Visual style: Qualitative and context-dependent categorization

JULIE JUPP[1] AND JOHN S. GERO[2]
[1]Engineering Design Centre, University of Cambridge, Cambridge, UK
[2]Key Centre of Design Computing and Cognition, School of Architecture, Design Science and Planning,
 University of Sydney, Sydney, New South Wales, Australia

Abstract

Style is an ordering principle by which to structure artifacts in a design domain. The application of a visual order entails some explicit grouping property that 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 semantic interpretation. We present a model of visual style based on the concept of similarity as a qualitative context-dependent categorization. The two core components of the model are semantic feature extraction and self-organizing maps (SOMs). The model proposes a method of categorizing two-dimensional unannotated design diagrams using both low-level geometric and high-level semantic features that are automatically derived from the pictorial content of the design. The operation of the initial model, called Q-SOM, is then extended to include relevance feedback (Q-SOM:RF). The extended model can be seen as a series of sequential processing stages, in which qualitative encoding and feature extraction are followed by iterative recategorization. Categorization 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: initial per feature detection and extraction, selection of feature sets corresponding to different spatial ontologies, unsupervised categorization of design diagrams based on appropriate feature subsets, and integration of design context via relevance feedback. From our experiments we compare different outcomes from consecutive stages of the model. The results show that the model provides a cognitively plausible and context-dependent method for characterizing visual style in design.

Keywords: Design Context; Qualitative Spatial Representation; Relevance Feedback; Similarity

1. 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 artifact, modality, society, culture, period, and so forth. In developing a model of visual style, we approach the problem from an "object" viewpoint, where an artifact's common physical characteristics are primary (Edwards, 1945; Ackerman, 1967). A useful and well-accepted definition of this view expresses style as a principle by which to provide

order, allowing a set of artifacts to be structured according to some set of criteria (Knight, 1994).

An object view of style essentially relies on the assessment criteria derived from a group of artifacts, its canonical rerepresentation, and a measure that enables comparison. Assessment criteria can be specified on a number of different dimensions. These dimensions can include structural or behavioral attributes and may be represented either quantitatively or qualitatively. The utility of qualitative rerepresentation is evident in its ability to derive design semantics (Gero & Park, 1997; Ding & Gero, 1998). However, because of the complexity of deriving semantics relevant to the domain, this approach is less prominent in models of visual style. A further restriction typical of such models is their inability to handle important cognitive properties of similarity assessment such as when the perception of a feature

or its semantic interpretation is influenced by the designer's context. As a result, computational models of style have inherited an emphasis on linear analysis, which focuses on distance measures that maintain a static world assumption in which style is treated as unrelated to its locus of application.

These limitations are critical to a digital characterization of visual style in design because designers are capable of analyzing artifacts using a number of different dimensions. The perception of features can be influenced by the design task, the designer's intentions, and their goals. In this way, interpreting a design and distinguishing some visual style can be dependent on a feature's relevance relative to context. Judging the visual similarity of design artifacts and identifying styles therefore depends on context. In Tversky (1977) and Tversky and Gati's (1982) investigations on the concept of similarity, they describe aspects of context in mathematical properties of similarity including asymmetry, minimality, and triangle inequality. These properties are significant in visual and spatial reasoning in design yet have largely been ignored in modeling an object view of style.

In exploring the deficiencies of current approaches to digital characterizations of visual style we ask, how can high-level semantic dimensions be derived efficiently from low-level structural ones and how may we compare them to provide a cognitively plausible and contextually dependent model of style? To address this two-part question we look to qualitative modeling and clustering techniques.

Spatial information directly available to human observers is typically qualitative in nature (Cohn, 1997). The benefits of utilizing qualitative encoding lie in the description of attributes that are significant to the preservation of salient design qualities. Encoding a design corpus qualitatively can provide meaningful data sets that can then be utilized by any clustering approach. However, in modeling style as qualitative and context-relevant similarity assessments (relative to the observer and the design task), two basic problems arise: the gap between high-level semantics used by designers to understand design content and low-level structural features extracted from the design artifact; and their relevance or significance relative to the design context. In providing a possible solution to this gap, three research issues are addressed here:

- qualitative rerepresentation of the design diagram's pictorial content,
- semantic feature mappings and feature salience, and
- context-sensitive/user-relevant similarity measures.

The approach to visual style presented here explores each of these issues and utilizes the application area of two-dimensional (2-D) architectural diagrams to develop a model and test its implementation. The system is based on an existing qualitative representation schema capable of describing a hierarchy of spatial ontologies (Jupp & Gero, 2004, 2006)

and self-organizing maps (SOMs) as a method for design comparison. The model described in the remainder of this article is capable of automatically structuring a design corpus according to selected feature semantics as an iterative process, which adapts to an observer's requirements and preferences. Adaptation is based on the relevance of clusters that are judged in relation to some design task. This approach is commonly used in text classification and retrieval systems and is known as relevance feedback (RF; Salton & McGill, 1983). Although it is a well-established technique in text categorization and retrieval systems, there appears to be no model of similarity in design that integrates RF in this way. Applied here, the implementation of SOM and RF in conjunction with qualitative feature-based rerepresentation presents the main contribution of this work. This flexible, open-ended approach to visual style automates qualitative rerepresentations, learns from the data set unsupervised, and modifies assessments according to design context. We posit that the creation of a model of qualitative context-dependent similarity assessment can provide a cognitively plausible and contextually relevant characterization of the visual style of design artifacts.

2. STYLE, SIMILARITY, AND VISUOSPATIAL REASONING

Identifying a visual style is a judgment process that requires an artifact to be decomposed into elements in which they are the same and elements in which they are different. Design artifacts 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. Human observers are able to recognize, interpret, and search for salient features in diagrams in order to detect visual similarities and ultimately identify members of a style. The last 40 years of research surrounding the concept of similarity has provided a variety of insights on both theoretical and empirical levels (see Tversky, 1977; Tversky & Gati, 1982; Love & Sloman, 1995; Sloman, 1996).

In design research, there is still a lack of understanding of how designers classify, form concepts, and make decisions based on the similarity perceived between two or more designs. In related research, 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 of complex diagrams largely depends on intuitive similarity assessments that are later revised after the initial assessment (Jupp & Gero, 2005). Such cognitive investigations are difficult because similarity assessments in design typically involve comparisons that rely on a variety of dimensions, and 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, 2006). In this regard, the high-level semantic content contained in design diagrams frequently results in the initial assessment being revised a number of times in light of the design context and design requirements.

Despite the lack of cognitive studies on similarity and visual style, a variety of computational models have been developed (Gross & Do, 1995; Park & Gero, 2000; Davies & Goel, 2001; Gero & Kazakov, 2001; Forbus et al., 2003; Burns, 2004). In general, these models are based on a rerepresentation of either 2-D or 3-D design artifacts and some function of similarity that allows those artifacts to be compared and ordered. Many 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 systems. Yet, most of the existing approaches to visual style are limited because a comparison 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 moving toward the idea that it is a process whose outcome can only be reported in a *post hoc* fashion.

We use the concept of similarity as the grouping principle (Wertheimer, 1977) by which 2-D design diagrams can be ordered. However, we assume that similarity is also related to the way information is processed and that the reporting of similarity judgments is a metacognitive process (Thomas & Mareschal, 1997) requiring explicit comparison of design information both prior and subsequent to processing by the system.

2.1. Toward a model of style

The visual style of a design may be described as transient, where the perceived similarities can depend on the corpus in question; the amount of *a priori* information available; and the order of design comparison. When considered in relation to a task-specific context, it may also depend on design objectives and requirements. The problem of modeling visual style largely reflects these 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. From the perspective of the application domain being tested here, this level of complexity translates to categorizing only a single architect's designs that encompass only residential diagrams. This level of complexity enables a relatively straightforward assessment based on the automatic segmentation of the pictorial content of the diagram. In the other extreme lies the problem of categorizing a design corpus that encompasses a variety of the following:

1. design domains (arts, architecture, engineering, industrial design, etc.);
2. design artifacts, taking the architectural domain as an example, which could refer to the kinds of building typologies such as residential, commercial, industrial, religious, and so forth;
3. design descriptions, for example, plans, sections, elevations, and so forth;
4. design rerepresentations, that is, the type and number of dimensions; and
5. design contexts, which can include perceptual constraints, such as the order of the diagrams being compared (to more task-specific related requirements, such as the design brief) to broader contextual aspects such as society, culture, region, and period.

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 (Sloman, 1996; Thomas & Mareschal, 1997). This view is particularly relevant in design because the designer operates within a context and his or her perception and judgment of design similarities are influenced by it.

In view of such complexities, similarity as a measure of visual style is defined here not as a fixed and irreducible concept, but as a process, which takes place over some more or less open and variable dimensions of the designs being compared. Under this definition, similarity is treated as contextually dependent because the attributes that the process uses as input can change according to certain criteria. This view gives less explanatory force to similarity because it demands analysis of the design attributes whose similarity it computes in relation to context. In addition, unannotated diagrams require rich rerepresentations at successive levels of abstraction (Marr & Nishihara, 1978). In cognition, this type of information processing is not only sensitive to context but is also typically intuitive and subjective and is not a function of strict mathematical models. In considering these aspects of visual style, a similarity measure should be capable of both incorporating design semantics and operating in relation to the design task and its requirements. Our approach therefore highlights the importance of both a qualitative and contextually dependent assessment.

2.1.1. Qualitative rerepresentation

Qualitative approaches to representation are a common analysis paradigm in design reasoning applications. There are multiple dimensions that could be modeled qualitatively, resulting in a variety of data sets. Here, we consider as essential two general criteria distinguished by Schapiro (1961) as being significant to the characterization of style:

shape elements and attributes and their spatial relationships. Using these criteria as our foundation, the objective of the qualitative schema used here is to provide a rich set of structural attributes (for open and closed shapes) as well as spatial relations incorporating adjacency and connectedness (for shape aggregations).

A canonical description of a corpus of design diagrams 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 because the use of common-sense descriptors supports the identification of designs that are not only structurally or spatially close but also conceptually close, while not being identical.

