120

Characterising place by scene depth

Adam Davis

University College London, Foster + Partners
adam.davis.10@ucl.ac.uk

Sean Hanna

University College London
s.hanna@ucl.ac.uk

Francis Aish

Foster + Partners
faish@fosterandpartners.com

Abstract

Turner and Penn introduced the notion of integration of isovist fields as a means to understand such fields syntactically -- as a set of components with a structural relationship to a global whole (1999). This research was further refined to put forward the concept of visibility graph analysis (VGA) as a tool for architectural analysis (Turner, Doxa, O'sullivan, & Penn, 2001), which has become widely used.

We suggest a complementary method of characterising place that does not make use of integration or a graph yet which allows -- as visibility graph analysis does -- discrete view points to be dimensioned in relation to a set of such viewpoints. In our method, Principal Component Analysis (PCA), a statistical technique, is employed to infer salient characteristics of a set of views and then to situate these component views within a low dimensional space in order to compare the extent to which each view corresponds to these characteristics. We demonstrate the method by reference to two distinct urban areas with differing spatial characteristics. Because PCA operates on vectors, order of the data has important implications. We consider some of these implications including view orientation and chirality (handedness) and assess the variance of results with regard to these factors.

Keywords

Isovist, place, classification.

1. Introduction

Benedikt first proposed the isovist as a computational tool for the study of architectural space and demonstrated the utility of 'isovist fields' for providing a comprehensive understanding of space (1979). Turner and Penn argue that in order to understand such fields syntactically, it is necessary to consider the integration of discrete points therein (1999). Building on this research, they and colleagues subsequently proposed the architectural intervisibility graph as a means to obtain similar global spatial metrics without the need to refer to isovists for constituent viewsheds (Turner et al., 2001). This paper proposes a way of inferring 'global' properties of a space without integrating – that is without taking account of the inter-relation of each sample point as a node in a graph. Instead, principal component analysis is employed to find the most salient differentiating characteristics between all views, and then to organize views according to their correspondence (or non-correspondence) to these characteristics.

PCA has been used in a similar fashion in computer face recognition. In what is known as the eigenface approach, PCA is employed to determine how like or unlike an individual human face is to a body of specimen faces used for comparison (Turk & Pentland, 1991). This statistical method (like other 'machine learning' methods) allows for a relative but precise metric of how 'far apart' or 'close together' various data points (isovists, or views in the present study) may be within the context of an entire data set.

PCA is well-suited to the problem of face recognition because of the relatively invariant geometric information of different faces. Most frontal face images display a high degree of symmetry about the vertical axis and disposition of facial features does not vary greatly. By contrast, a method for characterising urban places should allow for much greater variance in image geometry, reflecting various characteristics of views of urban morphology: differing building heights, street widths, and dimensions of open spaces, to name but a few. In addition to establishing the feasibility of the basic approach for analysing urban space, it became important to assess the extent to which the method was capable of determining associations between sample place images despite geometric variation between them.

2. Measuring scene views

In describing their method, Turner and Penn indicate that their computational apparatus 'Vista' is capable of capturing three-dimensional depth information but that they constrained this to two dimensions in the context of their study. The PCA method described in this study could be applied to two-dimensional visibility data, however for reasons external to the current study, we have used three-dimensional views similar to what Turner and Penn describe. Such views may be considered a subset of isovists. These might be termed depth views and they are similar to Benedikt's definition of isovists in that they record points visible from a given vantage point (Benedikt, 1979). Yet Benedikt's isovists are always bounded, with occluding and boundary surfaces comprising those portions of the isovist which are not occupied by real surfaces, whereas depth views capture only real surfaces. Views are further distinguished from isovists by virtue of the fact that in a view, visible points are transformed by perspective projection.

3. Methodology

Highly simplified three dimensional models of London and Manhattan were used for the study. These models represented buildings as simple extrusions of their plan extents, each according to surveyed height. This data was obtained from GIS databases and processed for import into a CAD system for further manipulation. Open spaces in the model were sampled at 2 metre intervals to create points from which to measure view.

Views were registered as a perspective projection of the model to a two dimensional plane with the

depth of each rasterised point recorded as an integer value in the interval $0 \dots 2^8$, thus preserving information about the third dimension, depth (Figure 1). This depth information had a resolution of approximately 1m. Because these views are projections to a 2D plane, they were closer to natural scenes than an isovist, which is typically understood by its plan representation. However, while many of these views are recognizable as a representation of an urban environment, they show depth of visible surfaces as greyscale values and thus contain locally significant information that would not be directly encoded in a natural scene. Also, because panoramic images were ultimately determined to provide the best measure of place within the context of PCA analysis, the individual representations diverge in this respect from natural perception of space.

Figure 1: Sample depth views produced from model (cropped from full panoramic view).

4. Analysis with principal components

In order to determine commonalities between views, principal component analysis was employed in a manner similar to that used in the "eigenfaces" approach to face recognition (Turk & Pentland, 1991). This approach takes advantage of the general commonality amongst faces -- most have eyes, nose, mouth, each set out in similar proportions across various faces (O'Toole, Deffenbacher, Valentin, & Abdi, 1993) -- to compose a useful set of measures indicating the greatest difference among the set members. Similarly, PCA as we have employed it, takes advantage of common features in urban space (a ground plane, streets and other passages, buildings), and certain proportions and relationships among these to identify what qualities of viewpoints express the greatest amount of variation.

The product of PCA are principal components, which are a set of vectors, each equal in dimension to the input vector, ordered according to the amount of variance they can explain amongst the input vectors (depth views, in the present study). The first principal component represents the greatest variance of data within the set. Successive principal components express the greatest amount of variance still possible within the set after taking into account the variance explained by previous vectors, always subject to the constraint of being orthogonal to the preceding component vector

(Figure 2). For any set of vectors of length L for which each vector is different, there will exist $L - 1$ components required to explain all variation within the set.

