Visual Style: Qualitative and Context Dependent Categorisation

Authors: Julie Jupp and John S. Gero
Postal Address: Key Centre of Design Computing and Cognition, School of Architecture, Design Science and Planning, University of Sydney NSW 2006 Australia
Courier Address: Room 290 Wilkinson Building (GO4) 148 City Road University of Sydney Chippendale NSW Australia
Phone: +61-2-9351 2328 Fax: +61-2-9351 3031 Email: jupp_j / john @arch.usyd.edu.au

Number of Pages: 26
Number of Tables: 11
Number of Figures: 10
Full Title

Visual Style: Qualitative and Context Dependent Categorisation

Abstract

Style is an ordering principle by which to structure artefacts in a design domain. The application of a visual order entails some explicit grouping property which is both cognitively plausible and contextually dependent. Central to cognitive-contextual notions are the type of representation used in analysis and the flexibility to allow high-level semantic interpretation. This paper presents a model of visual style as a qualitative context-dependent categorisation, based on semantic feature extraction and Self-Organising Maps, called Q-SOM. The model proposes a method of categorising 2D un-annotated digital design diagrams using both low-level geometric and high-level semantic features automatically derived from the pictorial content of the design. The operation of the Q-SOM model can be seen as a series of sequential processing stages, where qualitative encoding and feature extraction are followed by iterative re-categorisation. Categorisation is achieved using an unsupervised SOM and contextual-dependencies are integrated via cluster relevance, determined by the observer’s feedback. The following stages are presented: (i) the initial per feature detection and extraction, (ii) the selection of feature sets corresponding to different spatial ontologies, (iii) the unsupervised categorisation of design diagrams based on appropriate feature subsets, and (iv) the integration of design context via relevance feedback. We investigate a comparison of different outcomes from each consecutive stage of the model. The results of the experiments show that the model provides a cognitively plausible and context-dependent method for a digital characterization of style.

Keywords

Qualitative spatial representation, similarity, context dependency.
1.0 Introduction

In the design domain, an analysis of visual similarity leads to the concept of style. The term style is polysemous and in design literature refers to different ideas concerning the artefact, modality, society, culture, period, etc. In developing a model of visual style, we approach the problem from an ‘object’ viewpoint, where an artefact’s common characteristics are primary (Ackerman 1967, Edwards 1945). A useful and well accepted definition of this view expresses style as a principle by which to provide order, allowing a set of artefacts to be structured according to some set of criteria (Knight 1994).

An object view of style essentially relies on assessment criteria. Assessment criteria can be specified on a number of different dimensions. Typically these dimensions are defined by structural (i.e., spatio-visual) attributes, represented either quantitatively or qualitatively (Cha & Gero 1998, Ding and Gero 1998, Gero and Park 1997, Knight 1994). Utilising qualitative descriptions of structure, semantic and conceptual aspects have also been derived in some of these models (Park and Gero 1999, Ding and Gero 1998), yet due to the complexity of deriving relevant domain semantics, this approach is less prominent in models of visual style. These models are typically limited by their inability to handle important cognitive properties of similarity assessments including high-level semantic concepts influenced by the design context. The perception of visual features is influenced by the observer’s context and their interpretation also depends on their relevance to the current design task. Thus, judging the similarity of design artefacts is dependent on context where properties of similarity assessments such as asymmetry (Tversky 1977) can influence categorisation.

Many computational models of style have inherited an emphasis on quantitative and /or linear analysis, focusing on distance measures that maintain a static world assumption where style is treated as unrelated to its locus of application. In exploring these deficiencies we ask: How can high-level semantic and conceptual dimensions be derived efficiently from structural ones and how may we compare them to provide a cognitively plausible yet contextually dependent model of visual style?

To address this two-part question we look to qualitative modelling and clustering techniques. Typically spatial information directly available to human observers is qualitative in nature (Cohn 1997). By incorporating qualitative encoding methods that automatically re-represent a design, high-level semantics can be derived. The benefits of utilising qualitative encoding lie in the description of attributes which are significant to the preservation of salient design qualities. Encoding a design corpus qualitatively provides meaningful data sets that can be utilised by any clustering approach.

Since our objective is to model style as an intuitive yet relevant clustering method (relative to the observer and the design task), the basic problem for such an approach is the gap between the high-level and context-dependent semantics used by designers to understand design content and the low-level visual features extracted from the design artefact. In providing a possible solution to this gap, three important research issues relate to:

1) low-level qualitative re-representation of design artefacts and high-level semantic mappings,
2) features selection methods used prior to categorisation, and
3) integration of context-dependent comparisons so as to provide user-relevant categories.

The model of style presented here addresses each of these issues and utilises the application area of 2D architectural diagrams to test the models performance. Using a qualitative representation schema for a hierarchy of spatial features and a measure of similarity, we provide a method which automatically structures a design corpus according to selected feature semantics as an iterative process, adapting to the observer’s requirements and preferences. Adaptation is based on the relevance of the distinguished categories to the design task. This approach is commonly used in text classification and retrieval systems and is known as relevance feedback (Salton and McGill 1983). Applied here, the model of visual style is capable of categorising design
diagrams in a cognitively compatible manner and incorporates the observer’s feedback to refine subsequent design categorisations which may better satisfy some design requirement. This flexible, open-ended approach to similarity and categorisation has the ability to process qualitative representations and learn from the data set unsupervised as well as modify itself to reflect an observer’s intuitive and contextually dependent assessment.

It is claimed that through the creation of a model of qualitative context-dependent categorisation, that the semantics derived and utilised by the learning system in conjunction with an observer’s feedback can provide a cognitively plausible and relevant characterisation of visual design style.

2.0 Style, Similarity and Spatio-Visual Reasoning

In all visual domains, style is commonly used to describe consistencies among artefacts that are the product of an individual, culture, period, or region. In this way, the concept of style in design provides order within an otherwise apparently chaotic domain (Knight, 1994). Identifying a visual style is a judgment process that requires an artefact to be decomposed into elements in which they are the same and elements in which they are different. Design artefacts can be described as belonging to the same style to the degree that they have a particular dimension in common and are not differentiated by any distinctive one.

A variety of models of style and the related concept of similarity have been developed (Do and Gross 1995, Gero and Park 2000, Davis and Goel 2001, Gero and Kazakov 2001, Forbus et al. 2003, Burns 2004). Generally these models are based on a re-representation of either 2D or 3D design artefacts and some function of similarity that allows those artefacts to be compared and ordered. Many of these models have been developed as design support systems to aid in the perception of Gestalts, as well as decision making and analogy. Most have directly applied or adapted similarity functions from other fields of research, such as psychology and cognitive science as well as information analysis and retrieval.

Many models of similarity based style are limited since their analysis ultimately depends on quantifying common elements independent of their context. It is necessary to reformulate the approach to measuring design content by moving away from the idea of style as the outcome of a direct comparison and move towards the idea that it is a process whose outcome can only be reported in a post-hoc fashion.

Like the previous approaches cited above, we use the concept of similarity as the grouping principle (Wertheimer 1977) by which 2D design diagrams can be ordered, however we assume that it is related to the way information is processed and that the reporting of similarity judgements is a meta-cognitive process (Thomas and Mareschal 1997) requiring explicit comparison of design information both prior and subsequent to processing by the system.

2.1 Detecting Design Similarity

Human observers are able to recognize, interpret and search for salient features in diagrams in order to detect visual similarities. The last 40 years of research surrounding the concept of similarity has provided a variety of insights on both theoretical (Love and Sloman 1995) and empirical levels (Tversky and Gati 1982). In design, there is still a significant lack of understanding how observers classify design artefacts, form concepts and make decisions based on the similarity perceived between two or more designs. Tversky (1999) has shown in cognitive experiments that in reasoning about design diagrams, individuals are able to make comparisons across a variety of dimensions intuitively using abstraction, approximation, aggregation and other techniques to generate manageable spatial descriptions. Other recent cognitive studies in design research have shown that during designing, ad hoc visual sorting largely depends on intuitive similarity assessment (Jupp and Gero 2005).

