

Article

Fractal Dimension Calculation and Visual Attention Simulation: Assessing the Visual Character of an Architectural Façade

Ju Hyun Lee *  and Michael J. Ostwald 

UNSW Built Environment, Faculty of Arts, Design and Architecture, The University of New South Wales, Sydney, NSW 2052, Australia; m.ostwald@unsw.edu.au

* Correspondence: juhyun.lee@unsw.edu.au

Abstract: The design of a building façade has a significant impact on the way people respond to it physiologically and behaviourally. Few methods are available to assist an architect to understand such impacts during the design process. Thus, this paper examines the viability of using two computational methods to examine potential visual stimulus-sensation relationships in facade design. The first method, fractal analysis, is used to holistically measure the visual stimuli of a design. This paper describes both the box counting (density) and differential box counting (intensity) approaches to determining fractal dimension (D) in architecture. The second method, visual attention simulation, is used to explore pre-attentive processing and sensation in vision. Four measures— D -density (D_d), D -intensity (D_i), heat map and gaze sequence—are used to provide quantitative and qualitative indicators of the ways people read different design options. Using two façade designs as examples, the results of this application reveal that the D values of a façade image have a relationship with the pre-attentive processing shown in heat map and gaze sequence simulations. The findings are framed as a methodological contribution to the field, but also to the disciplinary knowledge gap about the stimulus-sensation relationship and visual reasoning in design.



Citation: Lee, J.H.; Ostwald, M.J. Fractal Dimension Calculation and Visual Attention Simulation: Assessing the Visual Character of an Architectural Façade. *Buildings* **2021**, *11*, 163. <https://doi.org/10.3390/buildings11040163>

Academic Editor: Ewa Janina Grabska

Received: 30 March 2021

Accepted: 13 April 2021

Published: 15 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: architectural design; visual complexity; visual attention; pre-attentive vision; conceptual design; differential box counting; eye-tracking

1. Introduction

The architectural façade, or external face of a building, is the first, physical interface most people engage with as they approach a building, providing important visual clues about its functionalities. People unconsciously see and continuously experience building façades and their elements (geometry, patterns, textures and details) [1]. Thus, creating the façade is an important part of the design process, not only for aesthetic reasons, but also because the façade is central to the relationship between visual character and human response. However, predicting the visual consequence of the façade using analytical tools is not formally investigated in the design process, which indicates a significant gap in disciplinary knowledge. As such, a major research question which this paper addresses is, “how can a designer proactively examine the visual character of a façade and its potential impacts during the early stages of the design process?” Before describing how this question will be answered, a brief overview of past research provides a context for this introduction.

Two key measurable properties of visual character are stimulus and perception. The former, stimulus, is associated with visual intensity whereas the latter, perception, relates to visual magnitude. The two are central to classical debates in psychophysics about visual character, with some scholars arguing that the relationship between them is linear, while others propose logarithmic (Weber-Fechner law) or exponential laws to describe them (Steven's power law) [2,3]. Despite the existence of this research, our understanding of the relationship between visual stimuli and perceptual parameters in building design is limited. A small number of studies have examined participants' evaluative responses to

façade design in a laboratory setting. Stamps [1], for example, uses participants' ratings of 16 computer-generated façade drawings to examine human visual preferences for architectural façades, with a focus on three types of stimuli (silhouette complexity, surface complexity and facade articulation). Justi and Eric [4] also examine the impact of urban façades on affective feelings. In their research, two levels (high and low) of façade stimuli that depict pedestrian-oriented façade designs and edges are evaluated by participants, using sliding scales that range from "not very pleasant" to "very pleasant". However, as authors acknowledge such rating methods have unavoidable limitations including innate and selection biases. Psychological evaluations of this type also depend on participants' life experiences and cultural backgrounds. Furthermore, participants' emotional ratings in the experimental research do not explain how they actually perceive the design or how visual stimuli direct human reactions and experiences. In contrast, observational and interview techniques can be employed to empirically investigate people's engagement with buildings [5]. Observational research potentially offers an insightful approach to understanding dynamic and emotional responses to building design and can also provide indicators about the impact of a façade on people's daily activities. For example, an observational study undertaken in Copenhagen captures the types of encounters people have with façades [6]. However, observer bias in such studies is difficult to control [7] and individuals' own perspectives are often ignored [8].

Importantly, most past research has largely examined built or existing façades [1,4,6], and the methods employed for this purpose are not appropriate for rapid exploration of stimulus-sensation relationships during the design process. Furthermore, the goal of understanding stimulus-sensation relationships during design is different to that of exploring its post-completion impact. During the design stage rapid methods are required to support design decision-making and visual reasoning while testing alternative façade compositions and expressions. Design decision-making and problem solving involve selecting a solution from viable solutions or options, a process that is largely reliant on designer's expertise and intuition [9]. Whereas there are multiple design decision support models and systems available for architects for other tasks [10], there are relatively few to support visual assessment during the design process. Specifically, a façade design contains a combination of perceptual and geometric information, which can be regarded as a set of visual stimuli. Human attention affects perception [11], because attention processes respond to the visual stimuli, shaping our visual experience and behaviour [5,12–15]. Thus, exploring the visual complexity of, and our visual attention to, a façade during the early conceptual or schematic design stages not only plays a significant role in visual design thinking processes, but also enables predicting visual perception of the design.

There are further options and issues to consider when looking at potential methods for examining visual stimulus and perception in façade design. For example, participatory design [16] is one approach that seeks to engage laypeople in the production of a design, thereby providing a broader response to a potential façade option. Participatory design is, however, challenging, time-consuming and potentially expensive, being generally inappropriate for a conceptual design stage. Some examples do exist of collective design at the conceptual design stage, with hundreds of motivated individuals working on a project [17], but this does not necessarily support design decision-making in response to visual perception. There is also, no clear or accepted way of formally investigating the relationship between visual stimuli of a building facade and human visual processing. In response to all of these issues, the present research introduces and assesses two efficient computational approaches to (i) quantifying the visual stimuli or complexity of a façade in a holistic way and (ii) simulating visual attention to the façade. Both approaches are applicable to early design stages and are introduced hereafter.

The first approach treats the formal and geometric characteristics of a building as indicators of visual stimulus. Multiple studies have sought to quantify the formal properties of building exteriors [1,4,18]—typically through examination of geometric and/or non-geometric variables—but they tend to focus on discrete architectural properties, rather

than an entire façade. One method, however, measures the visual complexity of an entire façade, providing a mathematical indicator of the volume of visual stimulus spread across it. This method of fractal dimension analysis has been used for the measurement of visual complexity in art and architecture [19,20], and it has been shown to be significant predictor of aesthetic preference. Fractal dimensions have been used to quantify the complex structure of both natural patterns and aesthetic works [21]. In addition, recent studies reveal that fractal dimensions play a role in the experience of architectural and urban form [20,22,23]. Thus, the first method examined in this paper is fractal dimension analysis, which provides an efficient means of understanding visual complexity and stimulus in a façade.

The second approach in this paper investigates visual attention, which is a key factor in human spatial experience. For over a century, visual attention has been studied in psychology, neurophysiology and perceptual science [11]. In the fields of architecture and urban design, visual attention has been explored through the use of eye-tracking technology [5,12–14]. Importantly, such research shows how pre-attentive processing in vision can reveal the ways the visual stimuli of a building facade instantly attract the eyes, and thereby predict visual experience and human behaviours [24]. Hollander et al. [25] also highlight how unconscious (pre-attentive) eye movements shape human responses to urban environments. Thus, the present paper focuses on “pre-attentive” or “first-glance” vision (the initial 3–5 s of viewing), that is automatically processed without consciousness. Past eye-tracking studies have, however, been time-consuming and labour-intensive to undertake, whereas recent developments in visual attention software (VAS) emulate the results of human eye-tracking experiments, providing fast and effective results [26,27]. VAS algorithmically captures and simulates human “pre-attentive processing” [28]. The present paper uses VAS to simulate potential unconscious visual responses to façade design.

This paper is divided into three main sections. The following section introduces fractal dimensions and visual attention in architectural design. Following this, the two computational methods are described and used to simulate the visual impacts of architectural drawings of two facades. The simulation results, using two applications (ImageJ (version 1.53e)+FracLac (version 2015Sep090313a9330) and VAS (version 1.0.6)), are then presented with a discussion about their implications and limitations.

2. Fractal Dimension and Visual Attention

2.1. Fractal Dimension Calculation

Fractal Dimension analysis is a rigorous approach to measuring “the relative density and diversity of geometric information in an image or object” [20] (p. 3). Whereas the close observation of fractal geometry reveals “a cascade of never-ending, self-similar, meandering detail as one observes them more closely. The fractal dimension is a mathematical measure of meandering of the texture displayed . . . [it] provides a quantifiable measure of the mixture of order and surprise in a rhythmic composition” [29] (p. 3). Even though there is often confusion about the difference between fractal dimension and fractal geometry, Ostwald and Vaughan [20] distinguish fractal geometry—particular geometric sets only existing in topological space and exhibiting high levels of self-similarity—from fractal dimension that is used to describe “the space-filing properties of irregular objects which may exist in either topological or the material world” (p. 9). Figure 1a shows an example of fractal geometry, the Sierpiński triangle.