Principal component analysis can in some circumstances produce measures that correlate to a degree with semantic categorization. For instance O'Toole et al. demonstrated in a face recognition task that the second eigenvector had strong correlation to the sex of input faces (O'Toole et al., 1993). However, the same study identified at least 12 out of 100 eigenvectors displaying significant correlation to sex. Accordingly it is infeasible to isolate one or even a limited set of principal components as representing differentiation along a specific semantic axis. In our study, this means that specific principal components do not correspond to a given space, street, neighbourhood or city but rather express qualities of space as registered in depth views that may, in conjunction with other components, have some degree of correlation to semantic categories.

In outline, the algorithm for computing the principal components involves:

- assembling the vectors N of dimension M in a matrix of dimension $M \times N$

- calculating the empirical mean of the matrix by finding the mean vector of each column of the matrix

- subtracting the mean from each column (vector) in the matrix (mean centring)

- finding the covariance matrix of the mean centred matrix

- computing the eigenvalues and eigenvectors of the covariance matrix

    (Jolliffe, 2002; Oliva & Torralba, 2001; Turk & Pentland, 1991)

The eigenfaces approach proposes dimensionality reduction as a means for interpreting likeness of an arbitrary image to a training set, based on the proximity of that image in the space defined by the principal components. By contrast, the current study is less concerned with view recognition -- that

is in understanding the relation of one sample view to some set of views -- and is instead focused on understanding how multiple sets of views, each of which represents a unique view of physical space, relate to one another.
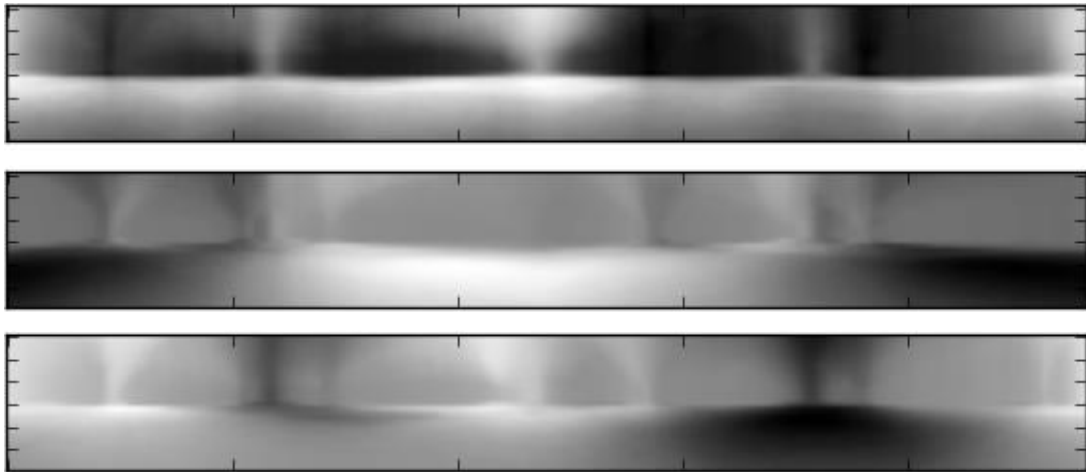

Figure 2: The first 3 principal components (eigenviews) of views of two neighbourhoods from London (Seven Dials) and New York (Manhattan West Side)

5. Eigenviews

Principal components can also be referred to as eigenvectors, since they represent the axes in high-dimensional space along which measured data varies most. The magnitude of each eigenvector is its eigenvalue, which is a measure of the amount of variance explained by that eigenvector. The ratio of any given eigenvalue to the sum of eigenvalues in the set gives the proportion of variance explained by that eigenvector. Because principal components in this study are used to express the variance in view data, they are referred to as eigenviews (Figure 2).

For purposes of visualization, once the principal components of the entire set of queried images were obtained, each view was plotted as a vector against the first three eigenviews: for a given

vector view, v, its position relative to eigenviews a, b, and c, expressed as matrix $\begin{vmatrix} a & b & c \end{vmatrix}$ is given

as $v \cdot \begin{vmatrix} a & b & c \end{vmatrix}$. This technique of plotting individual specimens against the principal components of their set has previously been demonstrated useful in categorizing space as measured by axial maps (Hanna, 2007; and Laskari, Hanna, & Derix, 2008). In these precedents similar groupings occupy dense regions of the plot space in much the same way that matter in space tends to coalesce around a gravitational centre. In the current study, by contrast, similarity is observed in linear and surfacial extensive groupings.

6. Comparison of panoramic views associated by streets

Because "place" is a concept that can be understood at multiple scales, the method assessed correspondence between place and view at the scale of individual spaces (eg streets or courtyards), and at the neighbourhood scale, using models of two cities with significantly different morphology: the Seven Dials neighbourhood of London (SD) and a portion of Manhattan's upper west side (MH). In terms of the organization of data for this study, then, place is fundamentally a question of set

membership: views (v) were grouped by place (P): $v_{1x} \ldots v_{nx} \subset P_{ix}$ and by city (C): $v_{1x} \ldots v_{nx} \subset C_x$

where x represents each of SD and MH and $C_{SD} \cap C_{MH} \quad \varnothing$.

Approximately 2000 views each were sampled from Seven Dials and the West side of Manhattan (Figure 3). Each view is looking directly Southward at the horizontal middle of the image, with North being at the left and right of the image. From these a training set was created by randomly sampling