However, such investigations are difficult since similarity assessments typically refer to comparisons among design artefacts using a variety of dimensions. Often when designers judge the similarity of two diagrams, the dimensions themselves or even the number of
dimensions are not known and what might appear intuitively to be a single dimension may be a complex of several (Jupp and Gero 2005).

A rich set of dimensions can be obtained as the basis for comparison by developing an information processing activity that involves constructing re-representations at successive levels of abstraction (Marr and Nishihara 1978). This type of information processing is typically intuitive and subjective, and is not a function of strict mathematical models (Tversky 1977). The importance of a qualitative approach is highlighted. Computational methods are needed that go beyond quantitative descriptions and which address the importance of design semantics.

2.2 The Role of Representation

Qualitative approaches to representation are a common analysis paradigm in design reasoning applications. There are multiple dimensions that could be modelled qualitatively, resulting in a variety of data sets. Here, two general criteria distinguished by Shapiro (1961) as being significant to the characterisation of style are considered: (i) shape elements and (ii) their spatial relationships. Using these criteria as our foundation, the objective of the qualitative schema is to define a rich set of structural attributes (for open and closed shapes) as well as relations of adjacency and connectedness (for shape aggregations).

A canonical description of a design corpus should also be capable of mapping on to a variety of design semantics. By encoding a diagram using multiple spatial ontologies, a variety of feature semantics can be obtained via pattern matching techniques to derive meaningful design concepts. This approach to representation plays an important role in simulating cognitively plausible comparisons since the use of common-sense descriptors supports the identification of designs that are not only structurally close but also conceptually close, whilst not being identical. This will be illustrated by establishing semantic mappings capable of identifying higher-level design concepts.

2.3 Similarity and Context

The criteria by which a set of design diagrams are distinguished as similar carry with them important contextual dependencies. Cognitive studies have demonstrated that similarity processing often depends on context (Medin et al. 1993) and an increasing consensus regards similarity as a dependent property that extends the focus of inquiry to include contextual aspects (Thomas and Mareschal 1997, Sloman 1996). This view is particularly relevant in design since the designer operates within a context and their perception and judgement of design similarities are influenced by it.

Contextual dependencies can include perceptual constraints, such as asymmetry resulting from the order of the diagrams being compared – through more specific task related requirements, such as the design brief – to broader contextual aspects such as society, culture, region and period. In considering these aspects of visual style, a similarity measure should not only be capable of incorporating design semantics but also of operating in relation to the design task and requirements. Under existing approaches to similarity assessment in design, the observer is unable to select those categorisations which are relevant to the current design task.

The approach to a measure of similarity presented here relies on a process that utilises initial unsupervised categorization which is then guided in successive iterations in a supervised fashion. The similarity measure treats the observer as an inseparable part of the categorisation process to provide categorisations that the observer is most likely to be interested in. Although a well established technique in text categorisation and retrieval systems, there appears to be no model of similarity in design that integrates relevance feedback in this way.

2.4 Self-Organising Maps (SOM)

By treating visual style as a multi-dimensional similarity measure it is possible to model this approach using an artificial neural network. The underlying mathematical properties of most neural networks used in categorisation are scalar distance-based algorithms. Of these, the Self-Organising Map (SOM) is a typical representative. As such SOMs can be used in a
variety of ways, with a number of different configurations being available (Kohonen, 1995). The main advantage of using SOMs in design comparison is that they do not require target values for their outputs and learning occurs unsupervised. Since there is no absolute definition of the commonalities between design artefacts in structural or semantic terms, there is no single definitive exemplar to establish reliable target outputs that can be used to train a supervised network. For this reason, SOMs are commonly used to find and construct classifiers, and hence provide a continuous topological mapping between the feature space and the 2D space. This is an important property of SOMs as they are able to represent a mapping, which preserves relations in the input space while at the same time performing a dimensionality reduction onto the 2D mesh of neural units in the competitive layer.

The competitive learning process in the SOM produces weight vectors that correspond to distinct clusters of the input vectors. The weight vectors can be considered to be the cluster centres of the probability density function of the input data. However, although extensively used in other fields of research on similarity and categorisation such as text and image retrieval, SOMs have not been widely utilized in design categorization systems. One application is a model for assessing the similarity of works of art proposed by Colagrossi et al. (2003). Colagrossi and his colleagues measured the similarity of Mondrian’s Neoplasticist paintings according to a selection of features. By consolidating algebraic functions a variety of parameters were processed with only a few neurons in both input and output of the SOM. Those parameters considered useful by the authors included line type, line weight and colour. Yet, the application of SOMs by Colagrossi et al. does not resolve those limitations identified above due to the lack of semantics, contextual relevance and most significantly, the highly restricted complexity of its application domain. The SOM implemented by Colagrossi et al. (2003) demonstrates relatively low-level complexity since a very distinct design corpus and only one type of design artefact is considered.

2.5 Towards a Model of Style

The visual style of a design is in a sense transient, where the perceived similarities can depend on the corpus in question, the amount of a priori information available, the order of design comparison and when considered in relation to a task specific context also depend on design objectives and requirements. The problem of modelling visual style using similarity assessment exhibits different levels of complexity. For example, a rather simple assessment problem occurs when the design corpus in question consists of diagrams from a strongly restricted set, like the one explored by Colagrossi et al. (2003). From the perspective of the application domain being tested here, this level of complexity translates to categorising only a single architect’s designs which encompass only residential plan diagrams. This level of complexity enables a relatively straightforward assessment based on automatic segmentation of the design diagrams. In the other extreme lies the problem of categorising a design corpus comprising a variety of:

(i) design domains, i.e., arts, architecture, industrial design, etc.,
(ii) design artefacts, taking the architectural domain as an example, this could refer to the kinds of building typologies such as residential, commercial, industrial, religious, etc.,
(iii) design representations, continuing with our example of the architectural domain this refers plans, sections, elevations, etc., and
(iv) design re-representation, i.e., the number and type of dimensions as well as the representational paradigm.

Similarity as a measure of visual style is not a fixed and irreducible concept, however most algorithms that compute similarity are. Similarity assessment should therefore result from a well-defined process, which takes place over some more-or-less open and variable dimensions of the designs being compared. Under this definition, similarity as visual style can be treated as context-dependent since the attributes that the process uses as input can change according to some set of criteria. This view gives less explanatory force to similarity because it demands analysis of the attributes of designs whose similarity it computes in relation to context.
3.0 Q-SOM: A Model of a Qualitative Context Dependent Style

The model of qualitative context-dependent categorisation is based on semantic feature extraction and uses Self-Organising Maps, called Q-SOM. The framework presented in Figure 1 is an iterative process which relies on the following consecutive stages:

- recognition, extraction and encoding of three different levels of spatial attributes,
- initial per feature selection of encoded spatial attributes and combination of feature lists,
- categorisation via unsupervised learning of design diagrams based on available features,
- positive and negative feedback processes via the observer’s input, and
- resulting weight adjustment and re-categorisation of design diagrams.

![Figure 1. Q-SOM, stages of diagram categorization.](image)

In the first stage of the framework in Figure 1, each diagram in the design corpus is described using a qualitative encoding schema capable of representing sets of higher-level semantics corresponding to three prescribed spatial ontologies. During the second stage, feature sets undergo a selection process as part of input pre-processing. A feature subset is produced using either principal component analysis or manual feature selection by the observer. The third stage utilises the feature subset as input to the SOM and categorisation occurs via unsupervised learning. How the distances in various feature spaces are weighted and combined to form a scalar suitable for minimisation, creates an opportunity to integrate contextual dependencies in the architecture of the SOM. The fourth stage continues as an interactive process that moves from unsupervised categorisation to one which is guided by the observer. The final stage re-categorises diagrams which are similar to the observer’s target diagram or meet some set of target criteria, by ordering those diagrams whose distance to the target is minimal in any or all feature sets.

