiao, Brasilia. (**a**,**d**) use of block method over Bouake and Brasilia, with DC component, without normalization, (**b**,**e**) use of block method over Bouake and Brasilia, without DC component, without normalization, (**c**,**f**) use of block method over Bouake and Brasilia, with DC component and normalization. The text in red highlights the modified parameter compared with image (**a**). (PX =pixel size, WS = window size, PM = partitioning method, MW = Moving Window, DC = DC component, N = normalize).

3.2.5. Comparison of Urban Units Extracted from FOTOTEX with Environmental Variables

As observed in the previous tests, the combination of a window size of 31, with DC component on a Pleiades image at 1m resolution allows us to distinguish urban units based on the textural signal. In this section, we investigate how variables such as building density, the percentage of built areas, the percentage of vegetated areas are related to the textural signature of these urban units. To do so, the contours of the urban units identified by different colors on the RGB composite (Figure 10a,c) are delineated using the urban footprint produced with FOTOTEX. Table 4 summarizes the mean values of texture from the PCA (PCA1, 2 and 3) used in the RGB composite, together with the percentages of built and vegetated areas (normalized by the surface of the unit) for each identified urban unit.

Table 4. Some metrics calculated for the urban units identified from texture using FOTOTEX on Pleiades images in Bouake and Brasilia (Sao Sebastiao). Correlations between the texture information represented by the mean of Principal Component PC 1, 2 and 3, and variables such as the percentage of built areas and vegetated areas are investigated.

	Urban Units	Mean PC1	Mean PC2	Mean PC3	Percentage of Built Area	Percentage of Vegetated Area
	1	7.13	−0.09	0.80	14.84	36.19
Sao Sebastiao	2	3.98	0.13	0.54	32.20	29.37
	3	3.62	0.37	0.00	38.87	33.51
	4	3.85	0.18	0.06	18.35	56.07
	5	3.79	0.89	0.20	10.30	59.69
	Urban Units	Mean PC1	Mean PC2	Mean PC3	Percentage of Built Area	Percentage of Vegetated Area
	1	6.24	−0.34	0.00	10.31	70.17
Northeastern part of Bouake	2	8.44	−0.37	−0.09	7.64	76.55
	3	5.47	0.02	−0.52	13.15	70.49
	4	6.29	−0.48	−0.27	12.00	72.52
	5	6.54	−2.59	−1.63	44.97	39.89

On Sao Sebastiao, texture shows 3 categories of urban units on the RGB composite: urban unit 1, urban units 2 and 3, and urban units 4 and 5 (Figure 10a). Urban units 2 and 3 correspond to the purple color textures (Figure 10a) associated with a high density of buildings (Figure 10b). They are also characterized by a percentage of built areas ranging from 32 to 38% and a percentage of vegetated areas ranging from 29 and 33% (Figure 10b, (Table 4). Urban units 4 and 5 present red textures interspersed with green textures reflecting a texture composed of more vegetation (Figure 10a) and associated with a lower density of buildings (Figure 10b). Urban units 4 and 5 have respectively between 10 and 18% of built area and 56 to 59% of vegetated area (Figure 10b, Table 4). Urban unit 1 presents a highly identifiable pink colour (Figure 10a) which is also associated with a very low density of buildings (Figure 10b). Urban units 1 is distinguished by a proportion of vegetation (36%) similar to units 2 and 3, but a percentage of built area (14%) comparable with urban units 4 and 5. For Sao Sebastiao, the textural discrimination of urban units is well explained by the use of landscape variables describing urban and vegetated areas.

On Bouake, 5 urban units are discriminated using texture (Figure 10c), although for unit 4, the texture reveals that 2 units could be extracted, and are not well represented but the contours of the urban footprint produced by FOTOTEX. However, building density is a good criteria for the subsetting of unit 4 into two different units, the northwestern part of unit 4 being characterized by a green colour on texture also corresponding to a high density of buildings (Figure 10d). As a consequence, texture information would lead us to gather urban units of Bouake into 3 groups: unit 1, 2 and the southeastern part of unit 4, unit 3