2.1.2. Method of comparison

In treating visual style as a multidimensional similarity measure we propose to utilize artificial neural networks. The underlying mathematical properties of most neural networks used in categorization are scalar distance-based algorithms. Of these, the SOM is a typical representative. 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. Because there is no absolute definition of the commonalities between design artifacts in terms of their spatial descriptions, 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 2-D space. This is an important property of SOMs because they are able to represent a mapping, which preserves relations in the input space while simultaneously performing a dimensionality reduction onto the 2-D 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 centers of the probability density function of the input data.

Although extensively used in other fields of research such as text and image retrieval, SOMs have not been widely utilized in design categorization systems. One application known to us is a model proposed by Colagrossi et al. (2003) for categorizing works of art. Colagrossi and 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 color. However, the application of the SOM by Colagrossi et al. does not address

the complexities identified in Section 2.1. This is partly attributable to its restricted application domain (i.e., distinct design corpus), as well as the lack of semantics and contextual relevancies. Under this and other existing approaches to similarity assessment in design, contextually relevant categorizations (e.g., relevant to the design task) cannot be identified. SOMs as a measure of design similarity can be improved by utilizing both qualitative descriptions and contextual input. By describing designs qualitatively using a number of different spatial ontologies and treating the designer as an inseparable part of the assessment process, it is possible to provide categorizations relative to some design task. The remainder of this paper presents a model capable of identifying the visual style of diagrams relative to a designer's context.

3. MODELING QUALITATIVE CONTEXT-DEPENDENT STYLE

The model of qualitative context-dependent visual style (Q-SOM:RF) has three main components: qualitative feature-based rerepresentation, SOMs, and RF. The framework presented in Figure 1 relies on the following consecutive stages:

1. recognition, extraction, and encoding of three different levels of spatial attributes;
2. initial per feature selection of encoded spatial attributes and a combination of feature lists;
3. categorization via unsupervised learning of design diagrams based on available features;
4. positive and negative feedback processes via the observer's input; and
5. resulting weight adjustment and recategorization of design diagrams.

In the first stage of the framework in Figure 1, each diagram in the design corpus is encoded using a qualitative schema capable of describing 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 preprocessing. A feature subset is produced using either principal component analysis or manual feature selection by the observer. The third stage utilizes the feature subset as input to the SOM and categorization occurs via unsupervised learning. The manner in which distances in various feature spaces are weighted and combined to form a scalar suitable for minimization 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 categorization to one that is guided by the observer. The final stage recategorizes diagrams that are similar to the observer's target diagram, meeting some set of target criteria, by ordering those diagrams whose distance to the target is minimal in any or all feature sets.
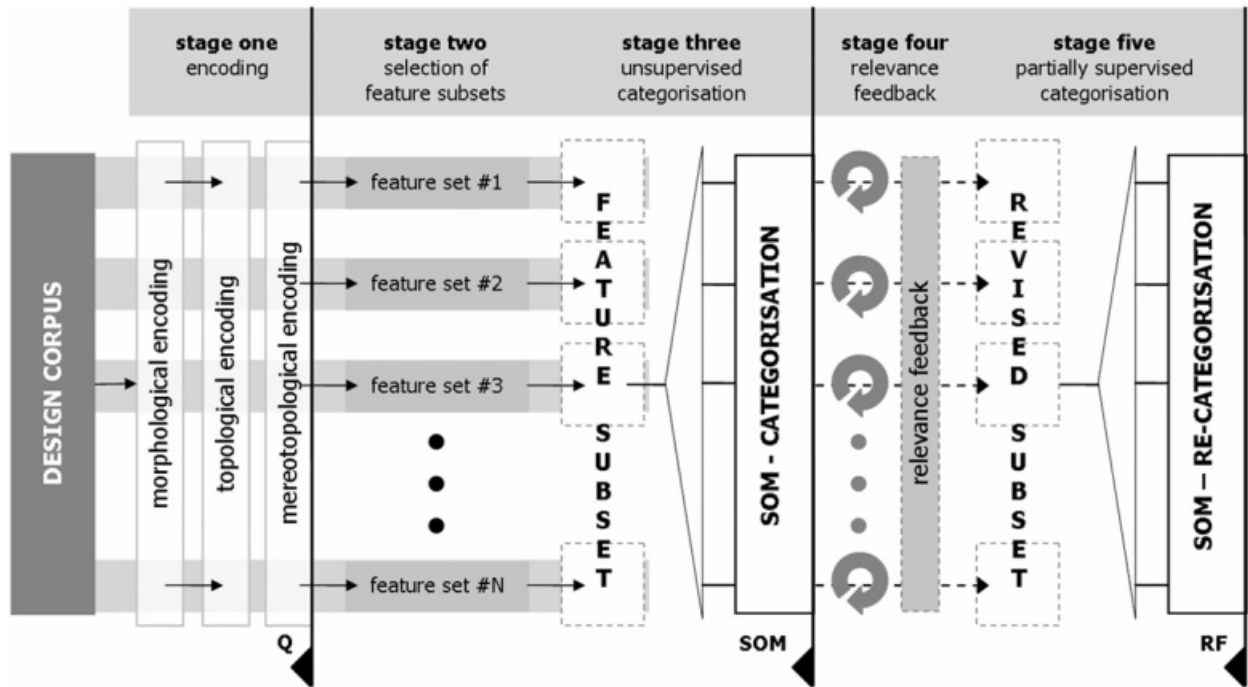
Fig. 1. The Q-SOM:RF stages of unannotated 2-D diagram categorization.

3.1. Rerepresenting multiple spatial ontologies

In utilizing SOMs for categorization, the homogeneity of a category is enforced by the appropriate choice of feature vectors. Because design diagrams are an explicit representation of the artifact's geometry, it is reasonable to expect that categorization be based on criteria that incorporate properties of 2-D geometry and have the following attributes:

- generic so that they have applicability over a wide spectrum of application domains;
- characterize as many physical dimensions of the diagram as possible, including orientation, distance, and topology;
- have both local and global support, that is, they should be computable on shape primitives and spatial relations around a landmark of significance;
- provide descriptions capable of higher level semantic mappings;
- stable and invariant over large ranges of viewpoints and scales; and
- reliable, robust, and readily detectable by procedures that are computationally stable.

To satisfy these requirements the process of encoding follows from physicality to symbol to regularity to feature. *Physicality* refers to the graphic descriptions of diagrams indicating the geometric information, and it is the prerepresentation upon which a process of information reduction is applied successively over three levels of abstraction. *Sym-*

bol refers to the unrefined symbolic encoding of graphic information; spatial attributes are recognized and converted into qualitative symbol values. *Regularity* is the syntactic matching stage in which regular or repetitious patterns of encodings are identified and grouped, where the detection of characteristics relies on "chunking" (Brown et al., 1995). *Feature* involves matching predetermined syntactic patterns with meaningful design semantics. In the following section we present a summary of the hierarchical encoding method used in the model. The ensuing sections present a summary of the encoding schema. For a more detailed account refer to our other work (Jupp & Gero, 2004, 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 & Park, 1997).

Where an angular change occurs, landmarks are initially set to P, separating convex and concave angles. The scanning order for each vertex is set to a counterclockwise direction and the magnitude of the vertex is also measured in this direction. The two primary codes L and Γ represent a vertex so that individual shapes can be described qualitatively. For complete descriptions and definitions of L and Γ, refer to Jupp and Gero (2006). The addition of codes cap-

Table 1. *Qualitative syntax for morphology*

	Angle Codes	Length Codes
Numeric value range	$0 \le \theta \le 2P$	$-\infty \le 1 \le \infty$
Landmark set	{0, P}	{−∞, 0, ∞}
Interval set	{(0, P), (P, 0)}	{(−∞, 0), [0, 0], (0, +∞)}
Q-code set	{L, Γ}	{L, Γ} ∧∨ {−, =, +}

turing 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 are used to group structural information in order 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; convexities can then be detected including indentation, protrusion, iteration, alternation, and symmetry.

Patterns of symbol sequences denoting specific categories of features are then identified that are familiar in contour or identify some particular shape semantic. There are two classes of defined semantics: simple and complex. Simple semantics include primary shape types that can be recognized and classed into subsequent labels such as rectangle, square, L-shape, U-shape, T-shape, and cruciform. Complex semantics incorporate domain-specific knowledge capable of mapping design concepts to features, for example, architectural design concepts such as chambers and niches. Because the patterns are derived from low-level structural primitives, they are defined as local shape features (LSFs). All LSFs are recognized by matching symbols with an existing database of features.

3.1.2. Topology

At the point of contact of two or more shapes there are specific extraction and embedding relationships for the intersection of line contours (Gero & Jupp, 2003). Where this occurs there is a transformation in representation that extends the encoding of two line intersections to include multiple lines. A distinction is made between the strictly morphological descriptions and topological ones, that is, L and Γ codes and T, ⊥, and C codes. T, ⊥, and C codes carry information regarding the diagram's topology and describe a point of contact of more than two contours. For complete descriptions and definitions of T, ⊥, and C codes, refer to Jupp and Gero (2006). 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 bottom-up development.

The symbols used to describe edges of graphs concern the disposition of physical intersections of lines that have been used to generate the polygon fields that are subsequently analyzed as graphs. Edges are labeled 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 definitions of dyad symbols are presented in Table 2.

The representation of dyad symbols reveals distinctive topological characteristics that are recognized 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 (ISFs), because the "neighborhood" of the description is based on local attributes as well as information describing topological properties. Like LSFs, ISFs are identified by matching an existing feature database.

3.1.3. Mereotopology

The dual 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 graph 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 2-D representation. The dual graph allows further derivation of spatial relationships.

Table 2. *Dyad symbols: Qualitative syntax for topology*

	Angle Codes	Area Codes
Numeric value range	$0 \le \theta \le 2P$	$-\infty \le 1 \le \infty$
Landmark set	{0, P}	{−∞, 0, ∞}
Interval set	{[0, 0], (0, P), [P, P], (P, 0)}	{(−∞, 0), [0, 0], (0, +∞)}
Q-code set	{L, Γ, T, ⊥, C}	{L, Γ, T, ⊥, C} ∧∨ {−, =, +}

By labeling the new dual edges, "tuples" composed of dyads (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. Labeled dual edges allow the identification of regularities and feature semantics describing part–whole relationships between two or more shapes. Because 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.

