



Toward a Computer Vision Perspective on the Visual Impact of Vegetation in Symmetries of Urban Environments

Salma Samiei, Pejman Rasti, Hervé Daniel, Etienne Belin, Paul Richard,
David Rousseau

► To cite this version:

Salma Samiei, Pejman Rasti, Hervé Daniel, Etienne Belin, Paul Richard, et al.. Toward a Computer Vision Perspective on the Visual Impact of Vegetation in Symmetries of Urban Environments. *Symmetry, Integrability and Geometry: Methods and Applications*, 2018, 10 (12), pp.1-18. 10.3390/sym10120666 . hal-01942463

HAL Id: hal-01942463

<https://institut-agro-rennes-angers.hal.science/hal-01942463>

Submitted on 3 Dec 2018

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.



Distributed under a Creative Commons Attribution 4.0 International License

Article

Toward a Computer Vision Perspective on the Visual Impact of Vegetation in Symmetries of Urban Environments

Salma Samiei ^{1,2} , Pejman Rasti ^{1,2}, Hervé Daniel ³, Etienne Belin ^{1,2} , Paul Richard ¹ 
and David Rousseau ^{1,2,*}

¹ LARIS, Université d'Angers, 62 avenue Notre-Dame-du-Lac, 49000 Angers, France; salma.samiei@etud.univ-angers.fr (S.S.); pejman.rasti@univ-angers.fr (P.R.); etienne.belin@univ-angers.fr (E.B.); paul.richard@univ-angers.fr (P.R.)

² UMR INRA IRHS, 42 Rue Georges Morel, 49070 Beaucouzé, France

³ UMR BAGAP, INRA, Agrocampus Ouest, ESA, 2 rue André-Le-Nôtre, 49045 Angers, France; herve.daniel@agrocampus-ouest.fr

* Correspondence: david.rousseau@univ-angers.fr

Received: 2 November 2018; Accepted: 17 November 2018; Published: 26 November 2018



Abstract: Rapid urbanization is a worldwide critical environmental challenge. With this urban migration soaring, we need to live far more efficiently than we currently do by incorporating the natural world in new and innovative ways. There are a lot of researches on ecological, architectural or aesthetic points of view to address this issue. We present a novel approach to assess the visual impact of vegetation in urban street pedestrian view with the assistance of computer vision metrics. We statistically evaluate the correlations of the amount of vegetation with objective computer vision traits such as Fourier domain, color histogram, and estimated depth from monocular view. We show that increasing vegetation in urban street views breaks the orthogonal symmetries of urban blocks, enriches the color space with fractal-like symmetries and decreases the cues of projective geometry in depth. These uncovered statistical facts are applied to predict the requested amount of vegetation to make urban street views appear like natural images. Interestingly, these amounts are found in accordance with the ecosystemic approach for urban planning. Also, the study opens new questions for the understanding of the link between geometry and depth perception.

Keywords: natural image statistics; urban views; fourier; RGB; fractals; depth map; symmetry; urban greenery; projective geometry

1. Introduction

At present, more than half of the world's population is estimated to live in the cities. Urban green spaces (UGS) [1,2] are an important factor of urban streetscape which provides aesthetic, economic, environmental, social, and health benefits to urban residents. Accordingly, societal benefits supplied by UGS to city dwellers are vital to maintain and increase urban citizen's quality of life. The study of the impact of UGS on public health [3–7], to manage the urban ecosystem [8,9] or to assess the aesthetic quality of UGS [10] can benefit from various computer vision based approaches. This includes computer vision to acquire the semantic information of every single pixel of an urban space [11–15] or analyzing the visual impact of vegetation in urban environments [16–18] from top view images in birds or satellite viewpoint.

In this article, we apply computer vision techniques in the urban landscape from the viewpoint of a pedestrian in an urban street. Urban street views are highly geometrical and symmetrical environments with orthogonal and parallel lines radically different from what is found in a natural environment

where structures following any orientations are more likely to occur. Also, from a color perspective, urban street views often offer few hue values due to their limited mineral content. This is very different from the richness of color found in nature. Characterizing the quantitative impact of vegetation on these visual symmetries would enable to assess the question of how much vegetation should be included in an urban street view to make it look more natural than man-made. We provide a computer vision quantification of the impact of vegetation in urban street views by determining their statistical properties.

Obviously, a very large set of potential descriptors could be used for this application. Instead of an exhaustive benchmark of existing computer vision tools of the literature, or an automatic selection of such tools with machine learning approaches, we proceed with an analytical approach where we select simple descriptors which have been successfully applied in the literature to identify statistical invariant features in pure man-made or purely natural scenes. We explore, for the first time to our knowledge, how this small set of historical descriptors behave in presence of various amount of vegetation. We also provide new experiments and dataset specially designed for this work.

As related works, one can recall that understanding the statistical properties of the natural environment is an important problem in computer vision [19–21]. Although explored for decades, characterization of statistical invariance and symmetries in natural images continues to progress. For instance, as recently reviewed in [20], this progress has been obtained by considering first and higher statistics of the luminance and color distributions of natural images, local orientation in images or statistical cues in the natural visual environment that is available to compute disparity. Other approaches have also contributed to the characterization of natural images by considering different categories of them. This includes gazing at natural scenes with new sensors (range camera [22,23], thermal camera [24], polarized light [25]...) or in various environments of interest for humans (underwater [26], man versus natural environment [27,28]). We adopt this approach of characterizing natural images by considering specific categories. We focus on urban street views with various amount of vegetation. These scenes are therefore important as already underlined for urban planning but also for vision understanding because they constitute an intermediate category between pure man-made scenes and purely natural scenes in which the visual system has evolved originally.

The two main novelties of the paper appear along the following lines : (i) none of the existing strategies for urban greenery has proposed so far a computer vision perspective at a pedestrian level. Computer vision was already used to analyze the statistics of man-made versus wild images but (ii) we propose the first application of these techniques for man-made block world with various amount of vegetation. The article is organized along the following structure. In the second section, we explore how vegetation in urban street view can be quantified with computer vision tools in the Fourier domain. In the following section, we investigate another aspect of the visual impact of urban vegetation in the color domain. The fourth section explains the last aspect of the visual impact of urban vegetation on depth. In conclusion, the application of these results is discussed in terms of urban planning, and we mention new questions now opened for further investigation by this work.

2. Impact of Vegetation in the Fourier Domain

Different categories of natural environments present different orientations. For instance buildings, for stability reasons, will provide orthogonal edges while plants [29], because they seek for maximum light, are more likely to propose edges in all directions. This has been demonstrated in [28] to statistically translate into a characteristic signature in the power Fourier spectrum of natural images with anisotropy patterns in the Fourier domain of urban views and isotropic patterns in the natural landscape. In this section, we reproduce the experiment of [28] with urban street views including a various amount of vegetation between pure man-made and purely natural scenes.

To this purpose we consider the cityscapes dataset [30] which includes 25,000 RGB annotated street pedestrian view images from 50 cities, mainly located in Germany. For experiments in this article, 5000 images from fine annotations category are used. The annotation includes the labeling of

vegetation. It is therefore straightforward to associate a percentage of vegetation with each image of this dataset. Three examples of this dataset are shown in Figure 1 with the RGB images, the annotated images and the percentage of vegetation which is evaluated through annotated images.

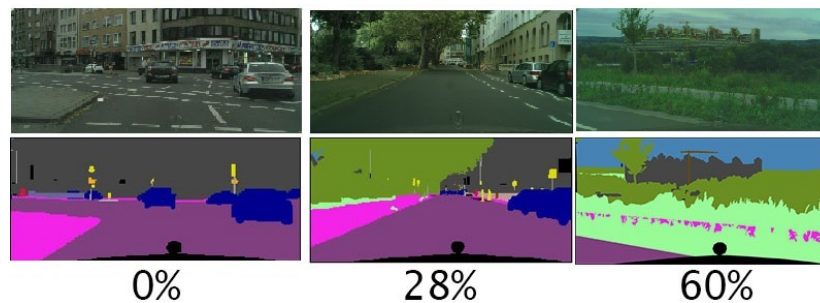


Figure 1. Three examples of images taken from the dataset considered in this study [30]. The first row shows the RGB images. The second row is the corresponding annotated images. The percentage value gives the percentage of vegetation measured from the annotated images.

Figure 2 illustrates spectral signature computed by Fourier transform for the same three images of the dataset as in Figure 1. As visible in Figure 2 and similarly to what was found in [28] the presence of plants tends to reduce the anisotropy between the spectral energy in horizontal and vertical frequencies.

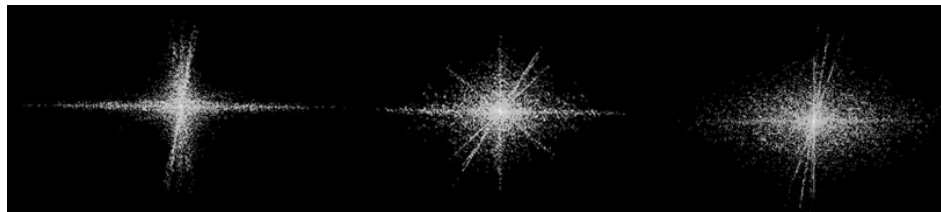


Figure 2. Three examples of the modulus of the Fourier transform of the RGB images of Figure 1 from left to right with 0%, 28% and 60% of vegetation respectively.