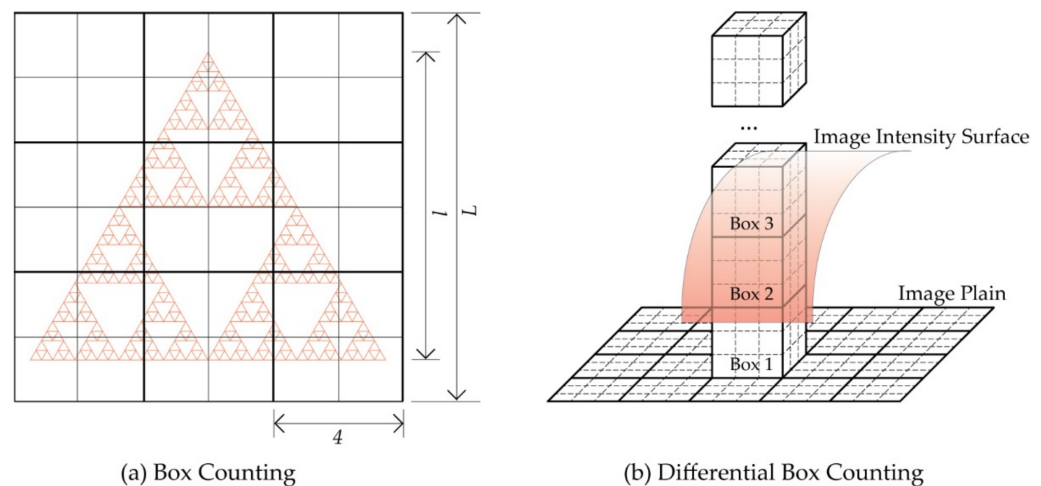


Figure 1. Determining the number of boxes (or squares) (N_s) by (a) box counting and (b) differential box counting (adapted from [20,30]).

Even though Mandelbrot identifies multiple approaches to measuring fractal dimensions [20,31], two types of box counting methods, (a) traditional box counting and (b) differential box counting, as shown in Figure 1, can be applied to determine the fractal dimension of an image, or of a façade. While both methods are used in different domains, the former has been explored in many architectural studies, while the latter is only rarely used of architectural or urban research. Theoretically, the traditional box counting method is reliable and accurate [29,32]. Since being introduced for architectural and urban research in the 1990s [29,32], a range of studies have used box counting to analyse various architectural forms: Mesoamerican architecture [33], Teotihuacan architecture [34], Ottoman houses [35], the Wright’s Robie House and Le Corbusier’s Villa Savoye [29], Sinan’s Kılıç Ali Paşa Mosque [36], 85 20th Century houses (designed by 16 famous architects) [20] and French Gothic cathedrals [37].

The fractal dimension (D) of an image, as calculated using the box counting method, is determined by the slope of the least square regression line (or trendline) for the log-log plot of box count (N_s) versus box size (s). The original box-counting theorem is:

$$D = \lim_{\epsilon \rightarrow 0} \frac{\log N_s}{\log(1/s)} \quad (1)$$

In essence, the dimension can be estimated by a comparison between two sequential grids [20], using this formula:

$$D = \frac{[\log N_{s2} - \log N_{s1}]}{[\log(1/s2) - \log(1/s1)]} \quad (2)$$

Where s_2 is a smaller (or magnified) grid than s_1 . For example, in the 3×3 grid (thick lines) of Figure 1a the image has 7 boxes with detail (or foreground pixel) contained in them, and its box size (scale) is 4. The second grid (6×6 grid) using thin lines identifies 24 boxes and its box size is 2. Thus, the fractal dimension using only two grids is:

$$D = \frac{\log 24 - \log 7}{\log \frac{1}{2} - \log \frac{1}{4}} = \frac{1.7776}{1.0000} = 1.7776$$

This simple calculation is useful for a quick determination, but the fractal dimension (D) using multiple box sizes should use the slope of the least square regression line. The choice of a logarithmic base is not important because fractal dimension is the slope of the line or the log-log relation. Theoretically, it is 2, because the box counting is basically involving binary decisions (bits).

While the box counting method is useful for determining the fractal dimension of architectural elevations and plans [20], it is only applicable for binary images. However, while most plans and elevations are binary images, this image type cannot capture the texture of a building, which is part of the visual complexity of a façade. Stamps' "surface complexity" (percentage of pixels covered by elements in the second *septave*) is the most important factor for visual preference in his study [1]. In addition, unlike experienced architectural designers, laypeople may find it difficult to interpret a line drawing of a façade that represents a complex, articulated form. Thus, using differential box counting, which enables the analysis of textural complexity of grey scale images, may compensate for this issue [30,38].

The differential box counting method can estimate the fractal dimension of grey scale images [30,38,39]. It is a generalization of the traditional box counting method to take into account an additional textural dimensional, image intensity: see Figure 1b. In other words, the third coordinate (z) encapsulates grey-levels ranging from 0 to 255 [39]. Such a box count (N_s) requires a different approach from the binary process of traditional box counting. Firstly, let the minimum and maximum grey-levels (intensity) of the surface in the image plain of a grid (i, j) fall in box number k and l , respectively [30,39]. The contribution (n_s) to the box counting is then:

$$n_s(i, j) = l - k + 1 \quad (3)$$

For example, the box counting of the image intensity surface in Figure 1b is $3 - 1 + 1$. In this way, the box counting that considers contributions from all blocks is:

$$N_s = \sum n_s(i, j) \quad (4)$$

The final fractal dimension is then estimated using the least square regression line of the log-log plot as described above. In this way, the fractal dimension of a grey scale image can be determined by both box counting methods, because it can be automatically converted to a binary image. Thus, this paper analysing grey scale façade images produces two types of fractal dimensions, (a) D -density (D_d) using the box counting method and (b) D -intensity (D_i) using the differential box counting method.

In essence, regardless of D_d and D_i , the values of fractal dimensions (D) of images are between 1 and 2, indicating the degree of complexity [20,40]. A lower D value indicates a less fractal or less complex form, while a higher D , indicates a steeper slope of the trendline, which is more complex, differentiable or irregular. The latter also means that complexity increases as the box size decreases. Thus, the determination of box sizes is very important for both methods. Previous studies suggest box sizes should be between $0.25l$ (maximum) and $0.03l$ (minimum), where l is the height of the shorter side of either the outline (or silhouette) of an image [40–42] or simply a bounding box (*bbox*)—the smallest rectangle enclosing the foreground pixels of a digital image [43]. For example, the maximum box size in a $L \times L$ image in Figure 1a is $0.25l$. This ideal maximum size is still valid for this research, but the minimum size of $0.03l$ may be less so for small size images like 480×360 pixel greyscale images [42]. However, in our multiple pilot tests, it is observed that even small boxes between 5 to 10 pixels can generate meaningful box counting without an error. Thus, this study uses a minimum box size of 10 pixels, regardless of the height (l).

In past research fractal dimensions have been correlated to human visual preference [19] and there is evidence that they can be used to develop insights into the aesthetic experiences of art, architecture and urban landscape [20,22]. Past studies have used various types of images—in terms of details, sizes and patterns—to identify different preference ranges of D [21]. For example, some researchers argue that D values in the range of 1.3–1.5 are the most preferred [44], while others propose that the most preferred D value for computer-generated grid-like patterns is around 1.8 [45]. In [20], the overall average D of more than 300 elevations is 1.4148 and its interquartile range is from 1.3698 to 1.5104. Thus, if appropriate knowledge about the preferable D range is developed, it can be used to evaluate the visual complexity of a design and predict human visual preferences. Thus,

it is very important for applications of fractal analysis to consistently use optimal settings (e.g., grid sizes, scaling coefficient and grid positions) and image processing approaches. In this way, D offers a powerful, repeatable measure for examining visual complexity in design, which can in turn be used to understand how architecture reflects human physical experience and response [23]. Consequently, D provides a connection between form, for example of a façade, and “sensual or cognitive participation” [46].

2.2. Visual Attention Simulation

Visual attention refers to a set of cognitive operations associated with sight. It enables us to selectively process visual scenes that consist of both relevant and irrelevant information [47]. This selectivity is significant, because human capacity to process visual information is limited in various ways [11,12]. Visual attention can be categorised into three types: spatial attention, feature-based attention and object-based attention. The common feature of all three, “attention”, relates to the way an observer is led to look at a particular location (spatial attention) or specific aspects (e.g., colour or orientation) of objects in an environment (feature-based attention). Attention can also be guided by object structure (object-based attention) [11]. In addition, people’s visual experiences depend on multiple factors, such as viewing distance, moving speed and visual permeability [6]. Furthermore, visual attention is related to visual working memory and pattern recognition [48]. As such, human cognitive processing depends on visual information including the geometrical characteristics of forms and images [26].

Traditionally, visual attention has been divided into two sequential types, pre-attentive and attentive [24,49]. According to a “feature-integration theory” of attention [50], pre-attentive processing in vision extracts object features (visual stimuli) such as colour, size, orientation and direction [24,49] and develops “proto-objects”, which are then selectively formed into a coherent object representation during subsequent attentive processing [51–53]. Figure 2 illustrates Rensink’s human visual system model consisting of (i) early visual, (ii) attentional and (iii) setting systems [51,52]. The early visual system is a high-capacity process that rapidly and unconsciously creates proto-objects across the visual field. The limited-capacity attentional system constitutes the set of proto-objects into a coherent representation, spatially and temporally (spatio-temporal coherence). The setting system is a non-attentional system that provides a context to guide attention, allocating focused attention based on long-term knowledge. These three processes are also reminiscent of aesthetic responses in Nasar’s probabilistic model [18], where perception of building attributes is followed by cognition (judgments of building attributes). Perception directly effects emotional reactions and cognition has affective appraisals and connotative meanings [18]. The combination of perception and cognition develops our aesthetic responses.