3.1 Choice of Features

In utilising an SOM for categorisation, the homogeneity of a category is enforced by the appropriate choice of feature vectors. Because design diagrams are an explicit representation of the artefact’s geometry, it is reasonable to expect that categorisation be based on a criterion that incorporates properties of 2D geometry which satisfy Shapiro’s (1961) characteristics mentioned in Section 2.3 where properties should:

1. be generic so that they have applicability over a wide spectrum of application domains,
2. characterise as many dimensions of the diagram as possible including orientation, distance and topology,
3. provide rich descriptions to be of use to higher-level semantic mappings,
4. have local support, that is, they should be computable on local primitives around a landmark of significance,
5. be stable and invariant over large ranges of viewpoint and scale, and
6. be reliable, robust, and readily detectable by procedures that are computationally stable.

Figure 2 presents the process of qualitative re-representation which satisfies the requirements outlined above for orthogonal geometries only (Jupp and Gero 2004).

![Figure 2. Model of a qualitative abstraction (after Jupp and Gero 2004)](image)

The processes of encoding and abstraction illustrated in Figure 2 follow from physicality to symbol to regularity to feature, where:

- **Physicality** – refers to the graphic descriptions of diagrams indicating the geometric information and is the pre-representation, upon which a process of information reduction is applied successively over three levels of abstraction.
- **Symbol** – refers to the unrefined symbolic encoding of graphic information. Spatial attributes are recognised and converted into qualitative symbol values.
- **Regularity** – is the syntactic matching stage in which regular or repetitious patterns of encodings are identified and grouped. The key concept of detecting characteristics from the syntactic encoding lies in ‘chunking’ (Brown et al. 1995).
- **Feature** – involves matching pre-determined syntactic patterns with meaningful design semantics.

In the following section a summary of the hierarchical encoding method used in the model is presented. For a more detailed account of the encoding schemata refer to Jupp and Gero (2006).

### 3.1.1 Morphology

Sign values for specifying specific qualities of isolated shape structures are based on a description of attributes encoded at a landmark vertex (intersection) where properties for line contours are divided into two separate codes. The first is a primary code and represents the relative angle. The second is an auxiliary code and represents the relative distance. The formal definitions of primary and auxiliary codes are presented in Table 1 (Gero and Park 1997).

<table>
<thead>
<tr>
<th>Table 1. Definition of Qualitative Syntax for Morphology.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Numeric value range</strong></td>
</tr>
<tr>
<td>$0 \leq \theta \leq 2\pi$</td>
</tr>
<tr>
<td><strong>Landmark set</strong></td>
</tr>
<tr>
<td>$[0, \pi]$</td>
</tr>
<tr>
<td>$((0,\pi), (\pi,0))$</td>
</tr>
<tr>
<td><strong>Q-code set</strong></td>
</tr>
<tr>
<td>${L, \bar{A}}$</td>
</tr>
</tbody>
</table>
Where an angular change occurs, landmarks are initially set to \( \pi \), separating convex and concave angles. The scanning order for each vertex is set to a counter-clockwise direction and the magnitude of the vertex is also measured in this direction. The two primary codes \( L \) and \( \bar{A} \) represent a vertex so that individuals can be described qualitatively. The addition of codes capturing the relative length of contours provides a description capable of distinguishing between shapes without increasing the number of primitives unnecessarily.

Encoding results in a symbol string and a syntactic handling method that employs a simple pattern recognition technique is used to group structural information so as to identify semantic aspects of regularities. The descriptions of patterns are based on the recognition of meaningful structural features and patterns that reflect basic repetitions and convexity are detected including: indentation, protrusion, (Gero and Park 1997), iteration, alternation and symmetry (Martinoli et al. 1988).

Patterns of symbol sequences denoting specific categories of feature classes that are familiar in contour or identify some particular shape semantic are also identified. There are two classes of semantics defined which can be divided into simple and complex. Simple semantics include primary shape types can be recognised and classed into subsequent labels such as “rectangle”, “square”, “L” -shape, “U” -shape, “T” -shape, and cruciform. Complex semantics incorporate domain specific knowledge which is capable of mapping design concepts to features. For example, design concepts from the architectural domain such as chambers and niches can be identified.

Since the patterns are derived from low level structural primitives they are defined as local shape features, LSF. All LSF are recognised by matching symbols with an existing feature knowledge base.

3.1.2 Topology

At the contact of two or more shapes, there are extraction and embedding relationships for the intersection of line contours. Where this occurs there is a transformation in representation which extends the encoding of two line intersections to include multiple lines. To encode multiple line attributes, graph diagrams derived from the original contour representation are used as the means by which to parse information in a consistent manner. In this way, graphs provide a notion of hierarchy and support a bottom-up development.

The symbols used to describe edges of graphs concern the disposition of physical intersections of lines which have been used to generate the polygon fields that are subsequently analysed as graphs. Edges are labelled according to the intersection type of the two vertices belonging to the line contour it crosses creating a dyad symbol. Dyad symbols are collapsed and later augmented by values describing the relative area of the shapes they bound. The specification method provides a description of spatial attributes in terms of shape adjacencies and area descriptors. The formal definition of dyad symbols are presented in Table 2.

<table>
<thead>
<tr>
<th>TABLE 2. Definition of Qualitative Syntax for Topology.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ANGLE CODES</strong></td>
</tr>
<tr>
<td>Numeric value range</td>
</tr>
<tr>
<td>Landmark set</td>
</tr>
<tr>
<td>Interval set</td>
</tr>
<tr>
<td>Q-code set</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

The representation of dyad symbols reveals distinctive topological characteristics that are recognised from syntactic regularities. Unlike morphological features, topological ones contain variations based on a reference frame. Using the dyad symbols in conjunction with a reference point, three types of adjacency semantics can be defined including: complete adjacency, partial adjacency and offset.

The regularities identified in dyad symbols produce feature semantics that are deemed
intermediary shape features, ISF, since the ‘neighbourhood’ of the description is based on local attributes as well as information describing topological properties. Like LSF, ISF are identified by matching an existing feature knowledge base.

### 3.1.3 Mereotopology

The dual of the graph diagram is used to derive composite symbol values in order to describe part-whole relations. Abstracting the initial graph to its corresponding dual graph ensures that unambiguous mappings can be derived. Once all mappings have been established, the dual is used to derive feature semantics. This results in transformations that are much clearer and easier to understand while still based, by virtue of the mapping, on the original 2D representation. The dual graph allows further abstraction and derivation of another level of spatial relationships.

By labelling the new dual-edges, tuples composed of dyad symbols (defined at the previous level of topology) are created. For each edge of the dual graph, labels are derived from the symbol values identified at the previous level. Labelled dual edges allow regularities to be identified and feature semantics describing part-to-whole relationships between two or more shapes are identified. Since dual graphs are undirected, regularities are identified from within the tuple itself and not from a string. Formal definitions of dyads are presented in Table 3.

<table>
<thead>
<tr>
<th>ADJACENCY CODES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numeric value range</td>
</tr>
<tr>
<td>Landmark set</td>
</tr>
<tr>
<td>Interval set</td>
</tr>
<tr>
<td>Q-code set</td>
</tr>
</tbody>
</table>

The semantic features identified at this level account more thoroughly for mereotopology. The regularities defined here are similar to Allen’s thirteen interval relations for the temporal domain (Allen, 1983). Mereotopological feature semantics include: meets/ met-by, overlaps/overlapped-by, starts/finishes, contains/ contained-by, equals and during. Allen’s interval calculus has previously been extended to other visual domains (Güsgen 1989, Mukerjee 1989) unlike previous approaches, here it is not restricted to rectangles and although is strictly based on orthogonal shapes, is still capable of handling arbitrary multi-sided forms.