and the northwestern part of unit 4, and unit 5. Urban units 1, 2 and the southeastern part of urban unit 4 present a low density of buildings. Urban units 3 and northwestern part of unit 4 have high density of buildings. Urban units 1 to 4 have similar proportions of vegetation (between 70 and 77%) and built areas (between 7 and 13%) (Table 4). Urban unit 5 has a very low density of buildings, combined with a high percentage of built-up area and a low percentage of vegetated area compared to urban units 1 to 4 (Table 4). For Bouake, texture information is well explained by the density of buildings while the percentages of urban and vegetated areas are not related to texture.
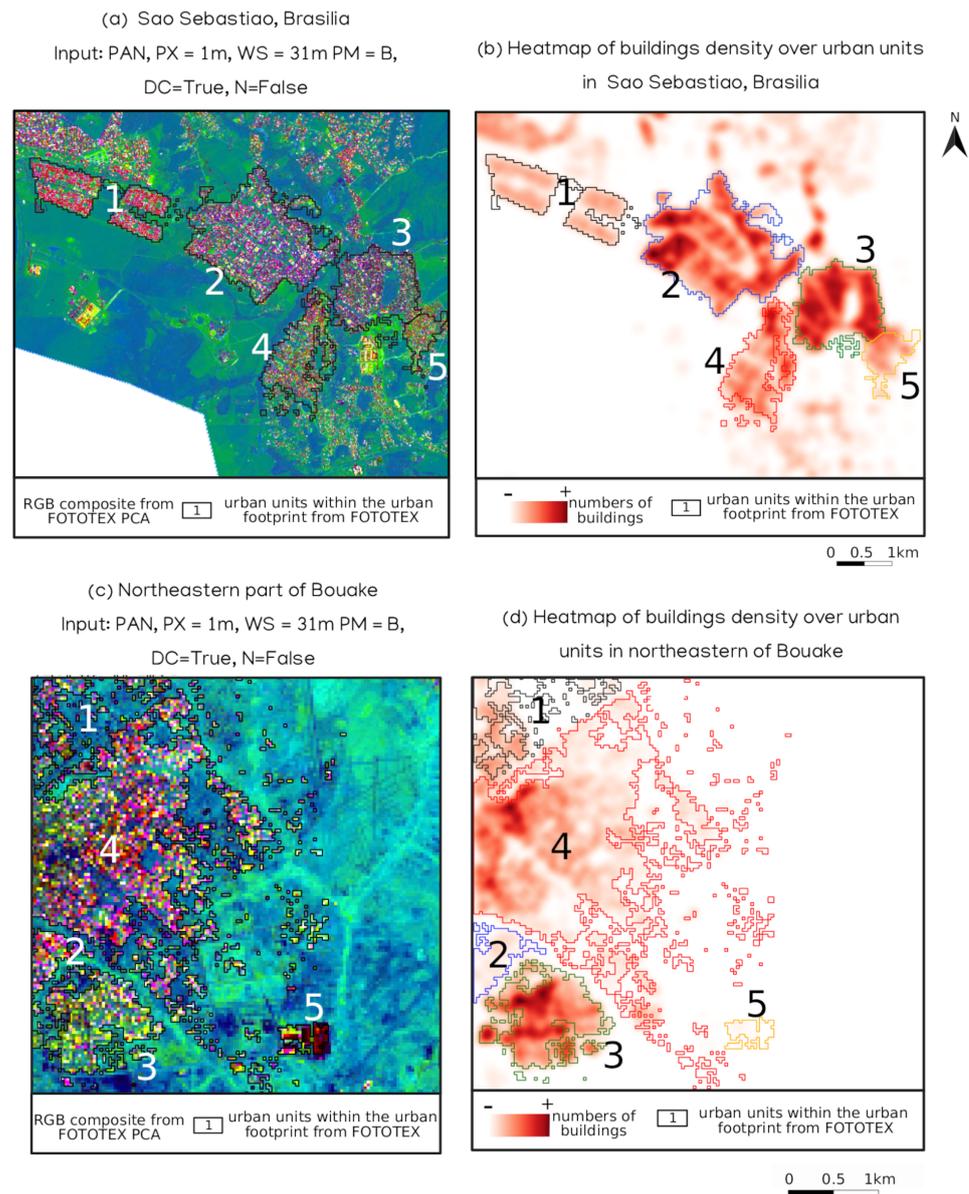

Figure 10. Study of the relationship between texture information at the meso-scale extracted using FOTOTEX on Pleiades images on Sao Sebastiao, Brasilia (**a**) and the northeastern part of Bouake (**c**), with the density of buildings represented as heatmaps using building contours from the SEDUH database (http://www.seduh.df.gov.br/ accessed on 20 September 2020) for Brasilia (**b**) and Open-StreetMap for Bouake (**d**). On this figure, black/colored lines and numbers are urban units extracted from the combination of texture and contours from the urban footprint produced with FOTOTEX and Sentinel 2.

### 3.3. Micro-Scale: Buildings Detection

3.3.1. Influence of Parameters

The results have shown previously that some parameters are useless and others very useful to identify urban landscapes at the macro-scale, meso-scale and therefore at the micro-scale. Normalization is therefore not necessary, but the DC component must be kept during the calculation. Then, the use of the moving window partitioning method is more efficient to identify small objects (Figure 11b,d). The delimitation of urban objects is more precise with the moving window method, both for large buildings as well as smaller houses surrounded by vegetation, or even cars (Figure 11b)
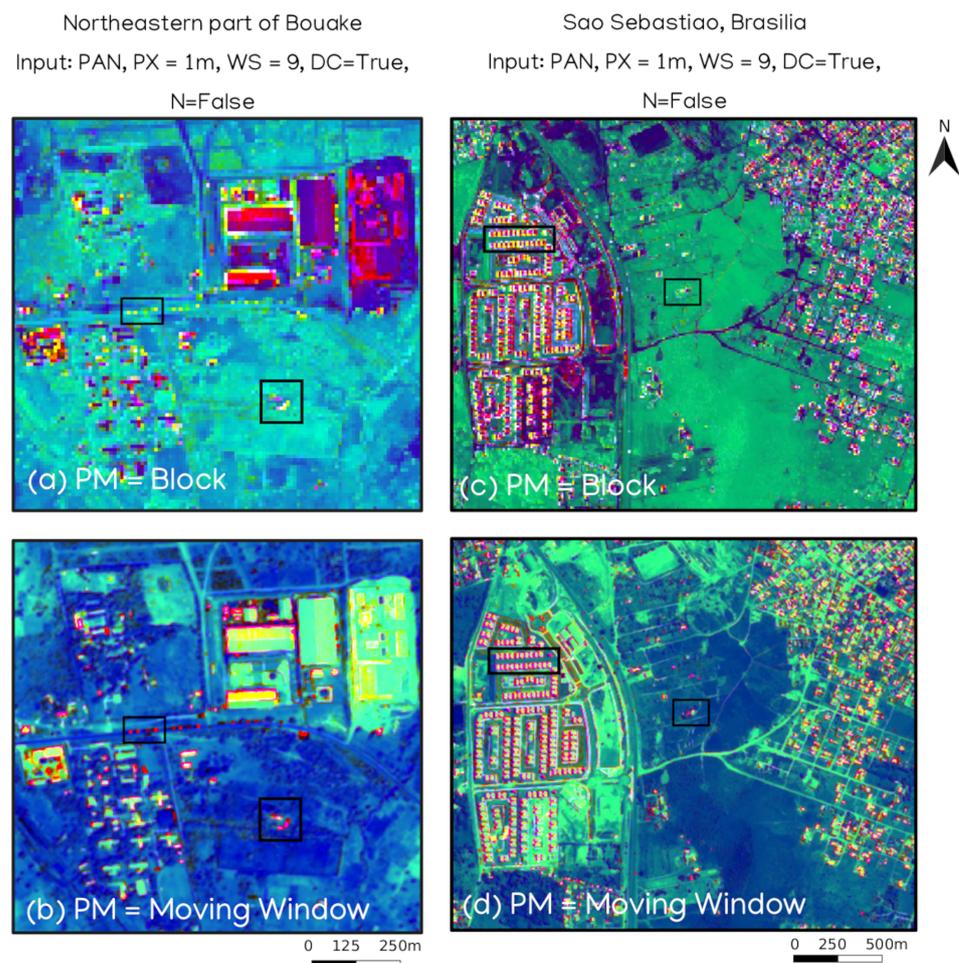
Figure 11. Influence of the partitioning method for the micro-scale characterization of urban objects on Bouake and Brasilia using Pleiades images (**a**) block method on Bouake, (**b**) block method on Brasilia, (**c**) moving window method on Bouake and (**d**) moving window method on Brasilia (PX = pixel size, WS = window size, PM = partitioning method, DC = DC component, N = normalize).

