Line Segmentation: A Computational Technique for Architectural Image Analysis

Michael J. Ostwald
The University of Newcastle, Australia

Chris Tucker
The University of Newcastle, Australia

Stephan Chalup
The University of Newcastle, Australia

ABSTRACT
Planar methods have typically dominated the computational analysis of architectural and urban space and form. For example, the majority of space syntax research that has been undertaken over the last three decades has been concerned almost exclusively with architectural and urban plans. In contrast, analytical methods that consider the formal, or visual, qualities of architectural façades, or images of buildings, are not only rare, but only a few have ever been repeated and adequately tested.

The present paper outlines a new method—derived from the Hough Transform algorithm—for the dissolution of architectural images into segmented lines that can be counted and charted, and that can have their spatial orientation determined. This method for investigating the visual qualities of buildings is demonstrated in an analysis of a series of images of suburban houses. The proposed method, line segmentation, is potentially significant because it is a method not commonly used for the quantitative analysis of the formal and textural character of real buildings, it is repeatable, and it delivers consistent results if a simple procedure is followed.

SOFTWARE

1 INTRODUCTION
This paper describes the application of a new computational method for the analysis of the visual qualities, or formal complexity, of architectural façades. Past works on element and boundary detection in façade analysis (Stamps III 2003; Malhis 2003) have inspired this new method, "line segmentation." However, while these past works have relied on manual techniques (typically the tracing of overlays on photographs), the present research describes the preliminary results of a computational extrapolation of these methods that uses the Hough Transform algorithm to automate the process of identifying patterns within complex or fragmented images. The authors have developed a program for line segmentation analysis that works as a module of their prototype Archimage software for the visual analysis of architecture.

This paper provides an overview of past attempts to quantify objectively the visual qualities of architecture. The paper then describes the new method, some of the limitations identified by the authors, and the line segmentation results for some typical suburban dwelling façades. The purpose of this paper is not to provide a comprehensive presentation of the method and its programming code, or to provide extensive evidence of its efficacy, but, rather, to offer an overview of the line segmentation concept and to suggest ways in which it might be usefully applied.

2 BACKGROUND TO QUANTITATIVE ARCHITECTURAL VISUAL ANALYSIS
Objective and quantifiable methods for the analysis of character and complexity in architectural façades are relatively rare (Stiny 1975; Stamps III 1999, 2003). A small number of methods have been proposed over the last four decades, but few of these have ever been repeated, and only one has been adopted and tested by multiple researchers. For example, Krampen (1975) famously used a technique known as parsing to calculate the diversity of elements within the façade of a building. By overlaying a grid on an image of the façade, he was able to manually count the instances wherein a particular element was present within each cell and, thereby, produce a mathematical determination of façade diversity (fig. 1).

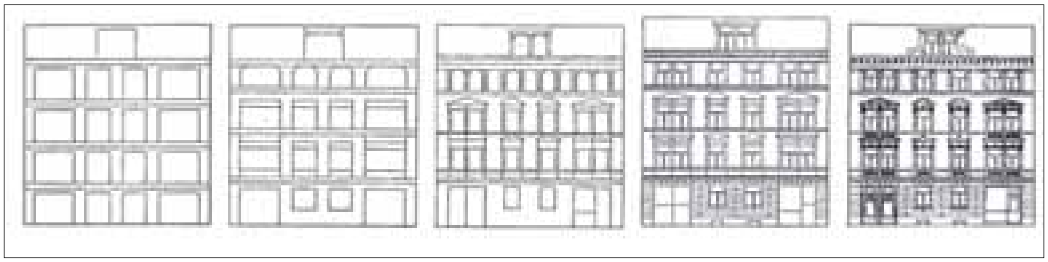


Figure 1 Dissolution of a Façade into Four Layers of Elements (Krampen 1975)

Rather than using an analytical grid, Elsheshtawy (1997) traced layers over photographs of buildings to uncover what he believed to be the "bounding surface" of each visual component of a façade, and then grouped these together, based on their relationship with other elements. He then derived a value for complexity within the façade from the number of groups formed at each layer or level. In a similar way, Imamoglu (2000) manually traced photographs of buildings and then manipulated them to produce drawings of the same building with different levels of visible detail. However, rather than producing a numerical calculation of visual character, Imamoglu surveyed people and asked them to rank each façade based on the desirability of its formal complexity. Designs that were highly complex or that had excessive numbers of unfamiliar elements were generally more negatively regarded.