Figure 3 illustrates the full pipeline followed for investigating the impact of vegetation in the Fourier domain.

Due to the underexposed aspect of the RGB images in this dataset, CLAHE [31] algorithm is used for adjusting the contrast of the image and intensity equalization. Then, the RGB images are converted into gray levels $L = 0.299 \times R + 0.587 \times G + 0.114 \times B$ and transferred from spatial domain to frequency domain through 2-D Fourier transform. The modulus of this Fourier transform is thresholded, as in [28], in order to keep 70% of the energy. To measure the vertical–horizontal anisotropy of this binarized spectrum the following ratio of orientation is computed as

$$R = \frac{D1 + D2}{\sqrt{H^2 + V^2}}, \quad (1)$$

where, as shown in Figure 4, $D1$, and $D2$ represent the diagonal and antidiagonal size of the spectral signature, H indicates the horizontal size of spectral signature and V represents the vertical size of it.

This anisotropy ratio is then plotted as a function of the percentage of vegetation. Examples of three cities from the dataset are given in Figure 5 where a linear trend clearly appears with decreasing anisotropy ratio as a function of the percentage of vegetation in the image. The average slope on the whole dataset [30] is weak. However, this trend is systematically found and statistically valid for all cities as demonstrated in Table 1.

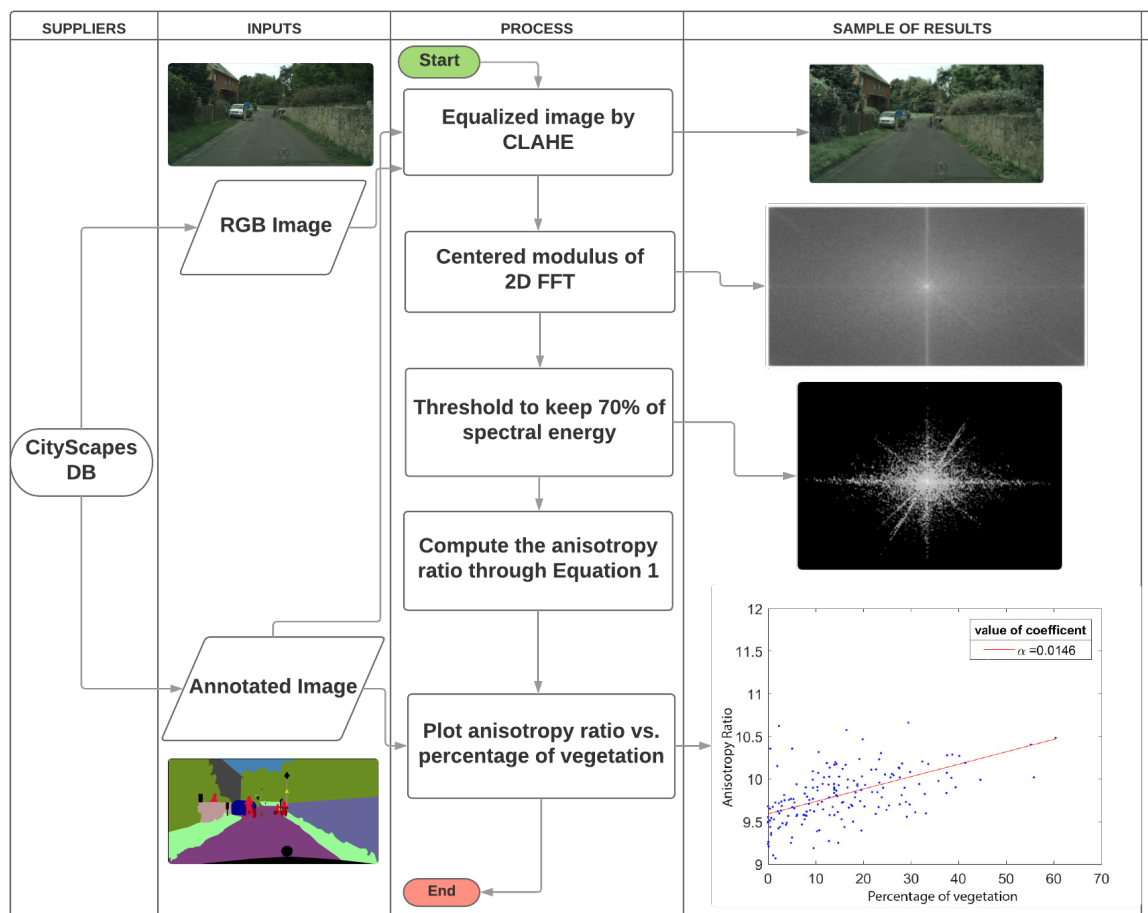


Figure 3. Image processing pipeline for the study of the impact of vegetation in the Fourier domain. CityScapes DB [30] corresponds to an available dataset as illustrated in Figure 1.

Table 1. The slope of the anisotropy ratio as a function of the percentage of vegetation and associated p -value.

City	Slope	p -Value
Aachen	1.3×10^{-3}	8.10×10^{-11}
Bochum	1.4×10^{-3}	1.73×10^{-7}
Bremen	1.3×10^{-3}	8.06×10^{17}
Cologne	1.1×10^{-3}	1.87×10^{-6}
Darmstadt	1.4×10^{-3}	9.72×10^{-6}
Dusseldorf	1.4×10^{-3}	8.67×10^{-17}
Hamburg	1.0×10^{-3}	4.44×10^{-8}
Hanover	1.4×10^{-3}	4.58×10^{-12}
Jena	1.0×10^{-3}	1.40×10^{-5}
Krefeld	1.5×10^{-3}	5.82×10^{-6}
Monchengladbach	1.5×10^{-3}	1.22×10^{-7}
Strasbourg	0.9×10^{-3}	1.92×10^{-8}
Stuttgart	0.8×10^{-3}	3.35×10^{-5}
Tubingen	0.5×10^{-3}	2.37×10^{-3}
Ulm	1.3×10^{-3}	7.71×10^{-7}
Weimar	1.2×10^{-3}	2.38×10^{-10}
Zurich	0.8×10^{-3}	3.76×10^{-5}

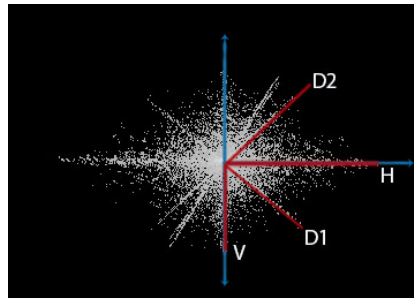


Figure 4. Elements computed in the anisotropy ratio of Equation (1) applied on the modulus of the Fourier transform.

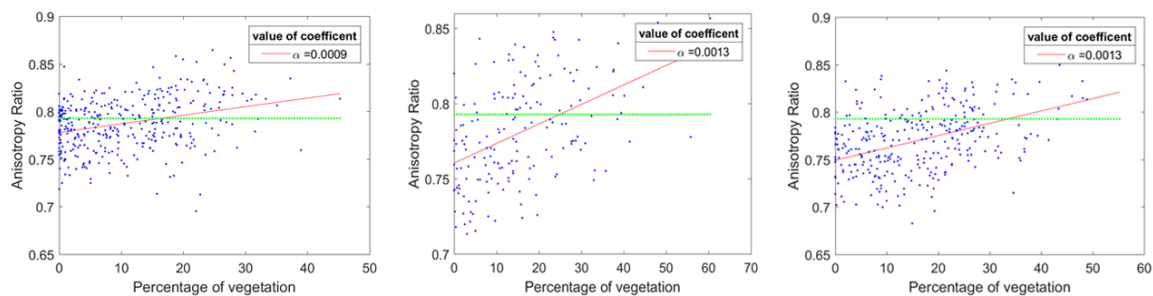


Figure 5. Anisotropy ratio as a function of the percentage of the vegetation of three examples of cities taken from the dataset considered in this study [30]. The left column is Strasbourg city. The column in the middle is for Aachen. The right column is Bremen city. The red line is the linear fit of the data. The green line is the reference 0.79 corresponding to the average anisotropy value between the pure man-made and pure natural environment.

Experimental results of this Fourier approach demonstrate, as one could intuitively expect from [28], that the presence of vegetation tends to break the horizontal-vertical symmetry in the images as quantified by the anisotropy ratio defined in this section.