2000 views from this set with an even probability of inclusion for each city in that set to create the matrix M from which the principal components were derived:

$$M = |V_{SD1}, V_{SD2} \ldots V_{SD1000}, V_{MH1}, V_{MH2}, V_{MH1000}|.$$

The images were 240 pixels wide by 30 pixels high, for an image-vector of 7200 dimensions. The principal components were extracted as outlined above and then all 4000 views were plotted within the view space.
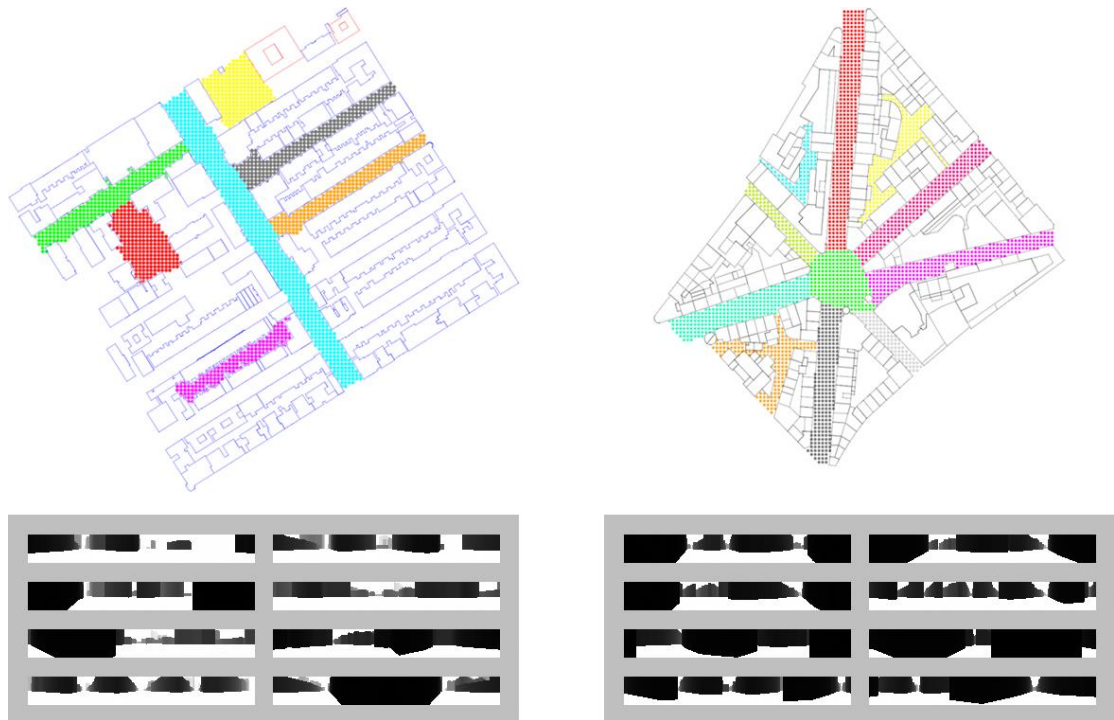

Figure 3: Plans of neighbourhoods in Manhattan (left) and London (right), with randomly sampled panoramic views of each below. Coloured dots in each plan show points from which views were sampled. Each colour represents a place for the purposes of the study.

At both of these scales the method provided for distinguishing between most places. Such distinction took the form of concentrations of eigenviews within the view space. At the scale of individual spaces (streets, etc) such concentration was typically expressed either by forms approximating a one-dimensional non-linear manifold (three-dimensional curve) or as a cluster of points in a particular region of the view space. At the neighbourhood scale, as the union of all such places within each neighbourhood, such concentration was represented by a point distribution roughly corresponding to a two-dimensional non-linear manifold.

Places differed in the extent of their order through plotting in the view space. Approximately half of the tested spaces clustered neatly along a parabola, with some views trailing off of these parabolas in sinusoidal curves. The remaining places represented a roughly ellipsoid cluster with even distribution and density of constituent views in the view space (Figure 4).

Overall, views from Manhattan were less ordered within the view space than those from Seven Dials. This is attributable to two factors:

- because of its rigorous grid, the urban morphology of Manhattan (as experienced from ground-level views) varies less and thus expresses less variance within the view space

- secondly, because buildings are considerably higher in Manhattan, there is less variation in the image where roof edges are visible as they recede toward the vanishing point.
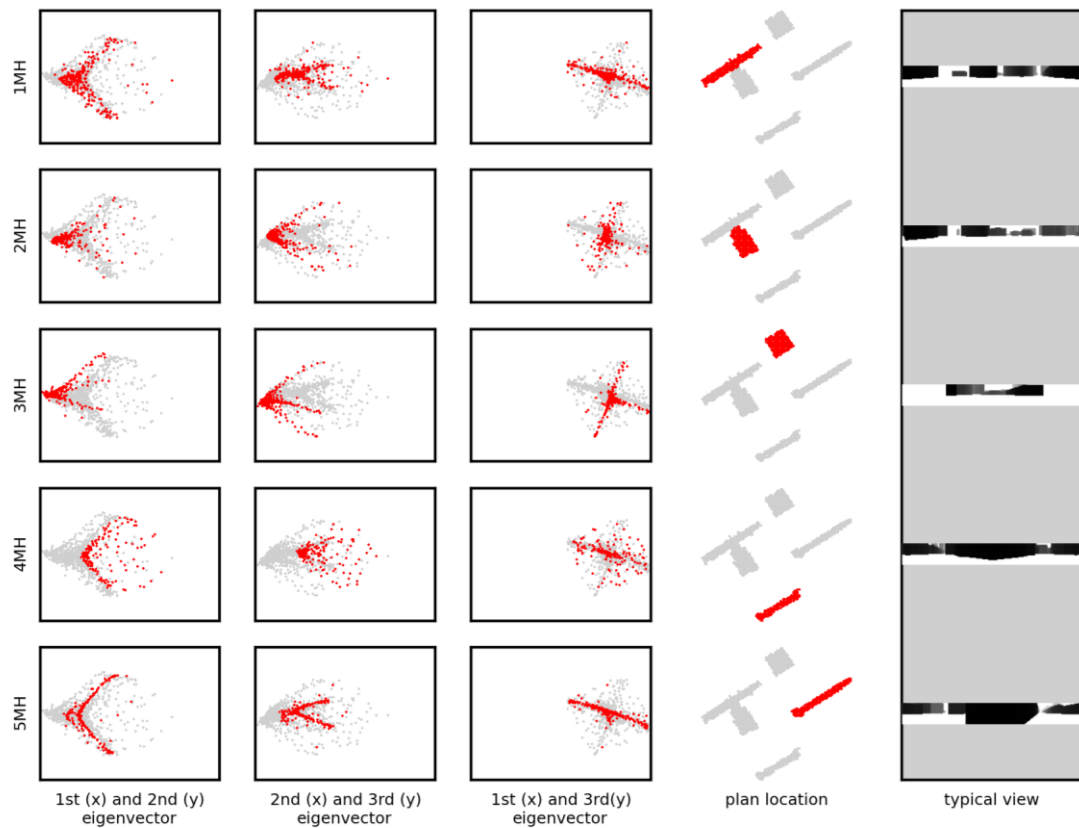
Figure 4: Manhattan West Side place plots against first three eigenviews with location plan and typical views.