Salingaros (1997) proposed a simple method to establish the visual character of a building's façade. Adopting, like Krampen (1979) before him, a variation of the concept of entropy, Salingaros proposed that façades might be measured by their "temperature." Salingaros then manually applied this method to twenty-five buildings; Klinger and Salingaros (2000) later refined this method to provide a measure of the "structure" or "complexity" of abstract visual arrays. Malhis (2003), like Elsheshtawy (1997), suggested that a more objective measure of the visual character of a building façade could be determined by recording the detail within the façade of each dwelling at different scales, or on what he calls, "layers." In his method, a thin line represents the shapes and forms within each layer, defining the feature. A professional assessor would then examine each façade to determine where the layers occur. Malhis used this manual system to classify 230 villas into eleven stylistic groups.

While the techniques of Krampen (1979), Elsheshtawy (1997), Salingaros (1997), Imamoglu (2000), and Malhis (2003) are all important, in that they seek to produce objective and quantitative methods for analyzing façades, they also all rely on time-consuming manual techniques and on various levels of personal judgment. However, a parallel set of approaches has been developed under the auspices of space syntax research, and while these methods have achieved a slightly higher degree of objectivity, they, too, have had only limited success and application.

While space syntax researchers have typically rejected the importance of the third dimension (the façade) of the built environment (Batty and Rana 2004), a few techniques have, nevertheless, arisen from this field. For example, Hillier (1996) proposes that building façades may be conceptualized as an arrangement of shapes that are defined by the ground-plane, but which are perpendicular to it. Hillier represents a building's façade as both a "metric tessellation" and as a diagram of "the dominant elements in the façade." Hillier then outlines the possibility of developing a "façade isovist" as a way of interpreting the relationship between a public space and the façades of buildings that define or control this space. In essence, the façade isovist is a supplementary technique for shaping and informing planar analysis. Teller (2003) has suggested an additional approach to the topic of façade analysis: He examines the three-dimensional openness of streetscapes and town squares by creating a two-dimensional image from a wide-angle view, looking vertically towards the sky. He then focuses this method on the "statistical distribution" of building heights and façade features, as observed from a fixed point.

One of the few analytical methods for investigating the visual qualities of façades that has been tested and repeated by multiple researchers is the "box-counting" approach to fractal analysis. This method, the architectural version of which was first presented by Bovill (1996), determines the fractal dimension—a measure of characteristic visual complexity—of an architectural façade. Since that time, the manual version of the method has been repeated multiple times (Bovill 1997; Bechhoefer and Appleby 1997; Burkle-Elizondo, Sala, and Valdez-Cepeda 2004), and a computational variation has also been developed and tested (Ostwald, Vaughan, Tucker 2008; Ostwald, Vaughan, Chalup 2009). While the benefits of this method are yet to be convincingly demonstrated, it, nonetheless, has the advantage that it is repeatable and, in its computational form, provides a robust determinant of the visual complexity of a façade.

3 LINE SEGMENTATION METHOD

Just as space syntax researchers (Steadman 1983; Hillier and Hanson 1984) have shown that there is much to be learned from the analysis of room relationships in a plan (regardless of dimensionality) or in spaces that are visible from a single point in space (the isovist), so, too, can a range of façade details be productively examined in isolation from their associated plan. Elements in an image are defined by edges, which are typically boundaries dividing similar sections of an image. Because much of the built environment is orthogonal in plan, the majority of edges identified in any visual analysis of architecture are effectively lines. While significant effort has previously been invested in various methods for understanding the elements in a façade, very little time has been focused on the nature of the lines that define these elements (Stamps III 2003). In particular, no one appears to have asked the simple question, how many lines are there in an architectural image and how long are they?