We now come to propose a possible application of this statistical result for urban planning. We considered the pure concrete-based images and pure natural images (in the wild) of [28] to determine the expected range of the anisotropy ratio. Natural images were found on average with an anisotropy ratio of 0.84 and pure concrete of 0.74. An intermediate value in this range could constitute a transition limit between an environment perceived as natural and an environment perceived as pure man-made in the Fourier domain. To test this hypothesis, we considered the middle value of this anisotropy ratio between natural and pure concrete environment. This average value (0.79), pointed in green in Figure 5, crosses the linear fit computed for each city and provides an associated percentage of vegetation. The average value of the requested percentage of vegetation to reach this anisotropy ratio is not found to trivially be 50% but rather 28% with a standard deviation of 8.32%. The requested percentage of vegetation in cities is also debated in other scientific fields. Percentage of vegetation below which fragmentation of urban ecosystems has consequences on the diversity and viability of these ecosystems have been highlighted in [32–34] for instance. Interestingly, it appears that the threshold around which processes are favored or not are found to be between 20 and 30% of vegetation, i.e., in a similar range as the one found here with our computer vision approach. However, it is to be noticed that the decrease of anisotropy ratio statistically recorded could also be obtained without any vegetation by simply using non-orthogonal building architectures promoting, for example, curved shapes with edges in all orientations.

3. Impact of Vegetation in the Color Domain

Another aspect of the visual impact of vegetation on the images is explored in this section. In a pure concrete urban environment, it is likely that the color embedded in the color histogram

of images will be limited to some blue in the sky, grey-black on the ground and a small number of colors correlated to the mineral content used for walls of the building in the image (generally not green). Adding vegetation in an urban concrete environment is therefore expected to enrich the color histogram. The statistics of the color histogram of natural images have been studied in [35–37] where scale invariant symmetries were observed in the organization of the color in the RGB 3D histogram. We reproduce similar experiments with the dataset of the previous section [30] to investigate the evolution of the RGB 3D color histogram as a function of the amount of vegetation in the image.

Figure 6 shows three examples of the 3D color histogram for the same images of Figure 1. As visible in Figure 6, the presence of plants tends to increase the complexity of the 3D color histogram.

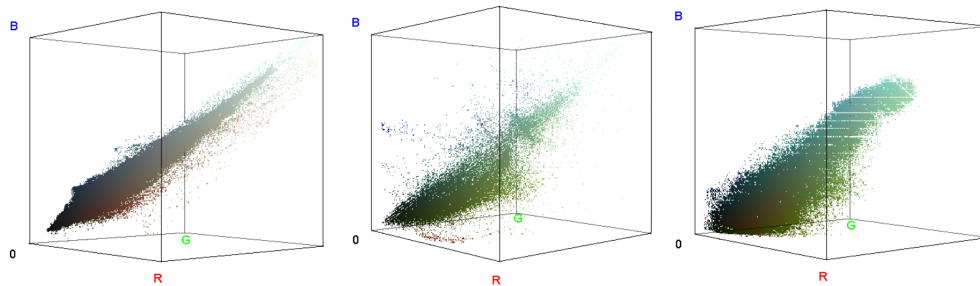


Figure 6. Three examples of the 3D color histogram of the RGB images in Figure 1. From left to right with 0%, 28% and 60% of vegetation respectively. The presence of plants tends to increase the richness of the 3D color histogram.

Figure 7 illustrates the full pipeline followed for investigating the impact of vegetation on the 3D color histogram. A box-counting procedure is applied as follows. The colorimetric cube is successively covered with boxes of side a and volume a^3 , with varying a . For each box of size a , one computes the number $N(a)$ of boxes which are needed to cover the support of the 3D histogram, i.e., to cover all the cells of the colorimetric cube which are occupied by pixels of the image. As observed in [35–37] the evolution of the number of counts $N(a)$ as a function of the color scale on a log-log scale is well approximated by straight lines with slope $-D$, over a significant range of colorimetric scales a . This scale invariant symmetry is associated with a fractal behavior with fractal dimension D . Then for all images in each city, the slope values D are plotted as a function of the percentage of vegetation. Examples for three cities of the dataset are given in Figure 8 where a linear systematic trend clearly appears over the whole set of cities with increasing fractal dimension D as a function of the percentage of vegetation in the image. Table 2 gives the slope of the fractal dimension and associated p -value for all cities in CityScape dataset.

Table 2. Slopes of fractal dimension and associated p -value in the colorimetric approach.

City	Slope	p -Value
Aachen	0.19×10^{-2}	3.47×10^{-11}
Bochum	0.23×10^{-2}	5.87×10^{-6}
Bremen	0.19×10^{-2}	1.59×10^{-14}
Cologne	0.10×10^{-2}	9.79×10^{-3}
Darmstadt	0.35×10^{-2}	1.45×10^{-9}
Dusseldorf	0.23×10^{-2}	5.67×10^{-14}
Erfurt	0.17×10^{-2}	3.94×10^{-7}
Hamburg	0.18×10^{-2}	1.54×10^{-6}
Hanover	0.14×10^{-2}	2.00×10^{-3}
Jena	0.25×10^{-2}	7.90×10^{-10}
Krefeld	0.16×10^{-2}	5.31×10^{-2}
Monchengladbach	0.20×10^{-2}	2.66×10^{-2}
Strasbourg	0.10×10^{-2}	1.35×10^{-2}

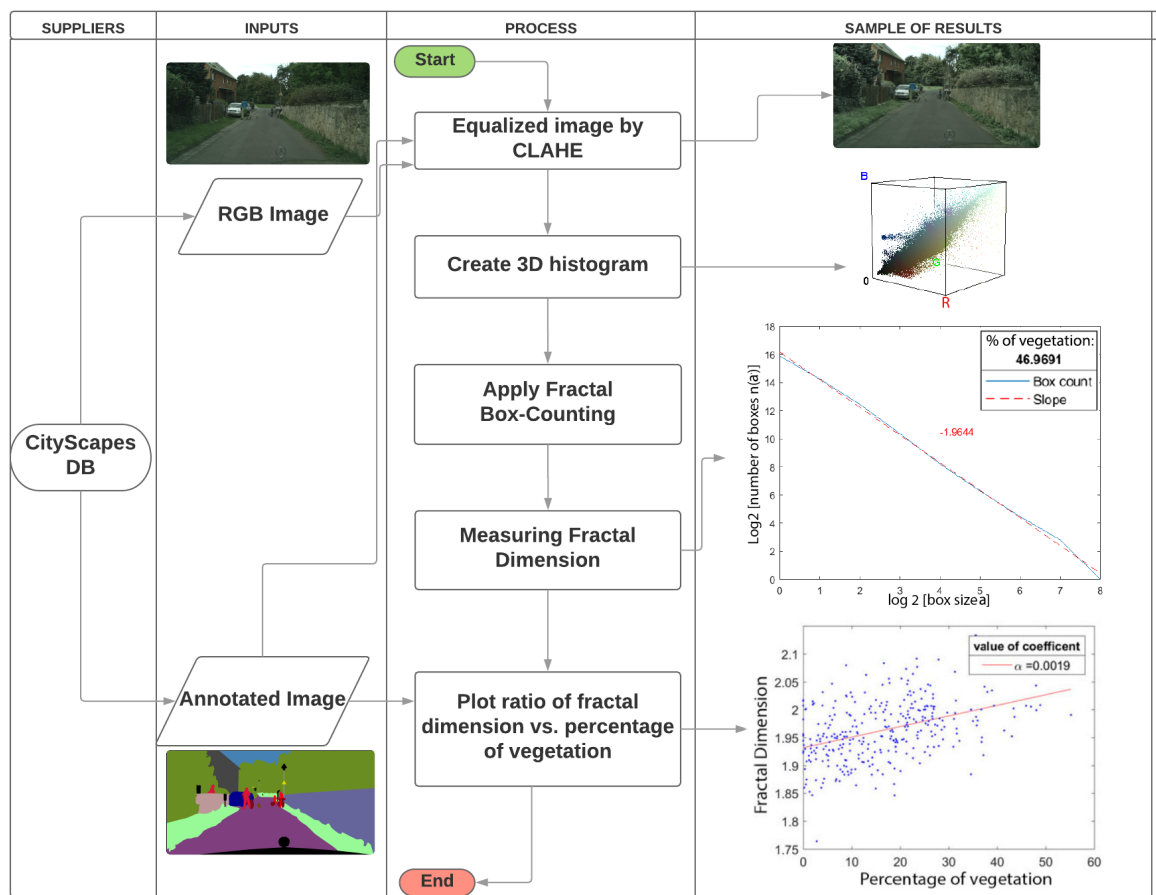


Figure 7. Image processing pipeline for the impact of vegetation in the color domain. CityScapes DB [30] corresponds to available dataset shows in Figure 1. All images of CityScapes dataset are processed and batched city by city.

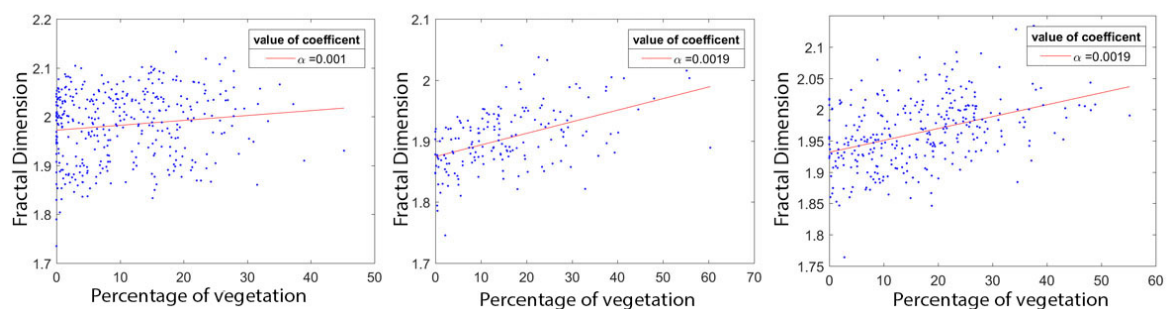


Figure 8. Fractal dimension values as a function of the percentage of the vegetation of three examples of cities taken from the dataset considered in this study [30]. The left column is Strasbourg city. The column in the middle is for Aachen. The right column is Bremen city.