7. Geometric variation

In order to establish a method which is sufficiently robust to potentially simulate human discernment, it was necessary to assess the extent to which the geometry of view images factored into the classifications produced by this unsupervised learning method. In order to test the effects of such geometric change on the overall classification, the analysis was run again for each of several geometric alterations, with the original analysed images serving as a control against which geometric variation was measured. The effects of such changes were measured initially through qualitative observation of the plotted data.

The first test, symmetry, represents a non-embodied transformation, in the sense that a subject would not experience such a transformation in the context of real urban space. Rather, the experiment regarding symmetry is included to test the associative nature of place recognition across different places: that we can perceive similar qualities of place despite mirroring about our normal axis of vision.

8. Symmetry

First the eigenviews method was tested to determine how mirrored views of a place would be classified. Mirror symmetry is one of the geometric transformations under which object recognition can be expected to remain invariant (Konen, Maurer, & Von Der Malsburg, 1994, p. 1019). The data from this experiment show translational symmetry (of a very minor offset) in the first and second dimension, while the third dimension shows a clear and obvious axis of mirror symmetry about the third eigenvector.

9. Rotation of viewpoint

In these experiments regarding geometric variation the control and altered image sets were sampled with an even probability of inclusion within the PCA training set. Thus, the translated views will have already factored into the principal components. Regarding the application of such a system for place recognition, it seems feasible to have expected transformations accounted for in the training set, to a limited extent. The initial experiments in single-orientation views described above show that where there is substantial variation between candidate views (for instance those that are mostly obscured by near objects), the eigenviews of the set do not provide satisfactory clustering of views common to a place. To provide for all possible transformations in the training set is clearly infeasible, then, as to do so would 'dilute' the usefulness of the principal components in providing definitive clustering within the eigenview space.

In the case of mirror symmetry, because there is only a single possible translation, a simple correlation operation about the axis of the third eigenvector could verify that a candidate view should be associated with a particular place. In the case of rotation, there are an infinite number of possible translations, the magnitude of which could be expected to have a proportional translation on the distance of correlated points in view space. Because of this high variability, it is difficult to allow for all possible rotations within the training set and accordingly it is not expected that the method could be supplemented to automatically check for chiral correspondence in the same way as proposed for reflection.

10. Visualization methods

The three dimensional view space against which place images were plotted provided a suitable method for classifying place information. Because of the amount of data being plotted, its hierarchical organization (city > neighborhood > place), and the high resolution of vector data being examined (relative to many other implementations of PCA in the literature), thorough examination of the results required implementing specialized visualizations. Preliminarily the results were plotted as data points in three dimensions. Plotting individual places in colour and plotting each of the cities in isolation allowed for basic discernment regarding some of the outcomes of the study. However, this basic representation made it difficult to associate the low-dimensional representation of the image with its original high-dimensional input (the source image).

A visualisation tool was created that plotted all images at their respective locations within the view space. This allowed for navigating quickly between looking at the image in sufficient resolution to compare individual features between that image and its neighbours, to looking at long chains of like images, to looking at larger patches of like surfaces, or even the entire set of sampled views (Figure 5).
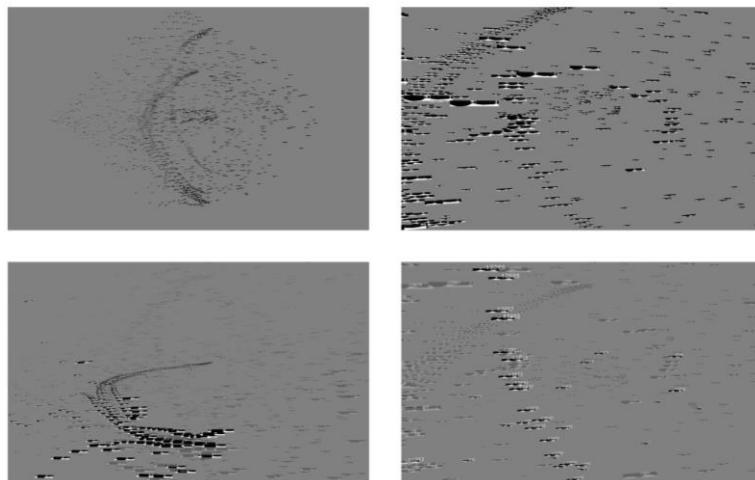

Figure 5: An interactive tool for navigating three dimensional plot data

11. Results and discussion

The method provides a means of differentiating between different places. Views of like places (defined by similar spatial structure, as reflected through depth) displayed similar morphology and/or regional clustering when organized by their principal components. Low-dimensional differentiation was in evidence between cities, and between different places within each of those cities. The most significant shortcoming of the method is that it fails to characterize reflected images or images taken from a rotated viewpoint as referring to identical places. The potential for supplementing the method to account for these failures to recognize identity are discussed below.

The method is computationally efficient and has the benefit that the classifiers (eigenviews) have some degree of human legibility.