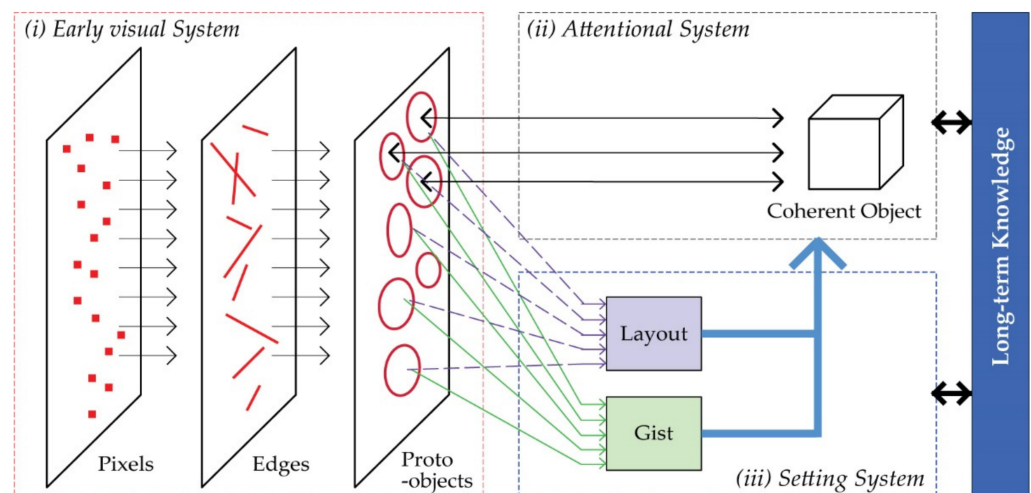


Figure 2. Visual attention mechanism (adapted from [51–53]).

As described in Figure 2, the attentional system in vision consciously refers to individual long-term knowledge, relying on individual intentions or goals. In contrast, the pre-attentive is unconsciously dealing with information (visual stimuli) from the environment. Importantly, the early visual system has three stages, transduction (pixels), primary processing (edges) and secondary processing (proto-objects) stages [53]. In the early vision, photoreception occurs and linear or quasi-linear filters measure image properties. Finally, proto-objects become the operands for the following attentional system [53]. More salient proto-objects attract more attention [54]. In this way, the pre-attentive stage unconsciously responds to the visual stimuli of architecture, formulating our visual experience in the built environment. Thus, this paper specifically addresses early vision, where stimulus salience (unconscious) can be more important than intentions or goals (conscious). Early vision offers an effective indicator of the stimulus-sensation relationship in design.

In the vision system, not all visual information (or stimuli) of a design can be processed at any one time, because of our limited processing capacity [47]. Visual attention is processed either overtly or covertly and deals with either a particular character of an object or of an entire object [47]. As past observation research has been unable to reveal these selective processes in vision or visual experience (e.g., [1,4,6]), eye-tracking technologies have been used to more precisely capture eye movements and fixations in response to varying façade designs [12–15].

Eye-tracking can be used to identify both the attentive elements of a design and the lengths of attention [12]. Thus, it captures participants' visual explorations and experiences, because gaze distributions can show how people visually engage with the design. In addition, mobile eye-tracking has been used to capture visual engagement with facades or urban street edges in everyday life. Past studies [5] reveal that social and spatial factors impact such visual engagement, and the amount of visual engagement is also dependent on pedestrians' everyday activities as well as street types (non-pedestrianised and pedestrianised). Eye-tracking technology is also used to capture people's perception of cityscape design in a virtual reality (VR) environment [55]. Furthermore, human interaction with 3D virtual models can be investigated through eye-tracking in terms of end-users participation in the design process or their satisfaction [56]. Thus, both screen-based and wearable eye-tracking technologies allow for the investigation of how architectural elements (geometric edges, contrasts, intensity and facial features) impact on visual attention [57]. Regardless of the different research themes being investigated in past research, two forms of graphical presentations of eye-tracking data, heat map (or hotspot map) and gaze sequence (fixation map or gaze plot) are common (Figure 3). Both of these graphical representations describe the participants' fixations and eye movements [14,15].

The heat map is a 2D density plot that represents where the participant "attentively observes" [12]. Thus, it identifies which part (or element) attracts the participant's attention. The heat map illustrates both the points that the participant looked at and the fixation time [15]. For example, a more intense warm colour indicates a longer fixation time on that location. That is, the dark "red" areas in Figure 3a are the most attractive ("hot") spots. The non-coloured areas are ones the participant ignores. As such, the heat map captures the intensity of the participant's attention and fixation activities. "As interest in a stimulus increases, the eye tends to fixate more (i.e., increased fixation count) and for longer durations" [25] (p.159). The second representation type, gaze sequence—also sometimes called a "fixation map", but "gaze sequence" better captures its content—addresses the sequence of gaze points (fixations) that occur during the observation. The gaze plot commonly illustrates a series of time-stamped, connected circles that represent the way the eye moves across an image. As shown in Figure 3b, the diameter of each circle can vary corresponding to the fixation time [15]. The number inside the circle shows the sequence of fixations and lines represent the paths between gaze points. Due to the limited human capacity of visual processing, attention is allocated one task at a time [12,58]. Thus, these data-based graphics are used in the fields of architecture and urban design to model and understand the allocation of visual attention over a visual

stimulus. Obviously, an “engaging” façade can better attract public interest, which is a factor that an architect or urban planner should take into account, depending on the type of building being designed [26]. In this context, the heat map and gaze sequence can help designers to examine the way their façade proposal will visually engage with people after it is constructed. Furthermore, the visual attention simulation method can be applied to examine and predict aesthetic experience in artworks such as painting and sculpture. In terms of dynamism, artists can examine visual saliency within a painting [59], and quantify how audiences will observe it [60]. Specifically, if an artwork has a clear purpose in attracting the eyes of viewers, knowing the eye fixation data in the artwork would be very useful for shaping its creation.



(a) Heat map



(b) Gaze sequence

Figure 3. Heat map and gaze sequence example (Medieval facades on the market square of Bremen, Germany).

3. Methodology

In this paper, both fractal dimensions and visual attention properties of façades are examined using two software programs, ImageJ+Fractal [43,61] and VAS [26,28], respectively. Using these programs, this paper presents four measures— D -density (D_d), D -intensity (D_i), heat map and gaze sequence—which capture the stimulus-sensation relationship in a facade. To present D_d and D_i , the first stage of this research (stimulus) measures two types of fractal dimensions, (i) a box-counting method for binary facade images and (ii) a differential box-counting method for grey scale images. The second stage (sensation) simulates the pre-attentive vision over each facade image used in the first stage, visualising its (iii) heat map and (iv) gaze sequence. Finally, this paper discusses the computational results, focusing on the visual character of a facade and its impacts.

3.1. Two Computational Applications (ImageJ+Fractal and VAS)

ImageJ is an image analysis program, developed by Wayne Rasband at the National Institutes of Health (NIH), which is largely used in the biological and physical sciences [61]. Karperien [43] developed Fractal, open-source software for fractal analysis using ImageJ's box counting algorithm. Even though there are other programs available for fractal analysis, the combination of ImageJ+Fractal is ideal for the purposes of this paper because it can calculate both D_d and D_i .

To consistently determine the fractal dimension of each facade image, this research uses Ostwald's and Vaughan's optimal architectural settings in Table 1 [20]. The scaling coefficient indicates the ratio between one grid and the next. In the architectural application of box counting method, a ratio of $\sqrt{2}:1$ (approximately, 1.4142:1) can be optimal [20]. In addition, the final value of fractal dimension is the average of four D values determined by four grid positions (four corners of the pixelated part (outline) of the image), because using a different grid position can result in a different box count (N_s).

Table 1. Variables and settings for fractal analysis.

Variable	Optimal Setting	Variable	Optimal Setting
Scaling coefficient	1.4142:1	Maximum grid size	0.25 l
Grid position	Four corners *	Minimum grid size	10 pixels

* Four corners of the pixelated part (outline) of the image ($bbox$).

Visual attention software (VAS) is used to mimic human perception and reproduce eye-tracking data without the use of real eye-tracking experiments [26]. Even though actual eye-tracking can produce more meaningful results, VAS provides a rapid and effective simulation of pre-attentive processing which is more useful during design. VAS models early vision caused by four elements (edges, intensity of colour hue, contrast among colour hues and similarity to a face). Thus, pre-attentive vision can be algorithmically predicted [26,28]. In addition, the VAS algorithm is theoretically based on Salingeros' mathematical model of design complexity [62] and cognitive Architecture [46]. Since the software has already been verified in cognitive psychology and neuroscience [26], this paper uses it to model pre-attentive processing in vision and to produce heat maps and gaze sequences. In VAS, heat maps show the probability that each part of the image is seen during the early vision, while gaze sequences visualise the four most likely gaze locations in a sequential order.

3.2. Image Processing

Since VAS requires a minimum image size of 600×600 pixels [28], the most important setting for both accurate VAS and fractal analysis is image size. In general, there is no singular optimal size for analysing architectural drawings or images, although high-quality images will typically produce more accurate D values [20,43,63]. For example, Mayrhofer-Reinhartshuber and Ahammer [63] test seven image sizes from 128×128 ($2^7 \times 2^7$) to 8192×8192 ($2^{13} \times 2^{13}$) pixels for fractal analysis. Their study on a new "pyramidal

fractal dimension" reveals 1024×1024 ($2^{10} \times 2^{10}$) pixels or larger images lead to increased accuracy of D values. Their comparison between multiple fractal dimension methods including the differential box counting then adopts a resolution of 4096×4096 ($2^{12} \times 2^{12}$) pixels. In this context, the present research tests two sizes of façade images (*bbox* sizes), 1024×1024 and 4096×4096 .