One reason for this lack of interest in this question may relate to the labor-intensive nature of the process of uncovering this information "by hand," and that past computational approaches to image analysis have had a relatively low success rate for line detection and segmentation of this type. This is because in the real world, a myriad of features complicate the computer's capacity to "read" a building. Every shadow cast by the sun, trees, shrubs, parked cars, power lines, and passing pedestrians will fragment the view of a building and its façade. However, a recent technique developed by computer scientists, known as the Hough Transform, has a previously unavailable capacity to extract discontinuous geometries and patterns from complex, real world images (Song and Lyu 2005).

The Hough Transform is important for line segmentation analysis of real world images, because it is at least partially able to mimic the human mind's capacity to seek spatial order in complex environments. For example, when confronted with a long white wall, with a thin, red pole standing in front of it (and continuing in height above the line of the wall), the human eye will easily determine that there are two elements in space: a wall and a pole. Even the most advanced image-detection software has problems with this simple situation. The computer will typically identify three elements: two white walls and a red pole. If the pole has cast a shadow, the computer may find four elements. These errors—or more correctly, "pedantic" readings—occur because computers cannot easily be programmed to make logical spatial connections when given discontinuous data.

The Hough Transform is different in this context because it uses a novel "voting" approach to determine if there is a high likelihood that there are deeper patterns in an image; it then identifies these in much the same way that the human eye does. However, the human eye can also be fooled, and the Hough Transform still possesses the typical computational flaw of being over literal in its readings. For example, the Hough Transform will record the head of a door that also lines up with the head of a nearby window as a single line of cumulative length. The human eye may also intuitively read a façade in this way, even if the mind understands that the door and the window are separate elements. Thus, the Hough Transform represents a major advance in complex image analysis, but it is still not perfect. It is, nevertheless, ideal for the consistent analyses of lines detected in photographs of building facades.

4 APPLICATION OF THE METHOD

Line segmentation is the process wherein the complete set of edges within an image are measured and counted, noting their orientation (vertical, horizontal, or diagonal), but ignoring their connectivity to other lines or elements. A line segmentation algorithm was authored for the present research and embedded in existing Archimage software. The computational approach for producing a line segmentation graph is as follows.

1) A digital photographic image is created as the primary data for the method. Typically, these images will provide a complete view of the façade of a building from a point perpendicular to the façade, and will have been taken in relatively neutral lighting conditions (to minimize the impact of shadows) (fig. 2).

Figure 2 Raw Data (Façade Image)

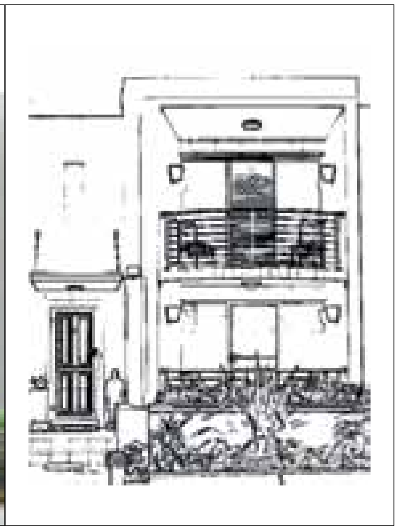


Figure 3 Edge Detection (Sobel Method)

