The Vocabulary and Grammar of Color Patterns

Aleksandra Mojsilović, *Member, IEEE*, Jelena Kovačević, *Senior Member, IEEE*, Darren Kall,
Robert J. Safranek, *Senior Member, IEEE*, and S. Kicha Ganapathy

***Abstract*—We determine the basic categories and the hierarchy of rules used by humans in judging similarity and matching of color patterns. The categories are**

1) **overall color;**
2) **directionality and orientation;**
3) **regularity and placement;**
4) **color purity;**
5) **complexity and heaviness.**

These categories form the pattern vocabulary which is governed by the grammar rules. Both the vocabulary and the grammar were obtained as a result of a subjective experiment. Experimental data were interpreted using multidimensional scaling techniques yielding the vocabulary and the hierarchical clustering analysis, yielding the grammar rules. Finally, we give a short overview of the existing techniques that can be used to extract and measure the elements of the vocabulary.

***Index Terms*—Color patterns, image databases, retrieval.**

I. Introduction

Together with color and shape, texture is the most important visual category in human perception, and has thus been extensively studied in computer vision, image processing and psychophysics [1]–[12]. By texture, we denote a visual phenomenon (such as grass, marble, brick) caused by the repetition of a structural element according to a certain rule. Textures generated by humans (such as textiles, ornaments or tiles) are usually called *patterns*. To specify that a pattern contains color, we will call it a *color pattern*. Unfortunately, our understanding of color textures and color patterns is very modest compared to our understanding of other visual phenomena such as color, contrast, or even gray-level textures or gray-level patterns. That is mainly due to the fact that the basic dimensions of color patterns have not yet been identified, a standardized vocabulary for addressing their important characteristics does not exist, nor is there a grammar defining how these dimensions are to be combined. Previous investigations in this field concentrated mainly on gray-level natural textures [9]–[11]. Particularly interesting is work of Rao and Lohse [11]: their research focused on how people classify textures in meaningful, hierarchically-structured categories, identifying relevant features used in the perception of gray-level textures. Similarly, here we determine the basic categories—*vocabulary*—used by humans in judging similarity of color patterns, their relative importance and relationships, as well as the hierarchy of rules—*grammar*. The attributes we extract are applicable to a broad range of textures, starting from simple patterns, all the way up to complex, high-level visual texture phenomena.

The paper is organized as follows. The first two sections present the basic concepts of multidimensional scaling and hierarchical clustering techniques used in the study for the analysis of subjective data. The third section describes the methodology of the data collection and analysis. Results are presented in Sections IV and V. A review of the feature extraction techniques for measuring the determined dimensions is given in Section VI. Discussion, conclusions and plans for further research are found in Section VII.

II. Multidimensional Scaling

Multidimensional scaling (MDS) is a set of mathematical techniques that enable researchers to uncover the hidden structures in data [13]. MDS is designed to analyze distance-like data called *similarity* or *proximity* data, that is, data indicating the degree of similarity between two items. Traditionally, similarity data is obtained via subjective measurement. It is acquired by asking people to judge similarity of pairs of objects—*stimuli*—on some scale. The obtained similarity value connecting stimulus i to stimulus j is denoted by δ_{ij}. Similarity values are arranged in a *similarity matrix* Δ, usually by averaging δ_{ij} obtained from all measurements. The aim of MDS is to place each stimulus from the input set into an L-dimensional *stimulus space* (the dimensionality of the space, L, is also determined as the result of the experiment). The points $\mathbf{x}_i = [x_{i1}x_{i2}\cdots x_{iL}]$ (note that this is a row vector) representing each stimulus are arranged so that the Euclidean distances d_{ij} between each pair of points in the stimulus space match as closely as possible the subjective similarities δ_{ij} between corresponding pairs of stimuli.

Here, we give a brief overview of two particular types of MDS used in this work: *classical MDS* (CMDS) and *weighted MDS* (WMDS, also called INDSCAL). CMDS analyzes only one similarity matrix obtained by averaging values from all the subjects. An important characteristic of CMDS is that once a configuration of points is obtained, it can be rotated, implying that the dimensions are not meaningful. Thus, when interpreting the results, higher-dimensional CMDS soon becomes impractical. On the other hand, WMDS analyzes several similarity matrices, one for each subject. This model assumes that individuals vary in the importance they attach to each dimension of the stimulus space. While one individual may perceive one dimension as being more important than another, another individual may have the opposite perception. In that way WMDS accounts

Manuscript received October 30, 1998; revised August 19, 1999. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Patrick Bouthemy.

A. Mojsilović, J. Kovačević, R. J. Safranek, and S. K. Ganapathy are with Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974 (e-mail: saska@research.bell-labs.com; jelena@research.bell-labs.com; rjs@research.bell-labs.com; skg@research.bell-labs.com).

D. Kall is with Bell Laboratories, Lucent Technologies, Holmdel, NJ 07733 USA (e-mail: dakall@lucent.com).

Publisher Item Identifier S 1057-7149(00)01509-8.

for individual differences in human responses. As opposed to CMDS, once the configuration is obtained, it cannot be rotated [31]. However, the stability of configuration depends heavily on the accuracy of the model; if the model fits that data well, the dimensions are meaningful which makes our job of interpreting them much easier.

In the rest of this paper, we will use the following notation.

$n = 1, \cdots, N$	index denoting stimulus n;
$k = 1, \cdots, K$	index denoting subject k;
$l = 1, \cdots, L$	index denoting dimension l;
δ_{ij}	similarity value attached to the pair of stimuli (i, j) obtained from all the subjects;
Δ	$N \times N$ similarity matrix having δ_{ij} as its elements (diagonal elements are ignored);
$\mathbf{x}_i$	row vector in L-dimensional space containing coordinates of the stimulus i in that space;
x_{il}	projection of the ith point (stimulus) along dimension l;
$\mathbf{X}$	$N \times L$ matrix containing x_{il} as its elements (*group configuration matrix*);
d_{ij}	Euclidean distance between $\mathbf{x}_i$ and $\mathbf{x}_j$;
$\mathbf{D}$	$N \times N$ matrix containing d_{ij}^2 as its elements;
δ_{ijk}	similarity value attached to the pair of stimuli (i, j) obtained from subject k;
Δ_k	$N \times N$ similarity matrix having δ_{ijk} as its elements (diagonal elements are ignored);
x_{ilk}	projection of the ith point (stimulus) along dimension l for subject k;
$\mathbf{X}_k$	$N \times L$ matrix containing x_{ilk} as its elements (*individual configuration matrix*);
w_{lk}	weight subject k gives to dimension l;
$\mathbf{W}_k$	$L \times L$ diagonal matrix containing w_{lk} along the diagonal;
d_{ijk}	Euclidean distance between $\mathbf{x}_i$ and $\mathbf{x}_j$ for subject k;
$\mathbf{D}_k$	$N \times N$ matrix containing d_{ijk}^2 as its elements.

A. Classical MDS

The central concept of CMDS is that the distance d_{ij} between points in an L-dimensional space will have the strongest possible relation to the similarities δ_{ij} from a single matrix Δ. The similarities are averaged original ratings obtained as the result of the experiment. The CMDS data analysis problem can be summarized as follows:

$$f(\mathbf{\Delta}) = \mathbf{D} + \mathbf{E} \tag{1}$$

where $\mathbf{D}$ contains d_{ij}^2 as its elements and

$$d_{ij}^2 = (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T. \tag{2}$$

In other words, Δ—the original similarity data—are equal, by transformation f, to the transformed similarity data, which in turn are equal to the obtained squared Euclidean distances $\mathbf{D}$ plus error $\mathbf{E}$. CMDS solves for $\mathbf{D}$ and transformation f so that the norm of $\mathbf{E}$ is minimized.

Most often, f is chosen to be linear

$$f(\mathbf{\Delta}) = a\mathbf{\Delta} + b, \tag{3}$$

where, for a given configuration, values a and b must be discovered using numerical optimization.

CMDS requires f to be defined before going into a computational procedure. Therefore, given an initial configuration $\mathbf{X}_I$, one first finds the best f, yielding $\mathbf{D}_I$ (the question of how to choose the initial configuration is addressed in [12]). Once the best f is found, we then search for the best configuration of points in the stimulus space and iterate. We repeat this procedure for different L's until further increase in the number of dimensions does not bring a reduction in the error functions. Finally, we are left with the task of interpreting and labeling the dimensions we have. Note that the computational procedure works on the points in the configuration, not on distances d_{ij}.