Like the structural features identified for morphology and topology, spatial semantics derived from visual patterns are identified and domain knowledge is integrated using design concepts that map onto spatial features. Continuing with the example application of architectural domain knowledge, spatial concepts relating to the use of behaviour of a space such as corridor, quadrangle, and courtyard are mapped to patterns detected in each tuple.

Since all features are derived from higher level spatial primitives they are defined as global shape features, GSF. The three level schema summarised here is characterised by the class aspect of handling and labelling design concepts and is useful when dealing with different design categorization scenarios. The concession of the approach is that it is essential to have a large knowledge base of concept-to-feature mappings. Despite this requirement it is still an efficient and robust representation method.

### 3.2 Example Encoding of a Design Diagram

Figure 3 shows an example of an encoded design diagram labelled according to the qualitative encoding schemata. The example chosen is a simple 2D residential plan drawing of the Farnsworth House, designed and completed by Mies van der Rohe.

Figure 3 illustrates the four stages of encoding, where three consecutive stages of abstraction are followed by the initial re-representation stage where the diagram is transformed in to vectorial format. The mapping from physicality $\rightarrow$ symbol $\rightarrow$ regularity $\rightarrow$ feature involves detecting regularities and matching features from the geometric information so that specific patterns correspond to topologies of known feature semantics (Brown et al 1995).
3.3 Categorisation Using SOMs

The SOM implemented in the model follows a process where each neuron is assigned a pattern to which it is sensitive. Appearance of the same or a similar pattern on the input results in a high activation of that neuron – similarity being considered therefore as the opposite of distance. The SOM tries to place or adapt neurons in such a way that they serve as good prototypes of the input data for which they are sensitive (Kohonen 1995).

The architecture of a typical SOM consists of two layers, a layer of input nodes and a competitive layer consisting of neural units or Kohonen’s units (Kohonen 1982). A weight vector is associated with each connection from the input layer to a neural unit. The neural units in the competitive (and cooperative) layer are organized in a regular geometric and the units are interconnected with their local neighbours.

Using a winner-take-all network, the input vector is broadcast in parallel to all neurons and for each input vector and the most responsive neuron is located. The weights of this neuron and those within a neighbourhood around it are adapted to reduce the distance between its weight vector and the current input vector. The output of a SOM takes the form of a two-dimensional array of output nodes.

The competitive phase of the learning algorithm employed in the SOM determines a winning neural unit whereas the cooperative phase of the learning algorithm updates the weights of the winner and the neural units in its neighbourhood. The SOM is able to learn to recognise different patterns in the input data and allocate them to appropriate ‘bins’ (nodes) in the output array, each bin representing a specific pattern. The output can be seen as an array of ‘classification bins’ (each representing a specific pattern in the input data) that are arranged in an ordered way such that near neighbours represent similar patterns and distant neighbours represent different patterns.

Since there does not and will not ever exist one single “correct” answer to the central issue of a definition of visual design style, the ability to combine the distances calculated in different feature spaces provides the critical point where relevance feedback can be incorporated. The SOM’s matching process can therefore also be driven by contextual considerations, the observer is able to determine the relative importance of distinguishing features by adjusting their weights. When contextual information is used for determining the

Figure 3. Model of a qualitative abstraction and physical to symbol to regularity to feature mapping for the Farnsworth Houses designed by Mies van der Rohe
importance of distinguishing features the correlation increases.

### 3.4 Relevance Feedback

The relevance feedback or RF method integrated here is the iterative refinement of the initial SOM categorisation. RF is provided using dynamic weight adjustments that allow the SOM to learn the optimal correspondence between the high-level concepts that the observer uses and the feature semantics automatically derived from 2D diagrams.

In text and image-based research, RF is an established approach that enables contextual-dependencies to be integrated for document and image retrieval. Recently this approach has been adopted by researchers using SOMs to retrieve information from large databases. The WebSOM (Honkela et al 1998) and PicSOM (Oja et al 1999) systems have implemented RF by adjusting the weights of different textual terms when matching text with the documents or images of a database.

In an analogous manner, we have aimed at integrating RF with a SOM in the design domain, by treating this process as a form of learning that moves from unsupervised learning to being partially supervised. The model tries to learn the observer’s visual preferences to subject and style by adjusting the feature weights accordingly. Feature weights in subsequent categorisations are adjusted using the information gathered from the observer’s feedback. The observer’s feedback guides the system in the following rounds of the assessment process to better approximate their present design requirements.

The task of assigning specific weights which coincide with the observer’s perception of each feature set is not feasible and therefore the initial results from the unsupervised clustering are displayed using the topographic map so that weights can be derived from user input. It is crucial that the results from the initial round are categorised in a plausible manner such that a level of visual similarity is evident to the observer – this being the primary objective of integrating qualitative encoding. The observer is not required to explicitly specify weights for different features and instead weights can be formed implicitly from the positive and negative values assigned to a diagram or cluster of diagrams. This process can be summarised whereby:

- an unsupervised SOM categorises a design corpus;
- the first round of results are displayed and stored to avoid the system entering a loop;
- the observer indicates which diagrams are to some extent relevant to the present design context and which are not and assigns positive and negative values accordingly;
- the adjusted weights are utilised in a re-initialised SOM and the design corpus is re-categorised;
- the second and any subsequent round of results are displayed to the user and stored; and
- the process continues until the observer is satisfied.

### 4.0 Implementation of Q-SOM

Q-SOM is implemented as five separate modular components: (i) automated encoding and feature extraction (using DeREP and AbREP operations – see Section 4.1), (ii) pre-processing and creation of feature subsets (iii) unsupervised SOM categorization (iv) observer feedback via positive and negative responses, (v) weight adjustment and updating to create new SOM for re-categorisation. The following sections describe each of the five stages.

#### 4.1 DeREP and AbREP

DeREP divides the problem into smaller more tightly constrained sub-problems by partitioning shapes into vertices and contours. To achieve this, the process eliminates the primary source of complexity by separating unrelated variables into distinct shapes. This process results in a compact and easily understandable description of the structure of the diagram.

The sequence of vertex labelling occurs as: contour traversal → vertex detection → value
The problem of computing all possible circuits in the diagram so that each circuit contains all vertices exactly once is achieved through finding all Hamiltonian circuits (Garey and Johnson 1983). A contour cycle (i.e., closed loop) algorithm is implemented where the agent starts the cycle from each point in the diagram and visits each adjacent vertex exactly once until a closed shape are generated or until a maximum branch limit is reached. This process iterates until all possible shapes starting from the selected point are found. Once all closed shapes are found starting from all points in the diagram, the final set is filtered to eliminate shapes containing other shapes so that the resulting set contains only the smallest shape units.

The perimeter shape is the sum of all of the smallest shape units. As geometric properties are scanned vertex by vertex, the angle and length magnitudes of the previous line segment become the landmark point for the following segment. Landmarks and intervals are set each time a new contour is compared. A landmark is set to the numeric value of the magnitude of the previous segment length and a ratio is provided to distinguish the relative difference. Figure 4 shows a sample diagram and the resulting closed shapes detected.

The pattern recognition system analyses, locates and registers specified sequences, or chunks, of syntactic structures. A systematic search for every possible pattern is necessary for the given shape or spatial description. Features already stored in a knowledge base identify syntactic patterns where the search and matching process examines the type and occurrence of ‘chunks’ and their structure as a sequence is then labelled.

The DeREP process is followed by AbREP which automates the encoding of all subsequent (graph) abstractions. The AbREP uses the array of symbols describing intersections of line contours. The systematic processing of topological and mereotopological attributes requires the mapping from physicality to symbol to follow a similar conversion process from the graphic state to the symbolic state as implemented in DeREP, i.e., contour traversal → vertex detection → value assignment → contour traversal...etc. Labelling also relies on the data structures built from the previous stage of encoding.