Experimental results of this approach in the color domain demonstrate that the presence of vegetation tends to enrich the complexity of the 3D color histogram. This enrichment is in the direction of higher dimension scale invariant symmetries as quantified by the box-counting fractal dimension defined in this section.

Similarly to what was proposed in the previous section on Fourier, there are possible applications of this statistical result for urban planning. Especially the increase of the fractal dimension could serve to control a requested amount of vegetation in cities where the uniform colors of concrete are used for buildings. However, it is to be noticed that the complexity of the 3D color histogram can

also be enriched in the direction of fractal signatures by simply painting concrete urban environment when there is no possibility of adding vegetation. This is shown in Figure 9 on a colorful urban image without any vegetation. This image has a high fractal dimension which would correspond to adding almost 100% of vegetation in the mono-color concrete cities of Figure 8.

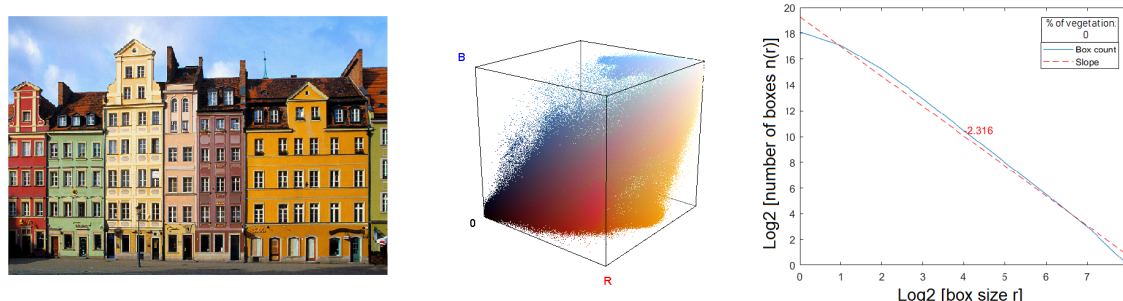


Figure 9. 3D color histogram and fractal dimension's plot of a colorful urban image without any vegetation.

4. Impact of Vegetation on Monocular Depth Cues

We explore the last aspect of how the impact of vegetation in pedestrian view can be quantified with computer vision tools and now focus on the depth, i.e., distance to a point of view. Humans and computers are known to be able to estimate depth from the stereo and monocular visions [38,39]. In this study, we restrict ourselves to monocular vision with single images of the streets. Different cues of depth can be present in monocular vision such as textures, shadows, defocus [40], parallel lines [41] producing in a projective geometry the presence of vanishing point on the horizon line, repetition of similar objects at different distances from the camera.

For humans, all these cues can contribute to the perception of depth in monocular vision and it is difficult to quantify the relative importance of each individual cue in a scene. With computers, it is possible to design a feature especially capturing the presence of a single depth cue. Also, with the current development of machine learning in computer vision, it is possible to produce a global quantitative estimation of depth incorporating all cues. In this section, we propose a quantification of the presence of vegetation on the quality of depth estimation with these two approaches including the detection of the vanishing point and a direct estimation of the depth map from a monocular RGB image.

Virtual environments in urban systems research are very useful to access to situations that do not (yet) exist in real environments [42–44]. In our case, we need to have a dataset with RGB images of urban street view with various amounts of annotated vegetation. Ground truth depth map should not be computed from the RGB images to be compared with the depth cue estimation or depth estimation. There are a lot of available datasets with this structure for indoor and outdoor environments [45]. The most related outdoor datasets found in the literature are listed in Table 3. It should be mentioned that for most of these datasets depths are not measured but estimated from RGB images in different ways such as semi-global matching (SGM) algorithm [46]. Therefore, they are not suitable depth for our purpose because we specifically want to extract depth from RGB and we need a depth map estimated from an independent distinct way as a ground truth. Also, the sole available outdoor dataset with a pedestrian urban street views in RGB and directly estimated depth [47] is not incorporating enough variation of percentage of vegetation. None of the available datasets listed in Table 3 incorporate all the required aspects of our study. Therefore, we had to produce our own virtual dataset.

Table 3. Available datasets for the analysis of the impact of vegetation on depth.

Dataset	Num. of Samples	RGB Images	Annotated Images	Depth Images Evaluated with
Make3D [48]	534	✓	-	Laser range data with ray position
KITTI [49]	93,000	✓	-	Lidar
Cityscapes [30]	25,000	✓	✓	SGM algorithm applied on RGB images [46]
WildDash [50]	70	✓	✓	-
Mapillary [51]	25,000	✓	✓	-
ApolloScape [47]	140,000	✓	✓	survey-grade dense 3D point cloud

We propose the virtual RGBD green-city dataset provided as Supplementary Material to this study where 300 high-resolution images (879×1680 pixels) generated from the virtual world in urban settings under the different percentage of vegetation. These virtual cities were created using the Unity game engine with available models of trees and urban blocks. The dataset includes the segmentation of the vegetation to compute the percentage of vegetation and the depth map. Figure 10 illustrates the content of this virtual RGBD green-city dataset with three examples of RGB images with different percentages of vegetation. One specific interest in working with simulated data is that it is possible to create datasets of the same street with various amounts of plants following different strategies of the positioning of the plants. We take benefit of this opportunity and design several (10) experiments. This includes positioning trees on one side or on both sides of the street, positioning trees with various orientations or a single vertical orientation, using a tree with different sizes or using a variety of trees. For further information related to the virtual RGBD green-city dataset see Supplementary Material.

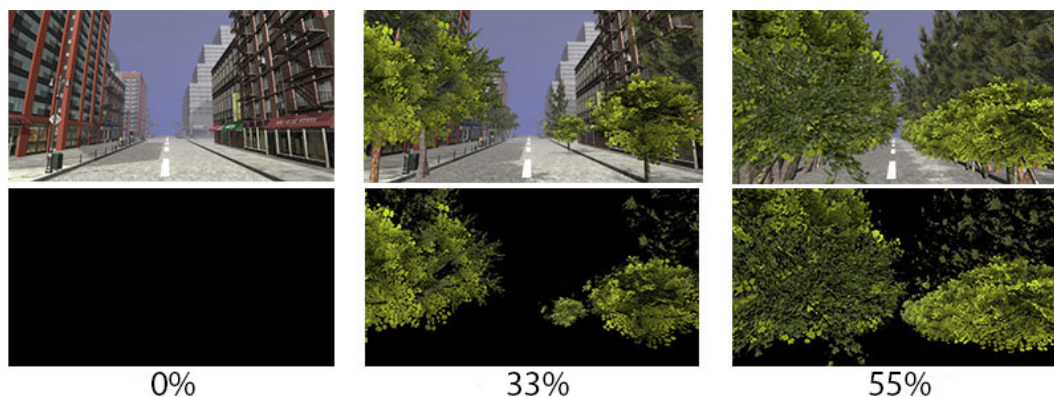


Figure 10. Three examples of virtual RGBD green-city dataset. The first row gives RGB images; the middle row is for the corresponding annotated images of the vegetation, the last row shows the percentage of vegetation.

4.1. Detection of Vanishing Point

The vanishing point corresponds to the place where parallel edges in a real scene produce lines which cross in the image due to the projective geometry created by the lens of a camera. There are different approaches to find the vanishing points in an image [39,52,53], in this study, our purpose is not to compare them or select any best method. Instead, we arbitrarily select one of the techniques which provides good results on similar images to ours and show how the performance of this technique evolves when the amount of vegetation is increased. To this purpose, the Hough transform-based technique of [54] is incorporated in the pipeline shown in Figure 11.