Unlike the use of scanned images in the biological sciences, this study analyses line drawings and rendered images produced using Building Information Modelling (BIM) software (Autodesk Revit and ArchiCAD), which exports a design image as a PDF or a JPG. For this purpose it is important that both line weight and image resolution are "1 pt" and "175 dpi", respectively, as suggested in [20]. However, the fractal analysis software processes bitmap images, not vector graphics. That is, any vector graphic should be converted to a bitmap image like a JPG. In addition, when comparing two or more different designs all images require, consistent image size and therefore additional image processing using a graphics editor like Adobe Photoshop is required. Furthermore, an image for fractal analysis is basically regarded as a scanned image that is full of grey scale pixelation. However, architectural drawings or rendered images can involve white space. Since the software used in this paper automatically detects the *bbox* of an image, the optimal white space (e.g., 40%–50% of image size) [20] is not necessary. Nonetheless, a marginal white space is considered for this image processing to make various samples a consistent image size. Given the maximum box counting size (0.25*l*), this paper considers white space of a quarter of the length of the maximum *bbox* size in Figure 4. Collectively, the final image (field) size developed in this paper is 1536×1536 or 6144×6144 pixels. Nevertheless, the maximum box size is still 0.25*l* and the minimum box size is 10 pixels. All of these settings ensure reliable, consistent image comparisons.

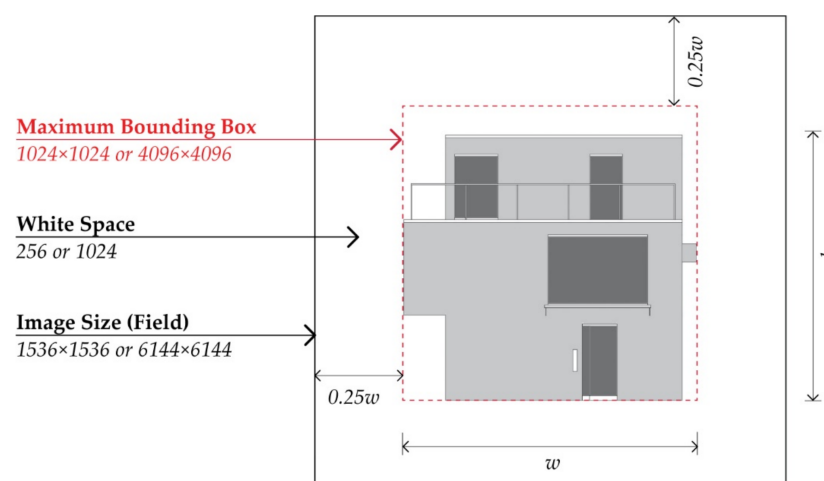


Figure 4. Image size processing example (Atelier Bow-Wow's Juicy House).

In addition to these methodological settings or parameters, this paper compares several different levels of representation and the impacts of two "entourage elements" (which are scale features architects add to drawings). Five "levels" of façade design representation used in the following early design stage tests are: level 1—a line drawing, level 2—a façade design rendered with one grey scale for a main body of building only (brightness: 80%), level 3—a façade design rendered with two grey scales for a main body of building (brightness: 80%) and openings (brightness: 40%), level 4—a façade design rendered with two grey scales and a tree, and level 5—a façade design rendered with two grey scales, a tree and three people.

Entourage elements, typically people, cars and trees, are added to architectural drawings to give a sense of human scale, and to assist people to imagine themselves within a building or streetscape. In practice, there is an assumption that such elements can shape the way people view and respond to an elevation or plan, which is why in most computational

studies they are excluded because they are confounding factors [34]. However, in a study which considers visual fixation and attention, their inclusion can be informative.

Figure 5 illustrates the five levels of representation of Frank Gehry's Spiller House and their binary images for the box counting method of fractal analysis. As shown in this figure, the binary images of levels 1 and 2 are the same because light grey colours (over 50% of brightness) become white in a binary image.

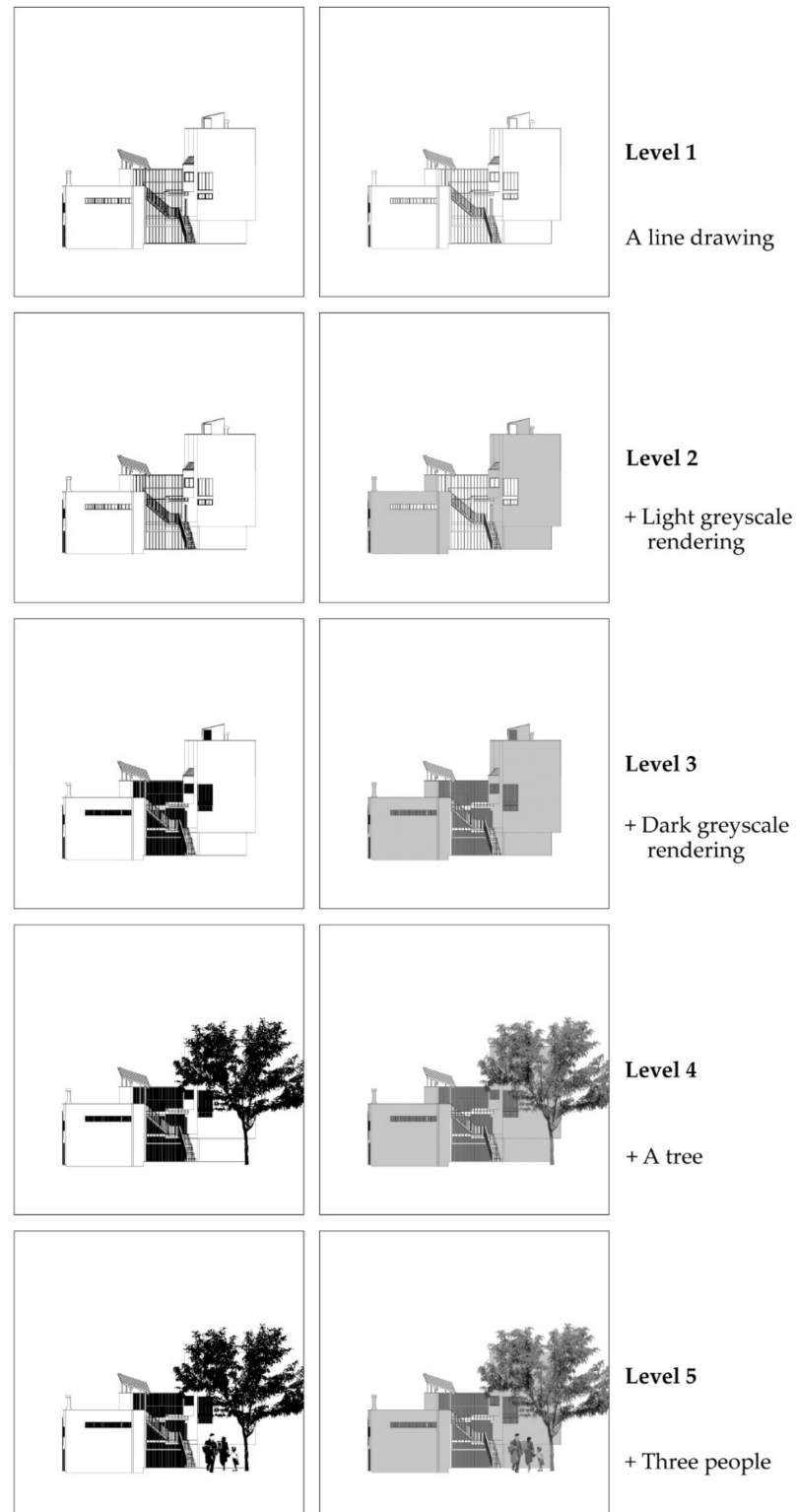


Figure 5. The five levels of representations of Frank Gehry's Spiller House.

The following section presents the result of four measures— D -density (D_d), D -intensity (D_i), heat map and gaze sequence—of five levels of representation for two façade designs, Spiller House and Juicy House. These facades (one Postmodernist and one Minimalist) were chosen based on past research which identifies them as possessing, respectively, complex and simple forms [34]. For example, Frank Gehry’s 1980 Spiller House (which is sometimes regarded as an early deconstructivist work) appears as a series of boxes and rectangles, with exposed stud frames creating multiple window mullions. Its scaffolding-like character gives it a complex visual quality. In contrast, Atelier Bow-Wow’s 2004 Juicy House is a relatively simple and compact design for a residence as shown in Figure 4. The same size and position of a tree and people are applied for the fourth and fifth levels of representation of the two designs.

The two different styles of architecture are selected only for the demonstration purpose of the methodology in this paper. It is, however, significant that the two computational approaches have been independently applied to characterise and compare architectural styles [20,26]. That is, the combined method presented in this paper can be used to holistically investigate how different architectural styles and their design elements work in terms of both visual complexity and engagement.

4. Results

4.1. Fractal Dimension

Table 2 describes the D -density (D_d) values of the Spiller House and Juicy House façades, along with the different levels of representation. The table also reports the D_d values of the different image sizes (maximum *bbox* sizes). When using the box counting method, the D_d values of levels 1 and 2 are theoretically the same due to the use of binary images (see Figure 5). The minor differences between the two levels in Table 2 are caused by the automated image conversion of the software (ImageJ+FracLac). Expectedly, the fractal dimension increases with each additional component (renderings, a tree and people) included in the representation. Thus, the last level develops the highest D_d value of each image. Notably, the fractal dimensions of the Spiller House are higher than those of the Juicy House, regardless of the levels of representation. This result confirms that the former façade is a more complex than the latter. However, from the third level, the disparities between the two are significantly reduced. Interestingly, the D_d values of the first levels of representation (line drawing only) of the low-quality images (1024×1024) are higher than the high-quality images (4096×4096). In contrast, the D_d values of the last three levels of the small images are lower. This anomaly may be caused by the image conversion process (from vector to bitmap) and can potentially lead to inconsistent results.

Table 2. D -density (D_d) values of two façades.