We now explain this in more detail. CMDS starts by defining an *error function* (also called *goodness-of-fit*, *objective function*). For any given set of data and for any given configuration, the error function yields a single number which shows how well the data fit into the configuration. One commonly used error function is referred to as "*stress formula 1*" or "*Kruskal's stress formula*." As explained earlier, we first try to find the best f given a configuration $\mathbf{X_I}$. CMDS does that by minimizing the error function. In principle, we could find the best f for any given $\mathbf{X}$; however, since f is linear, the choice of f is not crucial and we start from any $\mathbf{X}_I$ to obtain f [13]. The error (stress) formula used here is

$$stress(\mathbf{\Delta}, \mathbf{X_I}) = \min_{\text{all} f} fstress(\mathbf{\Delta}, \mathbf{X}_I, f) \tag{4}$$

where

$$fstress(\mathbf{\Delta}, \mathbf{X}_I, f) = \sqrt{\frac{\sum_i \sum_j [f(\delta_{ij}) - d_{ij}]^2}{\sum_i \sum_j f(\delta_{ij})^2}}. \tag{5}$$

Once the objective function f is obtained, we find the "best" configuration, that is, the configuration $\mathbf{X}$ which yields the lowest possible value of the error function

$$stress(\mathbf{\Delta}, \mathbf{X}) = \min_{\text{all}\mathbf{X}_I} fstress(\mathbf{\Delta}, \mathbf{X}_I, f). \tag{6}$$

One widely used procedure for finding the best configuration is the method of steepest descent [13]. Once the configuration is obtained, it is important how it is interpreted. A possible way to interpret each dimension in the resulting configuration is to examine peripheral objects, that is, objects that lay at the outermost edges of configuration. Then, it has to be established what is common to these objects and their nearest neighbors and how they differ from the stimuli at the opposite edges of the configuration. Since the configuration is based on the distances between the points (which do not change with the rotation of the space), rotation in CMDS is permissible leading to drastic changes in the projections. Consequently, it is possible that the coordinate axes which we could not interpret directly can be rotated revealing their true meaning. Usually, we aim to interpret each dimension of the space. However, the number of dimensions does not necessarily reflect all the relevant characteristics. Also, although a particular feature exists in the stimulus set, it may not contribute strongly enough to become visible as a separate dimension. This can be because the selected stimuli do not vary enough on that feature, because this characteristic is correlated with other dimensions, or because it was relevant only to a subset of subjects. The characteristics found by the algorithm are usually only a part of a much longer list of features. Therefore, one useful role of MDS is to indicate which particular features are important.

B. Weighted MDS

WMDS analyzes several similarity matrices, one for each of K subjects. In the WMDS model, δ_{ijk} indicates the similarity between stimuli i and j, as judged by the subject k. The notion of "individual taste" is incorporated into the model through weights w_{kl}, for each subject $k = 1, \cdots, K$ and each dimension $l = 1, \cdots, L$. Just as in CMDS, WMDS determines the configuration of points $\mathbf{X}$, called the *group stimulus space*. However, in order to find the best possible configuration, WMDS does not use distances among the points in the group space. Instead, for each subject k it creates a new configuration $\mathbf{X}_k$, and the distances in this configuration are used for finding an optimal solution. A configuration for each subject is made by altering the group configuration space according to the weights w_{lk}. Algebraically, given x_{il} (projection of the ith point along dimension l) from the group space, the points for subject k are obtained as

$$x_{ilk} = \sqrt{w_{lk}} \cdot x_{il}. \tag{7}$$

As in the CMDS, the WMDS data analysis problem can be summarized as follows:

$$f(\mathbf{\Delta}_k) = \mathbf{D}_k + \mathbf{E}_k \tag{8}$$

where $\mathbf{D}_k$ contains d_{ijk}^2 as its elements and

$$d_{ijk}^2 = (\mathbf{x}_i - \mathbf{x}_j)\mathbf{W}_k(\mathbf{x}_i - \mathbf{x}_j)^T. \tag{9}$$

In WMDS, the formula for stress is based on the squared distances calculated from each of K individual similarity matrices

$$fStress(\mathbf{\Delta}, \mathbf{X}_k, f) = \sqrt{\frac{1}{K}\sum_k \frac{\sum_i \sum_j [f(\delta_{ijk}) - d_{ijk}]^2}{\sum_i \sum_j f(\delta_{ijk})^2}}. \tag{10}$$

III. Hierarchical Cluster Analysis

Given a similarity matrix, hierarchical cluster analysis (HCA) organizes a set of stimuli into similar units [14]. Therefore, HCA helps us discover the rules and the hierarchy we use in judging similarity and pattern matching. This method starts from the stimulus set to build a tree. Before the procedure begins, all stimuli are considered as separate clusters, hence there are as many clusters as there are stimuli. The tree is formed by successively joining the most similar pairs of stimuli into new clusters. As the first step, two stimuli are combined into a single cluster. Then, either a third stimulus is added to that cluster, or two other clusters are merged. At every step, either individual stimulus is added to the existing clusters, or two existing clusters are merged. Splitting of clusters is forbidden. The grouping continues until all stimuli are members of a single cluster. Fig. 3 gives an example: there are 20 stimuli, each one being one cluster. The procedure ends with all the stimuli being members of a single cluster—cluster 39.

How the similarity matrix is updated at each stage of the tree is determined by the joining algorithm. There are many possible criteria for deciding how to merge clusters. Some of the simplest methods use *nearest neighbor technique*, where the first two objects combined are those that have the smallest distance between them. At every step, the distance between two clusters is obtained as the distance between their closest two points. Another commonly used technique is the *furthest neighbor technique* where the distance between two clusters is obtained as the distance between their furthest points. The *centroid* method calculates the distances between two clusters as the distance between their means. Note that, since the merging of clusters at each step depends on the distance measure, different distance measures can result in different clustering solutions for the same clustering method [14].

Clustering techniques are often used in combination with MDS, to clarify the dimensions and interpret the neighborhoods in an MDS configuration. However, in the same way as with the labeling of the dimensions in the MDS algorithm, interpretation of the clusters is usually done subjectively and strongly depends on the quality of the data.

IV. Experimental Setup and Data Analysis

A. Selection of Stimuli

We used 25 patterns from an interior design catalog. Twenty patterns were used in the actual study. Five patterns were used as a "warm-up" before each trial. This allowed the subjects to get comfortable with the testing procedure and to sharpen their own understanding of similarity. The digitized version of the 20 patterns selected are displayed in Fig. 1. We selected patterns that capture a variety of different image features and their combinations. As previously mentioned, the selection of stimuli is crucial for MDS. Since we postulated that visual similarity needs to be modeled by a higher number of dimensions, it was vital for this experiment to select the stimuli so that there is sufficient variation of potential dimensions.

B. Subjects and Ranking Procedure

Twenty-eight subjects (15 male and 13 female) participated in the study. They were selected from the staff members at Bell Labs with no background in image processing. The subjects had a mixture of technical and nontechnical background. Their ages ranged from 20 to 70. All the subjects had full color perception, that is, there were no instances of partial or full color-blindness.

The subjects were not familiar with the input data. They were presented with all 190 possible pairs of stimuli. For each pair, the subjects were asked to rate the degree of overall similarity on a scale ranging from 0 for "very different" to 100 for "very similar." There were no instructions concerning the characteristics on which these similarity judgments were to be made since

Fig. 1. Pattern set used in the experiment. They are obtained from an interior design catalog. Twenty were selected capturing a variety of different features. Another five were used as a "warm-up" in the study.

this was the very information we were trying to discover. The order of presentation was different for each subject and was determined through the use of a random number generator. This was done to minimize the effect on the subsequent ratings of both the same presentation order for all the subjects (group effect) as well as the presentation order for one subject (individual effect).

At the end of experiment, half of the subjects were presented with pairs they thought the most similar, and asked to explain why. Their explanations were used later as an aid in the interpretation of the MDS configurations.

C. Data Analysis

The first step in the data analysis was to compute the mean similarity rating for each of 190 pairs. These mean ratings were arranged into a similarity matrix Δ (given in Table I) to be an input to CMDS. Also, WMDS procedure was applied to the set of 28 individual similarity matrices. Note that in our experiment subjects' rankings represent similarities. However, since MDS methods are based on the idea that the scores are proportional to distances, it was necessary to preprocess the collected data according to the following relation:

$$dissimilarity = 100 - similarity. \qquad (11)$$

CMDS was performed in two and three dimensions, and WMDS was performed in two, three, four, five and six dimensions. We used the built-in MDS functions from the S+ and SPSS packages [31], [32].

There are important reasons why WMDS is used with higher number of dimensions. For WMDS, in most cases coordinate axes are meaningful. CMDS, since it cannot be rotated, yields configurations where axes of the space are not necessarily meaningful. Therefore, to label and interpret the dimensions, we first have to discover their orientation. As we go to higher-dimensional space. this becomes almost impossible due to our inability to visualize the configuration. Furthermore, due to the weighted Euclidean model, WMDS accommodates very large differences among the individual ratings, and even very different data from two subjects can fit into the same space.