The semantic features identified at this level account more thoroughly for mereotopology. The regularities defined here are similar to Allen's 13 interval relations for the temporal domain (Allen, 1984). 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 it is strictly based on orthogonal shapes, it is still capable of handling arbitrary multisided forms.

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

Because all features are derived from higher level spatial primitives, they are defined as global shape features (GSFs). The three-level schema summarized here is characterized by the class aspect of handling and labeling design concepts, and it is useful when dealing with different design categorization scenarios. The concession of the approach is that it is essential to have a large database of concept to feature mappings. Despite this requirement, it is still an efficient and robust method of rerepresentation.

3.1.4. Example encoding

Figure 2 shows an example of an encoded design diagram labeled according to the qualitative encoding schemata. The example chosen is a simple 2-D residential plan drawing of the Farnsworth House, which was designed and completed by Mies van der Rohe. The figure illustrates four stages, where the initial representation of the design diagram is transformed into vectorial format and is then followed by three consecutive stages of encoding. The mapping from physicality to symbol to regularity to 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).

3.2. Categorization using SOMs

SOMs follow a process in which 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. Thus, similarity is therefore considered as the opposite of distance. 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 neighbors.

In the case of identifying styles, SOMs are useful because they are able to adapt neurons in such a way that they serve as good prototypes of the input data for which they are sensitive. Using a winner-take-all network, the input vector is broadcast in parallel to all neurons and for each input vector the most responsive neuron is located. The weights of this neuron and those within a neighborhood around it are adapted to reduce the distance between its weight vector and the current input vector. Thus, 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 neighborhood. The SOM is able to learn to recognize different patterns in the input data and allocate them to appropriate "bins" (styles) in the output array, each bin representing a specific pattern. Therefore, we see the output 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 neighbors represent similar styles and distant neighbors represent different styles.

However, because no one single "correct" answer exists or will ever exist to the central issue of a definition of visual

Table 3. *Tuple symbols: Qualitative syntax for mereotopology*

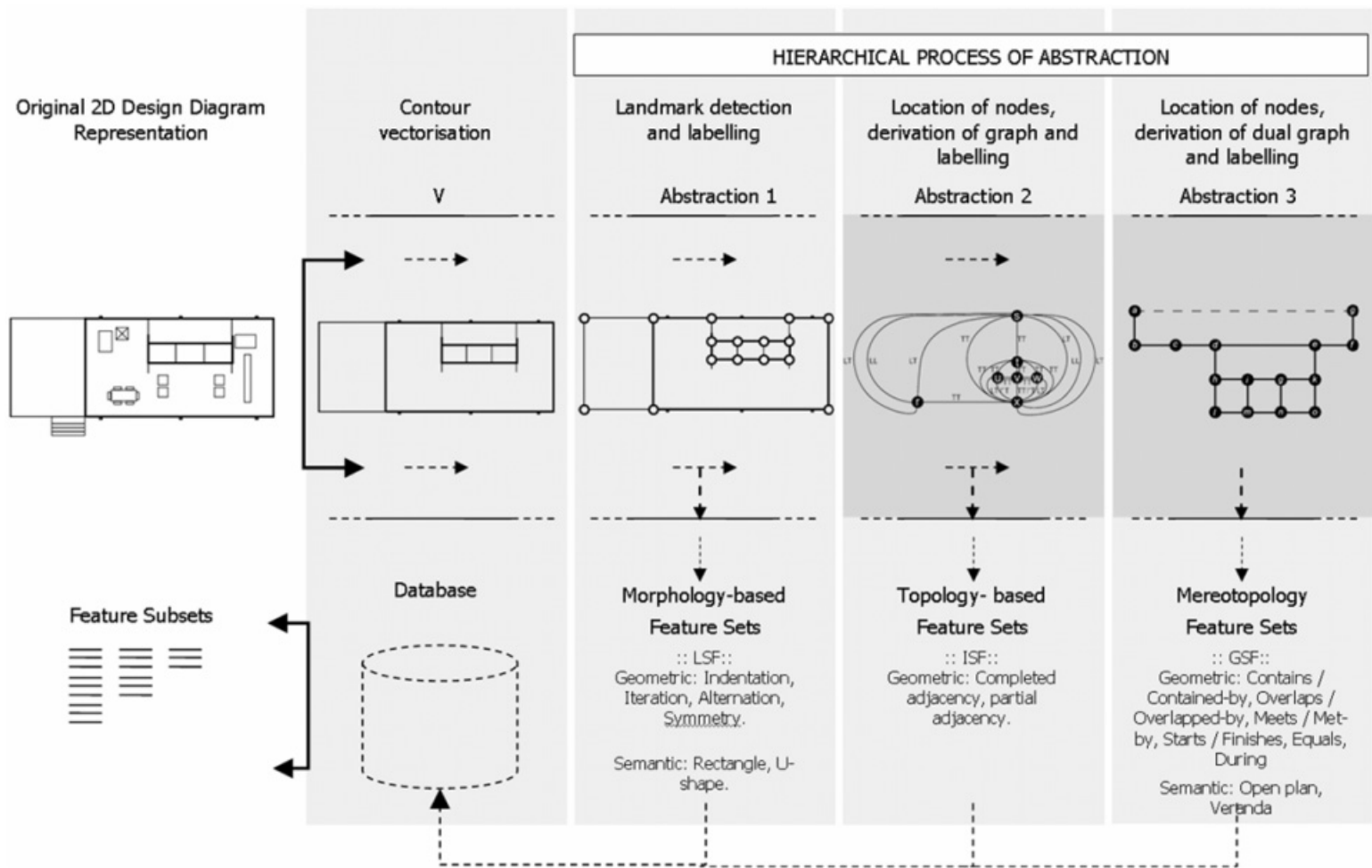|	Adjacency Codes
Numeric value range	$0 \leq \theta \leq 2P \wedge\vee 0 \leq \theta \leq 2P \wedge\vee -\infty \leq 1 \leq \infty$
Landmark set	$\{0, P\} \wedge\vee \{0, P\} \wedge\vee \{-\infty, 0, \infty\}$
Interval set	$\{[0, 0], (0, P), [P, P], (P, 0)\} \wedge\vee \{(-\infty, 0), [0, 0], (0, +\infty)\}$
Q-code set	$\{L, \Gamma, T, \bot, C\} \wedge\vee \{L, \Gamma, T, \bot, C\} \wedge\vee \{>, =, <\}$

Fig. 2. An example of a qualitative abstraction: physical, symbol, regularity, and feature mapping for Mies van der Rohe's Farnsworth House.

design style, the ability to combine the distances calculated in different feature spaces provides the critical point where RF can be incorporated. The SOM's matching process can therefore also be driven by contextual considerations, where the observer is able to determine the relative importance of distinguishing features by adjusting their weights. When contextual information is used for determining the importance of distinguishing features, the correlation between the designer's requirements and the styles identified can therefore increase.

3.2.1. RF approach

RF is the iterative refinement of an initial SOM categorization. It 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 2-D diagrams. In text and image-based research, RF is an established approach that enables contextual dependencies to be integrated for document and image retrieval. This approach has been adopted recently 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 sought to integrate 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 by adjusting the feature weights accordingly. Feature weights in subsequent categorizations 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 and preferences.

The task of assigning specific weights that coincide with the observer's perception of each feature set is not feasible. 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 categorized in a 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 follows, whereby

1. an unsupervised SOM categorizes a design corpus,
2. the first round of results are displayed and stored to avoid the system entering a loop,
3. 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,
4. the adjusted weights are utilized in a reinitialized SOM and the design corpus is recategorized,
5. the second and any subsequent round of results are displayed to the user and stored, and
6. the process continues until the observer is satisfied.

4. IMPLEMENTATION OF Q-SOM:RF

Q-SOM:RF is implemented as five separate modular components: automated encoding and feature extraction, preprocessing and identification of feature subsets, unsupervised categorization, observer feedback, and weight adjustment and updating for recategorization. The following sections describe each of the five stages.

4.1. Decomposition and representation (DeREP) and abstraction and representation (AbREP) operations

DeREP divides the problem into smaller more tightly constrained subproblems 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 labeling occurs as an iterative process: contour traversal, vertex detection, value assignment, contour traversal, and so forth until circuit completion (shape closure). The problem of computing all possible circuits in the diagram so that each circuit contains all vertices exactly once is achieved by finding all Hamiltonian circuits (Garey & Johnson, 1983). A contour cycle (i.e., closed loop) algorithm is implemented in which the agent starts the cycle from each point in the diagram and visits each adjacent vertex exactly once until a closed shape is generated or until a maximum branch limit is reached. This process iterates until all possible shapes 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 then found as the sum of all of the smallest shape units.

As line contours are scanned vertex by vertex, the angle and length magnitudes of the previous line segment become the landmark point for the following segment; that is, 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 angle or length and a ratio is provided to distinguish the relative difference. Figure 3 provides a sample diagram and the resulting closed shapes that are detected.