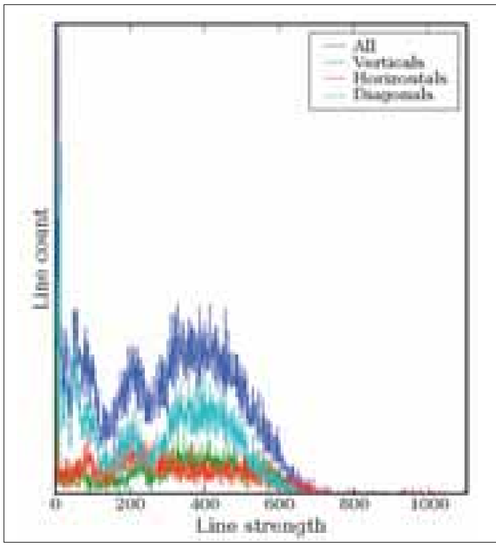
2) The image is imported into Archimage, which uses a Sobel edge-detection algorithm to remove color information from the data, and which produces an optimized "line drawing" of the original image (fig. 3).

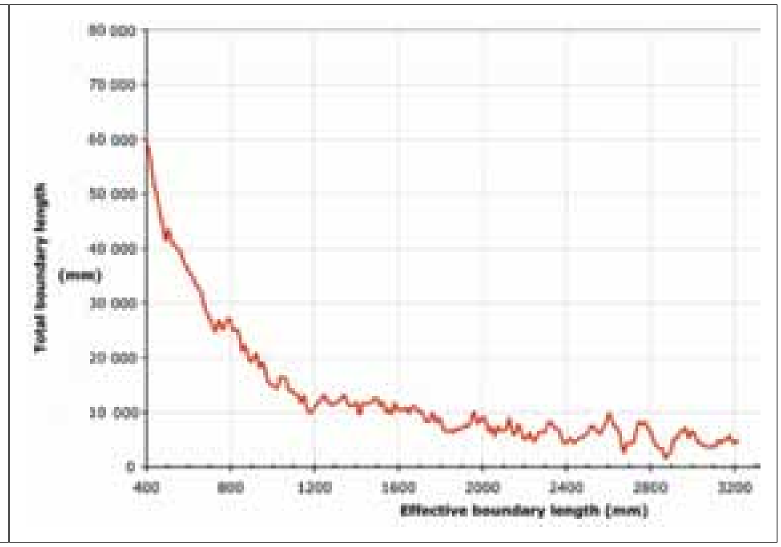
3) Archimage then applies the Hough Transform algorithm to the line drawing to identify discrete continuous sequences of pixels and seemingly connected discontinuous sequences of pixels (in essence, identifying and isolating every line in the data).

4) The number of lines in the image and their orientation are counted.

5) A graph is produced of the results, recording line count (frequency) on the vertical axis and line strength (length in pixels) on the horizontal. The data

SOFTWARE





can be reported either as a single set of results (the set of all lines) or divided into vertical lines, horizontal lines, and diagonal lines (fig. 4).

Figure 4 Line Count Graph (Subdivided by "All" and by Orientation)

Figure 5 Scaled Version of the Line Count Graph

To this point, the process is automated. However, to construct a comparison between multiple images, a scale must be manually set.

6) To set a scale, an element in the image must have a known dimension. For example, in figure 1, a typical dwelling image of 640 x 480 pixels, the door height is around 105 pixels, which means that, in this image, a single pixel is around 20 mm in length. This allows the operator to set an approximate scale for each image that, in turn, allows multiple images to be compared.

For a comparison between the results of different images, a modified chart has been developed (fig. 5). This graph records the intersection between the length of various isolated lines in the original image (x-axis of the graph), and the total length of such sets of lines (of the same length) when they are added together (y-axis); this is called the boundary length. The number of lines of a given length is found by dividing the total boundary length by the isolated line length. Variations on this approach for presenting the data have been tried, but few have the advantage of being able to be read so easily as this cumulative approach.