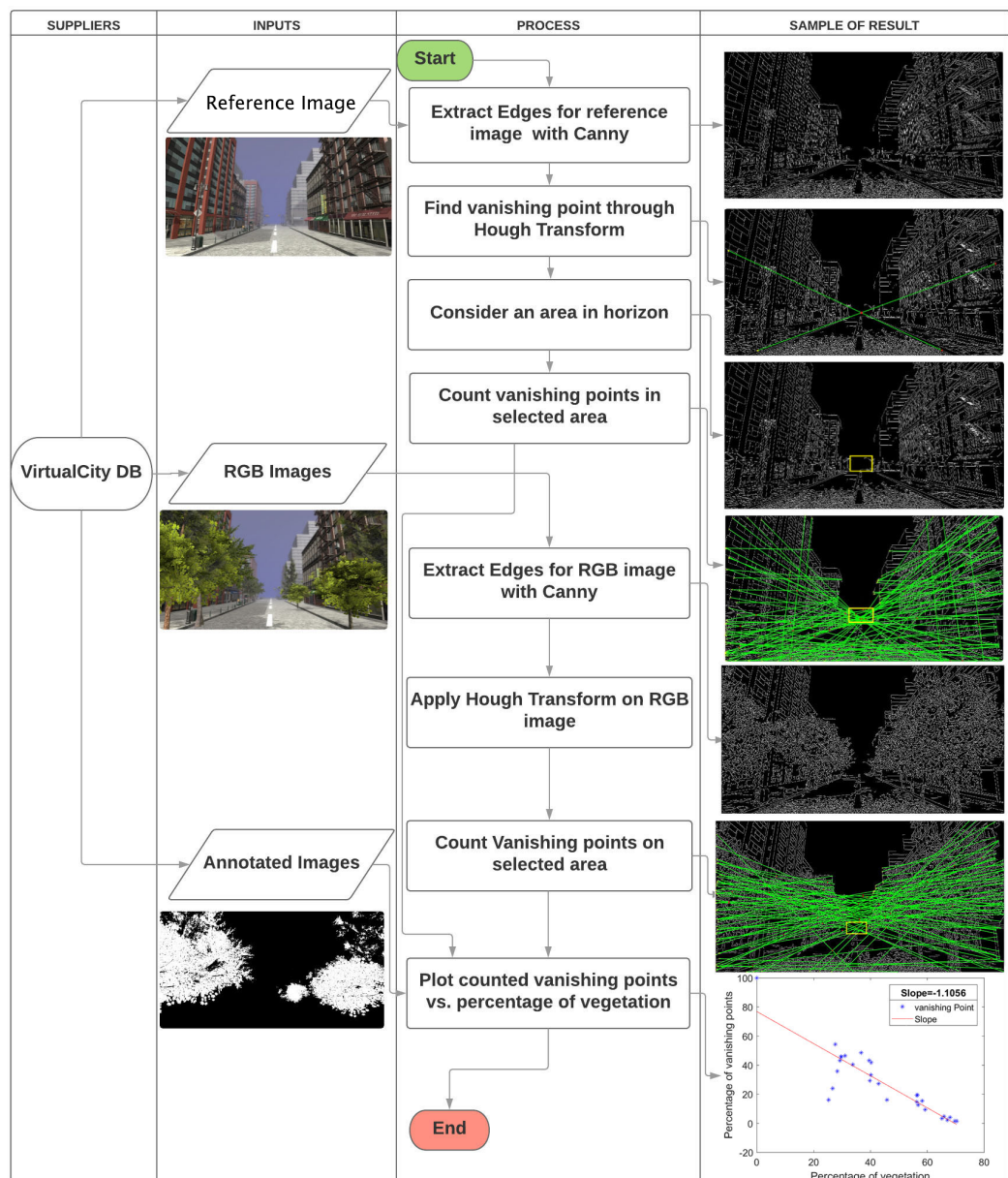


Figure 11. Image processing pipeline for the impact of vegetation on the detection of the vanishing point. VirtualCity DB refers to virtual RGBD green-city dataset which is provided for this study.

A Hough transform is first applied to estimate the vanishing point on a reference image taken as the image with 0% of vegetation. Edges are extracted with the Canny edge detector [55] of the reference image. The extracted edges are then transformed into Hough space. The vanishing point is chosen as an intersection point with a large number (empirically chosen to 70) of intersection in the Hough space. Then, the extracted line segments are merged if they are associated with the same Hough transform bin and the distance between them is less than the value of a threshold empirically set to 400 pixels. Afterward, the detected lines in the Hough transform are counted as a contribution to the vanishing point if they are crossing in a vanishing area. Indeed, due to the limited and discretized size of the virtual environment, the lines do not intersect exactly at a single point [56]. We define the vanishing area with the size of 55 by 45 pixels around the largest number of convergent lines in the reference image without vegetation. Finally, the percentage of convergent points in the vanishing area is recorded.

This pipeline (Figure 11) is applied to all the images of the dataset. The percentage of vanishing points is plotted with the comparison to the reference image free from vegetation as a function of the percentage of vegetation. Figure 12 shows the evolution of this percentage of vanishing points remaining as a function of the percentage of vegetation introduced. The values of the linear slope for all 10 experiments are represented in Tables 4 and 5. Table 4 represents slope values for scenes with trees located on both sides of the street or everywhere like a forest and Table 5 shows slope values for scenes which trees are located just on one side of the street.

From Figure 12 one can see that the presence of vegetation tends to change the position of vanishing points in the vanishing area and consequently fewer points are crossing in this area. An interpretation is that the increasing amount of leaf will increasingly cover the horizontal lines and thus affect the number of cues for the estimation of the vanishing point. However, it is to be noticed that other objects than vegetation could also occlude the horizontal lines and thus have an impact on the detection of vanishing points in a pure concrete urban environment. To further demonstrate the impact of vegetation on depth perception, we thus reproduced in the next section the experiment of this section using a completely different depth estimation method.

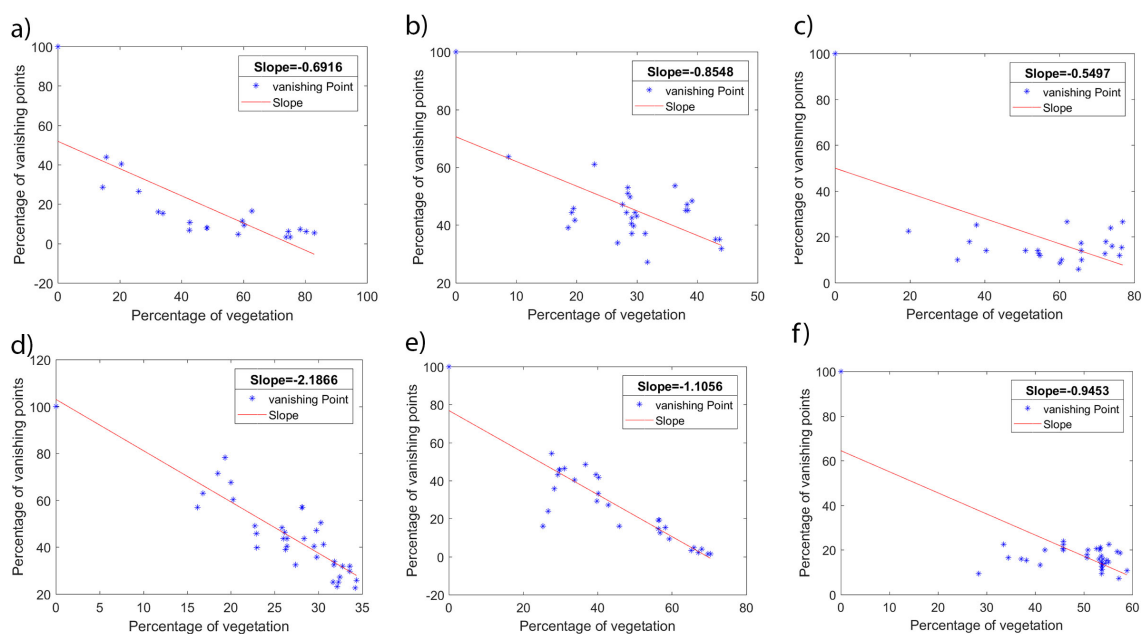


Figure 12. Percentage of detected vanishing points as a function of vegetation's percentage. (a) Forest, (b) Different-Size, One-Side (c) Different-Size, Two-Sides (d) Different-Orientation, One-Side (e) Different-Orientation, Two-Side (f) Different-Tree, Different-Size, Different-Orientation, Two-Side

Table 4. Result for detected vanishing point in experiments with trees placed on both sides of the street and forest.

