Representations of Streetscape Perceptions Through Manifold Learning in the Space of Hough Arrays

Stephan K. Chalup, Riley Clement, Joshua Marshall, Chris Tucker, and Michael J. Ostwald

Faculty of Engineering and Built Environment

The University of Newcastle, Callaghan 2308, Australia

Email: `chalup@cs.newcastle.edu.au`

Abstract— **This study is part of a project which investigates computational principles which underlie perception and representation of architectural streetscape character. Some of the principles can be associated with fundamental concepts in brain theory and Gestalt psychology. For the experimental analysis streetscapes were represented by sequences of digital images of house façades which were prepared by a team of researchers from architecture. Two methods for non-linear dimensionality reduction, isomap and maximum variance unfolding, were applied to a set of Hough arrays (for lines) of the given images. An analysis of the extracted "streetmanifolds" revealed groupings of house façades with similar visual character and proportions. Comparative tests were conducted on a simple cylinder shaped example manifold to evaluate the geometric stability of the two dimensionality reduction methods. All experiments addressed variations of the distance metric and the neighbourhood parameter.**

I. BACKGROUND AND MOTIVATION

One of the tasks architects are facing in urban design is determining aesthetic measures and guidelines, according to which houses in a street are built, such that the resulting streetscape is of appropriate visual complexity [16]. The aim of this study was to show that a combination of line extraction and manifold learning algorithms could find "architect-like" representations of digital image data of house façades and streetscapes in the form of a manifold. This may allow to simulate procedures of the human visual system essential for the aesthetic perception of architectural image data. Before an experimental approach is proposed some background information is provided to support the hypothesis that manifold learning can be applied in computer vision so that it parallels concepts from visual neuroscience and Gestalt theory [29], [47].

The visual cortex [20] is organised such that local signals received at the retina are processed through a hierarchy of layers into higher level representations which can be holistic, abstract, and multimodal. Another fundamental principle of the visual cortex is that it processes information in separate streams which can be specialised on, for example, spatial localisation, or form and object recognition.

Hubel and Wiesel [23], [24] discovered that so called simple cells in the striate cortex (V1) of cats and monkeys respond to oriented bars/edges in their receptive fields. Later it was found that simple cells are not only sensitive to the orientation and location of the bars but also to the spatial frequency of visual stimuli [14]. The ability to report the direction of movement is already available in retinal ganglion cells [2], [3], [33].

The perception and processing of line directions is a fundamental component of the brain's visual information processing system and regarded as essential for form recognition. Within the striate cortex, functionally separate pathways have been distinguished [32], [40]. Their corresponding neurons in V1 appeared to be associated with different categories of perceptual processing, namely coding of motion and stereopsis, coding of colour and contrast, and coding of form which is based on orientation selective neurons [36, p.48].

Orientation-selective adaptation in the human V1 was suggested in a study by Engel [17]. The results were supported by Larsson et al.'s work [31] who also showed that second-order stimulus orientation emerged through continued processing after V1. This means that orientation selective processing is a central concept of the human cortex and occurs in multiple visual areas including V1, V2, and others [31].

Evidence that local features are integrated into global shapes through processing in multiple visual areas comes from fMRI studies in humans and monkeys [30].

Gestalt psychology [29], [47] addresses some corresponding concepts and principles and argues, for example, that human visual perception is holistic ("the whole is more than the sum of its parts"). Several concepts found in modern visual neuroscience were predicted by Gestalt theorists about 60-80 years ago [48]. Gestalt theory assumes that the quality of a perceptual configuration depends on factors such as coherence, regularity, smooth continuity, unity, and simplicity. This implies that visual perception prefers continuous over broken or irregular transitions and there should be neurons sensitive to collinearity [41]. If a structure misses parts or is disturbed by noise, it is proposed the visual system tries to fill the gaps or to filter out noise. Similarity based grouping can occur driven by various features such as colour, texture, size, and shape [28], [34].

In architecture, the perception of lines, in particular the perception of horizontal and vertical lines, plays a dominant role in the visual evaluation of house façades and streetscapes. Therefore it should represent a significant component of the experience learned by an architect required for judging the aesthetic aspects of streetscapes.

Our hypothesis was based on the assumption that simulated processing in the form stream of the visual hierarchy, when