Even if satellite images allow us to identify urban objects, we tried the algorithm over a drone image (0.05 m spatial resolution) being available on Bouake pilot site. The available drone image was acquired on may 2020, with a DJI Phantom 4 Pro V2 drone and a sensor CMOS 1,2/3 in red, green and blue wavelengths. Figure 12 shows that FOTOTEX makes it possible to identify small urban objects very finely using texture. The principal components of the FOTOTEX (Figure 12c) output show a strong discrimination between landcover types (pervious and impervious surfaces)
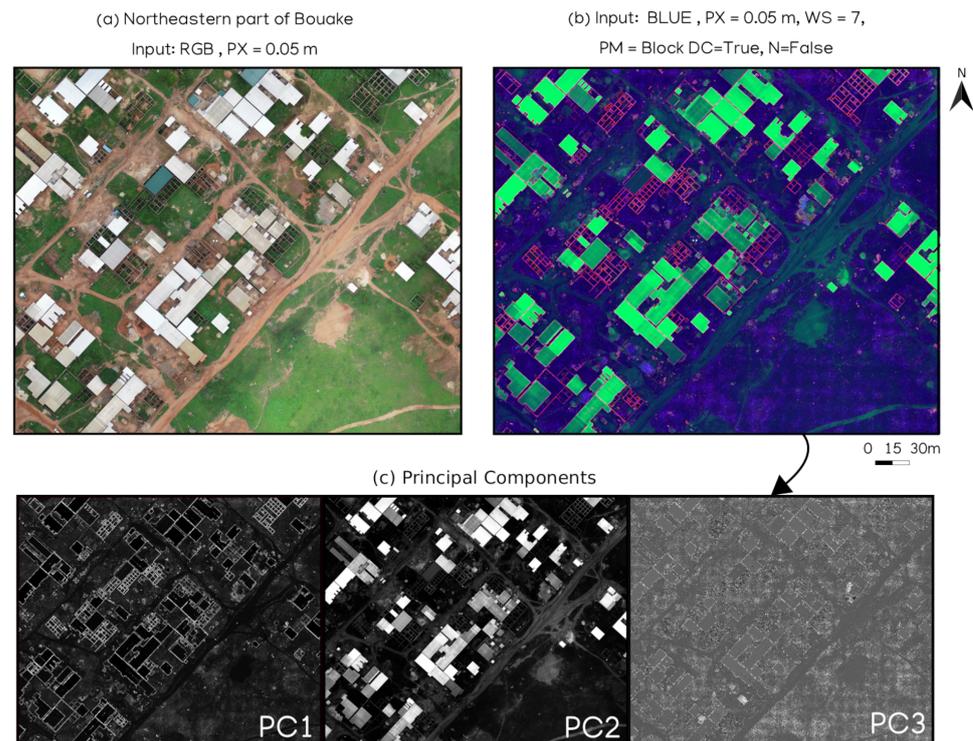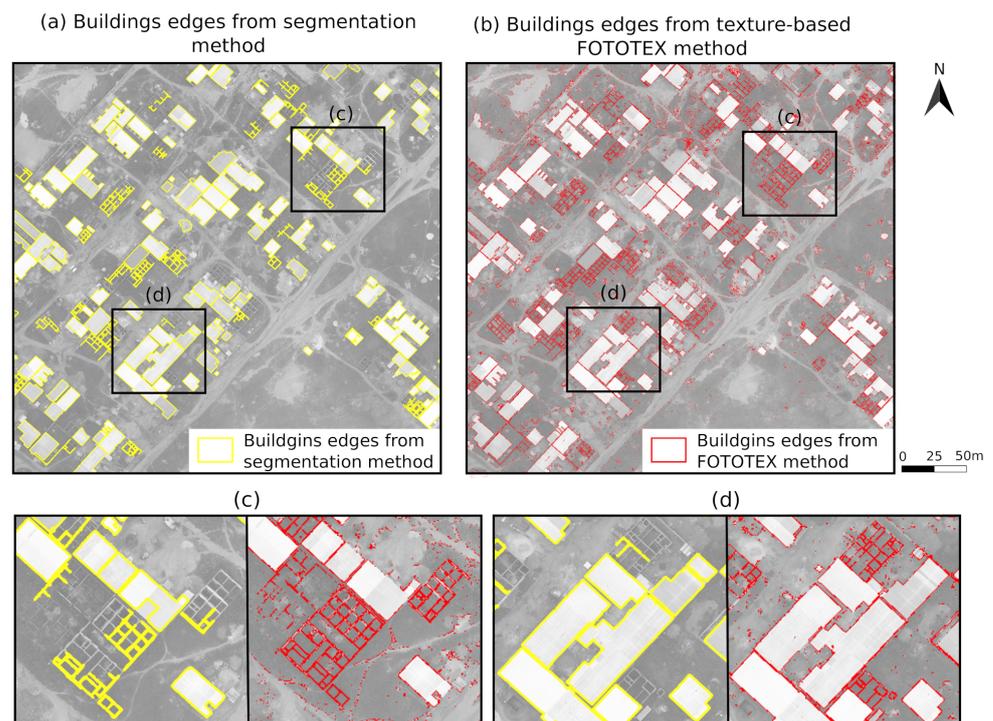
Figure 12. Implementation of the FOTOTEX algorithm on (**a**) a UAV image captured over the city of Bouake (**b**) showing the RGB composite produced by FOTOTEX (**c**) and the three Principal Components. (PX = pixel size, WS = window size, PM = partitioning method, DC = DC component, N = normalize).