When using MDS in psychophysical experiments, certain issues need to be addressed. First, for each individual, the scale is initially nonlinear. For example, if a subject ranks the first pair of patterns as 90, but the second pair looks twice as similar, on our scale it must be given a rating between 90 and 100. Typically, this problem is dealt with by normalizing the data. However, we observed that the nonlinearity disappeared after the subject's inner scale stabilized. This is why we performed a

TABLE I
SIMILARITY MATRIX Δ OBTAINED BY AVERAGING THE ORIGINAL INDIVIDUAL RATINGS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-	82.1	87.8	83.3	77.4	71.4	50.4	86.5	94.4	67.9	86.0	84.9	98.4	93.5	98.9	98.1	94.8	96.4	94.2	97.5
2	82.1	-	65.4	55.5	59.2	70.6	85.0	94.4	97.0	85.1	39.9	79.8	91.4	88.6	95.1	85.5	92.1	91.5	95.0	65.9
3	87.8	65.4	-	55.9	72.7	71.2	89.4	82.0	86.7	92.2	92.5	91.1	95.2	96.9	32.0	88.4	92.4	87.3	90.6	90.9
4	83.3	55.5	55.9	-	48.8	60.7	84.0	93.4	94.5	72.3	60.2	69.5	78.9	80.7	92.9	83.4	92.6	90.5	95.2	67.8
5	77.4	59.2	72.7	48.8	-	57.2	73.9	91.7	97.2	74.8	59.8	80.4	87.8	79.6	93.1	93.8	94.8	95.7	96.0	73.8
6	71.4	70.6	71.2	60.7	57.2	-	69.2	83.8	95.5	76.6	88.0	86.0	95.9	90.8	95.1	92.0	90.0	94.5	90.3	91.8
7	50.4	85.0	89.4	84.0	73.9	69.2	-	87.4	97.5	80.9	88.5	85.2	98.8	85.9	98.1	97.7	94.1	96.6	93.4	97.9
8	86.5	94.4	82.0	93.4	91.7	83.8	87.4	-	67.1	86.4	90.2	90.2	69.0	65.8	63.9	89.3	87.6	80.0	87.3	89.4
9	94.4	97.0	86.7	94.5	97.2	95.5	97.5	67.1	-	83.0	86.4	93.2	64.0	90.1	57.9	91.8	89.8	85.9	84.0	85.6
10	67.9	85.1	92.2	72.3	74.8	76.6	80.9	86.4	83.0	-	53.0	69.6	93.4	79.0	95.2	83.6	86.2	92.4	93.7	77.4
11	86.0	39.9	92.5	60.2	59.8	88.0	88.5	90.2	86.4	53.0	-	56.2	84.8	72.6	88.1	77.9	86.8	92.3	96.1	66.4
12	84.9	79.8	91.1	69.5	80.4	86.0	85.2	90.2	93.2	69.6	56.2	-	68.9	65.9	93.1	78.6	89.1	91.9	94.2	90.3
13	98.4	91.4	95.2	78.9	87.8	95.9	98.8	69.0	64.0	93.4	84.8	68.9	-	56.1	56.2	82.6	93.3	76.5	79.1	76.3
14	93.5	88.6	96.9	80.7	79.6	90.8	85.9	65.8	90.1	79.0	72.6	65.9	56.1	-	80.6	76.9	86.0	86.8	87.8	80.7
15	98.9	95.1	32.0	92.9	93.1	95.1	98.1	63.9	57.9	95.2	88.1	93.1	56.2	80.6	-	84.9	88.8	77.1	75.5	88.5
16	98.1	85.5	88.4	83.4	93.8	92.0	97.7	89.3	91.8	83.6	77.9	78.6	82.6	76.9	84.9	-	60.6	67.7	72.2	65.0
17	94.8	92.1	92.4	92.6	94.8	90.0	94.1	87.6	89.8	86.2	86.8	89.1	93.3	86.0	88.8	60.6	-	30.6	70.8	73.8
18	96.4	91.5	87.3	90.5	95.7	94.5	96.6	80.0	85.9	92.4	92.3	91.9	76.5	86.8	77.1	67.7	30.6	-	51.7	58.3
19	94.2	95.0	90.6	95.2	96.0	90.3	93.4	87.3	84.0	93.7	96.1	94.2	79.1	87.8	75.5	72.2	70.8	51.7	-	66.8
20	97.5	65.9	90.9	67.8	73.8	91.8	97.9	89.4	85.6	77.4	66.4	90.3	76.3	80.7	88.5	65.0	73.8	58.3	66.8	-

TABLE II
TWO-DIMENSIONAL CONFIGURATION OBTAINED AS A RESULT OF CMDS. EACH COLUMN CONTAINS THE COORDINATES OF A GIVEN STIMULUS. NOTE THAT THE OUTPUT OF THE CMDS IS ON A DIFFERENT SCALE THAN THE OUTPUTS OF WMDS GIVEN IN TABLES III–V

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	33.7	33.2	8.5	33.4	40.7	34.3	32.6	-21.0	-29.7	22.6	24.3	14.8	-30.9	-7.9	-36.7	-24.7	-30.6	-43.2	-39.1	-14.1
2	-16.2	15.2	-30.1	8.2	1.6	-12.4	-16.4	-36.5	-33.1	8.2	22.9	10.6	-15.0	-1.0	-40.5	33.1	30.8	22.4	11.5	36.7

warm-up experiment for each subject. Moreover, we performed several normalization of the individual ratings:

1) normalization with respect to zero mean and unit variance;
2) normalization to [0, 1] range;
3) normalization to a maximum magnitude of 1.

The fact that the results stayed the same for all the normalization confirms our hypothesis. Secondly, the algorithms we use for minimization in CMDS or WMDS do not guarantee a global minimum [13]. However, the fact the extracted attributes remain stable (see Section IV) over all solutions indicates that a local minimum we found is in fact the global minimum.

Finally, hierarchical cluster analysis aided us in verifying the results obtained with the MDS. Moreover, the HCA technique expresses the structure and groupings in the similarity matrix hierarchically; therefore, it allowed us to establish the rules and the hierarchy in which the MDS dimensions are combined in judging similarity.

V. MULTIDIMENSIONAL SCALING RESULTS: THE MOST IMPORTANT DIMENSIONS OF COLOR PATTERNS

The stress index (4) for the 2-D solution was 0.31, indicating that a higher-dimensional solution is necessary, that is, the error was still substantial. The stress values for the three-, four-, five-, and six-dimensional configurations were 0.26, 0.20, 0.18, and 0.16, respectively. We stopped at six dimensions since further increase did not result in a noticeable decrease of the stress value. Also, although we postulated that a fair number of dimensions is needed to model the human notion of pattern similarity, we used only a few. This is because some of these dimensions are individual, as well as because many of them are used only at a higher level of judgment. In other words, some dimensions are subject dependent and some are domain dependent. Therefore, our aim was to extract only the very basic perceptual attributes, allowing us to construct a general model for color pattern matching. Tables II–V give the two-, three-, four-, and five-dimensional configurations. Since the output of the MDS gives us only the configuration of points in N-dimensional space, without the associated dimensions, the interpretation of the dimensions is left to us. We used the subjects' explanations as an aid in accomplishing this.

A. Interpretation of the 2-D Solution

The 2-D CMDS configuration is shown in Fig. 2. Dimensions derived from this configuration are 1) presence/absence of a dominant color (or, as we are going to call it, "the dimension of

TABLE III
THREE-DIMENSIONAL CONFIGURATION OBTAINED AS A RESULT OF WMDS. EACH COLUMN CONTAINS THE COORDINATES OF A GIVEN STIMULUS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.82	0.63	0.28	0.46	0.85	1.37	1.91	-0.13	-0.99	0.80	0.30	0.23	-1.54	-0.53	-1.31	-0.88	-0.43	-1.02	-0.80	-1.02
2	-0.83	1.42	0.05	1.47	1.40	-0.25	-0.80	-0.6	-0.15	-0.11	1.21	0.78	0.77	0.48	0.07	-0.56	-1.78	-1.47	-1.79	0.61
3	-0.12	-0.29	-1.74	-0.07	-0.29	-0.80	0.23	-1.55	-1.63	1.26	0.82	1.34	-0.26	1.23	-1.39	1.36	0.80	0.30	-0.89	0.91

TABLE IV
FOUR-DIMENSIONAL CONFIGURATION OBTAINED AS A RESULT OF WMDS. EACH COLUMN CONTAINS THE COORDINATES OF A GIVEN STIMULUS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-1.84	-0.54	-0.06	-0.45	-0.90	-1.15	-1.90	-0.05	0.63	-0.88	-0.41	-0.44	1.35	0.24	1.24	1.08	0.61	1.21	1.06	1.12
2	1.05	-1.03	-0.05	-1.49	-1.41	0.40	1.07	0.64	0.31	0.11	-1.35	-0.90	-0.90	-0.43	-0.15	0.14	1.62	1..35	1.77	-0.75
3	0.22	-0.64	1.26	-0.10	0.28	0.31	0.32	1.64	1.88	-1.18	-0.88	-0.60	0.65	-0.16	1.61	-1.64	-1.30	-0.61	0.01	-1.07
4	-0.18	1.45	1.74	0.84	0.73	1.43	-0.11	-1.07	-0.94	-1.07	-0.50	-1.55	-1.02	-1.92	0.38	0.19	0.03	0.37	0.48	0.71

TABLE V
FIVE-DIMENSIONAL CONFIGURATION OBTAINED AS A RESULT OF WMDS. EACH COLUMN CONTAINS THE COORDINATES OF A GIVEN STIMULUS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.48	0.80	0.23	0.77	1.18	1.50	1.83	-0.30	-1.37	0.41	0.25	0.28	-1.45	-0.20	-1.26	-0.71	-0.69	-1.02	-0.79	-0.93
2	1.70	-0.77	0.49	-1.47	-1.29	0.04	1.21	0.95	0.81	0.35	-0.89	-0.95	-1.04	-0.84	0.19	-0.82	1.34	1.06	1.01	-1.08
3	-0.47	-0.68	-1.03	-0.48	-0.75	0.59	0.68	-0.49	-1.72	-0.81	-1.12	0.16	-0.33	0.98	-0.90	1.78	1.22	1.34	1.76	0.28
4	0.14	1.43	-0.35	0.02	-0.46	-0.72	-0.36	-1.96	-0.22	1.20	1.18	-0.16	-1.20	-1.26	-1.23	0.91	1.42	0.68	-0.52	1.44
5	-0.65	0.93	2.00	0.79	0.45	0.96	-0.51	-0.56	-0.46	-1.58	-0.96	-1.96	-0.39	-1.32	1.18	0.14	-0.17	0.51	0.94	0.68

overall color") and 2) color purity. It is interesting that both dimensions are purely color based, indicating that, at the coarsest level of judgment, people primarily use color to judge similarity and matching. As will be seen later, these dimensions remained in all solutions. Moreover, the 2-D configuration strongly resembles one of the perpendicular projections in the three-, four-, and five-dimensional solutions. The same holds for all three dimensions from the 3-D solution, indicating that these features could be the most general in human perception.