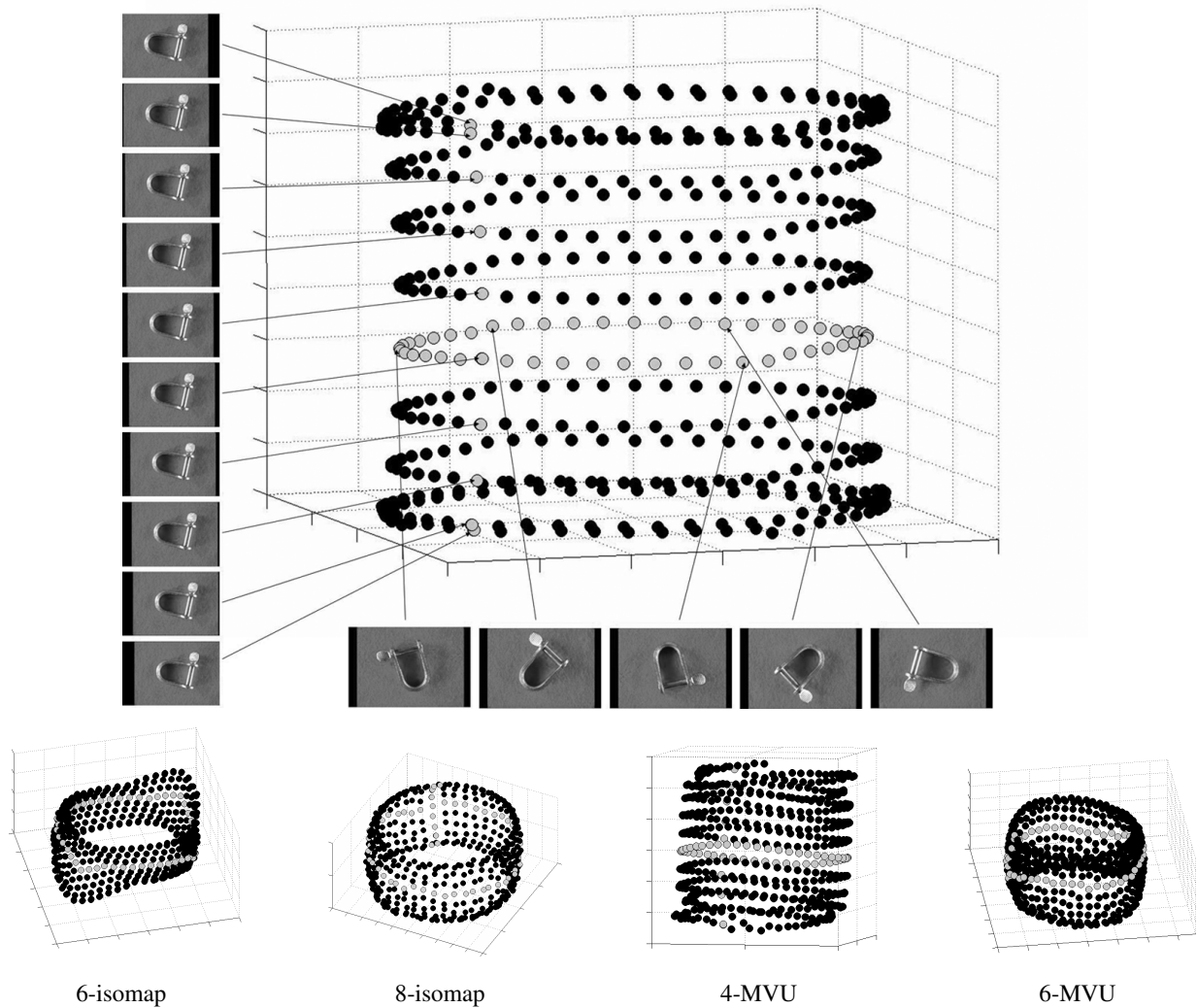6-isomap   8-isomap   4-MVU   6-MVU

Fig. 1. Cylinder shaped 2-dimensional manifold extracted from a 56,064-dimensional space of colour pixel. The data set consisted of 490 images taken of a shackle while it was rotated about 360° (in 49 steps) and translated sideways (in 9 steps) on a red backgound. The distance of translation is indicated as a black gap on the left and right side of each of the images for illustrative purposes. The large cylinder at the top was obtained using isomap with $k = 4$ and the standard Euclidean distance. Manifolds obtained using isomap with $k = 6, 8$ and MVU with $k = 4, 6$ are displayed in the bottom row.

applied to the images of a streetscape, should lead to a holistic representation of the combined information of the edge/line distributions extracted from all the house images in the data set. Following the suggestions from Gestalt theory the resulting representation should be a holistic, continuously connected object, such as a non-linear manifold. The latter will be called a "streetmanifold" within the context of this study.

Manifolds are locally Euclidean spaces with some other very general mathematical properties [42]. Visually, within dimensions 1 and 2 they can be seen as continuous, non-linear deformations of lines, circles, spheres, tori, pretzel surfaces, or similar objects.

Manifold learning describes a family of algorithms for non-linear dimensionality reduction [7]. Their purpose is to detect the essential underlying geometric structure of high-dimensional data sets and to extract it as a low-dimensional manifold. A simple and predictable example is shown in Figure 1 where a 2-dimensional cylinder embedded in $\mathbf{R^3}$ was extracted from a (192×292)-dimensional space of colour pixels. This example will be explained in more detail in Section II-C.

Previous publications which were related to the task of finding representations of the environment through manifold learning include, for example, Seung and Lee's [38] article on perception, Ham and Lee's work on robot localisation [21], and a paper by Peters and Jenkins [35] on robot sensor space representation.

The main steps of the approach taken in this study were:

1) Image data collection.
2) Application of a Hough transform for lines including appropriate parameter tuning.
3) Calculation of distances between Hough arrays.
4) Application of manifold learning.
5) Embedding of the manifold of Hough arrays into 2- or 3-dimensional space.