A typical graph is read as follows. In figure 5, at the left edge of the x-axis, when all of the 400 mm long lines in the original image are added together, they produce a 60m long boundary (60,000 mm in the y-axis); this means that there are around 150 lines in figure 1 that are around 400 mm in length. At the right-hand edge of the x-axis in figure 5, the total length of lines that are 3200 mm long, is 6400 mm; this means that there are only two lines of this length in the original figure.

5 DISCUSSION

Consider two suburban houses, the first, a "traditional" Federation-style house (approximately 70 years old), and the second, a "modern" terrace-style house (approximately 10 years old) (figs. 6 and 7, respectively). Both of these images are subjected to line segmentation analysis and then to manual scaling to overcome differences in size; the results are graphed together (fig. 8). The result for the former house is the upper line in the graph, and the latter is the lower line.

Figure 6 "Traditional" Dwelling

Figure 7 "Modern" Dwelling

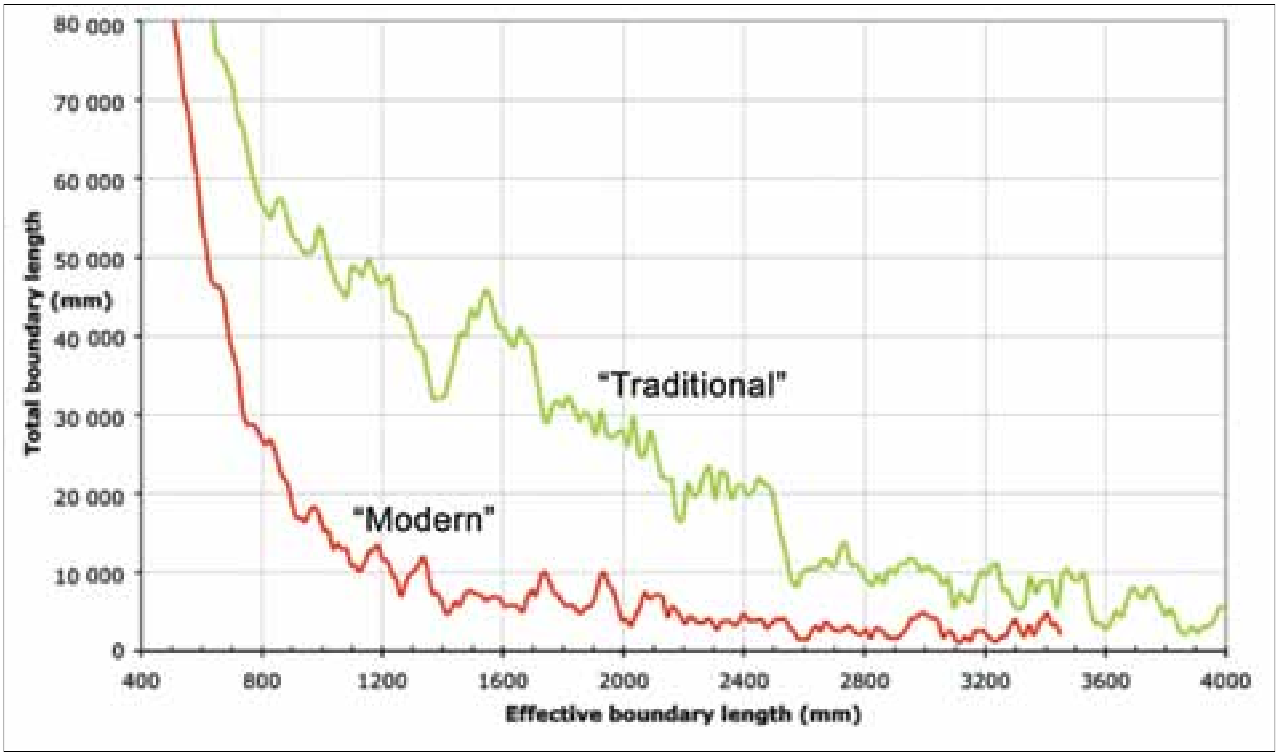


Figure 7
Line Segmentation Results for Traditional and Modern Dwellings

- For the traditional house, a line segmentation analysis reveals a total range of line lengths of between 600mm and 4000 mm. The shortest lines, the 600 mm ones, are repeated in the image around 133 times. Given that the image includes a traditional picket fence that is almost 600 mm high and has around 60 pickets (each of which generates two lines), this result conforms to expectations. There are also around 40 x 1200 mm lengths in the image, 25 x 1600 mm lengths, and around 15 x 2000 mm lengths. Some of these numbers are slightly higher than anticipated, but the density of detail, screening, and ornamentation in the traditional house probably leads to the Hough Transform finding slightly more discontinuous lines than the building actually possesses.

- For the modern house, the line segmentation analysis reveals a total range of lines between 500 mm and 3450 mm and a graph that is, at all points, below the graph for the traditional house. In essence, the modern house has fewer lines of all sizes in its façade and setting. Whereas the traditional house had 40, 25, and 15, respectively, of each of the lengths 1200 mm, 1600 mm, and 2000 mm, the modern house has 10, 4, and 2 times these lengths. This result reflects a much lower level of detail in the modern house, or a lower level of visual complexity across all scales of detail.