3.3.2. Extraction of Urban Objects

In order to assess the quality of building detection using FOTOTEX at the micro-scale, we compared the results obtained on Bouake with two different approaches: a segmentation (carried out with the Orfeo Toolbox using the Large Scale Generic Region Merging algorithm based on three criteria: spectral values, shape and radius) and the texture-based FOTOTEX method implemented on the blue band of UAV image with window size of 7, block method, with the dc component active and without normalization. The building edges where extracted with a threshold. Here we only provide a qualitative assessment based on a difference map of building edges from the two approaches.

Figure 13 shows the building edges from FOTOTEX in red and the building edges from the segmentation in yellow. We can see that both methods perform similarly on the detection of largest buildings, while FOTOTEX identifies a larger amount of small buildings. This is well illustrated by the fact that FOTOTEX is able to extract much more unfinished houses characterized by the presence of vertical walls without roofs than the segmentation (Figure 13c,d).

Figure 13. Extraction of buildings edges with (**a**) a segmentation method carried out with the Large Scale Generic Region Merging algorithm (Orfeo Toolbox) and (**b**) the FOTOTEX method implemented on the blue band of UAV image. (**c**,**d**) zooms on two different areas to compare detection by FOTOTEX and segmentation.

4. Discussion

4.1. Recommended Parameters for the Study of the Urban Landscapes at Three Spatial Scales

4.1.1. Macro-Scale

Although FOTOTEX relies on textural information, the use of frequency analysis on satellite imagery is also influenced by radiometry especially over heterogeneous areas where urban, bare soil and vegetation are present. A simple, homogeneous land cover (e.g., monotype vegetation) will produce an homogeneous signal in terms of texture and frequency content regardless of the spectral band used in the process. However, a more complex structure with heterogeneous land cover will result in several patterns (grass, bush, bare ground, trees) and will produce a different signal and a more complex frequency content depending on the radiometric band [46]. For complex and heterogeneous areas such as the cities of Brasilia and Bouake, several different frequency patterns are identified. Radiometry will hence have an influence over the textural signal and the related frequency content. As a consequence, for the identification of urban footprint using a Fourier transform-based textural approach, it is strongly recommended to use a spectral band where the radiometric discrimination between the urban footprint and other landcover types is strong. In the case of urban areas, our results show that Sentinel 2 blue (B2), green (B3), red (4) and red-edge (B5) bands are the most efficient to extract urban footprint, with similar results. Moreover, radiometric information also has an important influence on the types of patterns that are distinguished with FOTOTEX.

Both partitioning methods will allow to extract urban footprints. The user's choice regarding the partitioning method will be driven by the best compromise between calculation power available, calculation time and required output resolution. On one hand, the moving window method is time consuming but provides an output urban footprint at high resolution. On the other hand, the block method is faster, but produces a less resolved output. However, we showed that both types of urban footprints are usable and

that the high resolution provided by the moving window method is not determining in the assessment of the urban extent (Table 3). The extracted urban footprint is considered to be usable if the output resolution is not too coarse to avoid the overestimation of the urban extent (Table 3).

Here, the better parametrization is a compromise between the pixel size, the window size, and the expected output resolution, for a given partitioning method. In any case, the moving window method produces an output footprint at the resolution of the input image, and when using block method, the smallest window size has to be chosen (here, 5 pixels) to produce the best output resolution and minimize overestimation of urban extent. The result section over macro-scale shows that it is recommended to keep DC component active. The integration of the DC component will take into account the reflectance within the window and will combine it with the texture information. If the DC component is not taken into account, the discrimination is only function of the textural information, which can be similar between windows, therefore the integration of the average reflectance allows having an additional information to discriminate similar patterns. In this sense, the use of FOTOTEX with DC component is efficient in the extraction of urban footprint.

The r-spectra table must not be normalized for the extraction of the urban footprint. In the case of a heterogeneous area, with strong variance differences, it is preferable to be able to keep these differences and thus prevent smoothing by constraining the window values after PCA within a defined interval. The differences in values between the windows will allow the discrimination of the urban from other types of land cover patterns.

4.1.2. Meso-Scale

The main insight from the meso-scale results is the importance of the window size and the output resolution in order to identify urban units within the urban footprint. We showed that approximately 30 m is the most appropriate output resolution for this purpose. As a consequence the moving window method is not suitable for the meso-scale analysis. Theoretically, there are three ways to achieve a 30 m output resolution (in block method): (1) A 10 m resolution Sentinel 2 input image, with a window size of 3 pixels, (2) a 1 m pleiades image with a window size of 31 pixels or (3) a 3 m pleiades image with a window size of 9 pixels. At the meso-scale, unlike Sentinel 2, the use of Pleiades imagery at a finer resolution makes it possible to better distinguish the small objects that make up urban units. So solution 1 is ruled out. Urban units are then distinguished by the different urban objects that compose them (buildings, roads, bare soil, vegetation...) and by their arrangement in space (the texture). The only way of sampling the texture with a sufficient number of frequencies is to favor window size, to some extent in order to maintain a reasonable output resolution. Solutions 2 and 3 fullfill these requirements. The different tests for the identification of urban units by changing the pixel size or the window size show that these two parameters are extremely related. Adapting the window size can be used to compensate for a lack of access to an image at 1 m resolution.