The remainder of this article is structured as follows. Section II describes two methods for non-linear dimensionality reduction, isomap and maximum variance unfolding, which were applied in this study and evaluates their geometric stability on a simple example. Section III briefly addresses the streetscape image data. In Section IV the Hough transform and the Hough arrays are described. The experiments on the streetscape data and the results are described and discussed in Section V. Section VI is the conclusion.

II. MANIFOLD LEARNING

In dimensionality reduction tasks we have data vectors $x_i \in \mathbf{R}^d$, $i = 1, ..., n$ in a high-dimensional space. The aim is to find a mapping

$$\begin{aligned} \mathbf{R}^d &\longrightarrow \mathbf{R}^r \\ x_i &\mapsto y_i \end{aligned} \qquad (1)$$

for all $i = 1, ..., n$ such that $r < d$ and the new lower dimensional representation of the data still contains all essential structure and information.

Traditional methods for dimensionality reduction such as principal component analysis [26] or multidimensional scaling (MDS) [12] were able to reduce high-dimensional data to linear subsets. Recently developed methods for non-linear dimensionality reduction can extract non-linear submanifolds from high-dimensional data.

The first of a series of publications on new algorithms for non-linear dimensionality reduction were published in 2000 on the algorithms Isomap [43] and Locally Linear Embeddings (LLE) [37]. Since then a large number of variations, improvements and new algorithms followed such as Laplacian Eigenmaps [4], Local Tangent Space Alignment (LTSA) [49], and Maximum Variance Unfolding (MVU) [45], [46]. Two of them, isomap and MVU, have been employed in the present study to calculate streetmanifolds.

A. Isometric Feature Mapping

Isometric Feature Mapping (Isomap) [43] is a manifold learning method which can be seen as an extension of metric MDS [12] where instead of pairwise Euclidean distances approximate geodesic distances are used. In contrast to Euclidean distances which are measured straight through the surrounding space $\mathbf{R}^d$ geodesic distances can be longer because they are measured along (shortest) arcs within the manifold using its intrinsic metric.

Let $\{x_1, ..., x_n\}$ be a given set of data points $x_i \in \mathbf{R}^d$, $i = 1, ..., n$ which are assumed to be sampled from a low-dimensional manifold $\mathcal{M}$ which is embedded in the high-

dimensional input space $\mathbf{R}^d$. The basic version of Isomap can then be explained in three steps [43]:

1) Select a neighbourhood parameter $k \in \mathbf{N}$ and construct a neighbourhood graph $\mathcal{G}$ where each point of the manifold is connected only to its k nearest neighbours. Weight existing connections between two neighbouring vertices $x_p, x_q \in \mathcal{G}$ by their Euclidean distance $\|x_p - x_q\|_2$ in $\mathbf{R}^d$.

2) Generate a distance matrix $D = (d_{ij})_{i,j=1,...,n}$ where each coefficient d_{ij} is the shortest path distance in $\mathcal{G}$ between each pair of the initially given sample points $x_i, x_j \in \mathbf{R}^d$, $i, j = 1, ..., n$. The shortest paths in $\mathcal{G}$ can be calculated, for example, by Dijkstra's algorithm [11], [15]. The idea is that the path-length d_{ij} is an approximation of the geodesic distance between each pair of points $x_i, x_j \in \mathcal{M}$, $i, j = 1, ..., n$.

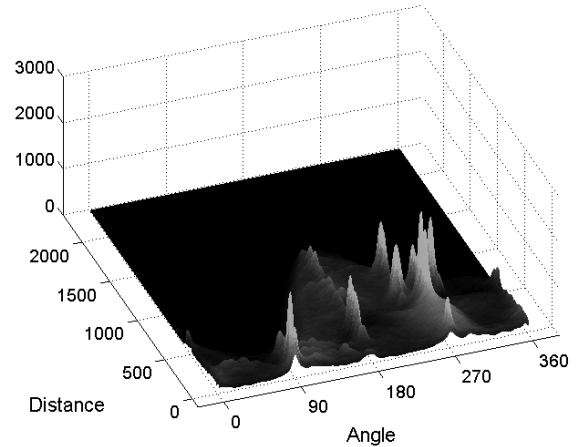3) Apply MDS using the d_{ij}, $i, j = 1, ..., n$ as inputs.

Fig. 2. *Bottom:* Example of a Hough array where line intensities are plotted as heights in z-direction above the corresponding ($Angle$, $Distance$)-point ($Angle$ is here the angle of a line's normal vector). Dominant entries can clearly be identified at points corresponding to horizontal and vertical lines. Note that in the displayed image each value in the original array was multiplied with a Gaussian function to increase radial spread. *Top:* The corresponding image of a house. The superimposed lines were obtained using an inverse Hough transform.

B. Maximum Variance Unfolding

Maximum variance unfolding (MVU) [45], [46] starts similarly to isomap from a weighted k-nearest neighbour graph. Then MVU aims to maximise the sum of pairwise distances of all data points, i.e.

$$\sum_{ij} \left(\|y_i - y_j\|^2 \cdot \delta_{NN}(x_i, x_j) \right) \qquad (2)$$

where $\delta_{NN}(x_i, x_j)$ is 1 if x_i and x_j are nearest neighbours and 0 otherwise; The maximisation is subject to two conditions

(I) $\quad \|y_i - y_j\|^2 = \|x_i - x_j\|^2$

(II) $\quad \sum_i y_i = 0$

which postulate that: (I) distances between nearest neighbour inputs should be the same as between the associated outputs and (II) the outputs should be centered at the origin.