Although insufficient, the 2-D solution was very helpful. For example, neighborhoods in the 2-D space have meanings associated with common similarity features and can be used to better interpret high-dimensional configurations. This is mainly because the neighborhood information is obtained primarily from small distances, thus revealing other patterns in the data. Four major similarity categories emerged from the neighborhood information.

1) Neighborhood 1: (Patterns 2, 4, 5, 10, 11, and 12) are directional patterns characterized by a single dominant local orientation within each portion of that pattern.

2) Neighborhood 2: (Patterns 8, 9, 13, and 15) are strongly uniform, repetitive patterns, obtained by replicating the primitive element according to strongly determined geometric placement rules. The quality that separates this group from Neighborhood 1 (which can also be seen as uniform and repetitive) is directionality. Namely, no pattern from Neighborhood 2 has a single preferred direction and all are equally repetitive in both horizontal and vertical directions.

3) Neighborhood 3: (Patterns 16, 17, 18, 19, and 20) contains patterns with similar color distributions. This cluster consists of different multicolored patterns. Still, all the patterns have almost identical color histograms. It is interesting that although completely different, these patterns were judged as very similar by almost all subjects. Another common feature for patterns in this group is the existence and repetition of a primitive element. However, due to a complex interaction of many colors, the placement rule or a dominant orientation can hardly be detected. Hence, these patterns are perceived as "busy," "heavy," and "complex," rather than "repetitive," "uniform," or "directional."

4) Neighborhood 4: (Patterns 1, 6, and 7) As opposed to the patterns from Neighborhood 3, patterns from this group are truly random; they possess no obvious structuring element. Hence, they can also be seen as complex, not due to the complex color distribution, but rather due to the complexity of the spatial pattern generation.

Patterns 3 and 14 were left isolated and could be annexed to more than one neighborhood (Neighborhood 2 would be a good choice).

B. Interpretation of the Higher-Dimensional Solutions

Both for CMDS and WMDS, the same three dimensions emerged from 3-D configurations. They are

1) overall color;
2) color purity;

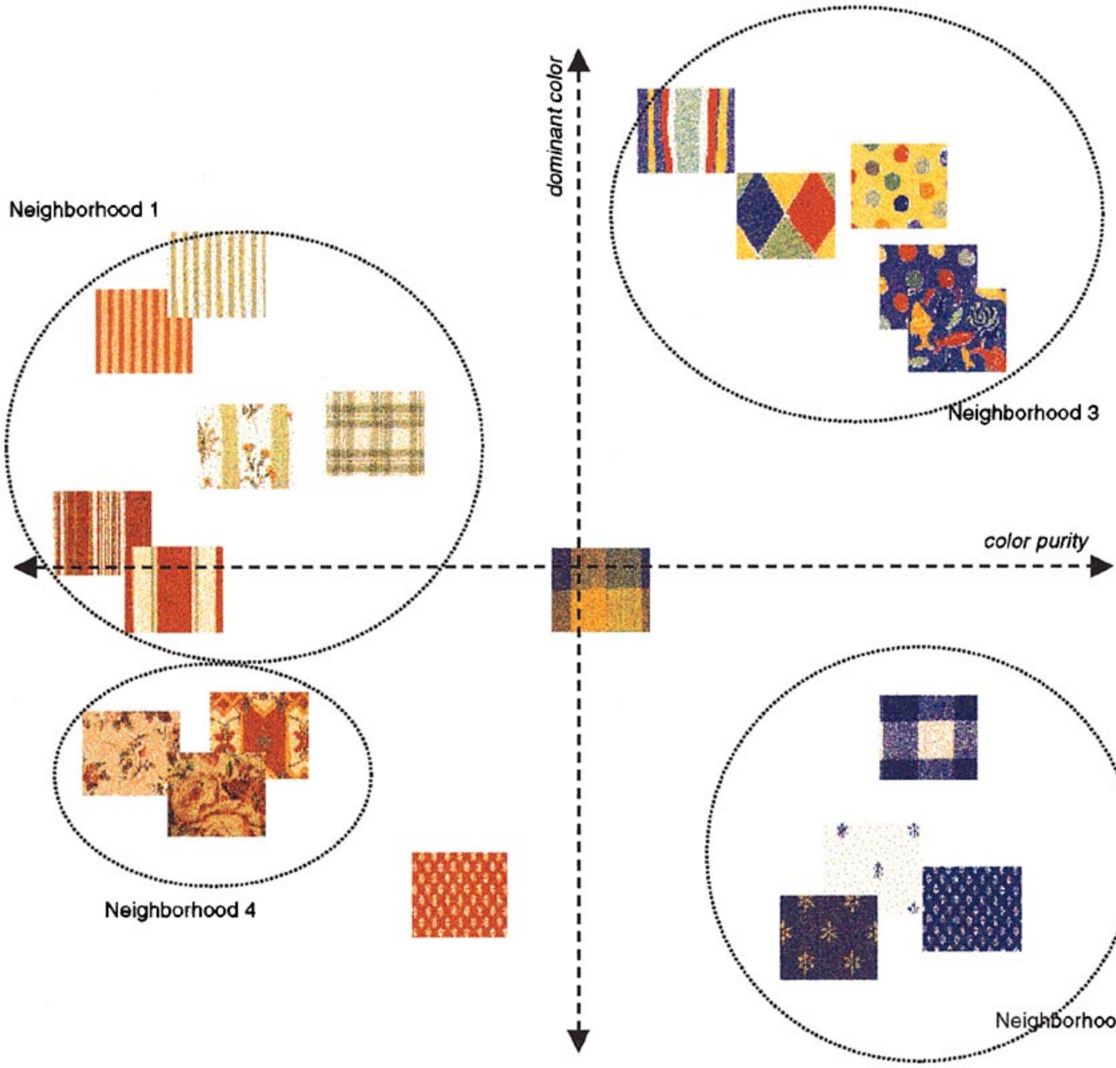


Fig. 2. Two-dimensional CMDS configuration. Horizontal axis represents the dimension of color purity whereas the vertical axis is the dimension of dominant color. Four major similarity categories emerged from the neighborhoods in this configuration: 1) directional patterns characterized by a single dominant local orientation (Neighborhood 1); 2) strongly uniform, repetitive patterns (Neighborhood 2); 3) patterns with similar color distributions (Neighborhood 3); and 4) random patterns (Neighborhood 4). These neighborhoods were useful in determining the rules people use in judging similarity.

3) regularity and placement.

The four-dimensional (4-D) WMDS solution revealed following dimensions:

1) overall color;
2) color purity;
3) regularity and placement;
4) directionality.

The five-dimensional (5-D) WMDS solution came with the same four dominant characteristics with the addition of a dimension which we called "pattern heaviness." The addition of this dimension did not improve the overall goodness-of-fit significantly, since it changed from 0.20 (for four dimensions) to 0.18 (for five dimensions). It appeared that this dimension was used with high weights by a few subjects (mostly women), while it was irrelevant to most of the others.

From the grouping of patterns along the five dimensions, we conclude that almost each of them represents a complex interplay between the pattern, color, tones, and contrast. Would an extra dimension from the six-dimensional (6-D) space contribute to the simplification of these concepts and a cleaner interpretation of the whole configuration? Although by going from five to six dimensions stress value is still decreasing, dimensions derived from the 6-D solution were not consistent with the interpretations above. In fact, the 6-D solution appears to be unstable and without a reasonable interpretation. This is in accordance with a rule of thumb connecting the number of stimuli N and dimensionality L [13]. Namely, if $N > 4L$, the interpretation of stress is not sensitive to N and L. On the other hand, as L gets closer to N, great changes can occur and the stress value is not a reliable indicator of the actual number of dimensions. Therefore, the inability to interpret the 6-D configuration can be due to the following factors.