12. Characterization of place clustering

The method creates discernible patterns within the eigenview space associated with each place. Several quantitative and qualitative characterizations of this data can be made, both through statistical sampling and empirical observation of the structures defined.
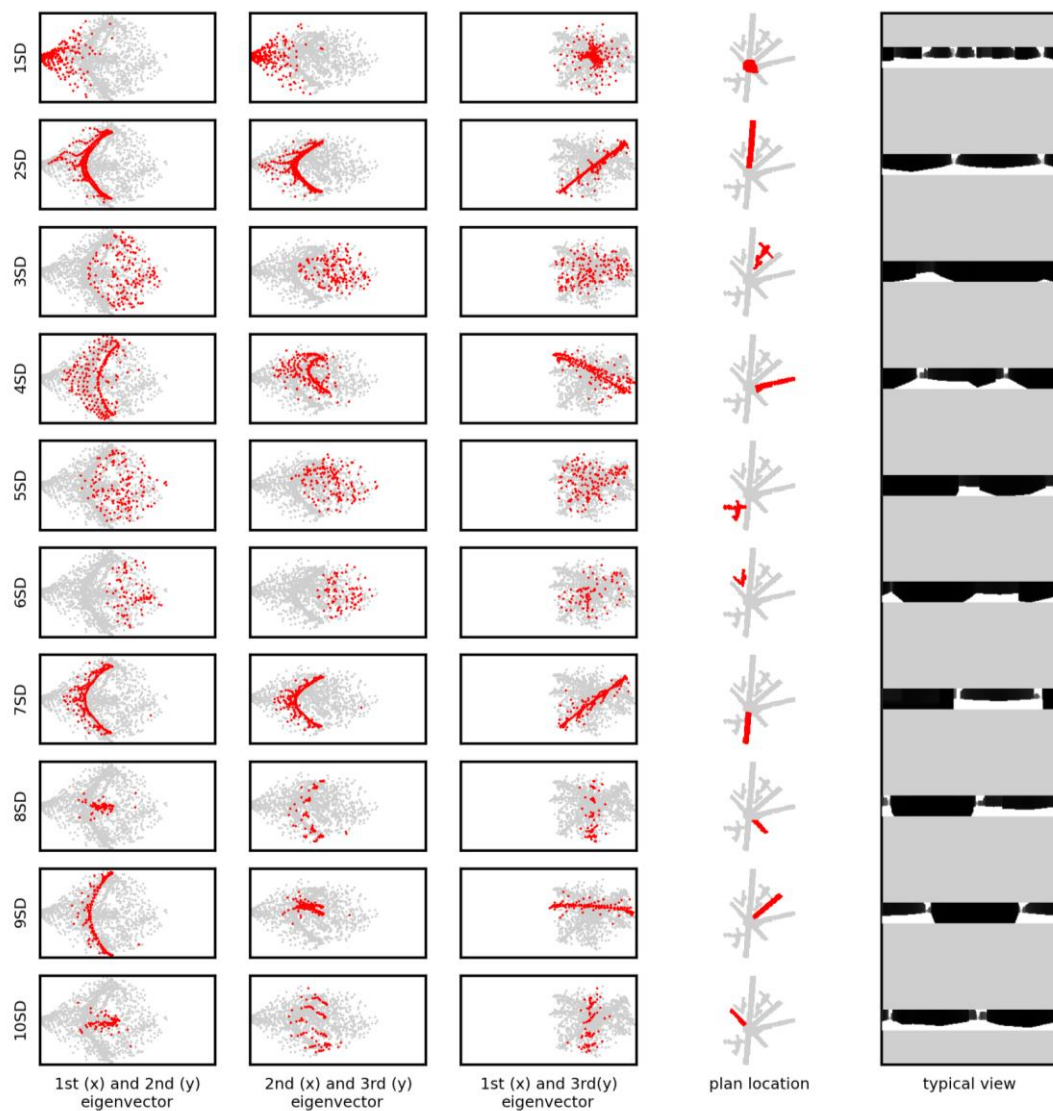

Figure 6: Seven Dials places plotted against the first three eigenviews of the view space (leftmost three columns with location plan and typical views.

- The amount of opening in a view of space (represented by lighter or white areas bordered by nearer surfaces) is inversely proportional to the first eigenvector. Places such as courtyards which have few or no immediately visible deep spaces (passages, avenues, etc.) are found further along that eigenvector, whilst those with many (or large) openings group at the opposite end of the data set.

- Because of the lack of contrast between near and deep spaces in such courtyard images, the plots of such points in the eigenview space have a less distinctive form. Such closed spaces usually manifest as a rather diffuse plot, whereas more open spaces typically have a recognizable form. Still, the distribution of points is informative about the nature of the space. For instance, Place 6SD is distributed farthest along the first eigenvector of any of the view space plots. Compared to other plots of enclosed space (Places 3SD and 5SD) which display similar clustering, Place 6SD is the smallest space (offers the least contrast in depth) in Seven Dials (and overall between the two cities) and also has no intersection of two passages, as with Places 3SD and 5SD.