An important aspect of AbREP is that sets of arrays describing shapes are analysed based on the relationship between each shape and the rest of the shapes in the set, Figure 5(a). By abstracting shapes using a description of vertex arrays, AbREP produces an abstract landmark and the two vertices are labelled according to specific characteristics of its endpoints (intersections). This is achieved by iterating over the set (taking two vertices belonging to one shape at a time) and traversing every pair of vertices of the shape. The arrays are checked so that if the vertex belongs to only one closed shape it is ignored. If the vertex is shared by two or more shapes then the vertex label is determined based on the number and direction of lines that compose the vertex as shown in Figure 5(b). The new set of codes replaces the representation of the shape. This is a reduction process whereby some
structural information about the shape is lost. Once the new arrays are derived and labelled, a representation of the shape’s adjacency is captured. At this level, arrays constitute an approximate representation of the topology of shapes.

The next stage of AbREP is a pattern matching process. The mapping from symbol to regularity to feature involves detecting regularities and matching features from the spatial relations so that sets of dyad symbols correspond to specific adjacency relations. At the implementation level the pattern recognition system locates and registers specified dyads stored in a knowledge base by a set of conditional or ‘if…else’ rules based on the columns of Table 4(a). A dyad symbol value is assigned to each set deriving a new label.

**TABLE 4. Examples of specified dyads and types based on conditional or ‘if else’ rules.**

<table>
<thead>
<tr>
<th>Set</th>
<th>(conditional or if else rules)</th>
<th>Dyad symbol</th>
<th>Tuple symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>([T,C])</td>
<td>(TT \land TT^*) (\land C)</td>
<td>(A_c)</td>
<td>(M_4)</td>
</tr>
<tr>
<td>([T,\bot, C])</td>
<td>(TT \land CC^*) (\land C)</td>
<td>(A_o)</td>
<td>(M_0)</td>
</tr>
<tr>
<td>([L, A, T, \bot, C])</td>
<td>(L\land \bot) (\land A) (\land TT)</td>
<td>(A_o)</td>
<td>(M_{1,2,3})</td>
</tr>
<tr>
<td>([L, A, T, \bot, C])</td>
<td>(L\land \bot) ((\neq, &gt;)) (\land A) (\land TT)</td>
<td>(A_o)</td>
<td>(M_{6,7})</td>
</tr>
</tbody>
</table>

Once the new set is formed the area of a shape is calculated and compared to the adjacent shape to obtain a description of the relative area. A list of adjacency and area features for each shape is created, and the combined symbols are themselves used to produce another set of codes, replacing the previous representation of adjacencies, i.e., the feature mapping at the level of topology results in a new set of codes. Again, specified features stored in a knowledge base are located and registered by a set of conditional or ‘if…else’ rules based on the columns of Table 4(b). Tuples are produced as a result of this process and the system iterates through the set and inspects both the type of all produced. Order outside the tuple itself is not considered and regularities in repetition are no longer determined via ‘chunking’ since each symbol value contains significant characteristics.

Using these specifications, qualitative descriptions of spatial relations are treated as the problem of representing distinctive characteristics of connectedness at the categorical level allowing spatial attributes to be detected and described qualitatively.

### 4.2 Pre-Processing Feature Sets

Each feature representation identified using the DeREP and AbREP operations can be used to create a meaningful subset of features. The Q-SOM model has two approaches to pre-processing of feature sets, where dimensionality reduction can be undertaken manually by the user or by using a statistical approach.

In manual selection of feature sets, subsets can be created directly by selecting those features of interest to categorisation. These may also be based on the features sets of a target diagram (if known). For example, an observer may wish to identify the visual style of design
precedents based on certain topological relationships, such as having complete adjacency, and in conjunction with certain morphological constraints such as a bounding cruciform shape and having all internal shapes defined squares.

Using the statistical approach, feature subsets can be created automatically using Correlation-based Feature Selection (CFS) (Hall 2000). CFS provides a filter based feature selection algorithm that uses correlation among features to select the best features for the given dataset. CFS evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy. Here, the CFS algorithm is used in conjunction with a Best First search method.

4.3 Unsupervised Categorisation

To interpret, categorise and visualize the multi-dimensional data sets obtained from the previous stages, a Self-Organizing Map (SOM) is implemented. In the third stage this is undertaken by the SOM using unsupervised learning. Initially, categorisation begins with a corpus of reference diagrams. This can be expressed by the following.

Let the design corpus, \( C \) containing \( K \) number of diagrams, \( d \), to be categorised be equal to:

\[
C = \{d'_1, d'_2, ... d'_i\} \subset c'
\]  

(1)

where \( c' \) denotes the initial categories found by the network.

By denoting each design diagram as \( d_n \), \( n = 1, 2, ... K \), we define a vector \( f_i \) consisting of features, is associated with each unit \( i \). If we have \( j \) different feature representations for each diagram, they can be written as:

\[
f_i^{(m)} = f_i^{(n)}, m = 1, 2, ...j
\]

(2)

The map consists of a regular (“city-block”) grid of neurons and unsupervised learning and categorisation to obtain \( c' \) follows the following steps:

1. The distances between the input vector \( x \) and all reference vectors (i.e., the weight vectors) are computed using a Euclidean distance measure.
2. A winner (i.e., a neural unit for which the corresponding weight vector is at a minimum distance from the input vector) is determined.
3. The weight vectors corresponding to the winner and the neural units in its topological neighbourhood are updated to align them towards the input vector.

The SOM attempts to represent the corpus of diagrams with optimal accuracy using the selected subset of features. When the network is executed, the winning node (i.e., the node with the lowest activation, indicating greatest proximity to the input case) is selected. For each diagram, the best-matching SOM unit is searched. At the same time, the diagrams become ordered on the grid so that similar diagrams are close to each other and dissimilar ones far from each other.

The learning rate is defined to be the constant of proportionality, indicating the extent to which the weight vectors are adjusted or updated towards a given input vector. The learning rate (updating rate) and the size of neighbourhood are reduced as the learning progresses over multiple iterations.

4.4 Refining Categories and Distinguishing Relevant Styles

The correspondence between high-level concepts and design features can often depend on the context of the observer and every design categorisation is different due to the hidden conceptions in the relevance of diagrams and their mutual similarity. This is the rationale behind the fourth stage, where if the design clusters selected by the observer map closely to each other on the SOM, then the corresponding feature performs well on the present categorisation and the relative weight of its opinion is increased.

By marking on the map, the categories the observer deems relevant, we are able to adjust each unit or node assigned a positive and negative value depending whether the observer has selected or rejected the corresponding design classification. The marking operation indicates
correctly classified design clusters as positive.

Diagrams are accumulated during the categorisation process into two sets and weights are adjusted in succeeding iterations, moving from an unsupervised SOM to one which is partially supervised or guided. The following section describes how the SOM modifies its behaviour depending on which design diagrams are included in the positive and negative design sets.

### 4.5 Integrating Positive and Negative Feedback

The selected and rejected diagrams result in positive and negative values on the best-matching units. Positive and negative responses are normalized so that their sum equals zero (Lassksonen et al. 2000). Since initial categorisation clusters diagrams as an unsupervised process, as soon as the observer’s feedback produces positive values, the diagrams are classified according to those feature subsets corresponding to the positive feedback.

The design diagrams associated with these units are then good candidates for next categorisation. This stage can be formalised, whereby the corpus’s nonintersecting subsets of positive $d^+$ and negative $d^-$ diagrams. The initial categories (defined in the previous section as $c$) lead to:

$$\frac{c^\prime}{d^+(d^+ \cup d^-)}$$

Categorising the remaining design corpus most similar to the positive marked designs can then be formally defined as:

$$\min!_C^c = \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{i=1}^{K} w_m E_m (x^* (d^*_n) x_n (d^*_m))$$

where $w_m$ are the weights for individual features and $E_m (...) \text{ is the Euclidean distance function with feature type } f^M$. Fitting of feature vectors is carried out by a sequential regression process, where $s = 1, 2, \ldots n$ is the step index such that for: $x(s)$, the first index $i = i(x)$ of the best-matching unit and all feature vectors or a subset of them that belong to those nodes centred around node $i = i(x)$ are updated (Lassksonen et al. 2000).