The pattern recognition system analyzes, locates, and registers specified sequences (chunks) of syntactic structures.
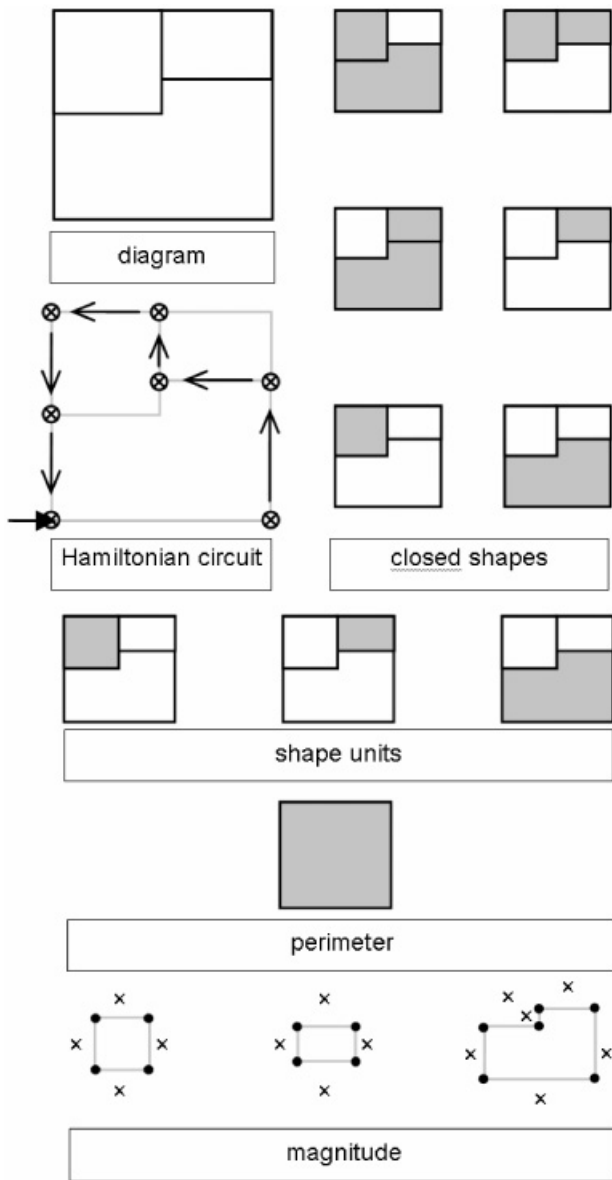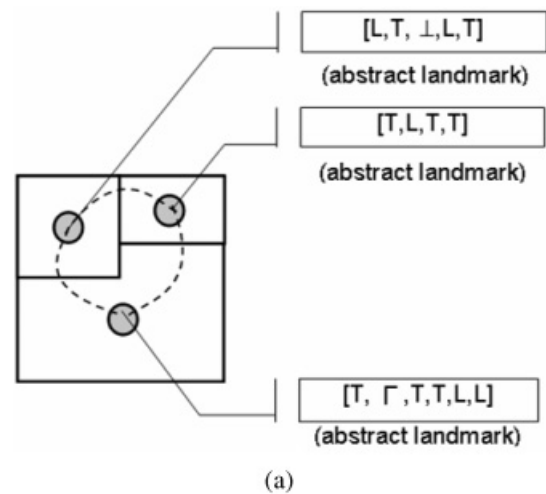
Fig. 3. An original diagram of a Hamiltonian circuit and the resulting closed shapes.

tour traversal, and so forth. AbREP uses an array of symbols describing intersections of line contours, and labeling relies on the data structures built from the previous stage of encoding.

An important aspect of AbREP is that sets of shape-describing arrays are analyzed based on the relationship between each shape. Shape adjacencies are captured using a description of vertex arrays. An "abstract landmark" (Fig. 4a) is created as an array and labeled according to a vertex's specific characteristics. This is achieved by iterating each set of vertices (shared by one line) one at a time and traversing every pair. Because this stage of encoding rerepresents shape adjacencies, arrays are checked, so that if the vertex belongs to only one closed shape, it is ignored (i.e., the external perimeter). 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 4b. This is a reduction process whereby some structural information about the shape is lost.

Fig. 4. Shape encoding for (a) an abstract landmark and (b) shape adjacency relations.

A systematic search for every possible pattern is necessary for the given shape or spatial description. Features already stored in a database identify syntactic patterns via the search and matching process examination of the type and occurrence of chunks, and their structure as a sequence is then labeled (Section 3.1.1, LSF).
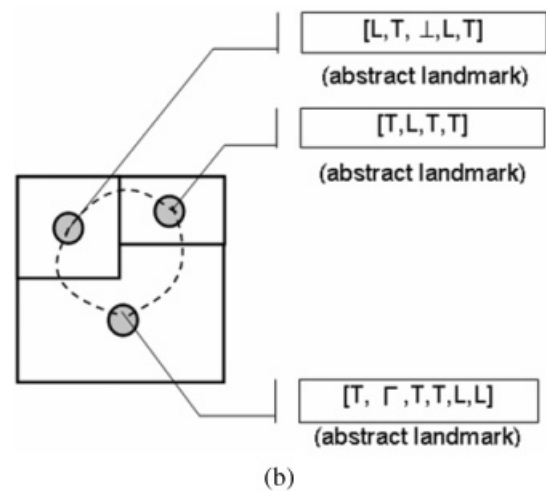
The DeREP process is followed by AbREP, which automates the encoding of all subsequent graph representations. The systematic processing of topological and mereotopological attributes requires the mapping from physicality to symbol to follow a similar conversion process from the graphic state to the symbolic state as implemented in DeREP: contour traversal, vertex detection, value assignment, con-

Once the new arrays are derived and labeled, a representation of adjacencies 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 symbols correspond to specific adjacency relations (Section 3.1.2, ISF). At the implementation level the pattern recognition system locates and registers specified dyad symbol values stored in a database by a set of conditional or if–else rules. The details of these algorithms are available in previous work (Jupp & Gero, 2006). A dyad symbol value is assigned to each set deriving a new label.

Once the new set is formed, the area of a shape is calculated and compared to the area of the adjacent shape to obtain a description of the relative area. As a result, the list of area magnitudes combined with their adjacency types is created for each abstract landmark. From the dyad of symbol values, specific mereotopological relations can be identified that produce another set of codes: tuple symbol values. The system iterates through the set of tuples inspecting their type and, like the previous stage, specified features stored in a database are located and registered by a set of conditional or if–else rules (see Jupp & Gero, 2006). Order outside the tuple itself is not considered and regularities in repetition are no longer determined via chunking because each symbol value has significance (Section 3.1.3, GSF). Using this schema, qualitative descriptions of three different spatial ontologies are represented as distinctive characteristics at the categorical level.

4.2. Preprocessing feature sets

Each type of feature representation identified using the DeREP and AbREP operations (morphological, topological, or mereotopological) can be used to create a meaningful subset of features. At this level of implementation the Q-SOM:RF model has two approaches to preprocessing of feature sets: 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 categorization. These may also be based on the features sets derived from 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 an external or bounding cruciform shape and containing all internal rectangular shapes.

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 subset. CFS evaluates the worth of a set 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 categorization

To interpret, categorize, and visualize the multidimensional data sets obtained from the previous stages, a SOM is implemented. The SOM learns unsupervised and initially categorization begins with a corpus of reference diagrams. This can be expressed by the following.

Let the design corpus, D_C containing k number of reference diagrams, d, to be categorized be equal to:

$$D_C = \{d_1, d_2 \cdots d_k\} \subset C_1, \tag{1}$$

where C_I denotes the initial categories found by the network.

By denoting each design diagram as d_n, $n = 1, 2, \ldots, k$, a feature vector $\mathbf{f}_i$ is associated with each unit i. If there are j different feature representations for each diagram, then,

$$\mathbf{f}^m(d_k) = \mathbf{f}_i^m, \quad m = 1, 2, 3, \ldots, j. \tag{2}$$

The map consists of a regular "city-block" grid of neurons and categorization (to obtain C_I) follows three steps:

1. the distances between the input vector x and all reference vectors (i.e., 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, and
3. the weight vectors corresponding to the winner and the neural units in its topological neighborhood are updated to align them toward the input vector.

Thus, the SOM then attempts to represent the corpus of diagrams with optimal accuracy using the selected subset of features.

4.4. Distinguishing relevant visual styles

The correspondence between high-level concepts and design features can often depend on the context of the observer, and every design categorization is different because of 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 categorization 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 on whether the

observer has selected or rejected the corresponding design classification. The marking operation indicates correctly classified design clusters as positive. Diagrams are accumulated into two sets during the categorization process and weights are adjusted in succeeding iterations, moving from an unsupervised SOM to one that is partially supervised or guided.

4.4.1. 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 (Laaksonen et al. 2000). Because initial categorization 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 deemed to be good candidates for the next categorization.

This stage can be formalized by the corpus' nonintersecting positive d^+ and negative d^- diagrams, and the initial categories [labeled in Eq. (1) as C_I] are defined by

$$\frac{C_I}{d(d^+ \cup d^-)}, \tag{3}$$

and categorization of the design corpus most similar to the positive marked designs is then defined as

$$\frac{\min!}{C_I} c = \sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{n=1}^{N+} w_m c_m(\mathbf{f}^m(d_I)\mathbf{f}^m(d_n^+)), \tag{4}$$

where w_m are the weights of individual features and $c_m(\cdots)$ is the Euclidean distance function with feature type $\mathbf{f}^m$. 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 centered around node $i = i(x)$ are updated (Laaksonen et al., 2000).

5. EXPERIMENTS

Two classes of experiments were carried out to assess the utility of the model. The first experiment tested the discriminatory power of qualitative schema combined with the SOM's ability to categorize encoded diagrams using specific feature sets (i.e., Q-SOM without RF). The second experiment tested the complete system (i.e., Q-SOM:RF) in which the relevance of categorizations is provided in an observer's feedback in the context of a design scenario.

5.1. Design corpus

The design corpus used in all experiments consists of 2-D architectural design diagrams. The corpus is relatively large,

totaling 131 diagrams and representing six architects: Palladio, Frank Lloyd Wright, Mies van der Rohe, Le Corbusier, Louis Kahn, and Mario Botta. The level of complexity of the corpus is considered to be relatively high because, although all diagrams are from a single domain (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.

The two studies undertaken used networks that have been trained using 36 diagrams, which comprises six designs randomly selected from each of the six architects. Exemplar diagrams from each architect and a sample of the feature sets extracted (as raw unprocessed data) are provided in Table 4.

There were 37,367 features extracted from the design corpus and an average of 287 features from 59 sets associated with each diagram. The characteristics of the feature sets in relation to each architect are shown in Table 5.