Façade Design	Image Size (<i>bbox</i> Size)	Level 1	Level 2	Level 3	Level 4	Level 5
Spiller House	1024×1024	1.5366	1.5333	1.5489	1.7251	1.7431
	4096×4096	1.5241	1.5245	1.6140	1.7678	1.7803
Juicy House	1024×1024	1.3076	1.3063	1.5255	1.6597	1.6859
	4096×4096	1.2856	1.2856	1.6033	1.7266	1.7357

The final D values are the average of four D values determined by four grid positions (four corners of the pixelated part of the image) as shown in Figure 6, because using a different grid position can result in a different box count (N_s). For example, the D value of the low-quality (1024×1024) image (level 1 of the Spiller House) is 1.5366, which is the average of four slopes of the least square regression lines in Figure 6a:

$$D = \frac{1.5495 + 1.5264 + 1.5372 + 1.5335}{4} = \frac{6.1466}{4} = 1.5366.$$

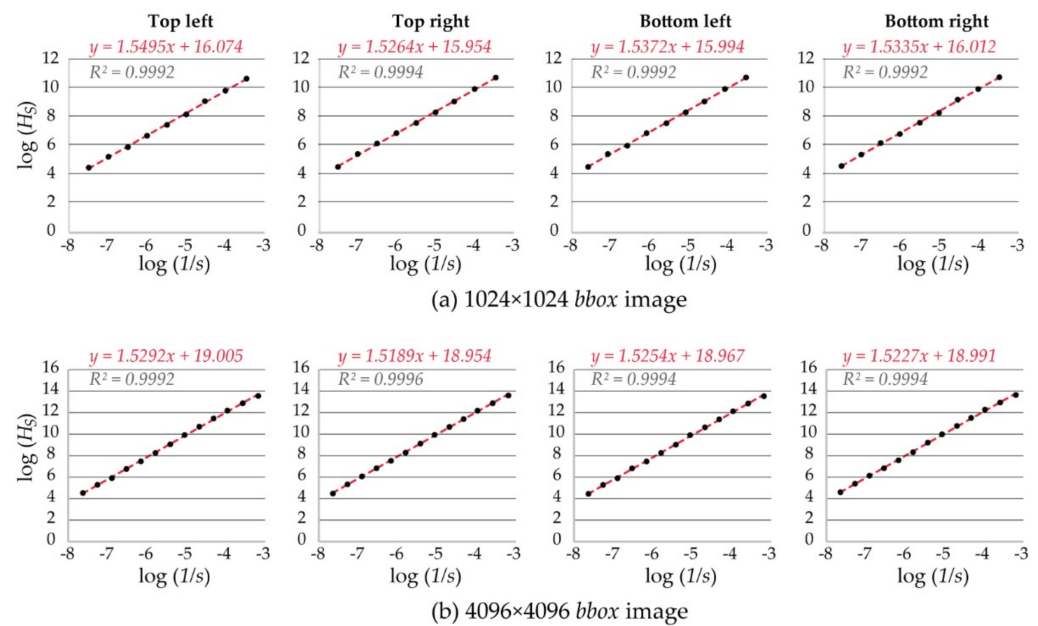


Figure 6. The four plots and trendlines of $\log(N_s)$ versus $\log(1/s)$ using four positions (two different sizes of façade images—(a) 1024×1024 and (b) 4096×4096 *bbox* sizes—of the Spiller House).

The determination of fractal dimension of the small image uses 9 box sizes from 11 to 180 pixels, while the fractal analysis of the large image uses 13 box sizes from 11 to 723 pixels. In this way, the large (higher quality) images can produce more accurate D values. In addition, the large image can depict details of each design more accurately than the small image. Thus, the high-quality images (4096×4096) should be used for fractal analysis of architectural drawings.

The D -intensity (D_i) values of the two façades determined by the differential box counting method are in Table 3. Like the D_d values in Table 2, the last level of representation provides the highest D_i value of each image. However, the third level develops the lowest D_i values. This is unexpected, because the fractal dimension commonly increases with additional layers of information [20]. The differential box counting method is, however, based on the difference between the minimum and maximum grey-levels of the image surface. As all façade images used in this measurement include black line drawings, the grey renderings (levels 2 and 3) give the images less contrast. Collectively, in terms of visual complexity D -density (D_d) tends to represent the formal complexity of the image, while D -intensity (D_i) identifies its textural complexity.

Table 3. D -intensity (D_i) values of two façades.

Façade Design	Image Size (<i>bbox</i> Size)	Level 1	Level 2	Level 3	Level 4	Level 5
Spiller House	1024×1024	1.4331	1.4160	1.3865	1.5410	1.5524
	4096×4096	1.4567	1.4465	1.4391	1.5478	1.5527
Juicy House	1024×1024	1.2462	1.2308	1.2090	1.4120	1.4306
	4096×4096	1.2723	1.2683	1.2653	1.4271	1.4375

Both D_d and D_i can be used to holistically examine visual stimuli at the early design stage. In spite of this, there are some differences between the two fractal dimension measures. Figure 7 shows the change in D -density (D_d) and D -intensity (D_i) with increase in level of representation. In the D_d measure of Figure 7a, rendering openings with the dark grey (40% of brightness) in level 3 and inserting a tree in level 4 are responsible for the biggest increases in the chart. In contrast, only the fourth level makes a notable increase in the D_i chart of Figure 7b. Because the tree itself has a complex fractal dimension— D_d is

1.7818 (1024×1024) and 1.8184 (4096×4096), and D_i is 1.5832 (1024×1024) and 1.6582 (4096×4096)—the result for level 4 makes sense.

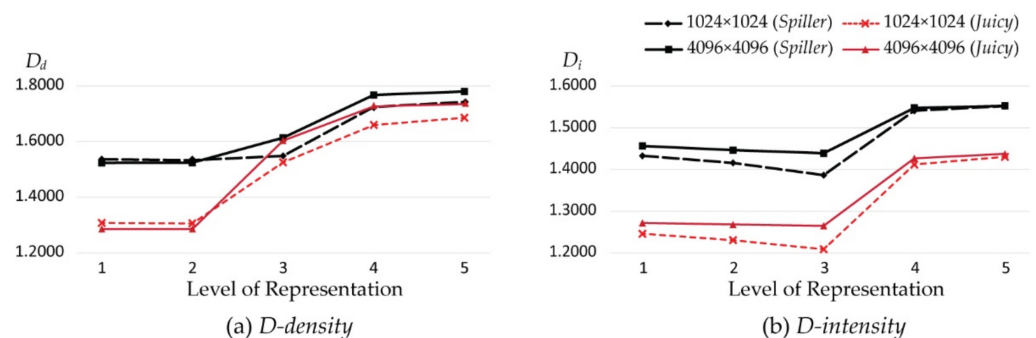


Figure 7. Change in D -density (D_d) and D -intensity (D_i) with increase in level of representation.

Unexpectedly, in the simple façade design for the Juicy House, the third level of representation develops the biggest increase of D_d value. As a result, from the third level of the D_d chart there is a little differentiation between the Spiller and Juicy designs, although the D values of the former design are continuously higher than the latter design in this chart. In contrast, the D_i chart is clearly identifying the complexity differences between the two designs, regardless of the level of representation. In addition, the D_i values of the high-quality images are also consistently higher than the low-quality images. Collectively, the differential box counting method can be regarded as an appropriate fractal analysis measure to examine visual complexity in various levels of representations in design process.

4.2. Visual Attention

The VAS results are depicted in two eye-fixation maps: heat map and gaze sequence. The heat map illustrates the intensity of attention—green (medium), red (considerable) and dark red (maximum)—while the gaze sequence map illustrates the first four fixation spots the eye moves to, without conscious awareness [26]. Firstly, the heat maps in Figure 8 show the different hotspots of each level of representation of the Juicy House. Specifically, the heat maps of the 1024×1024 *bbox* image in Figure 8a identify the different attractive spots from the ones of the 4096×4096 *bbox* image in Figure 8b. For example, major hotspots in the first two levels of the small image are limited to frames of openings and handrails, while the eyes are more distracted by multiple points in the first two levels of representation of the large image. This may be caused by the inaccurate representation of the low-quality image. One explanation for this might be that the frames or handrails are clearly represented as double lines in the high-quality image, but they are merged, shaded, or obscured in the smaller image. As a result, the merged lines become tightly packed pixels, which attracts the eyes. This black-white contrast generally draws visual attention [57]. For this reason, the 1024×1024 *bbox* images might be inappropriate for the visual attention simulation of line drawing images.

The heat maps and gaze sequences in Figure 8 clearly capture the sequential additions of visual stimuli in the level of representation. Significantly, in the third level of representation, both different *bbox* images show similar heat maps and gaze sequences, because the dark grey renderings are gaze attractors in the images. The third level also draws the eye to the entrance, whereas in the fourth and fifth levels, the tree attracts and holds the eyes significantly longer than the other areas. Biophilia is a visceral attraction [26,64] and although it is only an image of a natural object, it adds a distinct level of visual interest to the façade. In addition, the fourth level of representation develops the significant increases in both D_d and D_i charts. Thus, it is evident that these visual fixation patterns in Figure 8 are related to the visual complexity changes in Figure 7 in terms of the stimulus-sensation relationship.