### 5. Q-SOM Experiments

Two experiments have been carried out to assess the utility of the process. The first experiment tests the discriminatory power of encoding and the SOM’s ability to categorise diagrams using specific feature sets. The second experiment tests the complete Q-SOM system i.e., using relevance feedback (RF) in a design scenario.

#### 5.1 Design Corpus

The design corpus used in all experiments consists of 2D architectural design diagrams. The level of complexity of the corpus is considered to be relatively high since although all diagrams are from a single domain, i.e., architecture, the corpus consists of designs from a number of architects and several different building typologies including small and large scale residential, as well as public buildings. We have used a relatively large design corpus comprising a total of 130 plan diagrams.

The two studies undertaken use networks that have been trained using 36 plan diagrams comprised of six designs which have been randomly selected from each of the six architects. The architects are: Palladio, Frank Lloyd Wright, Mies van der Rohe, Le Corbusier, Louis Kahn, and Mario Botta. Exemplars of each architect and the feature sets extracted are presented in Table 5.

<table>
<thead>
<tr>
<th>Exemplar Diagram</th>
<th>Exemplar LSF and Semantics</th>
<th>Exemplar ISF</th>
<th>Exemplar GSF and Semantics</th>
<th>Architect</th>
<th>Building Typology</th>
<th>Period</th>
</tr>
</thead>
</table>
The number of features extracted from the diagrams totalled 37,367 and there is an average of 287 features from 59 sets associated with each diagram. The characteristics of feature sets in relation to each architect are shown in Table 6.

### Table 6. Characteristics of each feature set based on Architect

<table>
<thead>
<tr>
<th>Type (Architect)</th>
<th>Average No. of Features</th>
<th>Total No. of Features</th>
<th>No. Diagrams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palladio</td>
<td>254</td>
<td>3810</td>
<td>15</td>
</tr>
<tr>
<td>Frank Lloyd Wright</td>
<td>306</td>
<td>18360</td>
<td>60</td>
</tr>
<tr>
<td>Mies van der Rohe</td>
<td>268</td>
<td>4288</td>
<td>16</td>
</tr>
<tr>
<td>Le Corbusier</td>
<td>221</td>
<td>1547</td>
<td>7</td>
</tr>
<tr>
<td>Louis Kahn</td>
<td>327</td>
<td>6213</td>
<td>19</td>
</tr>
<tr>
<td>Mario Botta</td>
<td>243</td>
<td>3159</td>
<td>13</td>
</tr>
</tbody>
</table>

### 5.2 Experiment 1: SOM Studies 1 & 2

The first experiment is designed to evaluate the effectiveness of the derivation of semantic features and ascertain the benefits of dimensionality reduction in diagram categorisation. We trained, tested and evaluated two networks using a variety of network topologies and different feature subsets.

#### 5.2.1 Pre-Processing

Pre-processing of input data was undertaken using the statistical feature selection method outlined in Section 4.2. Using CFS we evaluated subsets of features by the correlation among characteristics of each feature set based on Architect.

In the first study, we used only the local structural descriptions (i.e., the LSF) extracted from the corpus in dimensionality reduction. Eight LSF including: Protrusion_0, Protrusion_3, Iteration_2, Alternation_1, Symmetry, Square, Cruciform and Niche were identified as significant. From this point, we shall refer to this network as SOM_1.

In the second study, dimensionality reduction included all feature sets extracted using the DeREP and AbREP operations and in addition to the eight LSFs identified above, four GSF: Contains/Contained_by, Overlaps/Overlapped_by, Equals, and Courtyard were evaluated as the optimal subset of attributes for clustering. Categorisation therefore relies on a combination of feature categories where the ratio of local to global features is 2:1. We shall refer to this network as SOM_2.
5.2.2 Training

Feature vectors were created based on the feature subsets. Neither SOM_1 nor SOM_2 utilised information regarding the type of diagram, i.e., the architect, building type, period, etc. Since there is no access to prior knowledge regarding the number of clusters in the data, the SOM proceeds unsupervised. A node may represent more than one diagram, but with different activation values.

In the SOM_1 network, each diagram had an average of 35 features based on the LSFs described in the previous section. The final vector model contained 1,239 feature instances, which created a unique feature vector for each 2D plan diagram.

SOM_1 was trained for 500 cycles and produced a clustered qualitative-based topographic map for the 36 diagrams as shown in Figure 6. Results observed in the map show categorisation of diagrams can be roughly linked to the architect as indicated by the colour map. The colour map in Figure 6 indicates each architect where: A ≈ Palladio; B ≈ Wright, C ≈ Van Der Rohe, D ≈ LeCorbusier, E ≈ Kahn, and F ≈ Botta. In some cases the node contains two architects (approximately 20%) and each label has been assigned on the basis of the dominant feature vector. Also observed from the map, the network appears to have clusters distributed separately corresponding to the same architect, including: Wright (C), Kahn (E) and Botta (F).

In the second study, diagram feature vectors in SOM_2 are based on features sets from both local and global attributes and each diagram had an average of 50 features. The final vector model contained 1812 feature instances comprising of feature vectors from 1,239 LSF and 573 GSF feature instances.

SOM_2 was also trained for 500 cycles and provides well defined clusters, shown in Figure 7. Like the results obtained for SOM_1, Figure 6, the results observed in the topographic map show categorisation of diagrams can be linked to the architect. The topological ordering of the diagrams shows a better result than obtained for SOM_1 training. This is evident from the separate clusters and the distinctive change of clusters across the map, where Kahn’s design diagrams E, are located in the upper left-hand corner of the map and the architect gradually changes towards the bottom-right corner to Le Corbusier, D. Although SOM_2 also distributes two clusters for Wright’s designs, B, clustering is more consistent with each architect.
Significantly, in SOM_2, all nodes except for the node marked “x” contain input vectors from the same architect. This can also be observed from the activation weights given to each individual input vector. SOM_1 input vectors also have much lower activations when compared to SOM_2 input vectors.

Testing is carried out to evaluate the clustering effectiveness of each trained network. The objective of testing is to evaluate the “success” of each trained network and the two different approaches to constructing feature vectors.

5.2.2 Testing

We tested SOM_1 and SOM_2 by looking at the clustering ability of each approach. In each test, a 3x3, a 5x5 and a 10x10 map were created. To analyse the results of categorisation between the topographic maps we utilised techniques from conventional text-based categorization analysis including: Precision (Slonim, Friedman and Tishby 2002), the Jacaard or JAC method (Downton and Brennon 1980), and the Fowlkes-Mallows or FM method (Fowlkes and Mallows 1983).

However, since classification is unsupervised it is not possible to apply evaluation methods directly as would be the case for supervised learning. To analyse the results of testing the unsupervised learning of SOM_1 and SOM_2 it is necessary to utilise the most dominant label of the cluster for all of the diagrams. For this reason we have maintained the labels used during training to assign categories, i.e., by architect and utilised the “micro-averaged” precision matrix method (Slonim, Friedman and Tishby 2002) to evaluate the results of each network, where:

\[
P(D) = \frac{\sum_{c'} \alpha(c', D)}{\sum_{c'} \alpha(c', T) + \beta(c', D)}
\]

where for each category \(c' \in \mathcal{C}\), \(\alpha(c', D)\) is defined by the number of diagrams “correctly” assigned to \(c'\), and \(\beta(c', D)\) defines the number of diagrams incorrectly assigned to \(c'\).