5.2. Experiment 1: Q-SOM

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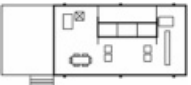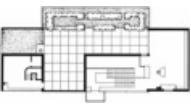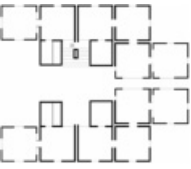
5.2.1. Preprocessing

Preprocessing of input data was undertaken using the statistical feature selection method outlined in Section 4.2, and we utilized CFS to evaluate subsets of features by the correlation among them. In the first study we used only a rerepresentation of morphology (i.e., LSF) extracted from the corpus for dimensionality reduction and eight LSFs were identified as significant by CFS: protrusion-0, protrusion-3, iteration-2, alternation-1, symmetry, square, cruciform, and niche. From this point, we refer to all networks created using this subset as SOM_L.

In the second study, dimensionality reduction included all feature sets extracted using the DeREP and AbREP operations. In addition to the eight LSFs identified above, four GSFs (contains/contained-by, overlaps/overlapped-by, equals, and courtyard) were evaluated as an optimal subset of attributes for clustering. It is interesting that no ISFs were identified. Categorization therefore relies on a combination of feature classes in which the ratio of local to global features is 2:1. We refer to all created networks using this subset as SOM_{L+G}.

5.2.2. Training

For all Q-SOM experiments, two different feature vector models were constructed using the subsets identified in the previous section. In all SOM_L networks each diagram had an average of 35 features and the final vector model contained 1239 feature instances, which created a unique feature vector for each 2-D plan diagram. In SOM_{L+G} networks

Table 4. *Exemplar diagrams based on Architect*

Exemplar Diagram	Exemplar LSF	Exemplar ISF	Exemplar GSF	Architect	Building Typology	Period
	Alternation, symmetry, square, rectangle, chamber, niche	Complete adjacency, partial adjacency	Meets/met-by, contains/contained-by, overlaps/overlapped-by, starts/finished-by, corridor, portico	Palladio	Residential, public	1528–1580
	Indentation, protrusion, iteration, alternation, square, rectangle, L-shape, T-shape, niche, stepped forward, hearth	Complete adjacency, partial adjacency, offset	Meets/met-by, contains/contained-by, overlaps/overlapped-by, starts/finished-by, equals, corridor, portico	Frank Lloyd Wright	Residential	1888–1959
	Indentation, protrusion, iteration, alternation, symmetry, rectangle, U-shape	Complete adjacency, partial adjacency	Contains/contained-by, overlaps/overlapped-by, starts/finished-by, during, equals, corridor, veranda	Mies van der Rohe	Residential, public (religious, library, theatre)	1912–1958
	Indentation, protrusion, iteration, alternation, symmetry, square, rectangle, U-shape, L-shape, niche, stepped backward	Complete adjacency, partial adjacency	Contains/contained-by, overlaps/overlapped-by, starts/finished-by, during, equals, portico, courtyard	Le Corbusier	Residential, public (religious)	1908–1965
	Indentation, protrusion, iteration, alternation, symmetry, square, rectangle, U-shape, chamber, locked space, niche, gallery, hearth	Complete adjacency, partial adjacency, offset	Meets/met-by, contains/contained-by, overlaps/overlapped-by, starts/finished-by, equals, courtyard, quadrangle	Louis Kahn	Residential, public (religious, library, theatre)	1951–1969
	Indentation, protrusion, iteration, alternation, symmetry, square, rectangle, L-shape, chamber, locked space	Complete adjacency, partial adjacency, offset	Meets/met-by, contains/contained-by, overlaps/overlapped-by, starts/finished-by, equals, courtyard	Mario Botta	Residential	1969–1996

each diagram contains an average of 50 features and the final vector model contained 1812 feature instances comprising feature vectors from 1239 LSF and 573 GSF instances.

In addition to the two different feature vector models, training was also varied in terms of the topology of the network and the number of cycles, where 500, 1000, and 1500 training cycles were used. Table 6 presents the training variables for each Q-SOM.

Neither SOM_L nor SOM_{L+G} utilize any other information about the diagram, such as the architect, building type, period, and so forth. Because there is no access to prior knowledge regarding the number of clusters in the data, the SOM proceeds unsupervised.

Based on a visual inspection, the 5×5 SOM_L trained for 500 cycles performed the best and resulted in the clusters shown in the topographic map in Figure 5. The results demonstrate that categorization of diagrams can be roughly linked to the architect as indicated by the map and labeled key in Figure 5, which shows each architect represented by a single letter label. A node may represent more than one diagram, but with different activation values. In some cases the node contains two architects (approximately 20%) and each label has been assigned on the basis of the dominant feature vector. It was also observed from the map that the network appears to have clusters distributed separately and corresponding to the same architect, including Wright (B), Kahn (E), and Botta (F).

The 5×5 SOM_{L+G} trained for 500 cycles also resulted in well-defined clusters, as shown in Figure 6. Like the results obtained for the 5×5 SOM_L, the results observed in the topographic map show that categorization of diagrams

Table 5. *Characteristics of each feature set based on Architect*

Type (Architect)	Ave. No. Features	Total No. Features	No. Diagrams
Palladio	254	3,810	15
Frank Lloyd Wright	306	18,360	61
Mies van der Rohe	268	4,288	16
Le Corbusier	221	1,547	7
Louis Kahn	327	6,213	19
Mario Botta	243	3,159	13

Table 6. *Characteristics of different Q-SOM networks used in diagram classification*

Vector Models		Network Characteristics	
Feature Subset	No. Input Nodes	Topology	Training Cycles
SOM_L	8	$3 \times 3, 5 \times 5, 10 \times 10$	500, 1000, 1200
SOM_{L+G}	12		

can be linked to the architect. The topological ordering of the diagrams in the 5×5 SOM_{L+G} shows a better result than obtained for the 5×5 SOM_L training. This is evident from the separate clusters and the distinctive change of clusters across the map, where Kahn's diagrams (E) are located in the upper left-hand corner of the map and the architect gradually changes toward the bottom-right corner to Le Corbusier (D). Although the 5×5 SOM_{L+G} also distributed two clusters for Wright's designs (B), clustering is more consistent across individual architects.

Significantly, in the 5×5 SOM_{L+G} 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. The SOM_L input vectors also have much lower activations when compared to the SOM_{L+G} ones. Testing was carried out to evaluate the clustering effectiveness of the trained networks. The objective of testing was to evaluate the success of each trained network using the two different approaches to constructing feature vectors (i.e., manual versus CFS selection).

5.2.3. Testing

The SOM_L and SOM_{L+G} networks were tested and their clustering ability was observed. As in training, the 3×3, 5×5, and 10×10 maps were all tested.

To analyze the results of categorization between the topographic maps, we utilized techniques from conventional text-based categorization analysis including Precision (Slonim et al., 2002), the Jacaard (JAC) method (Downton & Brennan, 1980), and the Fowlkes–Mallows (FM) method (Fowlkes & Mallows, 1983). Because classification was unsupervised, it was not possible to apply these evaluation methods directly as would be the case for supervised learning. To analyze the results of testing unsupervised SOMs, it is necessary to utilize the most dominant label of each cluster (obtained during training) for all diagrams. For this reason, the labels (architects) identified from training were maintained in order to assign categories. The "microaveraged" precision matrix method (Slonim et al., 2002) was first used to evaluate each network:
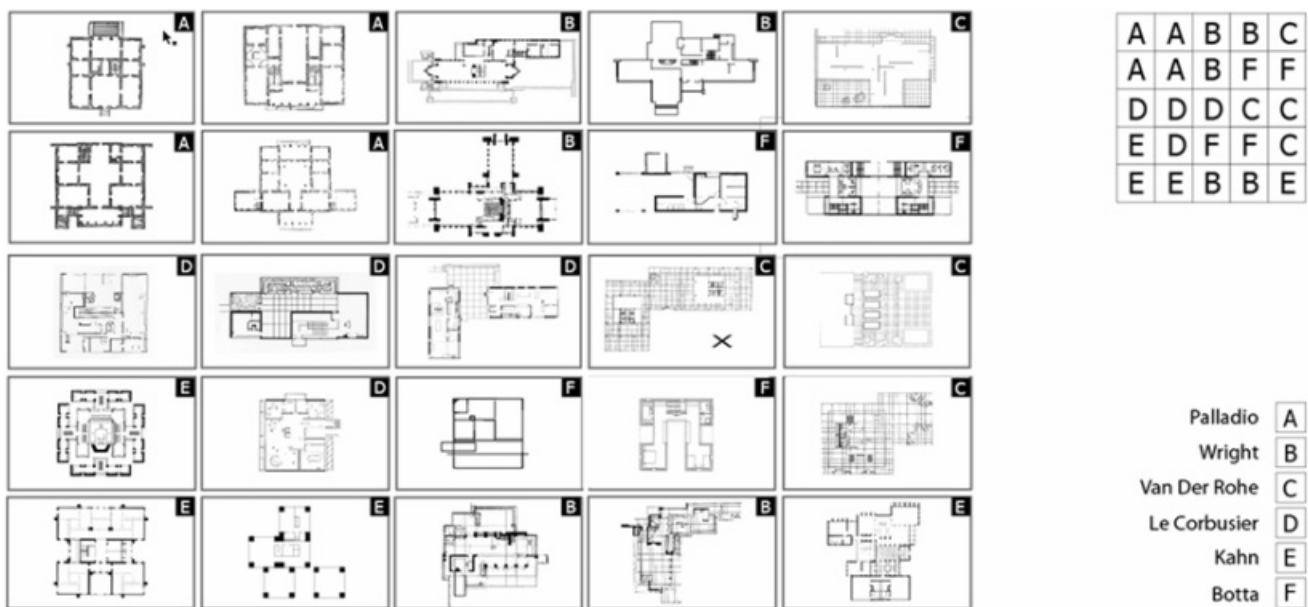
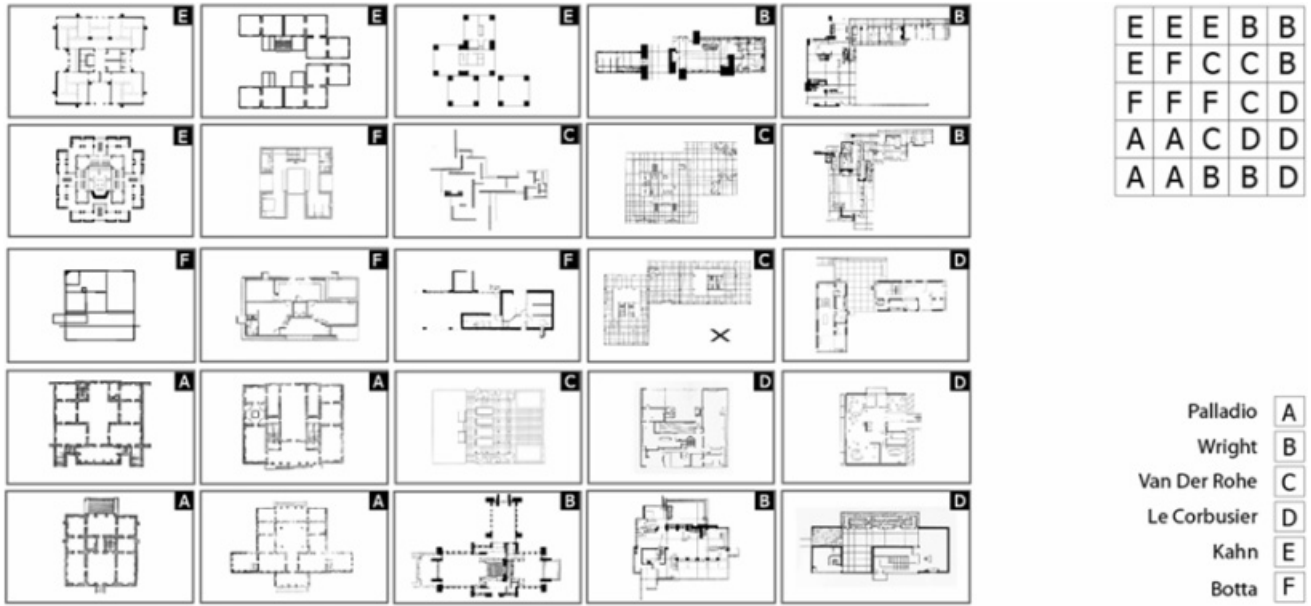**Fig. 5.** Training results for SOM_L clustering.