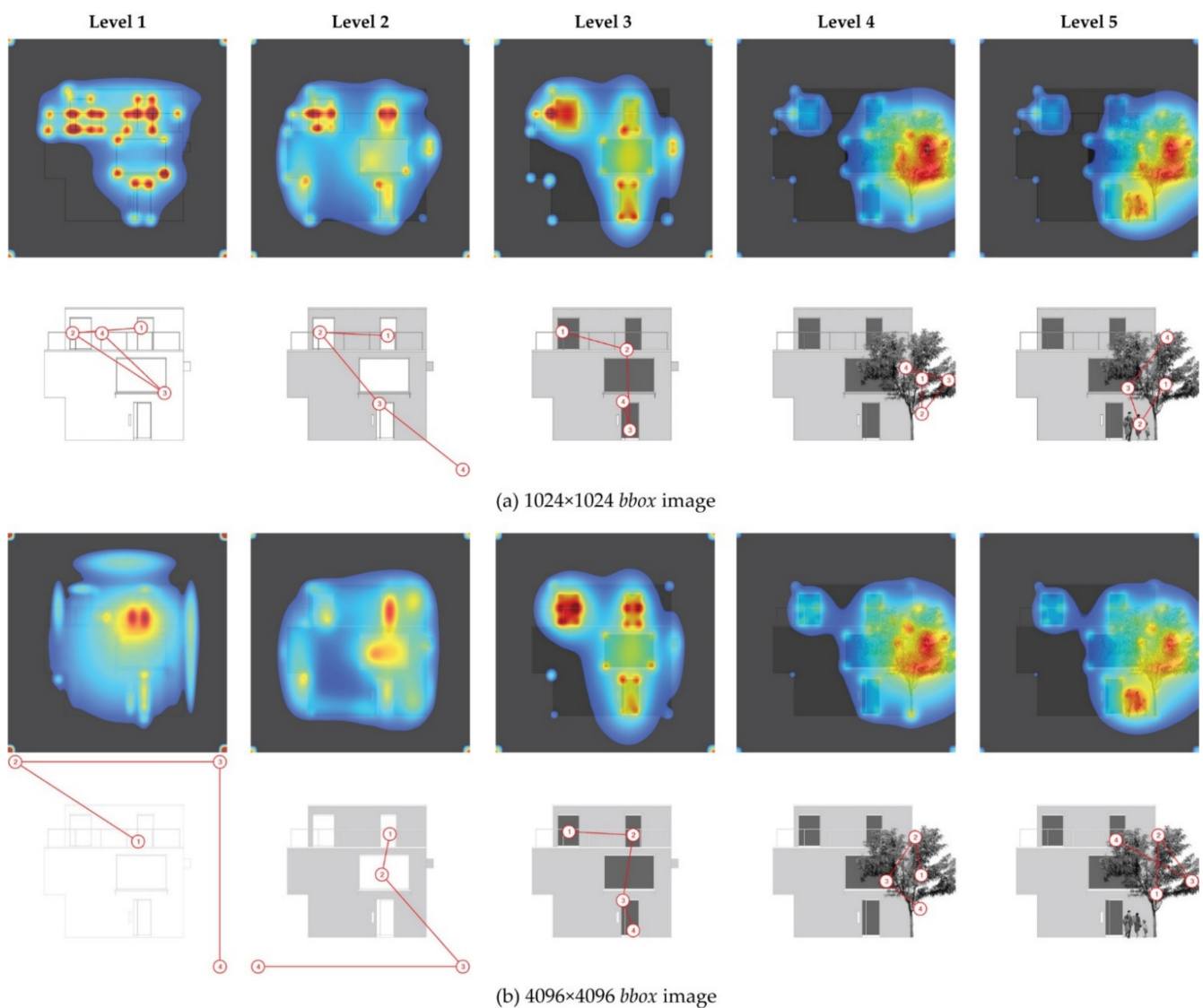


Figure 8. Heat maps and gaze sequences with increase in the level of representation of Juicy House.

The heat maps and gaze sequences of the Spiller House in Figure 9 show similar pre-attentive processing characteristics to the façade representations of the Juicy House (Figure 8). Since the Spiller House façade is regarded as a more complex design, the eye-fixation maps seem to present more attractive design elements than the ones of the Juicy House. For example, curtain wall mullions and balusters (or handrails) at the centre of façade and a stair enclosure are hot spots as well as one of four gaze locations in the first three levels of representation. In both *bbox* images, the third level of representation clearly captures major hotspots and their corresponding gaze sequences. It may also increase visual interest inside the façade design and keep the eyes fixated on it. Considering that a 4096×4096 *bbox* image is more appropriate for the fractal analysis of façade design at the early design stage, the third level of representation should be a best option to study both visual stimulus and visual attention at the same time.

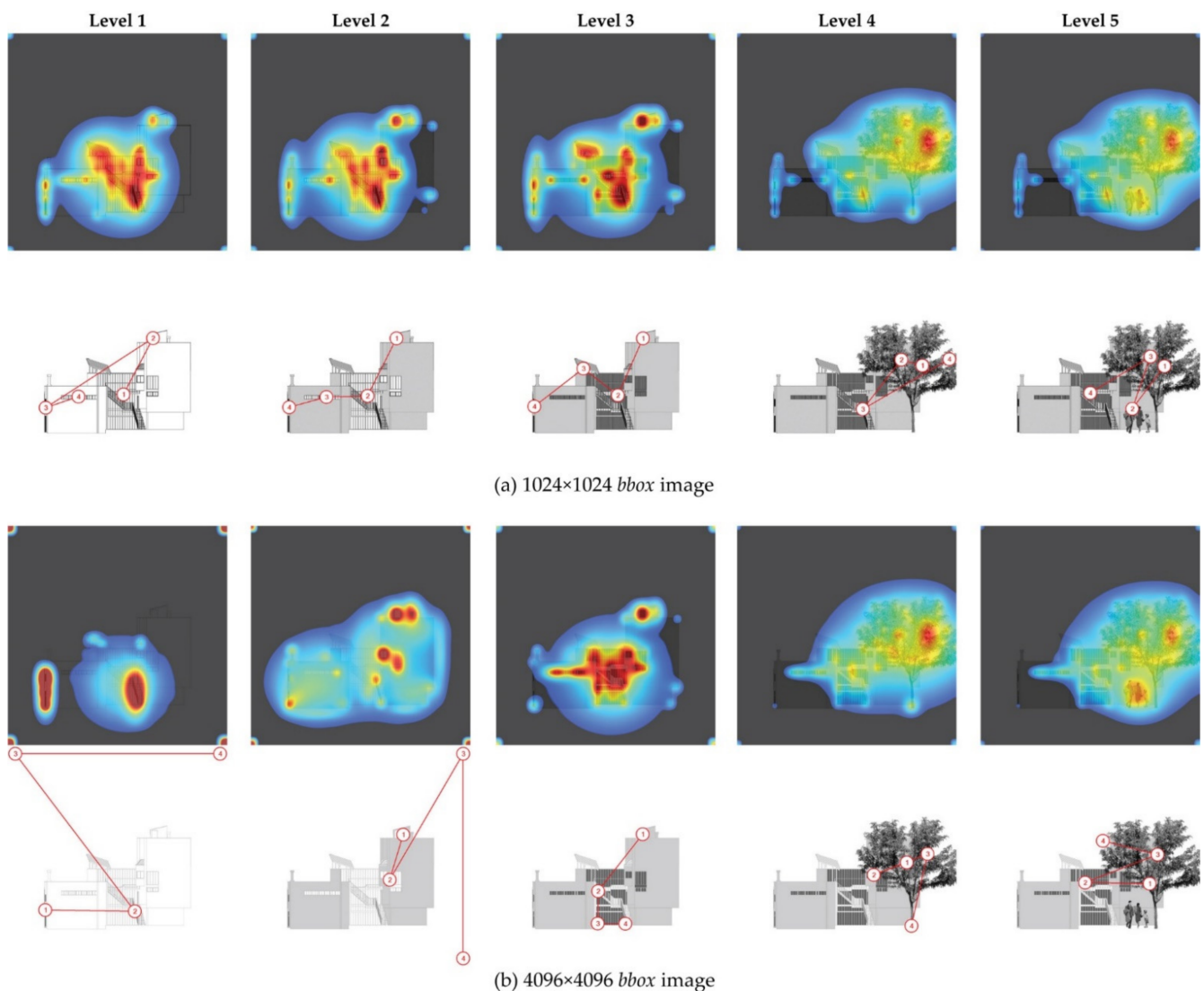


Figure 9. Heat maps and gaze sequences with increase in the level of representation of the Spiller House.

5. Discussion

This paper has explored the potential application of two computational approaches to design, and especially to the early design stages. Through a series of simulations using five levels of façade representation, this paper has demonstrated how to measure the visual character of a façade in terms of visual stimulus and its impact on early vision. The four visual properties— D -density (D_d), D -intensity (D_i), heat map and gaze sequence—of a façade design can be used as quantitative and qualitative indicators for architectural practice. Firstly, the D values can be applied to characterise and compare the visual complexities of design variations of a façade. With the development of knowledge about the D range [20], this approach can support the creation of a new designs consistent with an architectural type or style. For example, as demonstrated in Stamps' work [1], different styles or design elements (geometry, patterns, textures and details) can be examined at the conceptual design or schematic design stage. That is, like spatial and even structural configurations, the impacts of façade design decisions can be modelled and predicted during early stages, contributing to the development and refinement of the façade. To do this, D calculations of a collection of existing designs are necessary.

The analytical programmes used in this research are all dependent on the mode of representation of a façade. Again, multiple levels of representation should be considered before the calculations are undertaken. For example, a perspective rendering of a

façade may be useful for both D measurement and visual attention simulation. Moreover, colour-rendered images are more suitable for VAS, because its algorithm considers two colour contrasts (“red-green” and “blue-yellow”). Nonetheless, the combination of two computational approaches using greyscale images is still useful for providing predictability in architectural aesthetic experience and assessments.

Even though this paper has primarily addressed holistic ways of measuring architectural character (visual complexity), it is acknowledged that other architectural properties can also impact on our visual experience. For example, Nasar [18] suggests a probabilistic model of aesthetic response, highlighting three types of formal attributes of building exteriors: enclosure, complexity and order. The aesthetic variables are further extended to include symbolic aesthetics (style) and schemas (typicality). Interaction between the aesthetic variables of a building and an individual’s knowledge structures shapes the individual’s experience and responses to a building design [18]. Order, moderate complexity and popular styles can support pleasant responses, while high complexity, atypicality and low order may produce excitement for the public. Thus, the fractal dimension, visual complexity, presented in this paper can be regarded as just one of the formal design properties that shape visual attention.