- Several of the Places analysed in Seven Dials (2SD, 4SD, 7SD, 9SD) have a distinctive parabolic form visible along the axes of the second and third eigenviews (plot of first against second and first against third eigenvectors, respectively) (Figure 6). A survey of images from each of these places reveals that they are mostly characterized by having two openings, roughly half of the image width apart -- these two openings represent either end of a passage. The form of the parabola seems invariant across sample places to the orientation of the passage (notwithstanding rotation along the axis of the second eigenvector, described below).
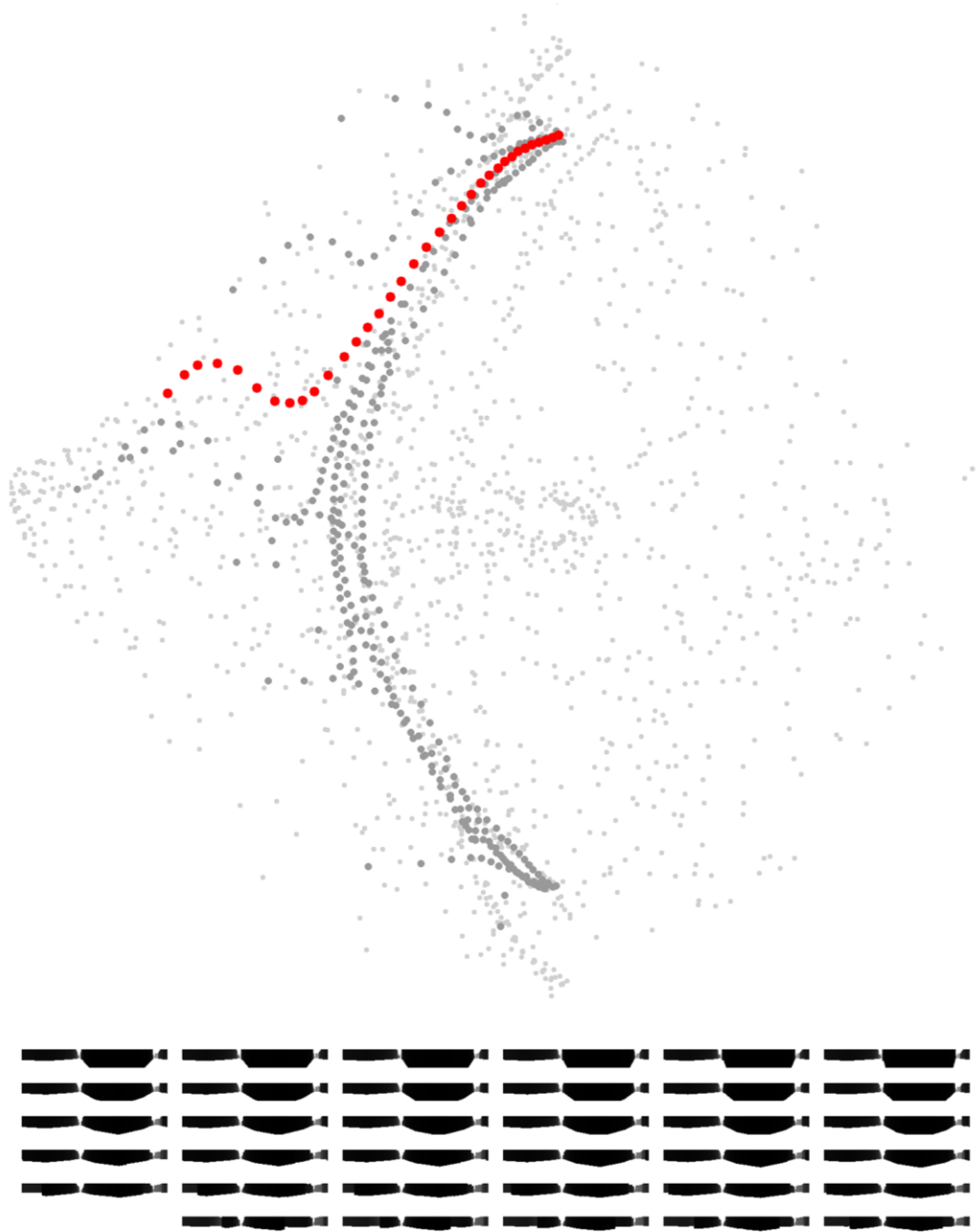
Figure 7: A plot of Seven Dials views in the view space, with a transect (red) along a string of related views. The views at bottom are ordered along the transect (from right to left and from top to bottom) and show the gradual increasing of open space (white and light colours).

- Those places which display a parabolic form in the view space typically have "tails" that bifurcate from the general trend of the parabola. A transect along one such tail demonstrates that these are an artefact of the transition between a passage and the more open, less directed Place 1SD (Figure 7). As one approaches the clearing of the round-about, the openings become wider, relative to the picture plane and the foreground occupies more of the overall image.

- The third eigenvector describes an axis of rotation about which the various passages (Places 2SD,4SD,7SD,8SD,9SD,10SD,) are oriented. This provides a consistent frame of reference against which orientation can be measured (allowing for the fact that separate iterations of

the analysis can yield a result that is the mirror opposite of a previous iteration because the selection of orthogonal vectors in PCA is not chiral).
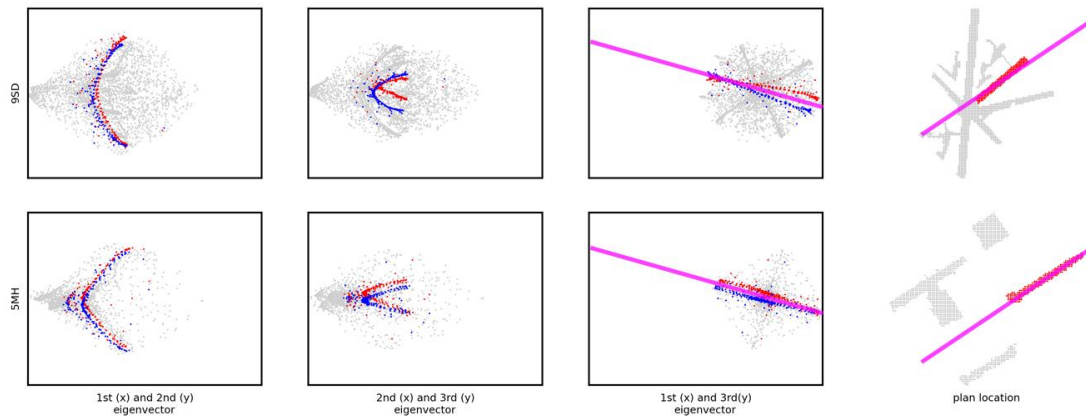

Figure 8: Two places of passage in Seven Dials and Manhattan, observed in the control (red) and view rotated 120 degrees Eastward(blue). The magenta lines at left describe axes of reflective symmetry on the plane of the 1st and 3rd eigenvectors. At right, they show the correspondence in cardinal orientation of the two places. Each vertical pair of magenta lines is parallel.