CONCLUSION

As part of the development of the line segmentation approach, the authors have tested over 100 photographs of dwelling facades. While these preliminary results are not yet ready for statistical analysis and publication, they begin to suggest that certain architectural styles and types are able to be identified, at least in part, from a calculation of the number and length of lines present in each façade.

The great strengths of this approach are that it will produce consistent, quantifiable, and objective results from a simple digital photograph. The key challenges that remain for the authors are in two areas:

1. The Hough transform, like the human eye, can find patterns that do not actually exist in an image; in essence, it is capable of connecting lines that are only accidentally adjacent. While this is not a critical problem—the lines are always consistently produced even if there are inaccuracies—it nevertheless calls for further refinement.

2. The graphical method for presenting and analyzing the results is still under review. A simple graph of line length (how long) to line number (how many) may be desirable, but it does not reveal a pattern that the eye can easily characterize or compare. The method presented here, of charting line length (how long) to composite boundary length (all of these lines added together), produces a graph that is more intuitive to read, but the authors are still seeking a better alternative.

ACKNOWLEDGMENTS

THIS RESEARCH WAS UNDERTAKEN WITH THE SUPPORT OF AN ARC DISCOVERY GRANT: DP0770106.

REFERENCES

BATTY, M. AND S. RANA. (2004). THE AUTOMATIC DEFINITION AND GENERATION OF AXIAL LINES AND AXIAL MAPS. ENVIRONMENT AND PLANNING B, PLANNING AND DESIGN 31: 615–640.

BECHHOEFER, W. AND M. APPLEBY. (1997). FRACTALS, MUSIC AND VERNACULAR ARCHITECTURE: AN EXPERIMENT IN CONTEXTUAL DESIGN. CRITICAL METHODOLOGIES IN THE STUDY OF TRADITIONAL ENVIRONMENTS. VOL. 97. UNPAG.

BOVILL, C. (1996). FRACTAL GEOMETRY IN ARCHITECTURE AND DESIGN. BOSTON: BIRKHÄUSER.

BOVILL, C. (1997). FRACTAL CALCULATIONS IN VERNACULAR DESIGN. CRITICAL METHODOLOGIES IN THE STUDY OF TRADITIONAL ENVIRONMENTS. VOL. 97. UNPAG.

BURKLE-ELIZONDO, G., N. SALA, AND R. D. VALDEZ-CEPEDA. (2004). GEOMETRIC AND COMPLEX ANALYSES OF MAYA ARCHITECTURE: SOME EXAMPLES. IN NEXUS V: ARCHITECTURE AND MATHEMATICS. ED. K. WILLIAMS AND F. DELGADO-CEPEDA, 57–68. K. W. BOOKS: FLORENCE.