Interestingly, the VAS algorithm (not its empirical basis) used in this paper is based on Klinger’s and Salingaros’s model [62], which is inspired by Christopher Alexander’s essay, “the nature of order” [65]. To predict a building’s emotional impact, their model highlights architectural temperature (T), harmony (H), life (L) and complexity (C) [62]. Each variable is quantified by several indexes. For example, T is a loose analogy of thermodynamic temperature, which is determined by design components such as intensity, density, curvature and colour, which are graphically presented in VAS (see Figures 8 and 9). The model then defines C as $T(10 - H)$ and L as TH . Finally, L corresponds to the emotional perception of a building design. In this way, the thermodynamic analogy model clearly presents one more important factor, harmony, which is not addressed in the present study.

Nonetheless, the methodological configurations suggested in this paper not only help us to use the two applications (ImageJ+FracLac and VAS), but also contribute to the investigation into fractal dimension and visual attention in the fields of architectural and urban design. First of all, this paper suggests starting with a high-resolution (4096×4096 *bbox*) image for both fractal analysis and visual attention simulation at the early design stage. In addition, a façade image rendered with two grey scales for a main body of building and openings is potentially optimal for both applications. In contrast, a colour-rendering may be more suitable for the visual attention simulation. An extended differential box counting method can then be applied to determine the fractal dimension of colour images [66], but there is no viable software available for this purpose. Thus, a colour-rendering for fractal analysis will need to be converted into binary or greyscale images. Fractal analysis has been regarded, for many years, as a repeatable and reliable method for determining visual complexity in design [20], while the series of D_d and D_i calculations in Tables 2 and 3 indicate the ways D is reliant on the quality and level of design representation. In addition, many studies in the fields of architecture and urban design have addressed the traditional box counting method, but this study reveals that the differential box counting method can be a better option to compare the fractal dimensions of some types or styles of façade designs regardless of the level of representation. Thus, a future study should examine the differential box counting method for fractal analysis in more detail.

Traditionally, the scale of the data developed from eye-tracking experiments has been too large to quantitatively correlate with the other formats of data. Nonetheless, numeric gaze data can be useful to statistically identify the stimulus-sensation relationship. For example, heat maps illustrate the intensity information of visual attention that is represented by the colour gradient. This feature is very similar to the D -intensity (D_i) value determined by the differential box counting method. Thus, extracting numeric values from the images can be valuable. That is, the eye-fixation images can also be investigated using image analysis applications like ImageJ+Color Histogram. On the other hand, even if an

eye-tracking experiment produces massive data, it might be more useful than the quick and simple VAS outcome in terms of the quantitative comparison.

The tree in the fourth and fifth levels of representation significantly influences both fractal dimension and visual attention, and arguably by these levels of representation the software is no longer measuring the façade, but the entourage elements. Biophilia and natural patterns in architecture are generally a big determinant in early vision [26,64]. Natural structures (or “non-linear fractal images” [40]) in the image are also close to fractal geometry, which can significantly affect the determination of its fractal dimension. In this study, the addition of fractal geometry to a façade image results in the notable increases of D_d and D_i values, which may also be related to the eye-fixations in the image. This interesting phenomenon should be further investigated in a future study with a sufficiently large sample size.

6. Conclusions

As noted at the start of this paper, it is widely understood that façade design involves important visual information which shapes our visual experience. Knowing this, the lack of tools, methods or systems to support architects to examine early design options is an obvious gap in disciplinary knowledge. To begin to fill this gap, this paper presents two computational approaches to examine the stimulus-sensation relationship at the early design stage: (i) measuring the visual characteristics of façade design (visual stimulus) and (ii) capturing intrinsic visual engagement without conscious awareness (pre-attentive visual attention). Whilst the focus of this paper has largely been on methodological exploration, the simulation results contribute to understanding the visual consequence of façade design and the ways people may respond to a façade.

Despite the various forms of façade analysis presented in this paper, it should also be clear that the fractal dimension measures, D -density (D_d) and D -intensity (D_i), are quantitative, while the results of eye-tracking simulation are only graphically presented by “heat map” and “gaze sequence”. The sample size is also limited, because this paper is focused on a test of the computational approaches, not on developing larger arguments about statistical validity. Thus, the actual stimulus-sensation relationship in a façade design is neither fully, nor empirically identified in this paper. Acknowledging that the scope and sample of this paper are limited, a future study will focus on a few potential design variables as visual stimulus as well as the advanced analysis on eye-fixation data.

Architectural aesthetic character can have a strong impact on people’s pre-conscious physiological and emotional reactions, shaping their behaviours. With the rapid integration of design and construction processes, decision-making and problem solving at the early design stage has become an increasingly prominent activity in this industry. With this in mind, it is not surprising that there is a growing need for tools or methods that can help architects predict the visual consequences of a façade design. This paper, examining visual character and predicting its impacts, suggests some ways future designers might employ to improve façade design. The computational approaches presented in this paper also enable them to consider laypeople’s perceptions in a conceptual computer-aided architectural design (CAAD) process. Furthermore, these approaches can be applied to characterise, compare and assess architectural styles or types. Even though this research has analysed the completed façades, the approaches can be used to examine and compare multiple designs and variations during the design process before the final façade is determined. Thus, this paper contributes to the disciplinary knowledge gap about visual reasoning in design.

Author Contributions: Conceptualization, J.H.L. and M.J.O.; methodology, J.H.L. and M.J.O.; data analysis, J.H.L. and M.J.O.; writing—original draft, J.H.L.; writing—review and text development, M.J.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by UNSW Scientia Academics (research) program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Stamps, A.E., III. Physical determinants of preferences for residential facades. *Environ. Behav.* **1999**, *31*, 723–751. [CrossRef]
2. Stevens, S.S. On the psychophysical law. *Psychol. Rev.* **1957**, *64*, 153–181. [CrossRef]
3. Dehaene, S. The neural basis of the Weber–Fechner law: A logarithmic mental number line. *Trends Cogn. Sci.* **2003**, *7*, 145–147. [CrossRef]
4. Hollander, J.B.; Anderson, E.C. The impact of urban façade quality on affective feelings. *Archmet-IJAR Int. J. Archit. Res.* **2020**, *14*, 219–232. [CrossRef]
5. Simpson, J.; Freeth, M.; Simpson, K.J.; Thwaites, K. Visual engagement with urban street edges: Insights using mobile eye-tracking. *J. Urban Int. Res. Placemaking Urban Sustain.* **2019**, *12*, 259–278. [CrossRef]
6. Gehl, J.; Kaefer, L.J.; Reigstad, S. Close encounters with buildings. *Urban Des. Int.* **2006**, *11*, 29–47. [CrossRef]
7. Kusenbach, M. Street phenomenology: The go-along as ethnographic research tool. *Ethnography* **2003**, *4*, 455–485. [CrossRef]
8. Heffernan, E.; Heffernan, T.; Pan, W. The relationship between the quality of active frontages and public perceptions of public spaces. *Urban Des. Int.* **2014**, *19*, 92–102. [CrossRef]
9. Lee, J.H.; Ostwald, M.J. Creative decision-making processes in parametric design. *Buildings* **2020**, *10*, 242. [CrossRef]
10. Olewnik, A.T.; Lewis, K. On validating engineering design decision support tools. *Concurrent Eng.* **2005**, *13*, 111–122. [CrossRef]
11. Carrasco, M. Visual attention: The past 25 years. *Vision Res.* **2011**, *51*, 1484–1525. [CrossRef]
12. De la Fuente Suárez, L.A. Subjective experience and visual attention to a historic building: A real-world eye-tracking study. *Front. Archit. Res.* **2020**, *9*, 774–804. [CrossRef]
13. Hollander, J.B.; Sussman, A.; Purdy Levering, A.; Foster-Karim, C. Using eye-tracking to understand human responses to traditional neighborhood designs. *Plann. Pract. Res.* **2020**, *35*, 485–509. [CrossRef]
14. Suurenbroek, F.; Spanjar, G. Exploring eye-tracking technology: Assessing how the design of densified built environments can promote inhabitants’ well-being. In *Urban Experience and Design: Contemporary Perspectives on Improving the Public Realm*; Hollander, J.B., Sussman, A., Eds.; Routledge: New York, NY, USA, 2021; pp. 91–103.
15. Lisińska-Kuśnierz, M.; Krupa, M. Suitability of eye tracking in assessing the visual perception of architecture—A case study concerning selected projects located in Cologne. *Buildings* **2020**, *10*, 20. [CrossRef]
16. Luck, R. Participatory design in architectural practice: Changing practices in future making in uncertain times. *Des. Stud.* **2018**, *59*, 139–157. [CrossRef]
17. Hocking, V.T.; Brown, V.A.; Harris, J.A. Tackling wicked problems through collective design. *Intell. Build. Int.* **2016**, *8*, 24–36. [CrossRef]
18. Nasar, J.L. Urban design aesthetics: The evaluative qualities of building exteriors. *Environ. Behav.* **1994**, *26*, 377–401. [CrossRef]
19. Forsythe, A.; Nadal, M.; Sheehy, N.; Cela-Conde, C.J.; Sawey, M.; Sawey, M. Predicting beauty: Fractal dimension and visual complexity in art. *Br. J. Psychol.* **2011**, *102*, 49–70. [CrossRef]
20. Ostwald, M.J.; Vaughan, J. *The Fractal Dimension of Architecture*; Birkhäuser: Cham, Switzerland, 2016.
21. Pickover, C.A. *Keys to Infinity*; John Wiley & Sons: New York, NY, USA, 1995.
22. Hagerhall, C.M.; Purcell, T.; Taylor, R. Fractal dimension of landscape silhouette outlines as a predictor of landscape preference. *J. Environ. Psychol.* **2004**, *24*, 247–255. [CrossRef]
23. Katona, V. Relief method: The analysis of architectonic façades by fractal geometry. *Buildings* **2021**, *11*, 16. [CrossRef]
24. Friedenber, J. *Visual Attention and Consciousness*; Psychology Press: New York, NY, USA; London, UK, 2013.
25. Hollander, J.B.; Purdy, A.; Wiley, A.; Foster, V.; Jacob, R.J.K.; Taylor, H.A.; Brunyé, T.T. Seeing the city: Using eye-tracking technology to explore cognitive responses to the built environment. *J. Urban. Int. Res. Placemaking Urban Sustain.* **2018**, *12*, 156–171. [CrossRef]
26. Salingeros, N.A.; Sussman, A. biometric pilot-studies reveal the arrangement and shape of windows on a traditional façade to be implicitly “engaging”, whereas contemporary façades are not. *Urban Sci.* **2020**, *4*, 26. [CrossRef]
27. Milliken, P.; Hollander, J.B.; Sussman, A.; Situ, M. Identifying biophilic design elements in streetscapes: A study of visual attention and sense of place. In *Urban Experience and Design: Contemporary Perspectives on Improving the Public Realm*; Hollander, J.B., Sussman, A., Eds.; Routledge: New York, NY, USA, 2021; pp. 75–90.
28. 3M. How Does VAS Work and What is First-Glance Vision? Available online: <https://vas.3m.com/support> (accessed on 8 February 2021).
29. Bovill, C. *Fractal Geometry in Architecture and Design*; Birkhäuser: Boston, MA, USA, 1996.
30. Sarkar, N.; Chaudhuri, B.B. An efficient differential box-counting approach to compute fractal dimension of image. *IEEE Trans. Syst. Man Cybern.* **1994**, *24*, 115–120. [CrossRef]
31. Mandelbrot, B.B. *The Fractal Geometry of Nature*; WH Freeman: New York, NY, USA, 1982; Volume 1.
32. Batty, M.; Longley, P.A. *Fractal Cities: A Geometry of Form and Function*; Academic Press: New York, NY, USA, 1994.