Fig. 6. Training results for SOM$_{L+G}$ clustering.

$$P(D) = \frac{\sum\limits_{c'} \alpha(C_\mathrm{I}, d)}{\sum\limits_{c'} \alpha(C_\mathrm{I}, d) + \beta(C_\mathrm{I}, d)}, \qquad (5)$$

where for each category $C_\mathrm{I} \in D_C$, $\alpha(C_\mathrm{I}, d)$ is defined by the number of diagrams correctly assigned to C_I, and $\beta(C_\mathrm{I}, d)$ defines the number of diagrams incorrectly assigned to C_I.

The well-established JAC and FM methods were also used to evaluate cluster quality:

$$\mathrm{JAC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \qquad (6)$$

and

$$\mathrm{FM} = \frac{\mathrm{TP}}{\sqrt{(\mathrm{TP} + \mathrm{FP}) \times (\mathrm{TP} + \mathrm{FN})}}, \qquad (7)$$

where TP is equal to the number of diagrams that the classifier correctly classified as belonging to a category (i.e., true positive), FP is the number of diagrams that the classifier classified as belonging to that category (i.e., true positives and false positives), and FN is the pairwise number of false negatives.

The three topologies of SOM$_L$ were tested and each network's ability to categorize the entire design corpus was analyzed. The 5×5 map was found to have the best results for all evaluation techniques measured, as shown in Table 7, with Precision and JAC results being comparable. The FM results also showed how the 5×5 map outperformed both the 3×3 and the 10×10 maps.

Then, we tested Q-SOM$_{L+G}$. Again the 5×5 map had the best results for all evaluation techniques measured. As expected, 5×5 SOM$_{L+G}$ produced better results for precision, JAC, and FM than SOM$_L$ (see Table 8).

5.2.4. Cluster evaluation

The nature of the categories produced by the two best performing networks (5×5 SOM$_L$ and SOM$_{L+G}$) are dif-

Table 7. *Clustering ability of different map topologies trained on SOM$_L$ feature subsets*

Study 1: Q-SOM	Precision	JAC	FM
SOM$_L$ 3×3	0.49	0.29	0.37
SOM$_L$ 5×5	0.62	0.38	0.45
SOM$_L$ 10×10	0.53	0.32	0.30

Table 8. *Clustering ability of different map topologies trained on SOM$_{L+G}$ feature subsets*

Study 2: Q-SOM	Precision	JAC	FM
SOM$_{L+G}$ 3×3	0.61	0.46	0.42
SOM$_{L+G}$ 5×5	0.74	0.53	0.50
SOM$_{L+G}$ 10×10	0.60	0.46	0.39

ficult to evaluate, except via visual (subjective) processes. Recently, conventional clustering techniques (e.g., K-means, EM, hierarchical, etc.) have been used to resolve this problem. Ahmad and Vrusias (2004) demonstrate the effectiveness of using conventional statistical clustering techniques in evaluating the output of maps of unsupervised networks. Such sequential clustering, which is first clustering using an unsupervised network and then clustering the output map, facilitates visualizing clusters that are otherwise implicit in the output map.

We used a sequential clustering method, Q-SOM followed by K-means, to examine the categories obtained. An application of K-means clustering on the output of the 5×5 SOM_L and SOM_{L+G} maps shows how they found data in the proximate types that are approximately defined by both architect and building type. Table 9 compares K-means clustering of all 131 plan diagrams for the 5×5 SOM_L and SOM_{L+G} networks. The table shows the distribution of plan diagrams where the feature vector models can be used to cluster diagrams according to an architect and their residential or public building types. Using the sequential clustering method, Table 9 shows that clustering of the design corpus based on the LSF subset (SOM_L) did not prove to be as defined. Diagrams associated with both architect and building type (represented using the combined LSF and GSF subset) shown in Table 9 were generally well clustered, except for Le Corbusier's designs where no distinct clusters are distinguishable. It is significant for both SOM_L and

Table 9. *Distribution of 131 plan diagrams using K-Means clustering on SOM_L and SOM_{L+G}*

Arch. & Bldg. Type	A	B	C	D	E	F	G	H	I	J
Clusters by K-Means on SOM_L										
A Palladio res. & pub.	9							2	3	1
B Wright res. (Prairie)	6	19	2		1			5		1
C Wright res. (Usonian)		7	14		4					
D Mvd Rohe res.	1		2	3	1	1				1
E Mvd Rohe Pub.			1		4			1	2	
F Le Corbusier Res.				1		3				
G Le Corbusier Pub.						3	0			
H Kahn Res.				2		1		8		
I Kahn Pub.		1		1	1				5	
J Botta Res.	1			2	1	2				7
Clusters by K-Means on SOM_{L+G}										
A Palladio res. & pub.	12				1			1		1
B Wright res. (Prairie)	1	25	2		1			5		1
C Wright res. (Usonian)		4	19		1		2	1		
D Mvd Rohe res.	1		1	6		1				
E Mvd Rohe pub.					5			1	2	
F Le Corbusier res.				1		3				
G Le Corbusier pub.				1		2				
H Kahn res.						1		10		
I Kahn pub.	1		1						7	
J Botta res.						2				11

SOM_{L+G} networks that Wright's designs are distinguished relative to two periods of Wright's work, the Prairie and Usonian houses, where clustering defines 87% of Wright's Prairie design diagrams and 84% of his Usonian.

5.3. Experiment 2: Q-SOM:RF

Based on the results obtained from the 5×5 SOM_{L+G}, the final experiment trained and tested a network's ability to categorize the same corpus using the complete Q-SOM:RF model to obtain clusters that are relevant to some design context. Preprocessing was again used in dimensionality reduction to create feature vectors, and categorization proceeded as a sequential process based on manual selection of diagrams. In this experiment, categorization was evaluated in the context of a design task where the observer (J.J.) provided positive and negative values to the units of the network.

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. The brief specified alterations and additions of an existing residential design to increase sleeping and living spaces, according to the following specifications of building layout: additional sleeping areas to accommodate two children; larger lounge, dining, and kitchen areas; and an outdoor living area.

A conceptual design sketch was then produced using the systems' digital drawing interface that constrains sketching to the orthogonal axis. Figure 7a presents the design sketch produced as a result of the brief's requirements, and Figure 7b shows the sketch as interpreted by the vectorization process of the system. The sketch was encoded (all labels are ignored) and included in the design corpus with the other 131 plan diagrams.

5.3.2. Preprocessing

In this experiment, the original (raw) feature set was processed manually. Manual processing was utilized because it enables the selection of a feature subset relevant to the design task. Because it was the aim for categorization to be grounded by the design task, the visual style of diagrams could be determined based on those features extracted from the design sketch shown in Figure 7. Therefore, from the 59 possible feature classes, selection of one or many feature subsets to create vector models was facilitated from the encoding of the design sketch. Salient features were identified from the local, intermediate, and GSFs contained within the design sketch. The feature classes within the design sketch were then used to create feature vector models to classify the design corpus. A total of 31 feature classes were extracted from the design sketch, which were distributed between local, intermediate, and global feature classes as 22, 3, and 6, respectively. All features extracted from the design sketch are provided in Table 10.
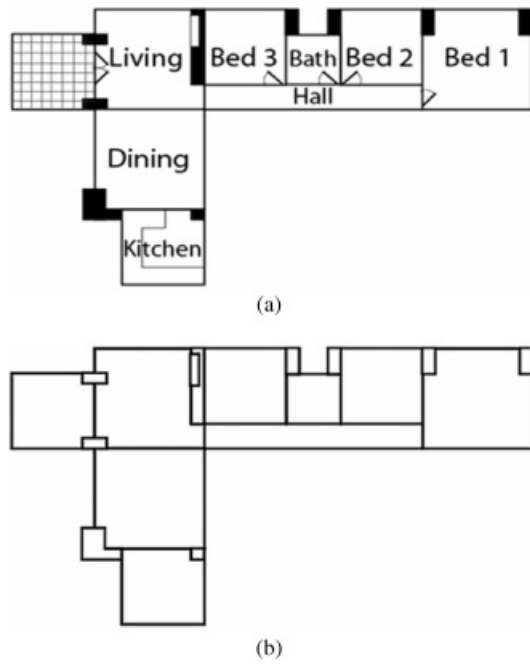
Fig. 7. The design sketch: (a) an original sketch of a residential dwelling and (b) contour vectorization.