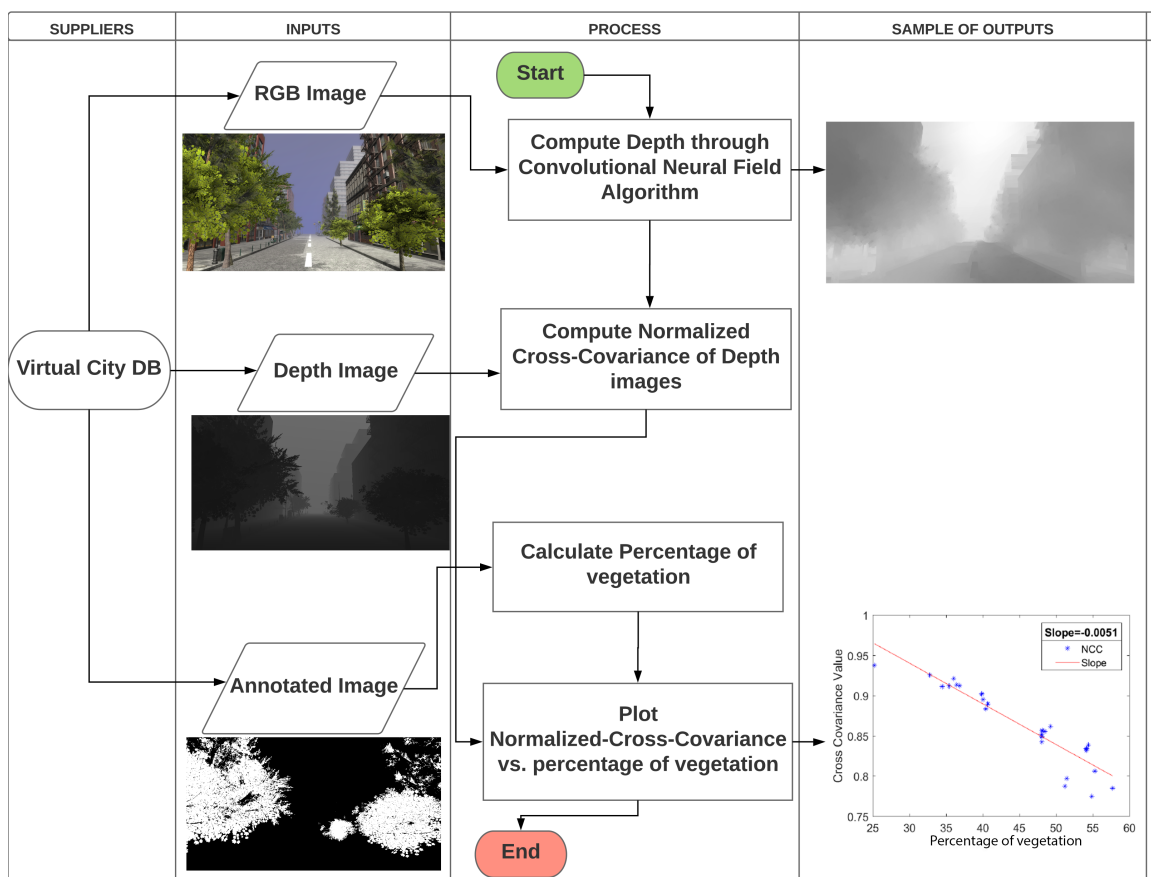
Category-Two Side	Slope
Same Tree-Same Size-Same Orientation	−0.9868
Same Tree-Same Size-Different Orientation	−1.1056
Same Tree-Different Size-Same Orientation	−0.5497
Different Tree-Same Size-Same Orientation	−1.3513
Different Tree-Different Size-Different Orientation	−0.9453
Forest	−0.6916

Table 5. Result for detected vanishing point in experiments with trees placed on one side of the street.

Category-One Side	Slope
Same Tree-Same Size-Same Orientation	−1.5522
Same Tree-Same Size-Different Orientation	−2.1866
Same Tree-Different Size-Same Orientation	−0.8548
Different Tree-Same Size-Same Orientation	−1.3543

4.2. Depth Estimation

We now propose to assess the impact of vegetation on the global perception of depth. To this purpose, we settled the pipeline described in Figure 13. Here again, as for the estimation of the vanishing point, there is a huge literature on estimation of depth with machine learning approaches. Our purpose is not to compare any of them but to show the impact of vegetation on one of them [57]. We arbitrarily select one of the recent methods performing well for depth estimation of similar images to ours and submit this to our original virtual RGBD green-city dataset which contains RGB, depth and annotated images.

**Figure 13.** Image processing pipeline for the impact of vegetation on normalized-cross-covariance approach. VirtualCity DB refers to virtual RGBD green-city dataset which is provided for this study.

At the first step, the estimated depth is calculated by Deep Convolutional Neural Field (DCNF) algorithm developed by [57]. DCNF model is a depth estimation approach exploring Convolutional Neural Network (CNN) and continuous Conditional Random Field (CRF). The whole networks of DCNF are trained on Make3D dataset [48] for outdoor scenes. Make3D dataset approximately includes 1000 outdoor street pedestrian views captured in good weather conditions with 50% of vegetation on

average. At the second step, the global similarity is calculated between the estimated depth images and ground-truth depth by normalized-cross-covariance (NCC) which reads

$$NCC(X, Y) = \frac{E \left[((X - E(X)) \cdot (Y - E(Y))) \right]}{\sigma_X \sigma_Y}, \quad (2)$$

where X and Y represent the estimated depth map and ground-truth depth map respectively, σ_X and σ_Y are the standard deviations and E stands for the 2D mean.

Figure 14 illustrates the value of normalized-cross-covariance as a function of the percentage of vegetation for the six different experiments included in the virtual RGBD green-city dataset. The impact of vegetation on the similarity of estimated depth map with the true depth map is systematic. The presence of the vegetation tends to decrease the quality of depth cues. This decrease modeled with a linear trend gives the slope provided in Table 6 for scenes with trees located on both sides of the street or everywhere like a forest and Table 7 for scenes with trees located just on one side of the street. Interestingly this result is similar to the one observed with the Hough transform in the previous section while it was obtained with a completely different approach.

From an applied urban planning perspective, if we discard the solution that would correspond to position trees everywhere (Forest case) these experiments demonstrate that the highest decreases of depth cue (as given in Tables 4–7) are obtained for urban streets when they include a variability of tree shape and/or tree size and/or various tree orientations. Projective geometry with parallel lines like the one produced by urban block world is not present in wild landscapes. Therefore, we come up with the interesting conclusion that, consistently with the well-known necessity of diversity in ecosystems, computer vision also suggests to add plant diversity as the most effective strategy to break the depth cues created by non-natural urban block worlds.

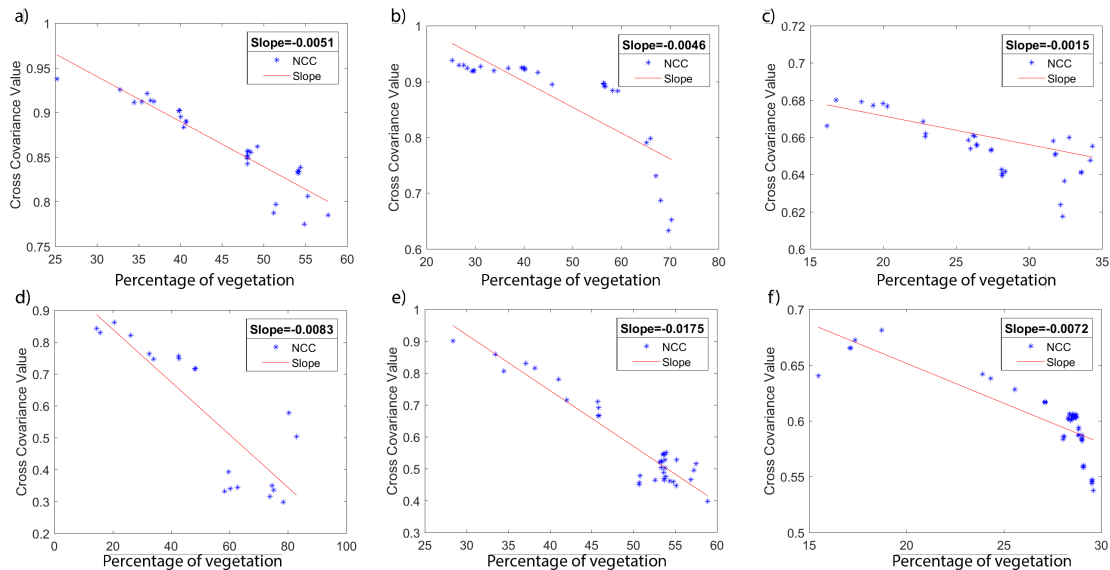


Figure 14. Normalized cross-covariance as a function of the percentage of vegetation. (a) Same tree-same size-Same orientation-Two side, (b) Different orientation-Two side, (c) Different orientation-One side, (d) Forest, (e) Different tree-Different size-Different orientation-Two side, (f) Same tree-Same size-Same orientation-One side.

Table 6. Result for global similarity in experiments with trees placed on both sides of the street and forest.

Category—Two Side	Slope
Same Tree-Same Size-Same Orientation	−0.0072
Same Tree-Same Size-Different Orientation	−0.0015
Same Tree-Different Size-Same Orientation	−0.0089
Different Tree-Same Size-Same Orientation	−0.0019
Different Tree-Different Size-Different Orientation	−0.0175
Forest	−0.0083

Table 7. Result for global similarity in experiments with trees placed one side of the street.

Category—One Side	Slope
Same Tree-Same Size-Same Orientation	−0.0051
Same Tree-Same Size-Different Orientation	−0.0046
Same Tree-Different Size-Same Orientation	−0.0118
Different Tree-Same Size-Same Orientation	−0.0016

5. Conclusions

In this pilot experimental study, we quantified, on large datasets, the impact of vegetation on visually perceptible symmetries in urban street pedestrian viewpoint. Correlation of this amount of vegetation with objective computer vision traits has been shown statistically in the Fourier domain, in the color histogram, and in depth from monocular view. This objectively quantifies the expected common-sense intuition that vegetation in street pedestrian views breaks the orthogonal symmetry of urban blocks, enriches the color space in the direction of higher dimension fractal symmetries and decreases the cues of depth included in projective geometry. The result obtained in the Fourier domain and the color histogram corresponded to existing experiments of the literature carried here for the first time on urban scene with various amount of vegetation while the experiment designed and carried on depth is, to the best of our knowledge, completely new.