1) Stimulus set is too small to allow the higher dimensionality. It could also be that the stimulus set did not vary enough on certain characteristics so these characteristics could not show up as a new dimension.

2) Actual number of dimensions is smaller than six. In that case, since the 6-D solution used too many input arguments, the configuration adapted itself to the random error in the data, preventing us for finding the actual dimensions.
3) Some of the perceptual criteria, although understood as a single characteristic, may correspond to two or more dimensions, making the whole configuration almost impossible to interpret.

C. Dimensions of Pattern Similarity

As a result of the experiment, five important similarity criteria emerged.

1) Dimension 1—Overall Color: can also be described in terms of the presence/absence of a dominant color. At the negative end of this axis are patterns with an overall impression of a single dominant color (such as 4, 5, 7, 8, 15 in Fig. 1). This impression is created mostly because the percentage of one color is truly dominant. However, a multicolored image can also create an impression of dominant color. This happens when all the colors within this image are similar, having similar hues but different intensities or saturation. At the positive end of this dimension are patterns where no single color is perceived as dominant (such as in true multicolored patterns 16, 17, 18, 19, and 20). Besides the color distribution or color histogram, impression of overall color is correlated with the spatial properties of color contrast and spatial rules in the primitive element repetition. For example, pattern 4 is judged as red by most of the subjects, although there is the same amount of beige in it. On the contrary, pattern 3, which has identical amounts of these colors and the same primitive element (stripe) was perceived as "without dominant color" mostly because of the high spatial frequency in the repetition of the structural element.

2) Dimension 2—Directionality and Orientation: This axis represents the dominant orientation in the edge distribution, or the dominant direction in the repetition of the structural element. The lowest values along this dimension have patterns with a single dominant orientation, such as stripes and then checkers (2, 4, 11, 12, and 13). Midvalues are assigned to patterns with a noticeable but not dominant orientation (5, 10), followed by those patterns where a repetition of the structural element is performed along two directions (3, 8, 9, and 15). Finally, completely nonoriented patterns (1 and 7) and patterns with uniform distribution of edges or nondirectional placement of the structural element (17, 18, and 19) are at the positive end of this dimension. This dimension highlights the sensitivity of the human visual system to horizontal and vertical directions [15]. For example, although it obviously has a single preferred direction of 45°, pattern 16 is not perceived to be as directional as patterns 2, 4, 5, and 11. This dimension also highlights the sensitivity of the human visual system to small perturbations in direction. Hence, although one single direction is prominent in both patterns 4 and 20, pattern 20 is perceived as less directional.

3) Dimension 3—Regularity and Placement Rules: This dimension describes the regularity in the placement of the structural element, its repetition and uniformity. At the negative end of this axis are regular, uniform, and repetitive patterns, where the repetition is completely determined by a certain set of placement rules (2, 3, 8, 9, 11, and 15), whereas at the opposite end are nonrepetitive (5 and 7) or nonuniform patterns (19). However, this dimension does not affect only regularity, uniformity or repetitiveness in terms of the placement of the structural element, it also applies to the spatial distribution of color. Hence, along this axis between the patterns 13 and 14 (which have exactly the same geometry but different color scheme and layout) pattern 13 is perceived as more uniform and regular than pattern 14. This dimension also reflects the human sensitivity to small perturbations in placement. For example, although patterns 3 and 17 follow a similar placement rule, pattern 17 is perceived as less regular, since the structural elements are slightly displaced from the position they would have had in a completely regular pattern.

4) Dimension 4—Color Purity: This dimension arose somehow unexpectedly, but it remained stable in all MDS configurations, clustering results, even in the subjects' explanations of their rankings. This dimension divides patterns according to the degree of their colorfulness. At the negative end are found pale patterns (1 and 10), patterns with unsaturated overtones (7), patterns with dominant "sandy" or "earthy" colors (5, 6, and 11). At the positive end are patterns with very saturated and very pure colors (9, 13, 19, etc.). Hence, this dimension can also be named the dimension of overall chroma or overall saturation within an image.

5) Dimension 5—Pattern Complexity and Heaviness: This dimension was the one most difficult to interpret. It showed only in the last, 5-D configuration, hence it can be seen as optional. Also, as we will show in the next section, it is not used in similarity judging until the very last level of comparison. For that reason, we have also named it "a dimension of general impression". At the one end of this dimension are patterns that are perceived as "light" and "soft" (1, 7, and 10) while at the other end are patterns described by subjects as "heavy," "busy," and "sharp" (2, 3, 5, 17, 18, and 19). According to the grouping of patterns along this axis and the results of HCA, texture heaviness is determined by one of the following factors: type of the overall color (light versus dark), overall chroma (unsaturated versus saturated), spatial frequency in the repetition of the structural element, and finally color contrast. Hence, this attribute reflects pattern values along other four dimensions. Still, it is obtained as their Boolean combination, rather than as a linear combination or a simple projection onto a new plane.

VI. Hierarchical Clustering Results: Rules for Judging Similarity

Fig. 3 shows the ordering of clusters obtained as a result of the HCA, whereas Fig. 4 shows the HCA tree, obtained from the complete similarity matrix for 20 patterns used in the study. By comparing this result to the result of the 2-D MDS shown in Fig. 2, there is an excellent correspondence between neighborhoods in the MDS configuration and clusters determined by the HCA. One simple way to confirm the stability of the dimensions we obtained and their combining rules is to split the original data in several ways and to perform separate HCA's for each part. As suggested in [13], we eliminated some of the stimuli from the data matrix and determined the HCA trees for the remaining

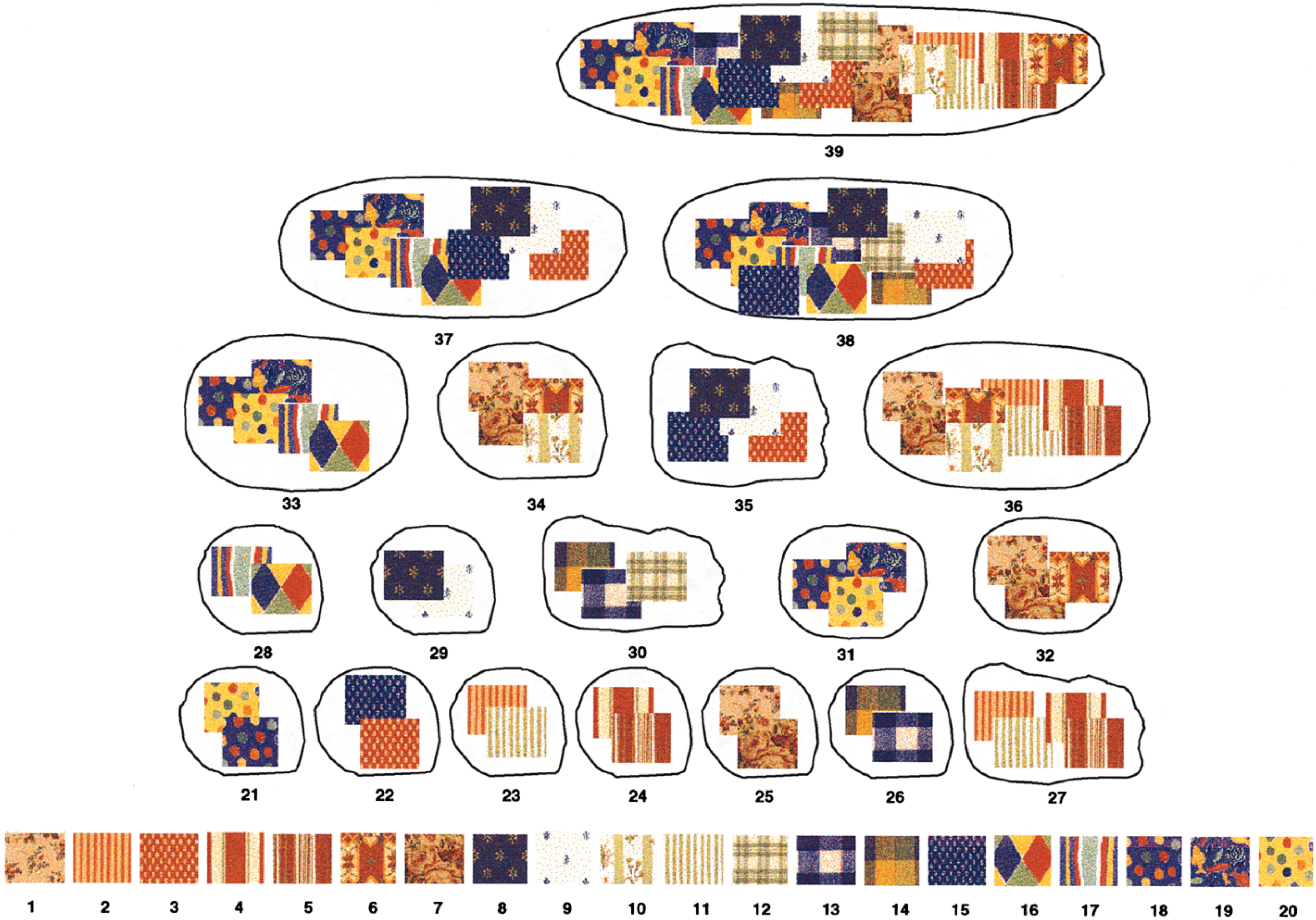


Fig. 3. Result of the HCA applied to the complete set of stimuli. Clusters 1–20 are original patterns and clusters 21 to 37 represent successive nodes of the tree. In the last step, clusters 36 and 38 are joined to form the top cluster. The ordering of clusters was used to determine the rules and the sequence of their application in pattern matching.