As explained in more detail by Weinberger [45], [46] the above (non-convex) constraint optimisation task can be reformulated as a semidefinite program [44] over the matrix $K = (K_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ where $K_{ij} = y_i \cdot y_j$. With the additional condition that K should be positive semidefinite this becomes a convex task.
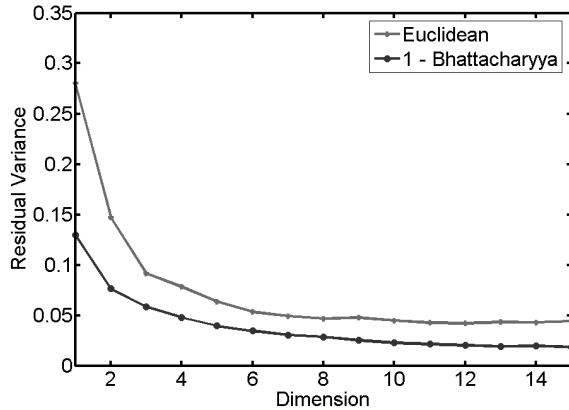

Fig. 3. The residual variance calculated by isomap [43] for the embedding based on the Euclidean distance matrix is higher than that for the Bhattacharyya measure based distance matrix.

C. The Cylinder Example

A digital image is a pixel array, that is, it can be represented by a vector in a d-dimensional space where d is the number of pixels of the image [5]. Hence a set of digital images can be seen as a point set in a d-dimensional vector space (more precisely it is a $(3 \times d)$-dimensional space because every RGB colour pixel consists of three values). If the set of images together describes a particular configuration or dynamics, of an object or agent, then manifold learning could theoretically be used to extract the essential structure of this configuration in the form of a low-dimensional manifold. For example, if the set of images describes an object moving along

a one-dimensional line then manifold learning would extract a one-dimensional line as the essential implicit geometric configuration that best captures the dynamics described by the set of images. If the object was rotated the algorithm should extract a circle. And if the object was rotated and translated at the same time the outcome should be a cylinder which was demonstrated by the experiments with the images of a shackle in Figure 1.

Experiments with isomap were conducted for $k = 4, 6$, and 8. For $k = 4$ the result was a very well-ordered cylinder (large top image in Figure 1). For $k = 6$ isomap exhibited a wave pattern which previously was reported as a typical feature of isomap when applied to a rotating object task such as the teapot in Weinberger and Saul (2006) [46]. The wave pattern seemed to disappear when k was increased from 6 to 8 and the cylinder eventually collapsed when the k value that was selected was too large. The same experiments with MVU led to very similar outcomes. The points of the resulting MVU manifolds were slightly less evenly distributed than those obtained with isomap. For $k = 6$ the wave pattern was less clearly expressed than with isomap. We also repeated all experiments using a Bhattacharyya similarity based distance. The outcomes were almost identical to those obtained using the Euclidean distance.

III. THE STREETSCAPE IMAGE DATA

The data for this study consisted of several hundred digital images of houses and streetscapes from selected areas in Newcastle. Due to space restrictions only a small number of examples can be displayed in the present paper to explain the experimental outcomes. The smallest image was 300×300 pixels and the largest image resolution was 3000×1500.

IV. THE SPACE OF HOUGH ARRAYS

The Hough transform (HT) for lines is a well established image processing method [22], [25], [39], [19], [18]. It takes a global view at an image of an architectural scene and can determine edge directions or lines including broken or virtual lines. These could be the roof of a house or a virtual horizontal line as a result of connecting horizontal components which are part of a row of several parked cars. Many variations and generalisations of the Hough transform for lines have been developed [1], [13], [8], [25].

A. Basic Hough Transform

Geometrically a line can be seen as a set of points $\mathbf{x} = (x_1, x_2)$ in a two dimensional space of real numbers $\mathbf{R}^2$. It can be determined by using the *Hessian Normal Form*

$$\{ \mathbf{x} \in \mathbf{R}^2 \; ; \; [\cos \varphi, \sin \varphi] \cdot \mathbf{x} - b = 0 \} \qquad (3)$$

where $\varphi \in [0, 360^o[$ determines the slope of the line's normal vector and $b \in \mathbf{R}$ is its perpendicular distance from the origin.