At the meso-scale, we recommend the use of DC component and the deactivation of normalization to better discriminate urban units, for the same reasons as at the macro-scale.

4.1.3. Micro-Scale

At the micro-scale, the same choices of parameters are also efficient. Our tests show that it is recommended to integrate the DC component and to avoid normalization. The pixel size and the window size are also closely related. We observe that the partitioning method allows better defining the outline of urban objects. The first tests on a UAV image show a very strong potential to delineate the smallest objects even better with a pixel size in centimeters.

4.1.4. Global Recommendation

Table 5 summarizes the choice of parameters to be used in FOTOTEX for the study of the urban environment at three different scales, from the examples of Brasilia and Bouake.

In view of the results, it is recommended to choose the block partitioning method because it is the best compromise between reliable output and minimal calculation time. In terms of pixel choice, even if we are limited by access to satellite data, a pixel size of 10 m is sufficient to delimit the urban footprint. To analyze urban meso-scale units, a pixel size of 1 m is relevant but a pixel size of 3 m is also usable, to overcome the non-access to Pleiades with the possibility for example to use Planets products. However, an image with a resolution of 10 m such as Sentinel 2 allows distinguishing a density gradient at the meso-scale. Concerning the choice of the window size, for the scale of the urban footprint, it is recommended to not have a too large window since the detection of the urban footprint will be too coarse. At the scale of the urban units, it is recommended to have an analysis window which allows the identification of urban objects, therefore not too small, but large enough to allow the observation of a repetition of the pattern created by the structure of the urban objects. At the micro-scale, as the pixel size can be in centimetres, it is preferable to have a relatively small window size in order to keep the advantages of a very high spatial resolution. Finally, over the three scales of analysis, it is recommended to keep the DC component in order to use reflectance information by associating them with textural information and to not-normalized to avoid smoothing frequencies informations when analyzing urban landscapes.

Table 5. Summary of the recommendations for the choice of the optimal parameters to be used in FOTOTEX for the study of the urban landscapes at three different scales.

Parameters	Macro Scale	Meso scale	Micro scale
Partitioning Method (PM)	Block	Block	Block
Pixel Size (PX)	10 m	Between 1 m and 3 m	Under 1 m
Window Size (in pixel) (WS)	To be adapted	To be adapted (better result with larger window)	To be adapted
Keep DC Component (DC)	True	True	True
Normalize (N)	Non-normalized	Non-normalized	Non-normalized

4.2. The Contribution of the Method to Characterize Urban Landscapes at Three Scales

The analysis of urban landscapes using remote sensing is complex since it is characterized by urban objects with different structures and materials, at different scales. In this study, we chose to work at three scales, the textural approach developed in FOTOTEX contributing in various ways depending on the scale.

A lot of work has been done to produce global urban footprints from remote sensing data, using various methodological approaches and data. Studies dedicated to the characterization of urban units ("neighbourhoods") and urban objects are more scarce. Currently, most of the studies of urban footprints using satellite imagery are based on large volumes of data, and follow supervised procedures (not always open) with training and validation data mainly available on european territories. These approaches require high computational capacities that are not always available in developing countries, where the lack of validation is critical. Our study provides an alternative to those issues in several manners:

- we provide an open algorithm implementing a fully unsupervised procedure to characterize urban areas at three scales. This is particularly important to work on areas where training data are missing.
- for all the analysis scales, our methodological framework relies on single date images, strongly reducing the need for downloading and storing large volumes of satellite images
- using single images, the computational power required to run the algorithm is strongly reduced.

We showed that the use of various types of images can provide valuable information at the three scales. For instance, very high resolution images such as Pleiades or Planetscope

are not available for free to the users while high resolution Sentinel data are easily available. We showed that both types of images can be useful to study urban at the meso-scale (density gradient, identification of urban units). This part of the work that aimed at analyzing the urban at the meso-scale is the main originality of the study.

High resolution Sentinel 2 allows a meso-scale analysis, especially on the morphology of the city. On Bouake, we identify a classic density gradient (Clark, 1951) of a monocentric city where there is a general decrease in densities from the center to the periphery. Figure 5a illustrates this notion of gradient through pink/purple colors in the center going towards yellow in the periphery. This decrease in density at the periphery is also associated with an increase in the proportion of vegetation and bare soil. In Brasilia, this density gradient is not observed in the same way since the creation and construction of the city are particular. The city first follows the Pilot Plan of urban planner Lucio Costa, which is organized around two perpendicular axes giving the city the shape of an airplane. It is gradually surrounded by some twenty "satellite cities" [54]. The city is therefore very heterogeneous with a fragmentation of urban units. Figure 5b thus shows the Pilot Plan, with yellow colors (associated rather with textures containing a proportion of vegetation) and the more or less dense "satellite cities" with variation of pink and purple colors. Figure 5 illustrates that a larger window size has two effects: in Bouake, it better discriminates urban gradients within the urban footprint at the meso-scale, while in Brasilia, it discriminates urban units with different densities.

Using very high resolution at the meso-scale, we investigated the relationship between texture of urban units and three variables (building density, percentage of built area and vegetated area). The morphology of the cities being the result of the physical arrangement of man-made and natural objects in space, the texture analyzes the physical expression of this arrangement, described here by the variables we calculated on Brasilia and Bouake. In Bouake, the city being homogeneous, texture is strongly related with building density, corresponding to the urban density gradient. In Brasilia, the three variables are related to texture since the morphology is more heterogeneous and more complex than just an urban density gradient. However, we only discuss the relationship with the three variables we tested. Indeed, we studied vegetation and urban percentage while at the meso-scale, urban units are more complex. Other variables should be investigated in relationship with texture.