- In the 120 degree plots of both Seven Dials and Manhattan, the third eigenvector produces symmetric reflection of the plot distributions (Figures 8,9). The orientation of this axis of symmetry heavily favours the Manhattan plots in terms of minimum distance to this symmetrical axis. The only plots in Seven Dials with similar proximity to this axis are those where the streets in questions are oriented at similar angles to those of the Manhattan grid. This suggests that the uniform nature of the Manhattan data with regard to street orientation and the coincidence of these street orientations across the two data sets created a strong weighting toward this orientation in the resulting plots.
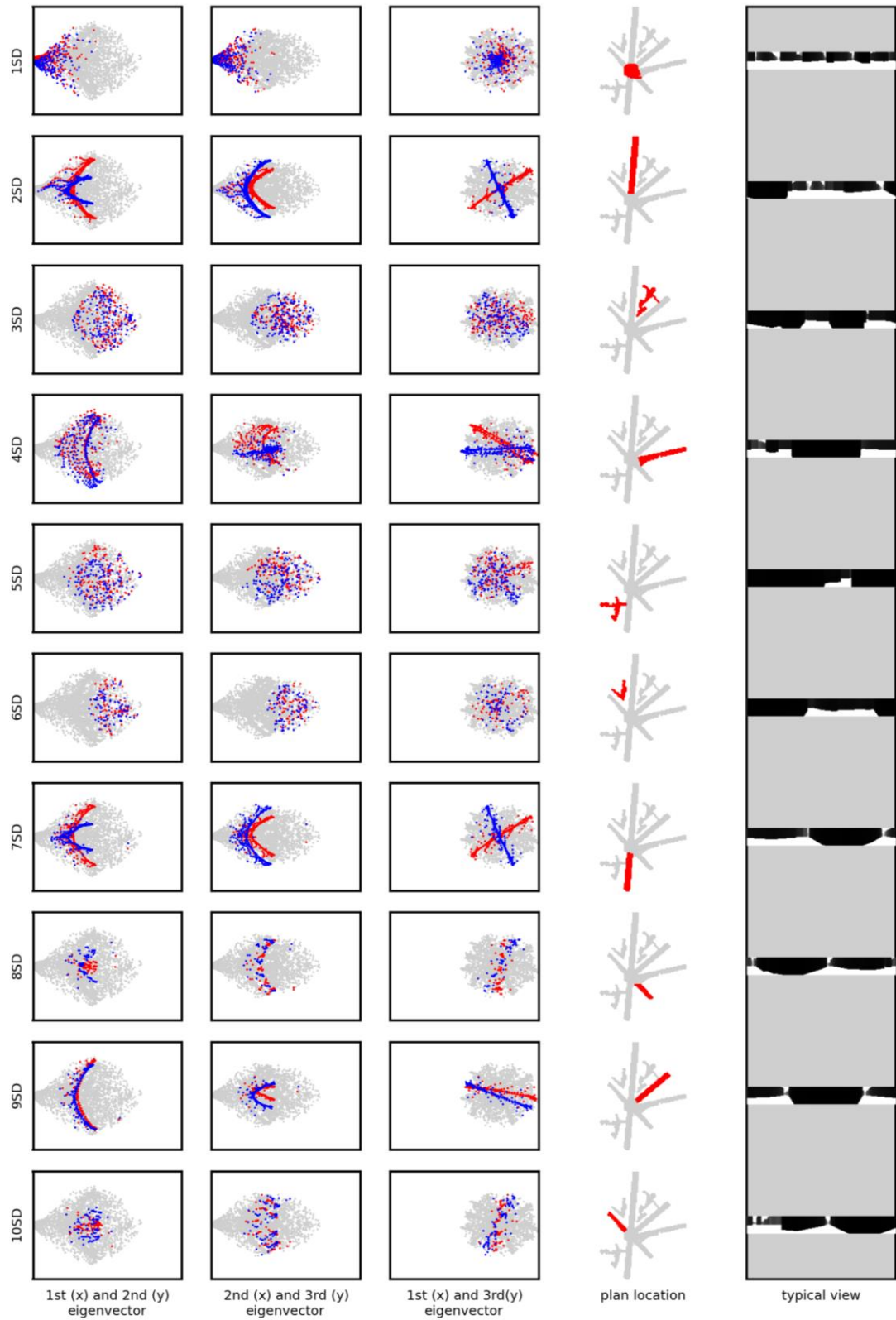
Figure 9: London views (red) and views thereof rotated 120 degree Eastward (blue) embedded in two city eigenview space

- Place 1SD (the Seven Dials roundabout) is recognizable, both in plan view and in the images generated therefrom as the place that is most unlike any other, and that is reflected in the separated distribution of its constituent views in the view space. It is at the extreme end of

the first eigenvector (defining open-ness), as could be expected because it is the most connected of the spaces measured. At the same time, it is diffuse along the third eigenview axis (defining orientation) because it does not have a primary orientation in geographic space. The set of all plotted views becomes pinched at the low end of the first eigenvector and it is this pole that Place 1SD occupies. It is probable that this pinched point is simply the extreme limit of the parabolic form expressed in other places as symmetry about the first eigenview. This seems to be confirmed by the New York place plots (1MH,2MH,3M), which display this quality where the measured area borders on wider space. The increased open-ness in images from this region of the view space comes both from additional openings but also from the more distant "horizon" where buildings meet the roundabout.
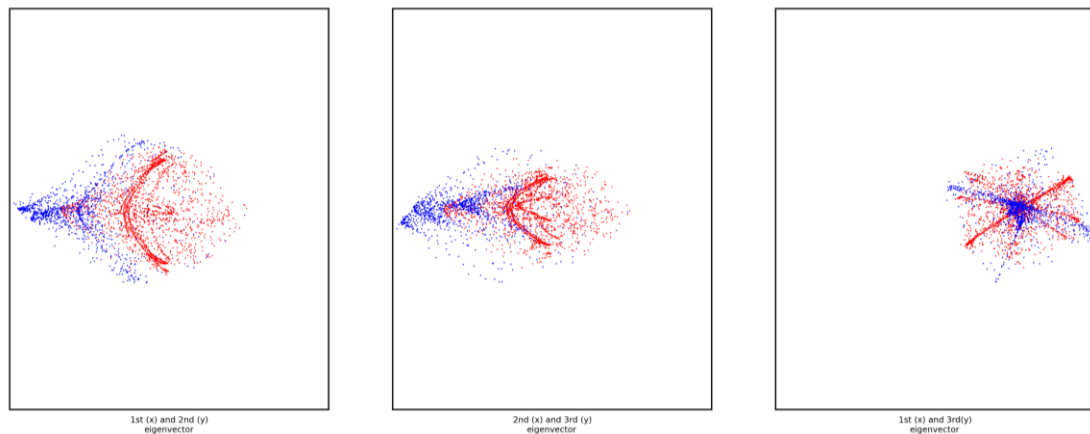