Using the reduced feature subset, it is also possible to identify other feature subsets that may be deemed to be more salient in relation to design requirements. To demonstrate the utility of selecting a user-specific subset from the design sketch, a feature subset is shown in the final column of Table 10, where 11 feature types have been selected based on the preferences of the designer. In addition, three other feature subsets were selected: LSF, ISF, and GSF.

Five feature vector models were then constructed and network training was also varied in terms of the number of

Table 10. *Reduced feature subset and user subset*

Features Extracted by DeREP & AbREP		Subset
LSF (geometry based)	Indentation 0, 1, 2, 3	✗, ✗, ✔, ✗
	Protrusion 0, 1, 2, 3	✗, ✗, ✔, ✔
	Iteration 0, 1, 2, 3	✗, ✗, ✗, ✗
	Alternation 0, 1, 2, 3	✗, ✗, ✔, ✗
	Symmetry 0, 1, 2, 3	✗, ✗, ✗, ✗
LSF (domain-based semantics)	Stepped forward	✔
	Niche	✔
ISF (geometry-based)	Complete adjacency	✗
	Partial adjacency	✔
	Offset	✔
GSF (geometry based)	Contains/contained-by	✔
	Overlaps/overlapped-by	✗
	Starts finished-by	✗
GSF (domain-based semantics)	Corridor	✔
	Portico	✔
	Stepped	✗

cycles, where 500, 1000, and 1500 training cycles were again used. Table 11 shows the final training variables for each of the five qualitative feature-based SOMs.

5.3.3. *Training and testing using RF*

All five feature vector models were trained using the 5×5 topography. However, before testing could commence it was necessary to make explicit the targeted categories. In order to demonstrate and then evaluate the performance of the five feature vector models in conjunction with RF, a category of designs had to first be defined within the corpus as the desired target(s). We selected two targets: Wright's Usonian period and Kahn's residential designs. The targets contained 27 and 11 plan diagrams, respectively. Each category was selected based simply on observer preferences. Neither target was necessarily more correct or valid than any other potential category of designs. However, because Wright's Usonian and Kahn's residential designs could now be explicitly targeted by the observer using RF, it was then possible to evaluate how well the system refined sequential categorization.

Based on positive and negative feedback, each tested SOM resulted in a recategorization of the corpus where the iterative process continued until the observer 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. Results from testing the five networks showed that the SOM that utilized "All Features" performed the best, and provided well-defined clusters. The remaining networks also resulted in well-defined clusters; however, the categorizations observed (on visual inspection) in these topographic maps did not appear to be as well defined.

The formal evaluation methods used here rely on JAC and FM measures to analyze the results of each SOM. The first map, All Features, returned the highest performance and the second map, "User Subset," also produced comparable results for JAC and FM measures. The performance of the remaining three maps was lower, as shown in Table 12.

The results of the best performing network, the 5×5 All Features, is illustrated in Figure 8, which was trained for a total of 1000 cycles. The figure shows that the categoriza-

Table 11. *Characteristics of different neural network systems used in diagram classification*

Vector Model		Network Characteristics	
Feature Subset	No. Input Nodes	Topology	Training Cycles
All features	31	5×5	500, 1000, 1500
User subset	11		
LSF	22		
ISF	3		
GSF	6		

Table 12. *Clustering ability of different feature vector models*

Experiment 2: Q-SOM:RF	JAC	FM
All features	0.53	0.50
User subset	0.46	0.42
LSF	0.46	0.39
ISF	0.38	0.45
GSF	0.32	0.30

tion of diagrams can be roughly linked to two architects as indicated by the map and labeled key.

The labeled key in Figure 8 indicates Wright and Kahn. It also shows other architects whose dominant feature vectors defined some labels. Unlike previous results shown in Experiment 1 (Section 5.2), there are multiple nodes of the map that contain input vectors from different architects. It is significant that only 5% of the nodes where Wright and Kahn's diagrams are clustered contain another architect whereas the majority of the remaining nodes contain more than one architect (approximately 50%) and include some of Wright's and Kahn's diagrams.

5.3.4. Cluster and feature subset evaluation

We evaluated the performance of the networks that used the Q-SOM:RF process using a method that resembles "target testing" developed by Cox et al. (1999). Here, instead of a single target, testing evaluated the two targeted categories: Wright's Usonian and Kahn's residential designs. To obtain the performance measure (τ), the targeted category (TC) of designs defined by the user's requirements (r) is used. For each diagram in the TC, the total number of different clusters categorized by the network until the final category is reached was recorded. From these data the average number of clusters formed before the final correct

response was divided by the total number of diagrams (k). The performance measure of the TC was then given by

$$\tau = \left[\frac{\varphi(C, A)}{2}, 1 - \frac{\varphi(C, A)}{2} c \right], \qquad (8)$$

where $\varphi(C, A)$ is the *a priori* probability of the TC, given by TC/k. In general the smaller τ is (<0.5), the better the performance.

The results of the performance measures for all five networks are provided in Table 13. The two feature subsets containing All Features and User Subset yielded better results than the LSF, ISF, or GSF subsets, which can be observed in the first two rows of the table.

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. Based on all performance measures, we observe that using more or all feature classes to create feature vectors yields better results than any one single feature class. Thus, a combination of all available morphological, topological, and mereotopological features in conjunction with RF resulted in the highest performance measure.

The implicit weight adjustments based on the relative importance of features contained in the diagrams showed that the model was capable of categorization using both geometric and semantic attributes contained in the corpus. This kind of automatic adaptation was desirable because it was generally not known which features would perform best in clustering the complex visuospatial information inherent in architectural diagrams.

The experiments demonstrated that utilizing Q-SOM:RF as a system for assessing the similarity of design diagrams to distinguish visual style not only provides a useful method for initial unsupervised categorization but also provides the flexibility to overcome a variety of problems resulting from context. The approach in this experiment provided a robust
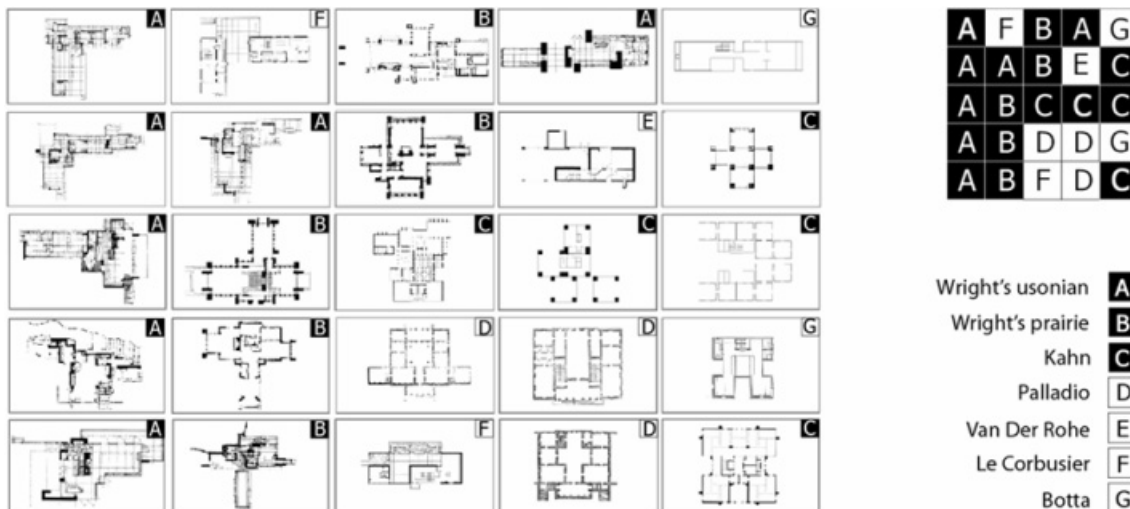

Fig. 8. The final categorization formed by All Features for Q-SOM:RF.

Table 13. *Resulting τ values in the Q-SOM:RF experiment*

Feature Subset	Target Cluster	
	Usonian	Kahn (Residential)
All features	0.23	0.26
User subset	0.29	0.31
LSF	0.36	0.33
ISF	0.54	0.62
GSF	0.39	0.43

method for defining the visual style based on a feature space that was capable of adapting to the contextual relevance of multiple spatial ontologies defined by both lower level geometric and higher level semantic features.

6. CONCLUSION AND FUTURE PLANS

We introduced the Q-SOM:RF approach to assessing visual styles by incorporating qualitative and contextual aspects in a formal model. Our approach to visual style focuses on modeling both a cognitively compatible and task-relevant clustering method. The model was shown to be capable of bridging the gap between the low-level visual features extracted from the design artifact and the high-level and context-dependent semantics used by designers to understand design content. 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 showed that the approach of the system was able to effectively select from sequential and iterative processing by adjusting weights of a SOM to coincide with an observer's conceptual view of diagrammatic similarity. In this way, initial unsupervised clustering can become more accurate according to a designer's intentions as new clusters are formed.

The model's strength lies in its utilization of a multidimensional qualitative encoding schema and its ability to simultaneously assess multiple reference diagrams. By providing a mapping from the physicality of the diagram using qualitative rerepresentation, geometric and semantic feature sets were obtained; as a result, meaningful feature subsets could be analyzed as input to the SOM. The RF technique makes the model differ from other content-based classification systems, such as feature-based distance measures of similarity. Assessment using such measures is based on processing only a single reference diagram at a time. Further, because an individual visual feature or even a class of features based on any one spatial ontology may not be sufficient in categorizing complex design diagrams, it is necessary to extract information from multiple dimensions so that the adaptation of categories relative to feedback and context are useful. Using target categories in the final experiment demonstrated that when contextual information was integrated to determine the relative importance of features, the correlation between the system's results and the observer's assessments increased. It was significant that this correlation resulted from the detailed definition, detection, and extraction of feature classes and user feedback determining feature salience.