Regarding urban object detection at the micro-scale, the example of Bouake shows that the FOTOTEX algorithm has several advantages, with respect to methods such as segmentation for instance. First, the implementation of FOTOTEX requires much less intervention from the user than a segmentation process which can be time consuming and not easily reproducible on different sites. Second, our results show that the performance of FOTOTEX in building extraction is not dependent on the size of the objects. Building of any size are equally identified by FOTOTEX. On the contrary, the very principle of segmentation makes it less efficient in the detection of various types of building since once the criteria of shape and radius have been determined, it is only suitable for some buildings corresponding to these criteria. Some limitations due to confusions between buildings and bare soils with FOTOTEX can alter the quality of detection at this scale. However, we consider that FOTOTEX provides a more realistic assessment of urban objects number and distribution at micro-scale.

Environmental variables extracted over urban areas usually are used as input variables in other thematic studies. For instance, descriptors of the meso-scale have been used by Zhou, G [55] to investigate the relationship between urban neighbourhoods and health issues. If knowledge about urban footprints is valuable, some processes such as exposure to pathogens, pathogens transmission or population well-being require studying urban areas following a multi-scale approach. Current global products over urban areas do not offer this possibility, while our approach represents a simple, unsupervised tool that can be implemented in any kind of urban context, given an appropriate set of parameters. We

focus on characterizing urban areas with the objective of providing urban variables for wider applications.

5. Conclusions

We developed the open FOTOTEX algorithm, available at https://pypi.org/project/fototex/, and demonstrated the potential of an unsupervised texture-based method for the multi-scale characterization of urban areas on Bouake and Brasilia. At macro-scale, the cross-comparison of the results between FOTOTEX, GHSL and GUF revealed that FOTOTEX allows a more realistic delineation of the urban surface with a single Sentinel 2 image than common global products over urban areas such as GHSL, and GUF. At meso-scale, the use of Pleiades images allows identifying urban units based on texture, which is the main originality of this work given the lack of studies at this scale of analysis. The textural information at meso-scale appeared to be strongly linked to metrics such as building density and the proportion of built-up and vegetated areas. At micro-scale, FOTOTEX proved its efficiency at detecting urban objects. For each scale of analysis, the parametrization of the algorithm was tested and the best configurations were determined. The method we propose at three scales can take advantage of the diversity of remote sensing data which is currently at the disposal of the community of urban experts. Global recommendations have been made, based on the different results, to guide future users of the FOTOTEX algorithm to characterize urban landscapes at the multi-scale. The present approach showed a strong potential to analyze the complexity of urban landscapes, on any kind of urban environment. This method answers the objective of studying complex objects using one single method while reducing calculation time, the number of images used, and overcoming the absence of training and validation datasets. Hence, FOTOTEX could be an interesting alternative, especially in southern countries.

References

1. Jensen, J.R.; Cowen, D.C. Remote Sensing of Urban/Suburban Infrastructure and Socio-Economic Attributes. In *The Map Reader*; John Wiley & Sons, Ltd.: Chichester, UK, 1999; pp. 153–163. [CrossRef]
2. Maktav, D.; Erbek, F.S.; Jürgens, C. Remote sensing of urban areas. *Int. J. Remote Sens.* **2005**, *26*, 655–659. [CrossRef]
3. Weeks, J.R. Defining Urban Areas. In *Remote Sensing of Urban and Suburban Areas*; Rashed, T., Jürgens, C., Eds.; Remote Sensing and Digital Image Processing; Springer Netherlands: Dordrecht, The Netherlands, 2010; pp. 33–45. [CrossRef]
4. Quattrochi, D.A.; Goodchild, M.F. *Scale in Remote Sensing and GIS*; CRC Press: Boca Raton, FL, USA, 1997.

5.   Gamba, P.; Herold, M. Global Mapping Of Human Settlement: Experiences, Datasets, and Prospects: Book Reviews. *Photogramm. Rec.* **2009**, *25*, 205–207. [CrossRef]

6.   Weng, Q. *Remote Sensing and GIS Integration: Theories, Methods, and Applications*; OCLC: 318409264; McGraw-Hill: New York, NY, USA, 2010.

7.   Zhu, Z.; Zhou, Y.; Seto, K.C.; Stokes, E.C.; Deng, C.; Pickett, S.T.A.; Taubenböck, H. Understanding an urbanizing planet: Strategic directions for remote sensing. *Remote Sens. Environ.* **2019**, *228*, 164–182. [CrossRef]

8.   Li, X.; Gong, P.; Zhou, Y.; Wang, J.; Bai, Y.; Chen, B.; Hu, T.; Xiao, Y.; Xu, B.; Yang, J.; et al. Mapping global urban boundaries from the global artificial impervious area (GAIA) data. *Environ. Res. Lett.* **2020**, *15*, 094044. [CrossRef]

9.   Reba, M.; Seto, K.C. A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change. *Remote Sens. Environ.* **2020**, *242*, 111739. [CrossRef]

10. Zhang, X.; Liu, L.; Wu, C.; Chen, X.; Gao, Y.; Xie, S.; Zhang, B. Development of a global 30 m impervious surface map using multisource and multitemporal remote sensing datasets with the Google Earth Engine platform. *Earth Syst. Sci. Data* **2020**, *12*, 1625–1648. [CrossRef]

11. Pesaresi, M.; Blaes, X.; Ehrlich, D.; Ferri, S.; Gueguen, L.; Halkia, M.; Kauffmann, M.; Kemper, T.; Lu, L.; Marin-Herrera, M.A.; et al. A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 2102–2131. [CrossRef]