ELSHESHTAWY, Y. (1997). URBAN COMPLEXITY: TOWARD THE MEASUREMENT OF THE PHYSICAL COMPLEXITY OF STREETSCAPES. JOURNAL OF ARCHITECTURAL AND PLANNING RESEARCH 14(4): 301–316.

HILLIER, B. (1996). SPACE IS THE MACHINE: A CONFIGURATIONAL THEORY OF ARCHITECTURE. CAMBRIDGE: CAMBRIDGE UNIVERSITY PRESS.

HILLIER, B. AND J. HANSON. (1984). THE SOCIAL LOGIC OF SPACE. CAMBRIDGE, NY: CAMBRIDGE UNIVERSITY PRESS.

IMAMOGLU, C. (2000). COMPLEXITY, LIKING AND FAMILIARITY: ARCHITECTURE AND NON ARCHITECTURE TURKISH STUDENTS' ASSESSMENTS OF TRADITIONAL AND MODERN HOUSE FAÇADES. JOURNAL OF ENVIRONMENTAL PSYCHOLOGY 20: 5–16.

KLINGER, A. AND N. SALINGAROS. (2000). A PATTERN MEASURE. ENVIRONMENT AND PLANNING B, 27: 537–547.

KRAMPEN, M. (1979). MEANING IN THE URBAN ENVIRONMENT. LONDON: PION.

MALHIS, S. (2003). THE MULTIPLICITY OF BUILT FORM MANIFESTATIONS. 4TH INTERNATIONAL SPACE SYNTAX SYMPOSIUM, LONDON.

OSTWALD, M. J., J. VAUGHAN AND S. K. CHALUP. (2009). A COMPUTATIONAL INVESTIGATION INTO THE FRACTAL DIMENSIONS OF THE ARCHITECTURE OF KAZUYO SEJIMA. DESIGN PRINCIPLES AND PRACTICES: AN INTERNATIONAL JOURNAL, 3, 1: 231–244

OSTWALD, M. J., J. VAUGHAN, AND C. TUCKER. (2008). CHARACTERISTIC VISUAL COMPLEXITY: FRACTAL DIMENSIONS IN THE ARCHITECTURE OF FRANK LLOYD WRIGHT AND LE CORBUSIER. IN NEXUS: ARCHITECTURE AND MATHEMATICS, ED. K. WILLIAMS, 217–232. TURIN: K. W. BOOKS AND BIRKHAUSER.

SALINGAROS, N. (1997). LIFE AND COMPLEXITY IN ARCHITECTURE FROM A THERMODYNAMIC ANALOGY. PHYSICS ESSAYS 10: 165–173.

SONG, J. AND M. LYU. (2005). A HOUGH TRANSFORM BASED LINE RECOGNITION METHOD UTILIZING BOTH PARAMETER SPACE AND IMAGE SPACE. PATTERN RECOGNITION 38: 539–552.

STAMPS III, A. (1999). ARCHITECTURAL DETAIL, VAN DER LAAN SEPTAVES AND PIXEL COUNTS. DESIGN STUDIES 20: 83–97.

STAMPS III, A. (2003). ADVANCES IN VISUAL DIVERSITY AND ENTROPY. ENVIRONMENT AND PLANNING B, 30: 449–463.

STEADMAN, P. (1983). ARCHITECTURAL MORPHOLOGY: AN INTRODUCTION TO THE GEOMETRY OF BUILDING PLANS. LONDON: PION.

STINY, G. (1975). PICTORIAL AND FORMAL ASPECTS OF SHAPE AND SHAPE GRAMMARS. BERLIN: BIRKHAUSER.

TELLER, J. (2003). A SPHERICAL METRIC FOR THE FIELD-ORIENTATED ANALYSIS OF COMPLEX URBAN OPEN SPACES. ENVIRONMENT AND PLANNING B, 30: 339–365.