33. Burkle-Elizondo, G.; Valdez-Cepeda, R.D. Do the mesoamerican artistic and architectural works have fractal dimension? In *Emergent Nature: Patterns, Growth and Scaling in the Sciences*; World Scientific: Singapore, 2001; pp. 431–432.
34. Oleschko, K.; Brambila, R.; Brambila, F.; Parrot, J.F.; López, P. Fractal analysis of teotihuacan, mexico. *J. Archaeolog. Sci.* **2000**, *27*, 1007–1016. [[CrossRef](#)]
35. Çağdaş, G.; Gözübüyük, G.; Ediz, Ö. Fractal based generative design for harmony between old and new. In *Generative Art 2005, Proceedings of the 8th GENERATIVE ART CONFERENCE(GA2005), Milan, Italy, 15–17 December 2005*; Domus Argenia: Milan, Italy, 2005; pp. 150–159.
36. Ostwald, M.J.; Ediz, Ö. Measuring form, ornament and materiality in sinan’s kiliç ali paşa mosque: An analysis using fractal dimensions. *Nexus. Netw. J.* **2015**, *17*, 5–22. [[CrossRef](#)]
37. Samper, A.; Herrera, B. The fractal pattern of the french gothic cathedrals. *Nexus. Netw. J.* **2014**, *16*, 251–271. [[CrossRef](#)]
38. Panigrahy, C.; Seal, A.; Mahato, N.K.; Bhattacharjee, D. Differential box counting methods for estimating fractal dimension of gray-scale images: A survey. *Chaos Solitons Fractals* **2019**, *126*, 178–202. [[CrossRef](#)]
39. Liu, Y.U.; Chen, L.; Wang, H.; Jiang, L.; Zhang, Y.; Zhao, J.; Wang, D.; Zhao, Y.; Song, Y. An improved differential box-counting method to estimate fractal dimensions of gray-level images. *J. Visual. Commun. Image. Represent.* **2014**, *25*, 1102–1111. [[CrossRef](#)]
40. Foroutan-pour, K.; Dutilleul, P.; Smith, D.L. Advances in the implementation of the box-counting method of fractal dimension estimation. *Appl. Math. Comput.* **1999**, *105*, 195–210. [[CrossRef](#)]
41. Koch, H.P. The concept of fractals in the pharmaceutical sciences. *Pharmazie* **1993**, *48*, 643–659. [[PubMed](#)]
42. Cooper, J.; Oskrochi, R. Fractal analysis of street vistas: A potential tool for assessing levels of visual variety in everyday street scenes. *Environ. Plann B Plann. Des.* **2008**, *35*, 349–363. [[CrossRef](#)]
43. Karperien, A. Fraclac for Imagej. Available online: <http://rsb.info.nih.gov/ij/plugins/fraclac/FLHelp/Introduction.htm> (accessed on 15 February 2021).
44. Taylor, R. Science in culture. *Nature* **2001**, *410*, 1038. [[CrossRef](#)]
45. Aks, D.J.; Sprott, J.C. Quantifying aesthetic preference for chaotic patterns. *Empirical. Stud. Arts* **1996**, *14*, 1–16. [[CrossRef](#)]
46. Sussman, A.; Hollander, J.B. *Cognitive Architecture: Designing for How We Respond to the Built Environment*; Routledge: New York, NY, USA, 2015.
47. McMains, S.A.; Kastner, S. Visual Attention. In *Encyclopedia of Neuroscience*; Binder, M.D., Hirokawa, N., Windhorst, U., Eds.; Springer: Berlin, Germany, 2009; pp. 4296–4302. [[CrossRef](#)]
48. Anderson, C.H.; Van Essen, D.C.; Olshausen, B.A. CHAPTER 3—Directed visual attention and the dynamic control of information flow. In *Neurobiology of Attention*; Itti, L., Rees, G., Tsotsos, J.K., Eds.; Academic Press: Burlington, VT, USA, 2005; pp. 11–17. [[CrossRef](#)]
49. Treisman, A. Features and objects in visual processing. *Sci. Am.* **1986**, *255*, 114B–125. [[CrossRef](#)]
50. Treisman, A.M.; Gelade, G. A feature-integration theory of attention. *Cognitive. Psychol.* **1980**, *12*, 97–136. [[CrossRef](#)]
51. Rensink, R.A. The management of visual attention in graphic displays. In *Human Attention in Digital Environments*; Roda, C., Ed.; Cambridge University Press: Cambridge, UK, 2011; pp. 63–92. [[CrossRef](#)]
52. Rensink, R.A. Commentary: Changes. In *Progress in Brain Research*; Hyona, J., Munoz, D.P., Heide, W., Radach, R., Eds.; Elsevier: Amsterdam, The Netherlands, 2002; Volume 140, pp. 197–207.
53. Rensink, R.A. The dynamic representation of scenes. *Visual Cognition* **2000**, *7*, 17–42. [[CrossRef](#)]
54. Yanulevskaya, V.; Uijlings, J.; Geusebroek, J.M.; Sebe, N.; Smeulders, A. A proto-object-based computational model for visual saliency. *J. Vision* **2013**, *13*, 27. [[CrossRef](#)]
55. Zhang, L.M.; Zhang, R.X.; Jeng, T.S.; Zeng, Z.Y. Cityscape protection using VR and eye tracking technology. *J. Visual Commun. Image Represent.* **2019**, *64*, 102639. [[CrossRef](#)]
56. Mohammadpour, A.; Karan, E.; Asadi, S.; Rothrock, L. Measuring end-user satisfaction in the design of building projects using eye-tracking technology. In *Computing in Civil Engineering 2015*; American Society of Civil Engineers (ASCE): Austin, TX, USA, 2015; pp. 564–571.
57. Tang, M. Analysis of signage using eye-tracking technology. *Interdisciplin. J. Signage Wayfinding* **2020**, *4*. [[CrossRef](#)]
58. Niebur, E.; Koch, C. Computational architectures for attention. In *The Attentive Brain*; Parasuraman, R., Ed.; The MIT Press: Cambridge, MA, USA, 1998; pp. 163–186.
59. Massaro, D.; Savazzi, F.; Di Dio, C.; Freedberg, D.; Gallese, V.; Gilli, G.; Marchetti, A. When art moves the eyes: A behavioral and eye-tracking study. *PLoS ONE* **2012**, *7*, e37285. [[CrossRef](#)] [[PubMed](#)]
60. Quian Quiroga, R.; Pedreira, C. How do we see art: An eye-tracker study. *Front. Hum. Neurosci.* **2011**, *5*, 98. [[CrossRef](#)] [[PubMed](#)]
61. Rueden, C.T.; Schindelin, J.; Hiner, M.C.; DeZonia, B.E.; Walter, A.E.; Arena, E.T.; Eliceiri, K.W. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinf.* **2017**, *18*, 1–26. [[CrossRef](#)] [[PubMed](#)]
62. Salinger, N.A. Life and complexity in architecture from a thermodynamic analogy. *Phys. Essays* **1997**, *10*, 165–173. [[CrossRef](#)]
63. Mayrhofer-Reinhartshuber, M.; Ahammer, H. Pyramidal fractal dimension for high resolution images. *Chaos Interdisciplin. J. Nonlinear Sci.* **2016**, *26*, 073109. [[CrossRef](#)]
64. Coburn, A.; Kardan, O.; Kotabe, H.; Steinberg, J.; Hout, M.C.; Robbins, A.; MacDonald, J.; Hayn-Leichsenring, G.; Berman, M.G. Psychological responses to natural patterns in architecture. *J. Environ. Psychol.* **2019**, *62*, 133–145. [[CrossRef](#)]

-
65. Alexander, C. *The Nature of Order: An Essay on the Art of Building and the Nature of the Universe*; Center for Environmental Structure: Berkeley, CA, USA, 2004; Volume 1–4.
 66. Nayak, S.R.; Mishra, J.; Khandual, A.; Palai, G. Fractal dimension of RGB color images. *Optik* **2018**, *162*, 196–205. [[CrossRef](#)]