The Hough transform [22], [25], [39] for lines associates each image with an array of discrete parameters $(\varphi, b) \in [0, 360^o[\times \mathbf{R}$ which is called the *Hough array*.
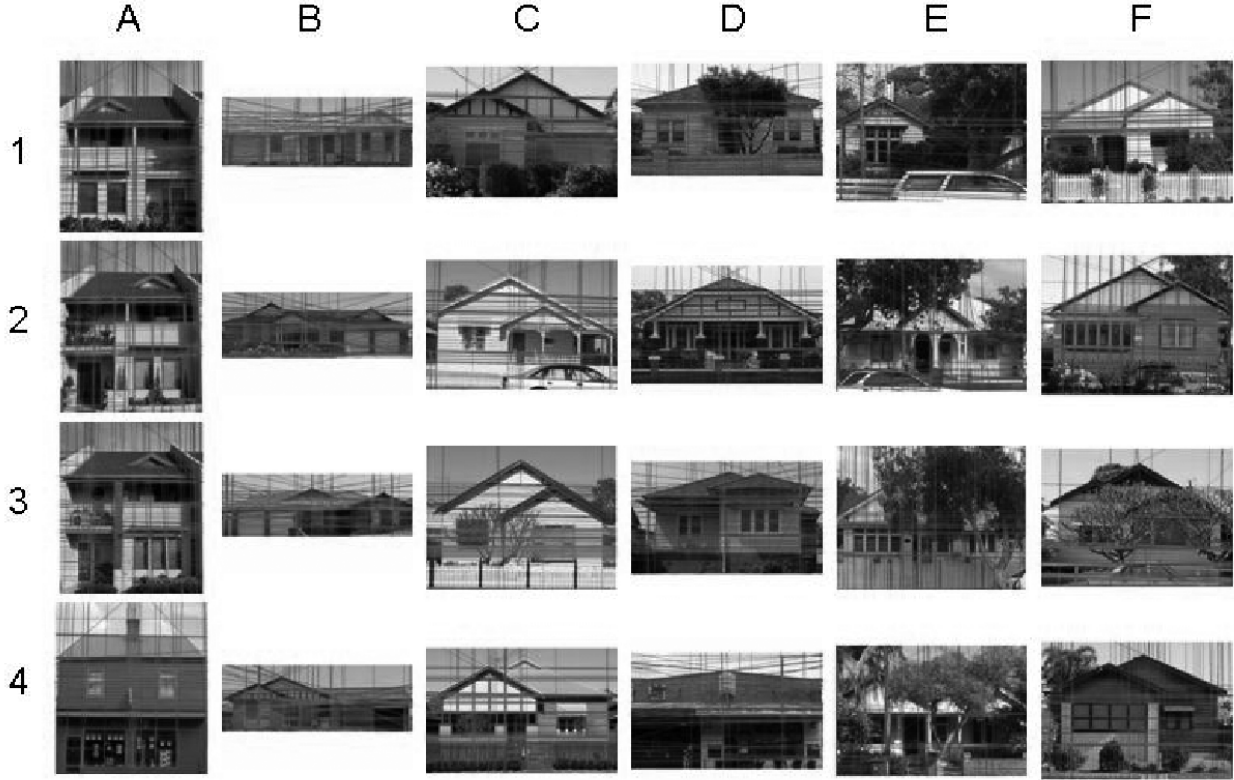
Fig. 4.   Selection of example images of house façades. The red lines were extracted with a Hough transform for lines. In the above images the number of displayed lines was limited to 60 where the strongest 20 lines were plotted thicker.

For each pixel and its neighbourhood in the original image the Hough transform examines whether a gradient for a potential line passing through that pixel can be estimated. It calculates the parameters φ and b for the potential line and stores them in the associated Hough array. For each gradient belonging to a particular line in the original image the corresponding entry in the Hough array is incremented. This is one of many possible variations of the Hough transform and is described in the book by Shapiro and Stockman [39].

B. Threshold Selection

The decision how many lines are detected and which directions are selected depends on at least two threshold parameters which had to be decided. The first parameter λ is a lower bound for the length of the gradients that should be considered as contributions to a particular line as an entry in the Hough array (either binary or as real value). The second parameter μ is a lower bound on the value of the coefficients of the Hough array. The size of the coefficients in the Hough array determine the 'strength' of a corresponding line. The parameter μ cuts out the weak lines. Those lines having support of only a few

pixels with gradients in their direction should not be counted as lines. For the present study λ and μ were set to fixed values which were determined in a series of pilot experiments.

C. Distances Between Hough Arrays

For the application of isomap and MVU in the space of Hough arrays the distances between each pair of Hough arrays had to be calculated. Given two $(m \times n)$-dimensional arrays $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ and $B = (b_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ their Euclidean distance can be calculated using

$$d_2(A, B) = \sqrt{\sum_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} (a_{ij} - b_{ij})^2}. \qquad (4)$$

Alternatively a distance based on the Bhattacharyya distance measure [6], [27] can be applied after normalisation of the arrays as follows

$$d_{Bhat}(A, B) = 1 - \sum_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \sqrt{a_{ij}} \sqrt{b_{ij}}. \qquad (5)$$

Point-wise calculation of the distances for sparse arrays would not lead to a meaningful result. Therefore the discrete set of point values was smoothened by multiplying each point in the array with a Gaussian function. This lead to a continuous surface as shown in Figure 2 which was used as basis for the calculation of the distance matrices.

V. Experimental Results and Discussion

The experiments with the cylinder data in Section II-C and Figure 1 demonstrated that it is possible to generate topologically and geometrically stable manifolds if the neigbourhood parameter k and the dimension of the embedding are well selected. In the cylinder example a choice of $4 \leq k \leq 8$ and an embedding into $\mathbf{R}^3$ produced satisfying results.

For the streetmanifold, although the resulting shape was not known in advance, the outcome with respect to stability was similar: The resulting manifolds did not change much if the neighbourhood parameter was selected within the range $5 \leq k \leq 10$. Therefore all final results were calculated with $k = 5$.