We also use the well-established JAC and FM methods which evaluate cluster quality:

\[
JAC = \frac{TP}{TP + FP + FN}
\]

and:

\[
FM = \frac{TP}{\sqrt{(TP + FP) \times (TP + FN)}}
\]

where \(TP\) is equal to the number of diagrams that the classifier correctly classified as belonging to a category (true positive), and \(FP\) is the number of diagrams that the classifier classified (true positives and false positives) as belonging to that category and \(FN\) is the pair-wise number of false negatives.
We first tested SOM_1 using the three different map topologies and analysed each network’s ability to categorise the entire design corpus. The 5x5 map has the best results for all evaluation techniques measured. The 5x5 SOM_1 produced comparable and good precision and JAC results, shown in Table 7. The FM levels show how the 5x5 map outperformed both the 3x3 and the 10x10 maps.

TABLE 7. Clustering ability of different map topologies trained on SOM_1 Feature subsets

<table>
<thead>
<tr>
<th>Study 1: SOM_1</th>
<th>Precision</th>
<th>JAC</th>
<th>FM</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOM_1 3x3</td>
<td>0.49</td>
<td>0.29</td>
<td>0.37</td>
</tr>
<tr>
<td>SOM_1 5x5</td>
<td>0.62</td>
<td>0.38</td>
<td>0.45</td>
</tr>
<tr>
<td>SOM_1 10x10</td>
<td>0.53</td>
<td>0.32</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Secondly, we tested SOM_2. Again the 5x5 map has the best results for all evaluation techniques measured. As expected the 5x5 SOM_2 produced better results for precision, JAC and FM than SOM_1 results, shown in Table 8.

TABLE 8. Clustering ability of different map topologies trained on SOM_2 Feature subsets

<table>
<thead>
<tr>
<th>Study 2: SOM_2</th>
<th>Precision</th>
<th>JAC</th>
<th>FM</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOM_2 3x3</td>
<td>0.61</td>
<td>0.46</td>
<td>0.42</td>
</tr>
<tr>
<td>SOM_2 5x5</td>
<td>0.74</td>
<td>0.53</td>
<td>0.50</td>
</tr>
<tr>
<td>SOM_2 10x10</td>
<td>0.60</td>
<td>0.46</td>
<td>0.39</td>
</tr>
</tbody>
</table>

5.2.3 Cluster Evaluation
The nature of the categories produced by the 5x5 SOM_1 and SOM_2 are difficult to identify and therefore evaluate, except via a visual (subjective) process. Recently, conventional clustering techniques (e.g. K-means, EM, Hierarchical, etc.) have been used to resolve this problem. Ahmad and Vrusias (2004) have demonstrated the effectiveness of using conventional statistical clustering techniques, in evaluating the output of maps of unsupervised networks. Such sequential clustering, i.e., first clustering using an unsupervised network and then clustering the output map, facilitates in visualising the clusters that are otherwise implicit in the output map.

We have used this sequential clustering method (SOM followed by K-means) to examine the categories obtained. An application of K-Means clustering on the output of the SOM_1 and SOM_2 maps shows how they have found data in the proximate types defined by both the architect combined with the residential or public building types. Tables 9(a) and 9(b) compare the K-Means clustering of the 130 plan diagrams by the best performing maps of SOM_1 and SOM_2.

The two tables show the distribution of plan diagrams where feature subsets can only be used to cluster diagrams according to architect in conjunction with either residential or public building type. Table 9(a) shows clustering of the design corpus based on the LSF subset has not proven to be clear cut using the sequential clustering method.

TABLE 9. Distribution of 130 plan diagrams using K-Means clustering: (a) SOM_1 and (b) SOM_2

(a) Clusters defined by K-Means on SOM_1

<table>
<thead>
<tr>
<th>Arch. &amp; Bld. Type</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Palladio Res. &amp; Pub.</td>
<td>9</td>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Wright Res. (Prairie)</td>
<td>6</td>
<td>19</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C Wright Res. (Usonian)</td>
<td>7</td>
<td>14</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D MVDB Rohe Res.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E MVDB Rohe Pub.</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F LeCorbusier Res.</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G LeCorbusier Pub.</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Clusters defined by K-Means on SOM_2

<table>
<thead>
<tr>
<th>Arch. &amp; Bld. Type</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Palladio Res. &amp; Pub.</td>
<td>12</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Wright Res. (Prairie)</td>
<td>1</td>
<td>25</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C Wright Res. (Usonian)</td>
<td>4</td>
<td>18</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D MVDB Rohe Res.</td>
<td>1</td>
<td></td>
<td>1</td>
<td>6</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E MVDB Rohe Pub.</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F LeCorbusier Res.</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G LeCorbusier Pub.</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The plan diagrams that are associated with the architect and building type and represented using the LSF and GSF subset, Table 9(b), have generally all been separately clustered except for Le Corbusier’s designs where no really distinct clusters can be distinguished. Significantly for both SOM_1 and SOM_2 networks, Wright’s designs were distinguished relative to two periods of Wright’s work – the Prairie and Usonian houses.

5.3 Experiment 2: Q-SOM Study

Based on the results obtained from SOM_2 with a 5x5 map topology tested in the previous experiment, the final experiment trains and tests the same network’s ability to categorise the corpus using the complete Q-SOM framework to obtain clusters which are relevant to some design context. In this experiment, we evaluated the relevance of the categorisations by the SOM in the context of a residential design task so as to apply positive and negative values to the units of the network.

Pre-processing is again used in dimensionality reduction to create feature vectors. Categorisation proceeds as a sequential process based on manual feature selection since it is not possible to order the design corpus by their mutual distances in advance using those features most appropriate to a visual assessment.

5.3.1 Design Context

A simple design task was formulated using a brief specifying the requirements of a residential plan design for a family of four. A conceptual design sketch was produced using Q-SOM’s digital drawing interface which constrains sketching to orthogonal axis. Figure 8 presents the design sketch produced as a result of the brief’s requirements.

The sketch was then encoded (all labels are ignored) and included in the design corpus with the other 130 plan diagrams.

5.3.2 Pre-Processing

In this experiment we pre-processed feature sets manually. Manual processing enables the selection of a feature subset relevant to the design task. From the 59 possible feature sets, selection is facilitated by an analysis of the design sketch itself. Since categorisation is grounded by the design context, diagrams which are similar can be determined based on those features extracted from the sketch. The number of relevant features is calculated by using the target sketch design as a hand-picked subset and applying this to create the feature vectors of the entire design corpus. This enables a visual style to be determined from specific features based on the design sketch. Those features extracted from the design in Figure 8 are shown in Table 10.

| TABLE 10. Reduced Feature List and Designer’s Feature Subset |
|---------------------------------|-----------------|
| FEATURES EXTRACTED BY DeREP & AdREP | SUBSET |
| LSF (Geometry-based Semantics) | Indentation 0,1,2,3 | x, x, x, x |
|                                  | Protrusion 0,1,2,3 | x, x, x, x |
|                                  | Iteration 0,1,2,3 | x, x, x, x |
|                                  | Alternation 0,1,2,3 | x, x, x, x |
A total of 21 feature sets were extracted from the sketch design, distributed between local, intermediate and global feature classes as: 6, 2 and 3. It is also possible to use a subset of the features detected the sketch. For example a feature subset can be chosen which is deemed to have some degree of relevance to the design requirements. An example subset is shown in the third column of Table 10 and where 11 feature sets have been selected (ticked).

For the experiments we selected five different feature vector models consisting of: (i) all sketch design feature classes, (ii) example subset (ticked) feature classes, (iii) LSF only, (iv) ISF only and (v) GSF only. These five subsets were used to train each of the five different networks.

### 5.3.3 Training and Testing Using Relevance Feedback

Testing was carried out to evaluate the clustering effectiveness of each network under the guidance of relevance feedback. To demonstrate and evaluate the performance of all feature subsets and the Q-SOM model with RF, separate visual design ‘styles’ must first be identified and selected from the corpus as the desired cluster targets. The targeted ‘styles’ are identified as (i) Wright’s Usonian period and (ii) Kahn’s residential designs. Each ‘style’ contains 26 and 11 plan diagrams respectively.