Possible applications in urban planning of the carried experiment have been proposed. The most interesting is that a percentage of the vegetation of 20–30% is found to be necessary to have an urban street which appear closer to natural images than pure man-made from the Fourier point of view. Interestingly this also meets the typical percentage of vegetation often mentioned as necessary to maintain ecological viability and diversity in urban ecosystemic studies [32–34]. The novel contribution on the impact of vegetation on depth could be extended in various directions to enable a deeper understanding of the recorded effect. For instance, it would be possible to reproduce the experiment carried here to objectively quantify the effect of vegetation on other computer vision traits. In particular, one could investigate the pedestrian viewpoint in urban streets with bio-inspired traits such as the estimation of depth from stereovision [58–60] or at a higher integrative level from visual saliency [61] or also with the heat map produced by eye-tracking systems when applied on large cohorts of human observers [62]. The effect of vegetation on pedestrian view in urban street was objectively quantified in this article. It could also be interesting to compare these results with aesthetic assessments [63–65] on the same scenes when perceived by the human. One could directly embed humans in virtual environments with grabbing tasks (similarly to [66] for instance) and assess their performance while varying the amount of vegetation. Finally, with the analytical approach followed in this preliminary work, the standard deviation from the simple linear trends identified on few descriptors, although already interesting, could be considered as still too high to serve as good predictors of a correct amount of vegetation to be placed in an urban planning. Using more expressive multivariate models with a learned and large feature space based on deep learning [67] could be a promising direction.

In another direction, one could seek to apply the computer vision traits used in this study on cognitive architecture [68] or urban landscape experiments carried in the domain of ecology.

In architecture, one could, for instance, investigate pedestrian view of the urban environment through the window at different floors of buildings (while the experiment in this article was limited to ground floor view). In the domain of ecology, one could study the relationship between the computer vision metrics designed in this study and the ecological diversity along the seasons (while the experiment in this article was limited to spring-like views with unlabeled ecological diversity) or in night vision as allowed by the simulation environment introduced in this work. Such architectural or ecological experiments had in the past been relying on qualitative psycho-visual studies [69,70] but also tend to use computer vision features in 2D dimensions [71,72] or as recently shown in [73] with 3D LIDAR data.

Supplementary Materials: The following are available at <http://www.mdpi.com/2073-8994/10/12/666/s1>.

Author Contributions: Conceptualization, S.S. and D.R.; methodology, S.S., P.R. and D.R.; software, S.S. and P.R.; validation, P.R., D.R. and H. D.; resources, P.R.; data curation, S.S.; writing—original draft preparation, S.S. and D. R.; writing—review and editing, S.S., D.R. E. B. and H. D.; visualization, S.S.; supervision, D.R.; project administration, D.R.; funding acquisition, D.R. and P. R.

Funding: Salma Samiei acknowledges Angers Loire Métropole for the funding of her PhD.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wolch, J.R.; Byrne, J.; Newell, J.P. Urban green space, public health, and environmental justice: The challenge of making cities ‘just green enough’. *Landsc. Urban Plan.* **2014**, *125*, 234–244. [CrossRef]
2. Li, X.; Zhang, C.; Li, W.; Ricard, R.; Meng, Q.; Zhang, W. Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban For. Urban Green.* **2015**, *14*, 675–685. [CrossRef]
3. Carpenter, M. From ‘healthful exercise’ to ‘nature on prescription’: The politics of urban green spaces and walking for health. *Landsc. Urban Plan.* **2013**, *118*, 120–127. [CrossRef]
4. Coppel, G.; Wüstemann, H. The impact of urban green space on health in Berlin, Germany: Empirical findings and implications for urban planning. *Landsc. Urban Plan.* **2017**, *167*, 410–418. [CrossRef]
5. Ekkel, E.D.; de Vries, S. Nearby green space and human health: Evaluating accessibility metrics. *Landsc. Urban Plan.* **2017**, *157*, 214–220. [CrossRef]
6. Sugiyama, T.; Carver, A.; Koohsari, M.J.; Veitch, J. Advantages of public green spaces in enhancing population health. *Landsc. Urban Plan.* **2018**, *178*, 12–17. [CrossRef]
7. Russo, A.; Cirella, G. Modern compact cities: How much greenery do we need? *Int. J. Environ. Res. Public Health* **2018**, *15*, 2180. [CrossRef] [PubMed]
8. Du Toit, M.J.; Cilliers, S.S.; Dallimer, M.; Goddard, M.; Guenat, S.; Cornelius, S.F. Urban green infrastructure and ecosystem services in sub-Saharan Africa. *Landsc. Urban Plan.* **2018**. [CrossRef]
9. Zinia, N.J.; McShane, P. Ecosystem services management: An evaluation of green adaptations for urban development in Dhaka, Bangladesh. *Landsc. Urban Plan.* **2018**, *173*, 23–32. [CrossRef]
10. Chen, B.; Adimo, O.A.; Bao, Z. Assessment of aesthetic quality and multiple functions of urban green space from the users’ perspective: The case of Hangzhou Flower Garden, China. *Landsc. Urban Plan.* **2009**, *93*, 76–82. [CrossRef]
11. Salesses, P.; Schechtner, K.; Hidalgo, C.A. The collaborative image of the city: Mapping the inequality of urban perception. *PLoS ONE* **2013**, *8*, e68400. [CrossRef] [PubMed]
12. Liu, L.; Silva, E.A.; Wu, C.; Wang, H. A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. *Comput. Environ. Urban Syst.* **2017**, *65*, 113–125. [CrossRef]
13. Li, X.; Ratti, C.; Seiferling, I. Quantifying the shade provision of street trees in urban landscape: A case study in Boston, USA, using Google Street View. *Landsc. Urban Plan.* **2018**, *169*, 81–91. [CrossRef]
14. Li, X.; Zhang, C.; Li, W.; Kuzovkina, Y.A. Environmental inequities in terms of different types of urban greenery in Hartford, Connecticut. *Urban For. Urban Green.* **2016**, *18*, 163–172. [CrossRef]
15. Long, Y.; Liu, L. How green are the streets? An analysis for central areas of Chinese cities using Tencent Street View. *PLoS ONE* **2017**, *12*, e0171110. [CrossRef] [PubMed]
16. Small, C. Estimation of urban vegetation abundance by spectral mixture analysis. *Int. J. Remote Sens.* **2001**, *22*, 1305–1334. [CrossRef]

17. Antczak, E. Urban Greenery in the Greatest Polish Cities: Analysis of Spatial Concentration. *World Acad. Sci. Eng. Technol. Int. J. Transp. Veh. Eng.* **2017**, *11*, 578–585.
18. McCool, C.; Beattie, J.; Milford, M.; Bakker, J.D.; Moore, J.L.; Firn, J. Automating analysis of vegetation with computer vision: Cover estimates and classification. *Ecol. Evol.* **2018**. [[CrossRef](#)] [[PubMed](#)]
19. Zhaoping, L.; Li, Z. *Understanding Vision: Theory, Models, and Data*; Oxford University Press: Evans Road Cary, NC, USA, 2014.
20. Elder, J.H.; Victor, J.; Zucker, S.W. Understanding the statistics of the natural environment and their implications for vision. *Vis. Res.* **2016**, *120*, 1–4. [[CrossRef](#)] [[PubMed](#)]
21. De Cesarei, A.; Loftus, G.R.; Mastria, S.; Codispoti, M. Understanding natural scenes: Contributions of image statistics. *Neurosci. Biobehav. Rev.* **2017**, *74*, 44–57. [[CrossRef](#)] [[PubMed](#)]
22. Chéné, Y.; Belin, É.; Rousseau, D.; Chapeau-Blondeau, F. Multiscale analysis of depth images from natural scenes: Scaling in the depth of the woods. *Chaos Solitons Fract.* **2013**, *54*, 135–149. [[CrossRef](#)]
23. Adams, W.J.; Elder, J.H.; Graf, E.W.; Leyland, J.; Lugtigheid, A.J.; Murry, A. The southampton-york natural scenes (syms) dataset: Statistics of surface attitude. *Sci. Rep.* **2016**, *6*, 35805. [[CrossRef](#)] [[PubMed](#)]
24. Morris, N.J.; Avidan, S.; Matusik, W.; Pfister, H. Statistics of infrared images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2007, CVPR'07, Minneapolis, MT, USA, 18–23 June 2007; pp. 1–7.
25. Vaughn, I.J.; Alenin, A.S.; Tyo, J.S. Statistical scene generation for polarimetric imaging systems. *arXiv* **2017**, arXiv:1707.02723.
26. Balboa, R.M.; Grzywacz, N.M. Power spectra and distribution of contrasts of natural images from different habitats. *Vis. Res.* **2003**, *43*, 2527–2537. [[CrossRef](#)]
27. Rosch, E.; Mervis, C.B.; Gray, W.D.; Johnson, D.M.; Boyes-Braem, P. Basic objects in natural categories. *Cognit. Psychol.* **1976**, *8*, 382–439. [[CrossRef](#)]
28. Torralba, A.; Oliva, A. Statistics of natural image categories. *Netw. Comput. Neural Syst.* **2003**, *14*, 391–412. [[CrossRef](#)]
29. Samavatekbatan, A.; Gholami, S.; Karimimoshaver, M. Assessing the visual impact of physical features of tall buildings: Height, top, color. *Environ. Impact Assess. Rev.* **2016**, *57*, 53–62. [[CrossRef](#)]
30. Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 3213–3223.
31. Reza, A.M. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *J. VLSI Signal Process.* **2004**, *38*, 35–44. [[CrossRef](#)]
32. Lindenmayer, D.; Luck, G. Synthesis: Thresholds in conservation and management. *Biol. Conserv.* **2005**, *124*, 351–354. [[CrossRef](#)]
33. Fahrig, L. Effect of habitat fragmentation on the extinction threshold: A synthesis. *Ecol. Appl.* **2002**, *12*, 346–353.
34. Huggett, A.J. The concept and utility of ‘ecological thresholds’ in biodiversity conservation. *Biol. Conserv.* **2005**, *124*, 301–310. [[CrossRef](#)]
35. Chapeau-Blondeau, F.; Chauveau, J.; Rousseau, D.; Richard, P. Fractal structure in the color distribution of natural images. *Chaos Solitons Fract.* **2009**, *42*, 472–482. [[CrossRef](#)]
36. Chauveau, J.; Rousseau, D.; Chapeau-Blondeau, F. Fractal capacity dimension of three-dimensional histogram from color images. *Multidimens. Syst. Signal Process.* **2010**, *21*, 197–211. [[CrossRef](#)]
37. Chauveau, J.; Rousseau, D.; Richard, P.; Chapeau-Blondeau, F. Multifractal analysis of three-dimensional histogram from color images. *Chaos Solitons Fract.* **2010**, *43*, 57–67. [[CrossRef](#)]
38. Torralba, A.; Oliva, A. Depth estimation from image structure. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 1226–1238. [[CrossRef](#)]
39. Szeliski, R. *Computer Vision: Algorithms and Applications*; Springer Science & Business Media: Berlin, Germany, 2010.
40. Ziou, D.; Deschenes, F. Depth from defocus estimation in spatial domain. *Comput. Vis. Image Underst.* **2001**, *81*, 143–165. [[CrossRef](#)]
41. Rogez, G.; Orrite, C.; Guerrero, J.; Torr, P.H. Exploiting projective geometry for view-invariant monocular human motion analysis in man-made environments. *Comput. Vis. Image Underst.* **2014**, *120*, 126–140. [[CrossRef](#)]