Four streetmanifolds were generated. Two of them used a Euclidean distance matrix and two of them a distance matrix based on the Bhattacharyya similarity measure. Each of the distance matrices was processed with isomap and MVU. The two resulting streetmanifolds obtained with isomap have been included in this paper and are displayed in Figure 5. The two MVU streetmanifolds were very similar to the corresponding isomap results and were not included.

The manifolds were best embedded in 3-dimensions and therefore a colour/greyscale gradient for coding the 3rd dimension was applied to the images of the two manifolds in Figure 5. The residual variance [43] of the Bhattacharyya measure based manifold turned out to be lower than that of the Euclidean distance based manifold (cf. Figure 3) and allowed an embedding into two dimensions in our example.

For our analysis of the streetmanifolds six categories of similar houses were identified in the Bhattacharyya measure based manifold in Figure 5. For display we selected four representative houses from each of the six categories and included them in Figure 4. A comparison with the streetmanifolds in Figure 5 revealed that the four houses from each category appeared as clusterstyle groups in both streetmanifolds.

The houses of category A were relatively narrow, had a relatively high percentage of vertical lines and not many diagonal lines. The houses of category B were wide, had strong horizontal components but also many vertical lines and some diagonal lines. The A and B "clusters" were well separated and could be identified in both manifolds.

The E category was special because its houses were hidden behind trees. The distribution of horizontal and vertical lines tended to be homogeneous and did not have very strong lines. The associated group of points in the manifold could be characterised as "undecided" and was located at a close to central position.

Categories C, D, and F were very similar. In D the distribution of horizontal lines was more homogeneous than in C

and F and it included strong lines. F seemed to have slightly more vertical and diagonal lines than C.

D and E could be regarded as transitional stages between A and B. Similarly F seemed to play the role of a transition between C and D which becomes more plausible if the 3rd dimension's colour/grayscale map is taken into account.

The results suggest that the streetmanifolds have captured and smoothly organised a variety of line based features of the data set. Isomap and MVU produced very similar outcomes. The manifolds using the Euclidean distance had a different geometry from the Bhattacharyya measure based manifolds. However, relative location of the selected test groups of houses were closely related. The difference of the geometries of the two manifolds is at least partially a consequence of the array normalisation which was conducted in connection with the calculation of the Bhattacharyya measure based distance matrix.

Future development of the presented approach and pilot experiments may involve other features such as texture, line segments and colour in addition to the lines. Image preprocessing may be applied such as cropping cars, trees, clouds, and compensation for shadows. There are a number of variations for the Hough transform and different parameter settings which could be evaluated in further detail. These may lead to slightly different results and could be used to emphasize or augment particular features such as the size of the houses, fine texture, and how similarity between lines is displayed. In the present study the origin of the Hough transform was placed at the midpoint of each image. However, it was proposed that for architectural analysis it may be better to put the origin centered at eye-height at the entrance door of each house.

The principles of visual neuroscience and Gestalt psychology discussed in the introduction can be used as basis for an interpretation of the process of generating the streetmanifolds: The extraction of local gradients in the digital image and identification of lines from incomplete local information (Hough transform for lines), can be interpreted as a form of hierarchical information processing, which achieves what Gestalt psychology would associate with principles of holism and continuity. The accumulation and nearby location of similar lines in the Hough arrays, and the distance based comparison of Hough arrays, may be interpreted as instances of functional specialisation (orientation selectivity) and would be associated with Gestalt principles such as similarity and coherence. Therefore the establishment of the streetmanifold as abstract, continuous, holistic representation of the streetscapes' visual information (the component based on lines), combines both of the two main principles of human visual information processing and associated Gestalt principles.