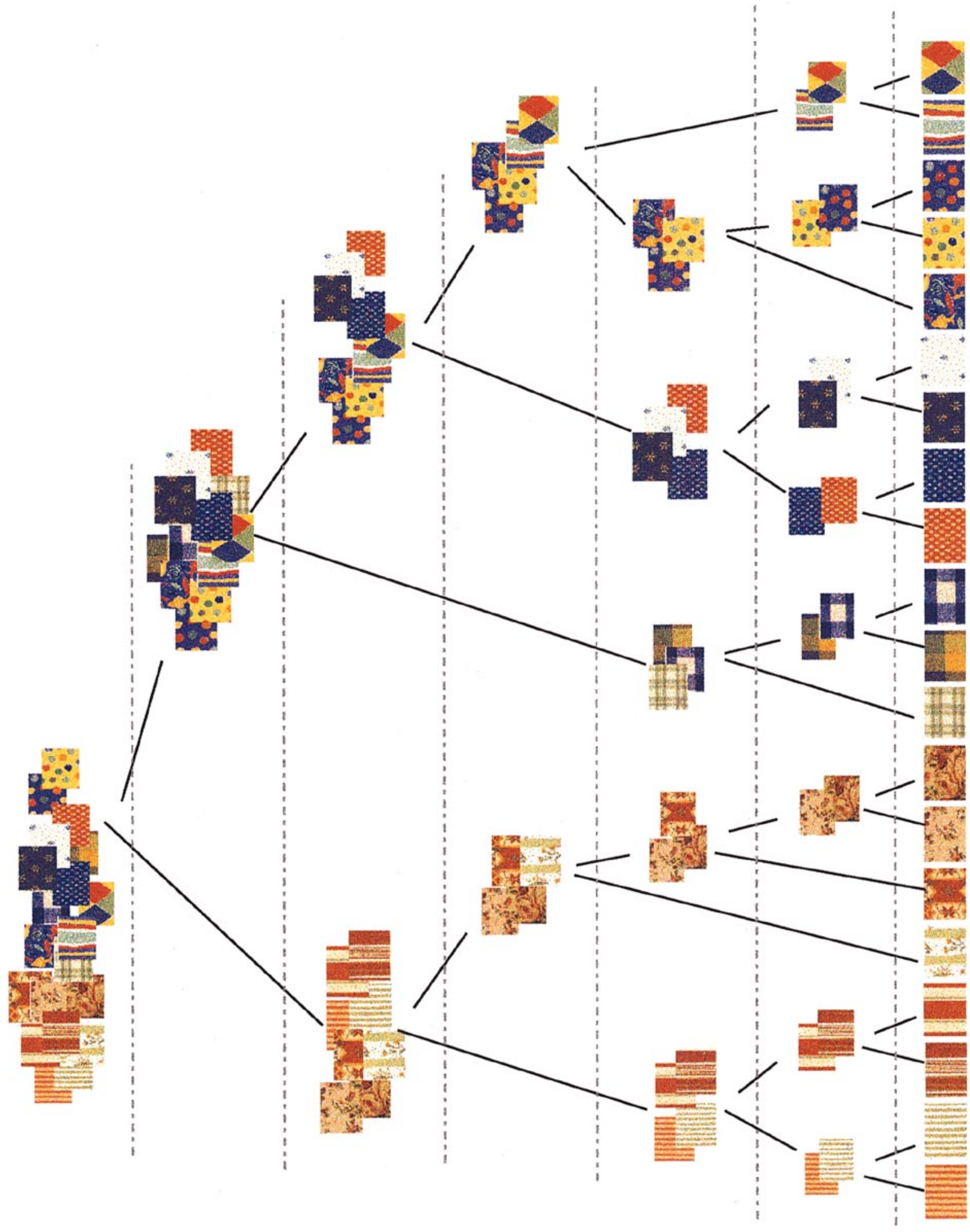


Fig. 4. The tree representation of the HCA result from Fig. 3.

stimuli. One such solution obtained by eliminating patterns 2, 11, and 17 is given in Fig. 5. The dimensions we enumerated in the previous section remained stable for various solutions; we thus conclude that the 5-D configuration can be used for modeling the similarity metrics of the human visual system. As a result of the HCA, we derived a list of similarity rules and the sequence of their application based on the analysis given below.

From the early stages of clustering, we were able to determine the initial rules used by humans in judging similarity (Rules 1 and 2). These were followed by rules emerging from the middle stages (Rules 3 and 4). Finally, at the coarsest level of comparison, we use Rule 5 (top nodes of the HCA tree in Figs. 4 and 5, or clusters 36–38 in Fig. 3).

1) Rule 1: The strongest similarity rule is that of *equal pattern*. Regardless of color, two textures with exactly the same pattern such as pairs (17, 18), (2, 11), and (3, 15) are always judged to be the most similar. Hence, this rule concerns the identity of Dimensions 3 and 2 (pattern regularity and directionality). However, once two patterns are perceived as similar along these two dimensions, color comes into play. The ordering of clusters as given in Fig. 3 illustrates this rule: the first three clusters obtained are 21, 22, and 23, but their internal ranking is determined by color (Dimensions 1, 4, and 5).

2) Rule 2: The second in the hierarchy of rules is the combination of Dimension 1 (dominant color) and Dimension 2 (directionality). Two patterns that have similar values in both dimensions, such as pairs (10, 11), (1, 7), and the triplet (2, 4, 5) are also perceived as similar.

3) Rule 3: The third rule concerns either Dimension 2 (directionality) or Dimension 3 (pattern regularity and placement rules). Hence, two patterns which are dominant along the same direction (or directions) are seen as similar, regardless of their color. One such example is the cluster (12, 13, 14). In the same manner, seen as similar are patterns with the same placement or repetition of the of the structural element, even if the structural element is not exactly the same (see patterns 8 and 9, or 17, 18, and 19).

4) Rule 4: In the middle of the hierarchy comes the rule of *dominant color*. Two multicolored patterns are perceived as similar if they posses the same color distributions regardless of their content, directionality, placement or repetition of a structural element (patterns 16–20). This also holds for patterns that have the same dominant or overall color (patterns 2–6). Hence, this rule concerns only identity along the Dimension 1 (dominant color).

5) Rule 5: Finally, at the very end of the hierarchy, comes the rule of *general impression* (Dimensions 4 and 5). This rule divides patterns into "dim," "smooth," "earthy," "romantic," or "pale" (laying at the one end of the corresponding dimension) as opposed to "bold," "bright," "strong," "pure," "sharp," "abstract," or "heavy" patterns laying on the opposite end. This rule represents the complex combination of color, contrast, saturation and spatial frequency, and therefore applies to patterns at the highest, abstract level of understanding.

This set of rules represents the basic grammar of pattern matching. How are we going to apply these rules? Our experiments demonstrated that individuals vary in the importance they attach to each dimension of the stimulus space. While some of the subjects perceived texture-related dimensions 2 and 3 as being more important, other individuals had just the opposite perception. Hence, individual weights for each user in WMDS were different. Therefore, pattern similarity can not be uniquely formulated as a linear combination of attributes. However, each rule can be expressed as a logical combination (logical OR, AND, XOR, NOT) of the pattern values along the dimensions involved in it. For example, consider cluster 24 composed of patterns 4 and 5 in Fig. 3. These patterns have similar overall color and dominant orientation, thus their values both along the Dimensions 1 and 2 are very close and the comparison using Rule 2 is expressed as follows.

[*Dimension*1(pattern 4) **similar to** *Dimension*1(pattern 5)] **AND** [*Dimension*2(pattern 4) **similar to** *Dimension*2(pattern 5)].

Furthermore, the rule of the equal pattern (being the strongest one in the hierarchy of rules) suggests that human perception of pattern is unrelated to the color content of an image. Namely, equal patterns are seen as the most similar, regardless of their color attributes, indicating the pattern-color separability. This observation is consistent with the data reported in [33] on the appearance of colored patterns. Moreover, previously reported neurophysiological studies [34] and perceptual models [28], demonstrate that the image signal is composed of luminance and a chrominance components, each being processed by separate pathways.

VII. How to Measure Dimensions of Color Patterns?

Having obtained the relevant dimensions of color patterns, we need image processing tools to extract and measure them.

1) Dimension 1—Overall Color: Since the importance of color is established in many pscychophysical studies there exist a vast number of techniques for extraction of color based information. One commonly used approach is based on a color histogram, representing the joint probability distribution of intensities in the three color channels. The features based on the color histogram can be used with various metrics to simulate human performance in judging image similarity [16], [17]. Besides the color histogram, moments of color distribution (such as mean, variances or higher-order statistics) can also be used as color features in image matching [18]. Finally, one can use features from color codebooks and color sets as quantized versions of the three-dimensional (3–D) color space [19], [20]. Note that the extraction of color features should be performed in any of the perceptually uniform systems (such as *CIE Lab*) so that metrics based on L^2 norms adequately describe perceptual differences among the colors compared.