Based on positive and negative feedback, each SOM presented a re-categorisation of the corpus where the iterative process continued until the observer (the author) was satisfied. The clustering of each model was produced by returning the best-scoring diagrams in each iteration step from the selections of the relevant designs among them. Using the JAC and FM measures we can examine the results of each SOM’s as shown in Table 11.

#### TABLE 11. Clustering Ability of 5x5 Map Topologies Trained on Different Feature Subsets

<table>
<thead>
<tr>
<th>Experiment 2: Q-SOM</th>
<th>JAC</th>
<th>FM</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x5 All Features</td>
<td>0.53</td>
<td>0.50</td>
</tr>
<tr>
<td>5x5 Example Subset</td>
<td>0.46</td>
<td>0.42</td>
</tr>
<tr>
<td>5x5 LSF</td>
<td>0.46</td>
<td>0.39</td>
</tr>
<tr>
<td>5x5 ISF</td>
<td>0.38</td>
<td>0.45</td>
</tr>
<tr>
<td>5x5 GSF</td>
<td>0.32</td>
<td>0.30</td>
</tr>
</tbody>
</table>

The first map whose feature vectors are based on all features from the sketch has the best results. The second map with the example subset also produced comparable results for JAC and FM measures. The performance of the remaining three maps was lower and result of the best map according to the statistical analysis is illustrated in Figure 10, which shows the 5x5 ‘All Features’ topographic map, which was trained for a total of 1200 cycles. Results observed in the map show categorisation of diagrams can be roughly linked to two architects as indicated by the colour map.
The colour map in Figure 10 indicates each architect where: A ≈ Wright (Usonian); B ≈ Wright (Prairie) and C ≈ Kahn. Other architects whose dominant feature vectors defined labelling included: d ≈ Palladio; e ≈ Van Der Rohe, f ≈ LeCorbusier, and g ≈ Botta. Unlike the previous experiments there are multiple nodes of the map that contain input vectors from different architects. Significantly, only 5% of the nodes where Wright and Kahn’s diagrams are clustered contained another architect whereas the majority of the remaining nodes contained more than one architect (approximately 50%).

5.2.4 Cluster and Feature Subset Evaluation

We evaluated the performance of the Q-SOM model using a method that resembles “target testing” developed by Cox et al. (1996). Here, instead of a single target, testing evaluates the targeted styles: Usonian and Kahn’s residential designs.

To obtain the performance measure \( \tau \), we use the targeted category \( T \hat{c} \), of designs defined by the user’s requirements \( r \). For each diagram in category \( T \hat{c} \), we record the total number of different clusters categorised by the network until the final category is reached. From this data, the average number of clusters formed before the final “correct” response is divided by the total number of diagrams \( K \). Then the performance measure of the target category is given by:

\[
\tau = \left[ \frac{\phi(c', A)}{2} - \frac{1}{2} \frac{\psi(c', A)}{c} \right]
\]

where \( \phi(c', A) \) is the \textit{a priori} probability of the category \( T \hat{c} \), given by \( T \hat{c} / K \). In general the smaller \( \tau \), i.e., \( \tau < 0.5 \), the better the performance.

The results of the performance measure with the Q-SOM networks are shown in Table 12. Results are presented for all five networks created based on their feature subsets. The two feature subsets containing all features and the example selection yielded better results than the LSF, ISF or GSF subsets, which can be seen from the first 2 rows in Table 12.

<table>
<thead>
<tr>
<th>FEATURE SUBSETS</th>
<th>TARGET CLUSTER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Usonian</td>
<td>Kahn (Residential)</td>
</tr>
<tr>
<td>All Features</td>
<td>0.23</td>
<td>0.26</td>
</tr>
<tr>
<td>Example Subset</td>
<td>0.29</td>
<td>0.31</td>
</tr>
<tr>
<td>LSF</td>
<td>0.36</td>
<td>0.33</td>
</tr>
<tr>
<td>ISF</td>
<td>0.54</td>
<td>0.62</td>
</tr>
<tr>
<td>GSF</td>
<td>0.39</td>
<td>0.43</td>
</tr>
</tbody>
</table>
Based on the performance measures we can observe that using all feature classes to create feature vectors yields better results than any one single dimension. The general trend observed from these results shows that using a larger set of features yields better results than using a smaller subset of features. The combination of all available morphological, topological and mereotopological features in conjunction with relevance feedback has resulted in the highest performance measure.

The implicit weight adjustment based on the relative importance of features shows that the model is capable of categorisation based on physical and semantic attributes of the design corpus. This kind of automatic adaptation is desirable as it is generally not known which feature combination would perform best in clustering the complex spatio-visual information inherent in diagrams of the architectural domain.

The experiment demonstrates that utilising RF in a SOM as a function for assessing the similarity of design diagrams not only provides a useful method for unsupervised categorisation but also provides the flexibility to overcome a variety of problems resulting from the complexities of context. The Q-SOM approach is able to provide a robust method for using multiple dimensions from both lower-level geometric and higher-level semantic features.

**Conclusion and Future Plans**

In this paper, we have introduced the Q-SOM approach to assessing visual styles by incorporating qualitative and contextual aspects in a formal model. Our approach to visual style focuses on modelling both a cognitively compatible and task relevant clustering method. The model has been shown to be capable of bridging the gap between the high-level and context-dependent semantics used by designers to understand design content and the low-level visual features extracted from the design artefact. This flexible approach to identifying visual styles uses an iterative and interactive process that enables the SOM to learn the desired categories or clusters. The results of our experiments show that the qualitative and context-dependent approach of the system is able to effectively select from sequential and iterative processing by adjusting weights of a SOM to coincide with an observer’s conceptual view of diagram similarity. In this way, initial unsupervised clustering becomes more accurate as new clusters are formed.

The model’s strength lies in its ability to use a multi-dimensional approach to qualitative encoding as well as multiple reference diagrams simultaneously. By providing a mapping from the physicality of the diagram using qualitative encoding, semantic feature sets can be obtained. As a result meaningful feature sets can be analysed as input to the SOM and a relevance feedback technique can be integrated. This feature makes this system differ from other content-based classification systems, such as feature-based distances measures of similarity which use only one reference diagram at a time. Further, since an individual visual feature or even a class of features (from any one single dimension) may not be sufficient in classifying complex design diagrams, it is necessary to extract information from multiple dimensions so that the categorisation performance can be improved relative to context.

Using a target ‘style’ in the final experiment, the model demonstrates that when contextual information is integrated to determine the relative importance of distinguishing features, the correlation between the system’s results and the observer’s similarity assessments increases. Significantly, this correlation can be seen to result from the detailed definition, detection and extraction of feature classes and user feedback determining feature relevance.

The approach provides a method to overcome the difficulties that arise from perceptual biases in the design domain. In other words, it becomes possible to integrate contextual dependencies where a visual style detected by a designer is influenced by design objectives. A key potential implication is that the isolated characteristics of features appear to be insufficient to formulate conclusions about the nature and distinctions of visual styles during designing. Instead, causality can also be attributed to context dependent factors that influence the perception and judgement of the entire corpus, individual diagram and feature. These complex aspects of spatio-visual information and cognition require further
exploration in relation to a computational model of visual design style.

Quantitative measures of the Q-SOM model’s performance are problematic due to human subjectivity and preference during the classification process. Although the results obtained so far are promising in terms of the potential of the Q-SOM method as a cognitively plausible classification system, additional studies are required in design scenarios based on multiple users of the system. In undertaking human-subject experiments, the Q-SOM results should correlate with both expert and novice designers’ judgements of diagram similarity. In future studies, the model must undergo further testing so that spatio-visual information can be inspected relative to design context and historically defined design styles. The initial objective of future research is to increase Q-SOM’s classification performance through further experimentation to find a set of well-balanced features, which on average, perform as well as possible. In this way, the effects of feature centrality in contextual terms can be investigated in the design domain.

References


This is a copy of the paper: Jupp, J and Gero, JS (2006) Visual style: Qualitative and context dependent categorisation, AIEDAM (to appear).