VI. Conclusion

Isomap and MVU were applied to architectural image data which is a novel application of non-linear dimensionality reduction [9]. Parallels of the procedure with principles in vision science were discussed. The experiments included comparative evaluations of MVU, isomap, different distance measures, and
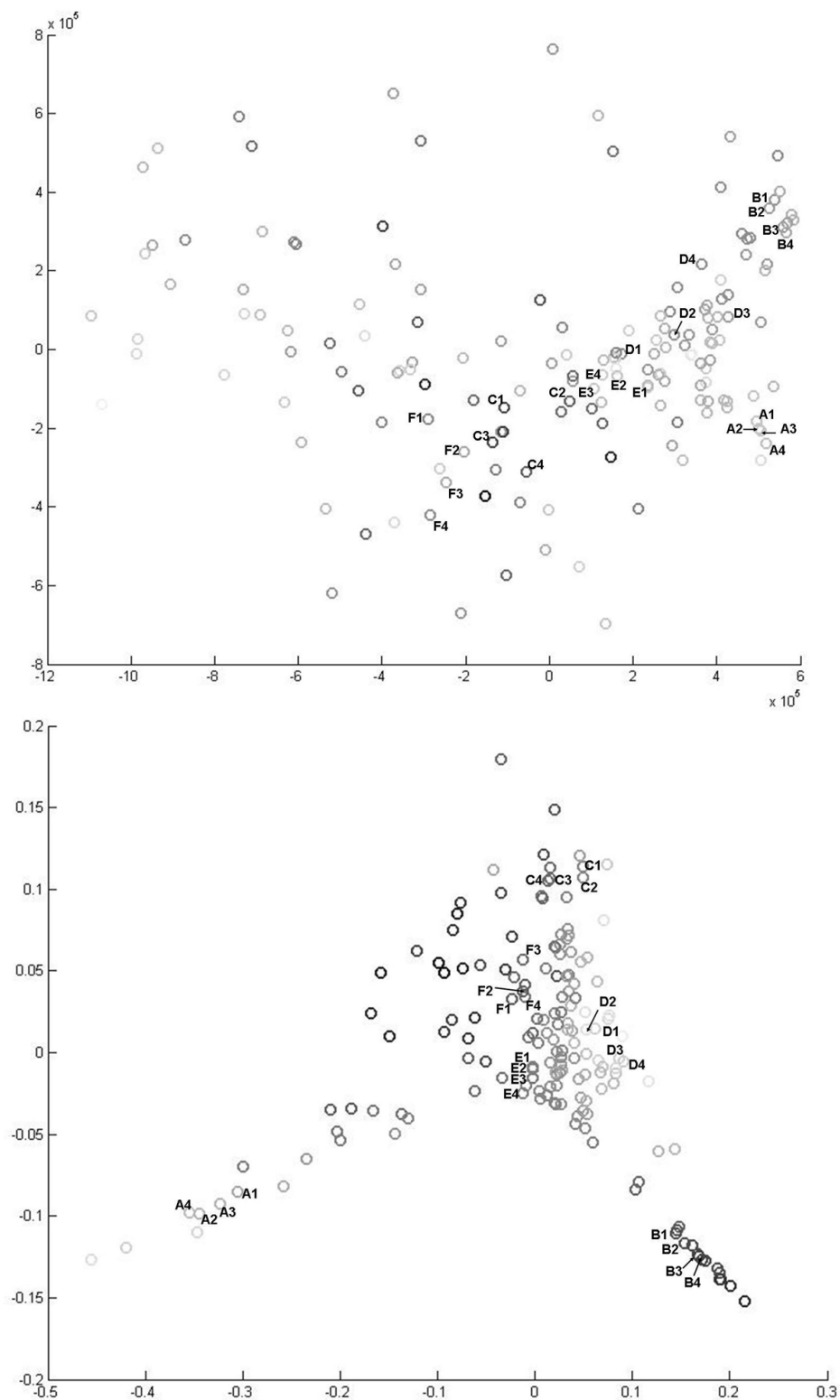
Fig. 5.   Isomap streetmanifolds based on Euclidean (top) and Bhattacharyya (bottom) distance calculation.

variations of the critical neighbourhood parameter k. The main contribution was a proof of concept of the streetmanifold. Its analysis unveiled a rich structure of line related visual features. The discussion of the present article addressed some basic aspects of this structure which appeared to be meaningfully represented in the streetmanifold. The geometric stability of the cylinder example (Figure 1) contributed to this conclusion.

REFERENCES

[1] D. H. Ballard, "Generalising the Hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.
[2] H. B. Barlow, R. M. Hill, and W. R. Levick, "Retinal ganglion cells responding selectively to direction and speed of image motion in rabbit," *Journal of Physiology (Lond.)*, vol. 173, pp. 377–407, 1964.
[3] H. B. Barlow and W. R. Levick, "The mechanism of directionally selective untis in rabbit's retina," *Journal of Physiology (Lond.)*, vol. 178, pp. 477–504, 1965.
[4] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.
[5] D. Beymer and T. Poggio, "Image representation for visual learning," *Science*, vol. 272, p. 1905, 1996.
[6] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bulletin of the Calcutta Mathematical Society*, vol. 35, pp. 99–109, 1943.
[7] C. J. C. Burges, *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Researchers and Practitioners*. Kluwer Academic Publishers, 2005, ch. Geometric Methods for Feature Extraction and Dimensional Reduction.
[8] M. J. Carreira, M. Mirmehdi, B. T. Thomas, and M. Penas, "Perceptual primitives from an extended 4d Hough transform," *Image and Vision Computing*, vol. 20, pp. 969–980, 2002.
[9] S. K. Chalup, R. Clement, M. J. Ostwald, and C. Tucker, "Applications of manifold learning in architectural façade and streetscape analysis," *Workshop on Novel Applications of Dimensionality Reduction at NIPS, Whistler CN 2006*, 2006, extended abstract.
[10] L. M. Chalupa and J. S. Werner, Eds., *The Visual Neurosciences*. The MIT Press, 2004.
[11] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. MIT Press and McGraw-Hill, 2001.
[12] T. F. Cox and M. A. A. Cox, *Multidimensional Scaling*, 2nd ed. Chapman & Hall/CRC, 2001.
[13] R. Cucchiara and F. Filicori, "The vector-gradient Hough transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 7, pp. 746–750, 1998.
[14] R. L. De Valois and K. K. De Valois, *Spatial Vision*. Oxford University Press, 1988.
[15] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.
[16] Y. Elsheshtawy, "Urban complexity: Toward the measurements of the physical complexity of street-scapes," *Journal of Architectural and Planning Research*, vol. 14, no. 4, pp. 301–316, 1997.
[17] S. A. Engel, "Adaptation of oriented and unoriented color-selective neurons in human visual areas," *Neuron*, vol. 45, pp. 613–623, 2005.
[18] D. A. Forsyth and J. Ponce, *Computer Vision A Modern Approach*. Prentice Hall, 2002.
[19] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Prentice Hall, 2002.
[20] K. Grill-Spector and R. Malach, "The human visual cortex," *Annual Reviews Neuroscience*, vol. 27, pp. 649–677, 2004.
[21] Y. L. J. Ham and D. D. Lee, "Learning nonlinear appearance manifolds for robot localization," in *IEEE/ISJ International Conference on Intelligent Robots and Systems*, 2005, pp. 1239–1244.
[22] P. V. C. Hough, "Methods and means for recognizing complex patterns," U.S. Patent 3,069,654, 1962.