12. Pesaresi, M.; Corbane, C.; Julea, A.; Florczyk, A.J.; Syrris, V.; Soille, P. Assessment of the Added-Value of Sentinel-2 for Detecting Built-up Areas. *Remote Sens.* **2016**, *8*, 299. [CrossRef]

13. Pesaresi, M.; Syrris, V.; Julea, A. A New Method for Earth Observation Data Analytics Based on Symbolic Machine Learning. *Remote Sens.* **2016**, *8*, 399. [CrossRef]

14. Corbane, C.; Pesaresi, M.; Politis, P.; Syrris, V.; Florczyk, A.J.; Soille, P.; Maffenini, L.; Burger, A.; Vasilev, V.; Rodriguez, D.; et al. Big earth data analytics on Sentinel-1 and Landsat imagery in support to global human settlements mapping. *Big Earth Data* **2017**, *1*, 118–144. [CrossRef]

15. Corbane, C.; Lemoine, G.; Pesaresi, M.; Kemper, T.; Sabo, F.; Ferri, S.; Syrris, V. Enhanced automatic detection of human settlements using Sentinel-1 interferometric coherence. *Int. J. Remote Sens.* **2018**, *39*, 842–853. [CrossRef]

16. Marconcini, M.; Metz-Marconcini, A.; Üreyen, S.; Palacios-Lopez, D.; Hanke, W.; Bachofer, F.; Zeidler, J.; Esch, T.; Gorelick, N.; Kakarla, A.; et al. Outlining where humans live - the world settlement footprint 2015. *Sci. Data* **2015**, *20*, 242. [CrossRef] [PubMed]

17. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [CrossRef]

18. Wang, P.; Huang, C.; de Colstoun, E.B.; Tilton, J.; Tan, B. *Global Human Built-up and Settlement Extent (HBASE) Dataset From Landsat*; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2017. [CrossRef]

19. Felbier, A.; Esch, T.; Heldens, W.; Marconcini, M.; Zeidler, J.; Roth, A.; Klotz, M.; Wurm, M.; Taubenböck, H. The global urban footprint—Processing status and cross comparison to existing human settlement products. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 4816–4819. [CrossRef]

20. Esch, T.; Heldens, W.; Hirner, A.; Keil, M.; Marconcini, M.; Roth, A.; Zeidler, J.; Dech, S.; Strano, E. Breaking new ground in mapping human settlements from space—The Global Urban Footprint. *ISPRS J. Photogramm. Remote Sens.* **2017**, *134*, 30–42. [CrossRef]

21. Lehner, A.; Naeimi, V.; Steinnocher, K. Sentinel-1 for Urban Areas—Comparison between Automatically Derived Settlement Layers from Sentinel-1 Data and Copernicus High Resolution Information Layers. In Proceedings of the 3rd International Conference on Geographical Information Systems Theory, Applications and Management, Porto, Portugal, 27–28 April 2017; SCITEPRESS—Science and Technology Publications: Porto, Portugal, 2017; pp. 43–49. [CrossRef]

22. Puissant, A.; Hirsch, J.; Weber, C. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *Int. J. Remote Sens.* **2005**, *26*, 733–745. [CrossRef]

23. Ouma, Y.; Tateishi, R.; Sri-Sumantyo, J. Urban features recognition and extraction from very-high resolution multi-spectral satellite imagery: a micro–macro texture determination and integration framework. *IET Image Process.* **2010**, *4*, 235. [CrossRef]

24. Feng, Q.; Liu, J.; Gong, J. UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis. *Remote Sens.* **2015**, *7*, 1074–1094. [CrossRef]

25. Lang, M.; Alleaume, S.; Luque, S.; Baghdadi, N.; Féret, J.B. Monitoring and Characterizing Heterogeneous Mediterranean Landscapes with Continuous Textural Indices Based on VHSR Imagery. *Remote Sens.* **2018**, *10*, 868. [CrossRef]

26. Dell'Acqua, F.; Gamba, P. Texture-based characterization of urban environments on satellite SAR images. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 153–159. [CrossRef]

27. Corbane, C.; Faure, J.F.; Baghdadi, N.; Villeneuve, N.; Petit, M. Rapid Urban Mapping Using SAR/Optical Imagery Synergy. *Sensors* **2008**, *8*, 7125–7143. [CrossRef]

28. Esch, T.; Marconcini, M.; Felbier, A.; Roth, A.; Heldens, W.; Huber, M.; Schwinger, M.; Taubenböck, H.; Müller, A.; Dech, S. Urban Footprint Processor-Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1617–1621. [CrossRef]

29. Ansari, R.A.; Buddhiraju, K.M.; Malhotra, R. Urban change detection analysis utilizing multiresolution texture features from polarimetric SAR images. *Remote Sens. Appl. Soc. Environ.* **2020**, *20*, 100418. [CrossRef]

30. Florczyk, A.J.; Ferri, S.; Syrris, V.; Kemper, T.; Halkia, M.; Soille, P.; Pesaresi, M. A New European Settlement Map From Optical Remotely Sensed Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1978–1992. [CrossRef]

31. Forget, Y.; Shimoni, M.; Gilbert, M.; Linard, C. Mapping 20 Years of Urban Expansion in 45 Urban Areas of Sub-Saharan Africa. *Remote Sens.* **2021**, *13*, 525. [CrossRef]

32. Sampson, R.J.; Morenoff, J.D.; Gannon-Rowley, T. Assessing "Neighborhood Effects": Social Processes and New Directions in Research. *Annu. Rev. Sociol.* **2002**, *28*, 443–478. [CrossRef]

33. Stewart, I.D.; Oke, T.R. Local Climate Zones for Urban Temperature Studies. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 1879–1900. [CrossRef]