2) Dimensions 2 and 3—Directionality, Orientation, Regularity, and Placement Rules: Besides color information, texture directionality and repetitiveness appear to be among the most important features used by humans in distinguishing color patterns. This is consistent with the conclusions from the experimental studies conducted by Rao and Lohse for gray-level textures [11], as well as with the set of features corresponding to the perceptional criteria detected by Tamura *et al.* [10]. Repetitiveness can be modeled by a primitive element and placement rules that specify how this element is to be replicated [1]. The feature of repetitiveness is correlated with regularity, uniformity

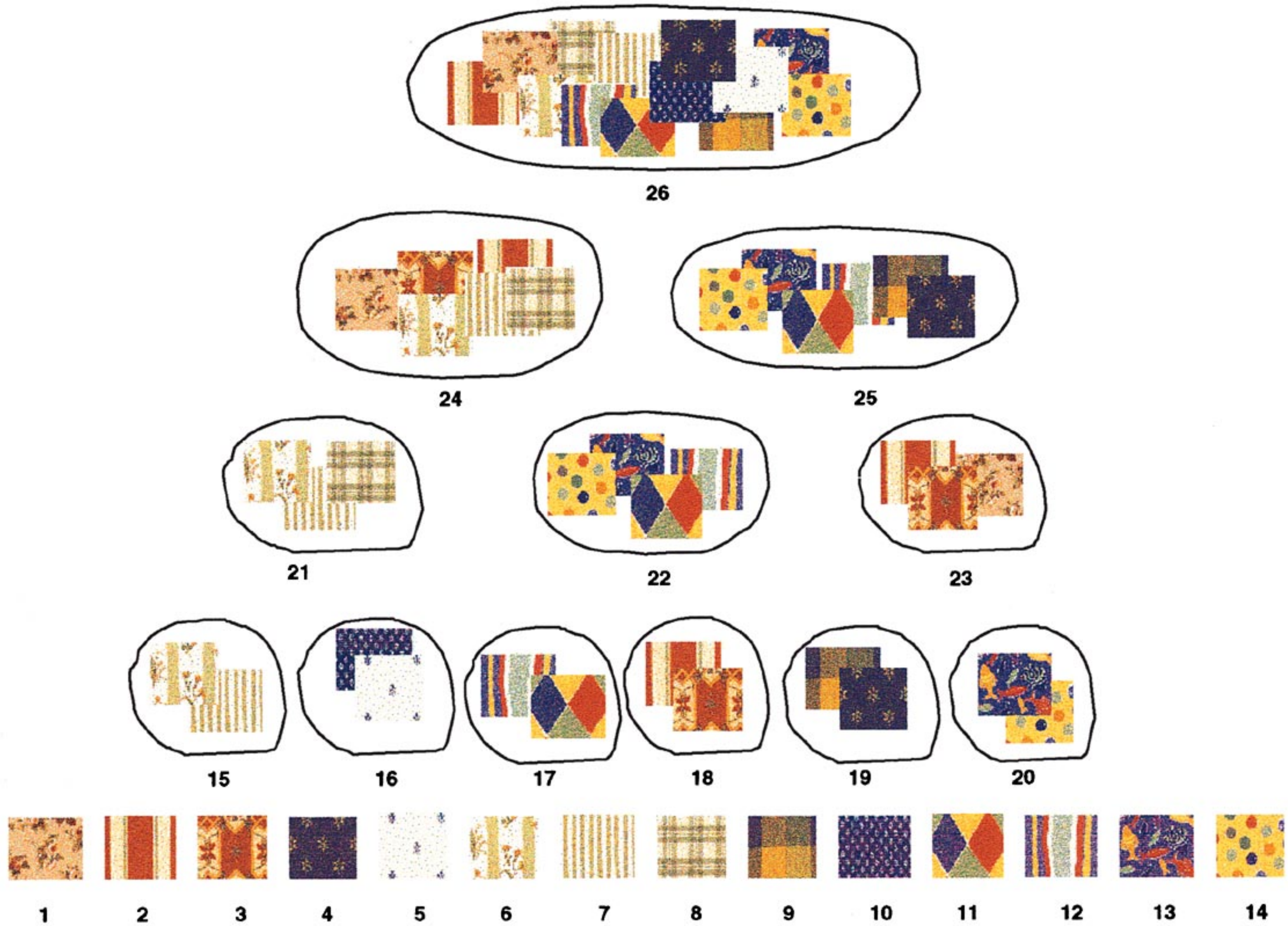


Fig. 5. To confirm and refine the rules following the *rule of equal pattern*, we applied the the HCA to reduced sets of stimuli. An example where patterns 2, 11, and 17 were removed is given.

and nonrandomness; it can thus be assessed by statistical parameters calculated at the optimal scale where the repetitiveness is maximum. The optimal resolution is determined as the one with the minimal variance of texture features. In that case, the resolution used contains the information about the size of the primitive element. Moreover, a statistical approach is adequate for describing nonrepetitive patterns, since the mean value addresses the roughness while the variance represents the degree of nonuniformity.

Directionality in patterns can be described in various ways. For example, edge maps at different resolutions contain the information about both global directionality and dominant local orientation. Global directionality can be assessed by calculating the statistical parameters (means, variances, and higher-order moments) of angular distribution. Dominant orientation within each portion of texture requires averaging operations within each local textured region, as suggested by Rao in [21]. A similar concept can be explored within any orientation sensitive multiresolution decompositions, such as wavelet and Gabor transforms [22], [23] or decomposition with steereable filters [24]. Texture directionality and regularity are also among features extracted in Tamura's representation [10] motivated by pscychophysiological studies in human perception of texture. This makes Tamura's representation very attractive for practical applications involving human interaction or human understanding of texture. This representation is further improved and implemented in the QBIC [25] and MARS image retrieval systems [26].

3) Dimension 4—Color Purity: To quantify this dimension, we convert the image into the *CIE HSV color space* (hue, saturation, value) where the S coordinate represents the saturation (purity) of the color. In that case, a histogram of the S channel or its first-, second-, and higher-order moments can be used to describe the overall color purity for the particular image.

4) Dimension 5—Pattern Complexity and Heaviness: This dimension is probably the most difficult to capture. At the simplest level, pattern complexity is perceived as a combination of color contrast, spatial complexity, and spatial frequency in the repetition and placement of the structural element. Hence, patterns with high values of at least one (usually two) of these attributes are considered as "heavy". Unfortunately, each of these attributes is difficult to assess. Spatial complexity and spatial frequency in the repetition of the structural element are connected to the structural description of texture. This approach requires an identification of a texture primitive as *a group of pixels having certain invariant properties that repeat in the given image*. Texture primitives may be defined by the color distribution, shape, homogeneity of any local property (orientation, second- or higher-order histogram parameters, and even micro texture). Once the primitives have been identified their placement determines the spatial relationship. The spatial relationship may be expressed in terms of adjacency, closest distance, periodicity, etc. However, without prior determination of texture primitives, pattern heaviness may be expressed by measuring the edge density, runlenghts of maximally connected pixels or relative extrema density [1]. Another approach toward capturing texture complexity is to consider complexity in terms of the length of a suitable description or to consider the values along the other four dimensions. For example, it is very likely that patterns with high values along the Dimensions 1 and 3 would be perceived as "heavy."

Color contrast is another important indicator of texture heaviness. The simplest way to measure it is through the *contrast ratio function*, which is defined as the ratio of luminance between the lightest and the darkest elements in a scene [27]. The contrast ratio function can be calculated on the entire pattern, or within the area defining a texture primitive. However, due to the importance of color in image similarity perception, this concept of achromatic contrast has to be extended to color contrast as well [28]. One contrast measure suited to the image similarity problem can be found in [29]. It starts from multiscale representation of oriented local contrast as proposed by Peli [30], and calculates the power-law contrast as a nonlinear difference between the lowpass channels at different levels of the multiresolution pyramid.

VIII. Discussion and Conclusions

Based on a subjective study, we identified the five most relevant dimensions of color patterns:

1) overall color;
2) directionality and orientation;
3) regularity and placement;
4) color purity;
5) complexity and heaviness.

These categories constitute the basic vocabulary of color patterns. We also determined the hierarchy of rules governing the use of the dimensions in the vocabulary. These rules can be seen as the basic grammar of the color pattern language. Our work can be seen as an extension of the work by Rao and Lohtse, who conducted a similar experiment in the domain of gray-level textures [11]. They identified three most significant dimensions of gray-level textures: repetitive versus nonrepetitive, high-contrast and nondirectional versus low-contrast and directional; granular, coarse and low-complexity versus nongranular, fine and high-complexity. Although our results apply to the domain of color patterns, they seem to prove findings in [11]. For example, both attributes of regularity and directionality within the domain of gray-level textures, directly translate into our model, as purely pattern–based Dimensions 2 and 3 (directionality, orientation, regularity, and placement rules). Also, the attributes of coarseness and complexity (being parts of the third dimension in the gray-level model) are included in the perception of texture complexity and heaviness (as the Dimension 5 in our model). However, it is extremely important to note that although the perception and understanding of these attributes remain the same in both domains, the mechanism that creates the perception is very different—hence the definition of the dimensions is the major distinctions between the models. For example, in the gray-level world regularity is perceived through the repetition of the primitive element on the regular grid, whereas in the color domain, color features and color appearance of a pattern are responsible for the regularity as well.