[23] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex," *Journal of Physiology (London)*, vol. 160, pp. 106–154, 1962.
[24] ——, "Receptive fields and functional architecture of monkey striate cortex," *Journal of Physiology (London)*, vol. 195, pp. 215–243, 1968.
[25] J. Illingworth and J. Kittler, "A survey of the Hough transform," *Comput. Vision, Graphics, Image Process*, vol. 44, no. 1, pp. 87–116, 1988.
[26] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 1986.
[27] T. Kailath, "The divergence and bhattacharyya distance measures in signal selection," *IEEE Transactions on Communication Technology*, vol. 15, no. 1, pp. 52–60, February 1967.
[28] G. Kanizsa, *Organization in Vision: Essays on Gestalt Perception*. New York: Praeger, 1979.
[29] K. Koffka, *Principles of Gestalt Psychology*. New York: Harcourt Brace, 1935.
[30] Z. Kourtzi, A. S. Tolias, C. F. Altmann, M. Augath, and N. K. Logothetis, "Integration of local features into global shapes - monkey and human fMRI studies," *Neuron*, vol. 37, no. 2, pp. 333–346, January 2003.
[31] J. Larsson, M. S. Landy, and D. J. Heeger, "Orientation-selective adaptation to first- and second-order patterns in human visual cortex," *Journal of Neurophysiology*, vol. 95, no. 2, pp. 862–881, October 2005.
[32] M. S. Livingstone and D. Hubel, "Segregation of form, colour, movement, and depth: Anatomy, physiology, and perception," *Science*, vol. 240, pp. 740–749, 1988.
[33] R. H. Masland, "Direction selectivity in retinal ganglion cells," in *The Visual Neurosciences*, L. M. Chalupa and J. S. Werner, Eds. The MIT Press, 2004, ch. 28, pp. 451–462.
[34] S. E. Palmer and I. Rock, "Rethinking perceptual organization: the role of uniform connectedness," *Psychonomic Bulletin and Review*, vol. 1, pp. 29–55, 1994.
[35] R. A. Peters and O. C. Jenkins, "Uncovering manifold structures in robonauts sensory-data state space," in *IEEE International Conference on Humanoid Robotics, Tsukuba, Japan*, December 2005.
[36] E. T. Rolls and G. Deco, *Computational Neuroscience of Vision*. Oxford University Press, 2002.
[37] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
[38] H. S. Seung and D. D. Lee, "The manifold ways of perception," *Science*, vol. 290, no. 5500, pp. 2268–2269, 2000.
[39] L. G. Shapiro and G. C. Stockman, *Computer Vision*. Prentice Hall, 2001.
[40] L. C. Sincich and J. C. Horton, "The circuitry of V1 and V2: Integration of color, form, and motion," *Annual Review of Neuroscience*, vol. 28, pp. 303–326, 2005.
[41] L. Spillmann and W. H. Ehrenstein, "Gestalt factors in the visual neurosciences," in *The Visual Neurosciences*, L. M. Chalupa and J. S. Werner, Eds. The MIT Press, 2004, vol. 2, ch. 106, pp. 1573–1589.
[42] M. Spivac, *A Comprehensive Introduction to Differential Geometry*, 2nd ed. Publish or Perish, Inc., 1979.
[43] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
[44] L. Vandenberghe and S. P. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.
[45] K. Q. Weinberger and L. K. Saul, "An introduction to nonlinear dimensionality reduction by maximum variance unfolding," in *Proceedings of the National Conference on Artificial Intelligence (AAAI)*. Nectar paper, Boston MA, 2006.
[46] ——, "Unsupervised learning of image manifolds by semidefinite programming," *International Journal of Computer Vision*, vol. 70(1), pp. 77–90, 2006.
[47] M. Wertheimer, "Untersuchungen zur Lehre von der Gestalt II," *Psychologische Forschung*, vol. 4, pp. 301–350, 1923.
[48] G. Westheimer, "Gestalt theory reconfigured: Max Wertheimer's anticipation of recent developments in visual neuroscience," *Perception*, vol. 28, pp. 5–15, 1999.
[49] Z. Zhang and H. Zha, "Principal manifolds and nonlinear dimensionality reduction by local tangent space alignment," *SIAM Journal of Scientific Computing*, vol. 26, no. 1, pp. 313–338, 2004.