1st (x) and 2nd (y) eigenvector   2nd (x) and 3rd (y) eigenvector   1st (x) and 3rd(y) eigenvector

Figure 10: Views of all London places (red) and all Manhattan places (blue) embedded in the view space.

- When the London places and Manhattan places are plotted together, the distinctive parabolic forms of each are still visible (Figure 10). There is significant overlap in the view space among the two distributions but the trend of each is distinctive along the first two plot axes, with London generally occupying points closer to the centre and Manhattan distributed further away from the centre.

13. Eigenviews

As mentioned earlier, one of the advantages of using eigenviews as a method for classification of images is that the eigenviews themselves have some degree of legibility as images of space. The results have demonstrated certain characteristics in the data by virtue of position in the view space. Some of these same characteristics are discernible in the eigenviews as well. For instance, the trend of decreasing open area along the axis of the first eigenview discussed earlier can be inferred from the eigenview image. The high contrast between the black building surfaces and the white bands indicates that these characteristics of the images express the greatest variance within the total set of images.

14. Depth views as a method for place distinction

The results show that views of depth information are capable of distinguishing places from which these views were taken and that this classification is successful at multiple, nested scales. Yet the notion of what defines 'place' in this context is only partially defined. This research has demonstrated a utility for a notion of place, but has not proposed specific terms for defining place (either ontologically or operationally). Further research is needed to test the appropriateness of the scale and units chosen against other possible configurations:

- how broadly or narrowly (in geographic terms) can place be defined before it loses the potential for classification demonstrated herein?

- what other (non-depth-specific) ways of encoding views of place might offer similar possibilities for classifying space?

Many of these refinements and adjustments are likely to grow out of the attempt to correlate these findings with perceptual experiments (described above).

15. Principal Component Analysis compared to visibility graph analysis

By the nature of the method, the results of principal component analysis are always contingent upon the data points sampled. In order to compare two or more places, eigenviews from each must be included in the vectors considered for analysis. Therefore, this method does not provide metrics independent of particular spaces, as mean shortest path does for visibility graphs (Turner et al., 2001, p. 114).

Where visibility graphs are one dimensional (the intervisibility of any two edges in the graph being the only data that is encoded), PCA offers the possibility of analysing an arbitrary number of dimensions up to n-1, where n is the number of vectors input to the method. Practically speaking, this would not be useful and in typical application of the method, the eigenvalues will determine how many dimensions should be used to interpret the data as they indicate marginal differentiation of each principal component.

16. Further research

While the method has demonstrated a capability to differentiate between distinct places and spatial characteristics, much further research is needed to test the method. The method presented here for interpreting the results of the analysis in a low dimensional space are still largely qualitative and repeated application of the method would benefit from quantitative summary of the relationship of constituent views in low-dimensional space.

Turner et al. raise the question of how to determine the positioning and spatial resolution of points sampled in visibility graph analysis (Turner et al., 2001). Because PCA can encode results of arbitrary dimension, it could potentially act as a complement to VGA, informing this issue of sample placement and resolution. Sparsely placed sample views that are determined to differ the most could prove to be the best candidates for visibility graph nodes.

17. Conclusions

We have proposed the use of principal component analysis applied to depth views (views encoding the depth of all visible surfaces) to recognise and classify places based on views thereof. This research is related to, and influenced by, other applications of the computational analysis of space such as isovist integration and intervisibility graph analysis. The method differs from these precedents most significantly in that it does not require modelling the direct inter-relation of each possible pair of component views (as does intervisibility calculation).

Principal Component Analysis has shown promise as an efficient method for distinguishing between places and demonstrating similar results for similar spatial conditions. Yet it fails to provide an appropriate solution for an important factor of our perception and cognition of place: that we tend to recognize place invariant to transformations of perspective such as reflection and rotation. This preservation of information about orientation and location of objects in the visual field is potentially both a strength of the method and a liability, as compared to visibility graph analysis.

References

Benedikt, M. (1979). To take hold of space: isovists and isovist fields. *Environment and Planning B*, *6*(1), 47–65.

Hanna, S. (2007). Defining implicit objective functions for design problems. In *Proceedings of the 9th annual conference on genetic and evolutionary computation*, ACM, 2013-2020.

Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). Springer.

Konen, W., Maurer, T., & Von Der Malsburg, C. (1994). A fast dynamic link matching algorithm for invariant pattern recognition*. *Neural Networks*, *7*(6-7), 1019–1030.

Laskari, A., Hanna, S., & Derix, C. (2008). Urban Identity Through Quantifiable Spatial Attributes. *Design Computing and Cognition'08*, 615–634.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, *42*(3), 145–175.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., & Abdi, H. (1993). Low-dimensional representation of faces in higher dimensions of the face space. *Journal of the Optical Society of America A*, *10*(3), 405–411. doi:10.1364/JOSAA.10.000405

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86.

Turner, A., & Penn, A. (1999). Making isovists syntactic: Isovist integration analysis. In M. Major & L. Amorim (Eds.), *Proceedings of the second international space syntax symposium*, 01.1–01.9. Brasilia: Universidade de Brasilia.

Turner, A., Doxa, M., O'sullivan, D., & Penn, A. (2001). From isovists to visibility graphs: A methodology for the analysis of architectural space. *Environment and Planning B: Planning and Design*, *28*(1), 103–121.