42. Portman, M.E.; Natapov, A.; Fisher-Gewirtzman, D. To go where no man has gone before: Virtual reality in architecture, landscape architecture and environmental planning. *Comput. Environ. Urban Syst.* **2015**, *54*, 376–384. [\[CrossRef\]](#)
43. Kuliga, S.F.; Thrash, T.; Dalton, R.C.; Hoelscher, C. Virtual reality as an empirical research tool—Exploring user experience in a real building and a corresponding virtual model. *Comput. Environ. Urban Syst.* **2015**, *54*, 363–375. [\[CrossRef\]](#)
44. Fisher-Gewirtzman, D.; Portman, M.; Natapov, A.; Hölscher, C. Special electronic issue: “The use of virtual reality for environmental representations”. *Comput. Environ. Urban Syst.* **2017**, *62*, 97–98. [\[CrossRef\]](#)
45. Stamos, I.; Pollefeys, M.; Quan, L.; Mordohai, P.; Furukawa, Y. Special Issue on Large-Scale 3D Modeling of Urban Indoor or Outdoor Scenes from Images and Range Scans. *Comput. Vis. Image Underst.* **2017**, *157*, 1–2. [\[CrossRef\]](#)
46. Hirschmuller, H. Accurate and efficient stereo processing by semi-global matching and mutual information. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), San Diego, CA, USA, 20–26 June 2005; Volume 2, pp. 807–814.
47. Huang, X.; Cheng, X.; Geng, Q.; Cao, B.; Zhou, D.; Wang, P.; Lin, Y.; Yang, R. The ApolloScape Dataset for Autonomous Driving. *arXiv* **2018**, arXiv:1803.06184.
48. Saxena, A.; Chung, S.H.; Ng, A.Y. Learning depth from single monocular images. In *Advances in Neural Information Processing Systems*; The MIT Press: Cambridge, MA, USA, 2006; pp. 1161–1168.
49. Geiger, A.; Lenz, P.; Stiller, C.; Urtasun, R. Vision meets Robotics: The KITTI Dataset. *Int. J. Robot. Res. (IJRR)* **2013**, *32*, 1231–1237. [\[CrossRef\]](#)
50. Zendel, O.; Murschitz, M.; Humenberger, M.; Herzner, W. How Good Is My Test Data? Introducing Safety Analysis for Computer Vision. *Int. J. Comput. Vis.* **2017**, *125*, 95–109. [\[CrossRef\]](#)
51. Neuhold, G.; Ollmann, T.; Bulò, S.R.; Kotschieder, P. The mapillary vistas dataset for semantic understanding of street scenes. In Proceedings of the International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 22–29.
52. Simond, N.; Rives, P. Homography from a vanishing point in urban scenes. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2003, Las Vegas, NV, USA, 27–31 October 2003; Volume 1, pp. 1005–1010.
53. Zhou, Z.; He, S.; Li, J.; Wang, J.Z. Modeling perspective effects in photographic composition. In Proceedings of the 23rd ACM International Conference on Multimedia, ACM, Reykjavik, Iceland, 4–6 January 2015; pp. 301–310.
54. Li, B.; Peng, K.; Ying, X.; Zha, H. Vanishing point detection using cascaded 1D Hough Transform from single images. *Pattern Recognit. Lett.* **2012**, *33*, 1–8. [\[CrossRef\]](#)
55. Canny, J. A computational approach to edge detection. In *Readings in Computer Vision*; Elsevier: Amsterdam, The Netherlands, 1987; pp. 184–203.
56. Chang, H.; Tsai, F. Reconstructing Three-Dimensional Specific Curve Building Models from a Single Perspective View Image. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *39*, 101–106. [\[CrossRef\]](#)
57. Liu, F.; Shen, C.; Lin, G.; Reid, I. Learning depth from single monocular images using deep convolutional neural fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 2024–2039. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Bertamini, M.; Martinovic, J.; Wuerger, S.M. Integration of ordinal and metric cues in depth processing. *J. Vis.* **2008**, *8*, 1–12. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Rzeszutek, R.; Androutsos, D. A framework for estimating relative depth in video. *Comput. Vis. Image Underst.* **2015**, *133*, 15–29. [\[CrossRef\]](#)
60. Turski, J. The conformal camera in modeling active binocular vision. *Symmetry* **2016**, *8*, 88. [\[CrossRef\]](#)
61. Itti, L.; Koch, C.; Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 1254–1259. [\[CrossRef\]](#)
62. Duchowski, A.T. *Eye Tracking Methodology*; Springer: London, UK, 2007.
63. Rigau, J.; Feixas, M.; Sbert, M. Informational aesthetics measures. *IEEE Comput. Graph. Appl.* **2008**, *28*, 24–34. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Dresch-Langley, B. Affine geometry, visual sensation, and preference for symmetry of things in a thing. *Symmetry* **2016**, *8*, 127. [\[CrossRef\]](#)
65. Chen, C.C.; Wu, J.H.; Wu, C.C. Reduction of image complexity explains aesthetic preference for symmetry. *Symmetry* **2011**, *3*, 443–456. [\[CrossRef\]](#)

66. Batmaz, A.U.; de Mathelin, M.; Dresch-Langley, B. Effects of Image Size and Structural Complexity on Time and Precision of Hand Movements in Head Mounted Virtual Reality. In Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Stadthalle, Austria, 18–22 March 2018; pp. 167–174.
67. Goodfellow, I.; Bengio, Y.; Courville, A.; Bengio, Y. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016; Volume 1.
68. Sussman, A.; Hollander, J.B. *Cognitive Architecture: Designing for How We Respond to The Built Environment*; Routledge: Abingdon, UK, 2014.
69. Zacharias, J. Preferences for view corridors through the urban environment. *Landsc. Urban Plan.* **1999**, *43*, 217–225. [[CrossRef](#)]
70. Danahy, J.W. Technology for dynamic viewing and peripheral vision in landscape visualization. *Landsc. Urban Plan.* **2001**, *54*, 127–138. [[CrossRef](#)]
71. Stamps, A.E. Fractals, skylines, nature and beauty. *Landsc. Urban Plan.* **2002**, *60*, 163–184. [[CrossRef](#)]
72. Van den Berg, A.E.; Joye, Y.; Koole, S.L. Why viewing nature is more fascinating and restorative than viewing buildings: A closer look at perceived complexity. *Urban For. Urban Green.* **2016**, *20*, 397–401. [[CrossRef](#)]
73. Casalegno, S.; Anderson, K.; Hancock, S.; Gaston, K.J. Improving models of urban greenspace: From vegetation surface cover to volumetric survey using waveform laser scanning. *Methods Ecol. Evol.* **2017**, *8*, 1443–1452. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).