34. Taubenbock, H.; Pengler, I.; Schwaiger, B.; Cypra, S.; Hiete, M.; Roth, A. A multi-scale urban analysis of the Hyderabad metropolitan area using remote sensing and GIS. In Proceedings of the 2007 Urban Remote Sensing Joint Event, Paris, France, 11–13 April 2007; pp. 1–6; ISSN: 2334-0932. [CrossRef]

35. Krayenhoff, E.S.; Voogt, J.A. Daytime Thermal Anisotropy of Urban Neighbourhoods: Morphological Causation. *Remote Sens.* **2016**, *8*, 108. [CrossRef]

36. Ejiagha, I.R.; Ahmed, M.R.; Hassan, Q.K.; Dewan, A.; Gupta, A.; Rangelova, E. Use of Remote Sensing in Comprehending the Influence of Urban Landscape's Composition and Configuration on Land Surface Temperature at Neighbourhood Scale. *Remote Sens.* **2020**, *12*, 2508. [CrossRef]

37. Stokes, E.C.; Seto, K.C. Characterizing and measuring urban landscapes for sustainability. *Environ. Res. Lett.* **2019**, *14*, 045002. [CrossRef]

38. Dell'Acqua, F.; Gamba, P. Discriminating urban environments using multiscale texture and multiple SAR images. *Int. J. Remote Sens.* **2006**, *27*, 3797–3812. [CrossRef]

39. Zhai, W.; Shen, H.; Huang, C.; Pei, W. Fusion of polarimetric and texture information for urban building extraction from fully polarimetric SAR imagery. *Remote Sens. Lett.* **2016**, *7*, 31–40. [CrossRef]

40. Shao, Z.; Tang, P.; Wang, Z.; Saleem, N.; Yam, S.; Sommai, C. BRRNet: A Fully Convolutional Neural Network for Automatic Building Extraction From High-Resolution Remote Sensing Images. *Remote Sens.* **2020**, *12*, 1050. [CrossRef]

41. Pandey, G.; Sharma, V.K.; Chaudhary, P.; Chowdary, V.M.; Udayraj. Integration of Texture and Spectral Response with AI Techniques for Buildings Footprint Identification Using High-Resolution Satellite Images. *J. Indian Soc. Remote Sens.* **2021**. [CrossRef]

42. Melchiorri, M.; Florczyk, A.J.; Freire, S.; Schiavina, M.; Pesaresi, M.; Kemper, T. Unveiling 25 Years of Planetary Urbanization with Remote Sensing: Perspectives from the Global Human Settlement Layer. *Remote Sens.* **2018**, *10*, 768. [CrossRef]

43. Gong, P.; Li, X.; Wang, J.; Bai, Y.; Chen, B.; Hu, T.; Liu, X.; Xu, B.; Yang, J.; Zhang, W.; et al. Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. *Remote Sens. Environ.* **2020**, *236*, 111510. [CrossRef]

44. Couteron, P.; Barbier, N.; Gautier, D. Textural ordination based on Fourier spectral decomposition: a method to analyze and compare landscape patterns. *Landsc. Ecol.* **2006**, *21*, 555–567. [CrossRef]

45. Proisy, C.; Couteron, P.; Fromard, F. Predicting and mapping mangrove biomass from canopy grain analysis using Fourier-based textural ordination of IKONOS images. *Remote Sens. Environ.* **2007**, *109*, 379–392. [CrossRef]

46. Lang, M. Caractérisation de l'hétérogénéité Spatiale de Milieux Naturels à Partir d'imagerie Optique Très Haute Résolution Spatiale: cas d'application aux Milieux Méditerranéens de Garrigue. Ph.D. Thesis, AgroParisTech, Paris, France, 2019.

47. Zhou, W.; Cadenasso, M.L.; Schwarz, K.; Pickett, S.T.A. Quantifying Spatial Heterogeneity in Urban Landscapes: Integrating Visual Interpretation and Object-Based Classification. *Remote Sens.* **2014**, *6*, 3369–3386. [CrossRef]

48. Serra, M.; Psarra, S.; O'Brien, J. Social and Physical Characterization of Urban Contexts: Techniques and Methods for Quantification, Classification and Purposive Sampling. *Urban Plan.* **2018**, *3*, 58–74. [CrossRef]

49. Qian, Y.; Zhou, W.; Pickett, S.T.A.; Yu, W.; Xiong, D.; Wang, W.; Jing, C. Integrating structure and function: mapping the hierarchical spatial heterogeneity of urban landscapes. *Ecol. Process.* **2020**, *9*, 59. [CrossRef]

50. Mugglestone, M.A.; Renshaw, E. Detection of geological lineations on aerial photographs using two-dimensional spectral analysis. *Comput. Geosci.* **1998**, *24*, 771–784. [CrossRef]

51. Couteron, P. Quantifying change in patterned semi-arid vegetation by Fourier analysis of digitized aerial photographs. *Int. J. Remote Sens.* **2002**, *23*, 3407–3425. [CrossRef]

52. Caloz, R.; Collet, C. *Précis de Télédétection: Traitements Numériques d'Images de Télédétection*; Google-Books-ID: 9bepLPU5udQC; PUQ: Quebec City, QC, Canada, 2001.

53. Lang, M.; Alleaume, S.; Luque, S.; Baghdadi, N.; Feret, J.B. Landscape Structure Estimation using Fourier-Based Textural Ordination of High Resolution Airborne Optical Image. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 6600–6603. [CrossRef]

54. Nunes, B.F.; Bandeira, L. Brasília: L'urbanité Dans une Ville Nouvelle. *Espac. Soc.* **2004**, 93–111.

55. Zhou, G. Detecting the Socioeconomic Conditions of Urban Neighborhoods through Wavelet Analysis of Remotely Sensed Imagery. Ph.D. Thesis, Louisiana State University, Baton Rouge, Louisiana, 2006; p. 172.