The approach provided 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 his or her design objectives and intentions. A 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 judgment of a feature, a diagram, and, as a result, a design corpus.

It is such complex aspects of context in relation to visuospatial information and cognition that also require further exploration in modeling visual style in design. Further, quantitative evaluation of the Q-SOM:RF model's performance are problematic because of human subjectivity and preference during the classification process. Although the results obtained thus far are promising (in terms of the potential of the model as a cognitively plausible classification system), additional studies are required in a variety of design scenarios and based on multiple users. We anticipate that in human-subject experiments the Q-SOM:RF system should perform equally well using either expert or novice designers. The performance of the model should not differ substantially between these two groups because their unique similarity assessments will be reflected by the model's flexibility to adapt to feedback. In such future studies, the model must also undergo further testing so that the subsets used to create feature vector models can be inspected relative to historically defined design styles. Therefore, the initial objective of future research is to increase the performance of the Q-SOM:RF model to find a set of well-balanced features, which on average, perform as well as possible. In this way, feature centrality can be investigated in the design domain relative to the effects of context as well as existing historically defined styles.

REFERENCES

Ackerman, J.S. (1967). A theory of style. In *Aesthetic Inquiry: Essays on Art Criticism and the Philosophy of Art* (Beardsley, M.C., & Schueller H.M., Eds.), pp. 54–66. Belmont, CA: Dickenson.

Ahmad, K., & Vrusias, B. (2004). Learning to visualise high-dimensional data. *Proc. 8th Int. Conf. Information Retrieval*.

Allen, J.F. (1984). Towards a general theory of action and time. *Artificial Intelligence*, *23(1)*, 123–154.

Brown, K., Sims, N., Williams, J.H., & McMahon, C.A. (1995). Conceptual geometric reasoning by the manipulation of models based on prototypes. *Artificial Intelligence in Engineering Design, Analysis and Manufacturing*, *9(5)*, 367–385.

Burns, K. (2004). Creature double feature: on style and subject in the art of caricature. *AAAI Fall Symp. Style and Meaning in Language, Art, Music, and Design*. Washington, DC: AAAI Press.

Cohn, A.G. (1997). Qualitative spatial representation and reasoning techniques. *Proc. KI-97* (Brewka, G., Habel, C., & Nebel, B., Eds.), pp. 1–30. New York: Springer–Verlag.

Colagrossi, A., Sciarrone, F., & Seccaroni, C.A. (2003). Methodology for automating the classification of works of art using neural networks. *Leonardo*, *36(1)*, 69–96.

Cox, I.J., Miller, M.L., Omohundro, S.M., & Yianilos, P.N. (1999). Target testing and the PicHunter Baysian multimedia retrieval system. *Proc. 3rd Forum on Research Technology Advances in Digital Libraries*, pp. 66–75, Washington, DC.

Davies, J., & Goel, A.K. (2001). Visual analogy in problem solving. *Proc. Int. Joint Conf. Artificial Intelligence*, pp. 37–382. San Francisco, CA: Morgan Kaufmann.

Ding, L., & Gero, J.S. (1998). Emerging representations of Chinese traditional architectural style using genetic engineering. *Int. Conf. Artificial Intelligence for Engineering* (Huang, X., Yang, S., & Wu, H., Eds.), pp. 493–498. Wuhan, China: HUST Press.

Downton, M., & Brennan, T. (1980). Comparing classifications: an evaluation of several coefficients of partition agreement. *Proc. Meeting of the Classification Society*, Boulder, CO.

Edwards, A.T. (1945). *Style and Composition in Architecture*. London: John Tiranti.

Forbus, K., Tomai, E., & Usher, J. (2003). Qualitative spatial reasoning for visual grouping in sketches. *Proc. 16th Int. Workshop on Qualitative Reasoning*, pp. 79–86, Brasilia.

Fowlkes, E., & Mallows, C. (1983). A method for comparing two hierarchical clusterings. *Journal of the American Statistical Association*, *78(4)*, 553–569.

Garey, M.R., & Johnson, D.S. (1983). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York: W.H. Freeman.

Gero, J.S., & Jupp, J. (2003). Feature-based qualitative representations of plans. *CAADRIA03*, pp. 117–128, Rangsit University, Bangkok.

Gero, J.S., & Kazakov, V. (2001). Entropic similarity and complexity measures for architectural drawings. *Visual and Spatial Reasoning in Design II* (Gero, J.S., Tversky, B., & Purcell, T., Eds.), pp. 147–161, University of Sydney.

Gero, J.S., & Park, S.-H. (1997). Qualitative representation of shape and space for computer-aided architectural design. *CAADRIA'97* (Liu, Y.-T., Tsou, J.-Y., & Hou, J.-H., Eds.), pp. 323–334. Taipei, Taiwan: Hu Publishers.

Gross, M., & Do, E. (1995). Drawing analogies—supporting creative architectural design with visual references. *3rd Int. Conf. Computational Models of Creative Design* (Maher, M.-L., & Gero, J., Eds.), pp. 37–58, University of Sydney.

Güsgen, H.W. (1989). *Spatial Reasoning Based on Allen's Temporal Logic*. Technical Report TR-89-049. Berkeley, CA: International Computer Science Institute.

Hall, M. (2000). Correlation-based feature selection for discrete and numeric class machine learning. *17th Int. Conf. Machine Learning*, pp. 359–366.

Honkela, T., Kaski, S., Kohonen, T., & Lagus, K. (1998). Self-organizing maps of very large document collections: justification for the WEBSOM method. In *Classification, Data Analysis, and Data Highways* (Balderjahn, I., Mathar, R., & Schader, M., Eds.), pp. 245–252. Berlin: Springer.

Jupp, J. (2006). *Diagrammatic reasoning in design: computational and cognitive studies in similarity assessment*. PhD Thesis. University of Sydney, Key Centre of Design Computing and Cognition.

Jupp, J., & Gero, J.S. (2004). Qualitative representation and reasoning in design: a hierarchy of shape and spatial languages. *Visual and Spatial Reasoning in Design III* (Gero, J.S., Tversky, B., & Knight, T., Eds.), pp. 139–163, University of Sydney, Key Centre of Design Computing and Cognition.

Jupp, J., & Gero, J.S. (2005). Cognitive studies in similarity assessment: evaluating a neural network model of visuospatial design similarity. *Int. Workshop on Human Behaviour in Designing (HBiD05)*, Aix en Provence, France, September 2005.

Jupp, J., & Gero, J.S. (2006). Towards computational analysis of style in architectural design. *Journal of the American Society for Information Science*, *57(5)*, 45–61.

Knight, T.W. (1994). *Transformations in Design. A Formal Approach to Stylistic Change and Innovation in the Visual Arts*. Cambridge: Cambridge University Press.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43(1)*, 59–69.

Kohonen, T. (1995). *Self Organising Maps*. New York: Springer–Verlag.

Laaksonen, J., Koskela, M., Laakso, S., & Oja, E. (2000). PicSOM—content-based image retrieval with self-organizing maps. *Pattern Recognition Letters*, *21(13–14)*, 1199–1207.

Love, B.C., & Sloman, S.A. (1995). Mutability and the determinants of conceptual transformability. *Proc. 17th Annual Conf. Cognitive Science Society*, pp. 654–659, Pittsburgh, PA.

Marr, D., & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society*, *B200(2)*, 269–294.

Medin, D.L., Goldstone, R.L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100(2)*, 254–278.

Mukerjee, A. (1989). Getting beneath the geometry: a systematic approach to modeling spatial relations. *Proc. Workshop on Model-Based Reasoning, IJCAI-89*, pp. 140–141, Detroit.

Oja, E., Lassksonen, J., Koskela, M., & Brandt, S. (1999). Self-organizing maps for content-based image database retrieval. In *Kohonen Maps* (Oja, E., & Kaski, S., Eds.). New York: Elsevier.

Park, S.-H., & Gero, J.S. (2000). Categorisation of shapes using shape features. *Artificial Intelligence in Design '00* (Gero, J.S., Ed.), pp. 203–223. Dordrecht: Kluwer.

Salton, G., & McGill, M.J. (1983). *Introduction to Modern Information Retrieval*. New York: McGraw–Hill.

Schapiro, M. (1961). Style. In *Aesthetics Today* (Philipson, M., & Gudel, P.J., Eds.), pp. 137–171. New York: New American Library.

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119(1)*, 3–22.

Slonim, N., Friedman, N., & Tishby, N. (2002). Unsupervised document classification using sequential information maximization. *Proc. SIGIR'02, 25th ACM Int. Conf. Research and Development of Information Retrieval*, Tampere, Finland. New York: ACM Press.

Thomas, M.S.C., & Mareschal, D. (1997). Connectionism and psychological notions of similarity. *Proc. 19th Annual Conf. Cognitive Science Society*, pp. 757–762. London: Erlbaum.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84(3)*, 327–352.

Tversky, A., & Gati, I. (1982). Similarity, separability, and the triangle inequality. *Psychological Review*, *89(1)*, 123–154.

Tversky, B. (1999). What does drawing reveal about thinking? *Visual and Spatial Reasoning in Design I* (Gero, J.S., Tversky, B., & Knight, T., Eds.), pp. 271–282, University of Sydney, Key Centre of Design Computing and Cognition.

Wertheimer, M. (1977). Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschung*, *4(3)*, 301–350.

Julie Jupp is a Research Associate in the Engineering Design Centre at the University of Cambridge. She recently completed her PhD in architecture at the Key Centre of Design Computing and Cognition at the University of Sydney. Julie has participated in developing prototype and commercial design tools with the Australian Collaborative Research Centre for Construction Innovation as well as the Commonwealth Scientific and Industrial Research Organisation for Manufacturing and Infrastructure Technology. Dr. Jupp's research work is aimed at the understanding of visuospatial reasoning as a way to improve the practice and education of design.

John Gero is Professor of design science and Co-Director of the Key Centre of Design Computing and Cognition at the University of Sydney. He is the author or editor of 39 books and over 500 papers in the fields of design science, artificial intelligence, optimization, and computer-aided design. Professor Gero is also an international consultant in the field of computer-aided design and artificial intelligence.