This work was the first part in building the system for matching and retrieval of color patterns. In the follow-up to this study we have actually implemented a system based on our vocabulary and grammar. The details of the implementation are given in [35]. To measure the dimensions from the vocabulary and implement the rules from the grammar, we followed the basic guidelines outlined in the previous sections. We tested our system on the following databases: Corel database (2000 images); stones (600 images); ornaments (110 images); oriental carpets (100 images); interior design (9000 images); architectural surfaces (500 images); and paintings (600 images). The vocabulary and grammar proved stable through all of our experiments.

Note that we obtained our experimental set from the database of fabrics, hence, there was no meaning attached to any of them. In that way, the hidden dimension of content within an image was eliminated (for example, two distinct patterns could be judged as similar if their meanings are related such as leaves and flowers, or pebbles and sand, whereas resembling textures such as marble and lace are perceived as dissimilar since they represent two different things). However, even when we tested our algorithms on the paintings database in which image content was apparent, the vocabulary and grammar remained stable, and the results were excellent. Still, when building any system dealing with image similarity, one should be aware of the importance of image content, and additional studies addressing this issue need to be conducted.

We expect that the color pattern understanding model presented in this paper will help researchers in image processing, computer vision, and computer graphics to build effective and elegant algorithms for texture analysis, manipulation, and display. Moreover, we expect that this model can be used in the development of more natural, human-like interfaces between users and machine.

ACKNOWLEDGMENT

The authors would like to thank F. Juang for initiating the research. Their special thanks go to J. Hall for his tremendous help with the multidimensional scaling and for many useful suggestions, D. Davis for providing the software for the subjective experiment, J. Hu for providing the toolkit for hierarchical cluster analysis, and J. Pinheiro for his help with the statistical analysis of the data. Finally, they are grateful to J. Mazo for his thorough and insightful comments.

REFERENCES

[1] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, pp. 786–804, May 1979.

[2] R. Chellappa and S. Chatterjee, "Classification of textures using Gaussian Markov random fields," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 959–963, Aug. 1985.

[3] J. Keller, S. Chen, and R. Crownover, "Texture description and segmentation through fractal geometry," *Comput. Vis., Graph., Image Process.*, vol. 45, pp. 150–166, 1989.

[4] K. I. Laws, "Texture image segmentation," Ph.D. dissertation, Univ. Southern California, Los Angeles, 1980.

[5] M. Unser, "Local linear transforms for texture measurements," *Signal Process.*, vol. 11, pp. 61–79, July 1986.

[6] D. Dunn and W. E. Higgins, "Optimal Gabor filters for texture segmentation," *EEE Trans. Image Processing*, vol. 4, pp. 947–964, July 1995.

[7] T. Chang and C. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Trans. Image Processing*, vol. 2, pp. 429–441, Oct. 1993.

[8] B. Julesz and J. R. Bergen, "Textons: The fundamental elements in preattentive vision and perception of textures," *Bell Syst. Tech. J.*, vol. 62, pp. 1619–1645, 1983.

[9] J. R. Amadsun and R. King, "Textural features corresponding to texture properties," *IEEE Trans. Syst., Man, Cybern.*, vol. 19, pp. 1264–1274, 1989.

[10] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 460–473, 1978.

[11] A. R. Rao and G. L. Lohse, "Toward a texture naming system: Identifying relevant dimensions of texture," *Vis. Res.*, vol. 36, no. 11, pp. 1649–1669, 1996.

[12] J. Kruskal and M. Wish, *Multidimensional Scaling*, London, U.K.: Sage, 1978.

[13] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*, New York: Wiley, 1973.

[14] A. Gagalowitz, "A new method for texture field syntehesis: Some applications to study of human vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 3, pp. 520–533, 1981.

[15] T. N. Cornsweet, *Visual Perception*. Orlando, FL: Academic, 1970.

[16] M. Ioka, "A method of defining the similarity of images on the basis of color information," IBM Res., Tokyo Res. Lab., Tokyo, Japan, Tech. Rep. RT-0030, Nov. 1989.

[17] M. Swain and D. Ballard, "Color indexing," *Int. J. Comput. Vis.*, vol. 7, no. 1, 1991.

[18] M. Stricker and M. Orengo, "Similarity of color images," in *Proc. SPIE Storage Retrieval Image Video Databases*, vol. 1908, 1993.

[19] W. Y. Ma, Y. Deng, and B. S. Manjunath, "Tools for texture/color base search of images," in *Proc. SPIE*, vol. 3016, 1997.

[20] J. R. Smith and S.-F. Chang, "Single color extraction and image query," in *Proc. IEEE Int. Conf. Image Processing*, 1994.

[21] A. R. Rao and R. Jain, "Computerized flow field analysis: Oriented texture fields," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 693–709, 1992.

[22] W. Y. Ma and B. S. Manjunath, "A comparison of wavelet transform features for texture image annotation," in *Proc. IEEE Conf. Image Processing*, 1995.

[23] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 837–842, Aug. 1996.

[24] W. T. Freeman and E. H. Adelson, "The design and use of steereable filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 891–906, Sept. 1991.

[25] W. Equitz and W. Niblack, "retrieving images from a database using texture—algorithms from the QBIC system," IBM Comput. Sci., Tech. Rep. RJ 9805, May 1994.

[26] T. S. Huang, S. Mehrotra, and K. Ramchandran, "Multimedia analysis and retrieval system (MARS) project," in *Proc. 33rd Annu. Clinic Library Application Data Processing—Digital Image Access Retrieval*, 1996.

[27] C. A. Poynton, *A Technical Introduction to Digital Video*, New York: Wiley, 1996.

[28] T. V. Papathomas, R. S. Kashi, and A. Gorea, "A human vision based computational model for chromatic texture segregation," *IEEE Trans. Syst., Man, Cybern. B*, vol. 27, pp. 428–440, June 1997.

[29] T. Frese, C. Bouman, and J. P. Allebach, "A methodology for designing image similarity metrics based on human visual system models," in *Proc. SPIE*, vol. 3016, 1997.

[30] E. Peli, "Contrast in complex images," *J. Opt. Soc. Amer. A*, vol. 7, pp. 2032–2040, Oct. 1990.

[31] *SPSS Professional Statistics*. Chicago, IL: SPSS, 1997.

[32] *S+ for Unix User's Guide*. Seattle, WA: MathSoft Inc., 1996.

[33] A. B. Poirson and B. A. Wandell, "Appearance of colored patterns: Pattern-color separability," *J. Opt. Soc. Amer. A*, vol. 10, no. 12, Dec. 1993.

[34] E. DeYoe and D. C. VanEssen, "Concurent processing streams in monkey visual cortex," *Trends Neurosci.*, vol. 11, pp. 219–226, 1988.

[35] A. Mojsilovic, J. Kovacevic, J. Hu, R. J. Safranek, and S. K. Ganapthy, "Matching and retrieval based on the vocabulary and grammar of color patterns," *IEEE Trans. Image Processing*, vol. 9, pp. 38–54, Jan. 2000.

Aleksandra Mojsilović (S'93–M'98) was born in Belgrade, Yugoslavia, in 1968. She received the B.S.E.E., M.S.E.E., and Ph.D. degrees from the University of Belgrade.

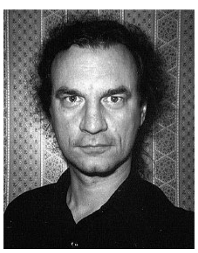
From 1994 to 1998, she was a Member of Academic Staff with the Department of Electrical Engineering, University of Belgrade. Since 1998, she has been a Member of Technical Staff at Bell Laboratories, Lucent Technologies, Murray Hill, NJ. Her main research interests include multidimensional signal processing, computer vision, and image analysis.

Dr. Mojsilović is a member of the IEEE Signal Processing society.

Jelena Kovačević (S'88–M'91–SM'96) received the Dipl.Elect.Eng. degree from the Electrical Engineering Department, University of Belgrade, Belgrade, Yugoslavia, in 1986, and the M.S. and Ph.D. degrees from Columbia University, New York, NY, in 1988 and 1991, respectively.

In November 1991, she joined AT&T Bell Laboratories (now Lucent Technologies), Murray Hill, NJ, as Member of Technical Staff. In the Fall of 1986, she was a Teaching Assistant at the University of Belgrade. From 1987 to 1991, she was a Graduate Research Assistant at Columbia University. In the Summer of 1985, she was with Gaz de France, Paris, France. During the Summer of 1987, she was with INTELSAT, Washington, DC, and in the Summer of 1998, with Pacific Bell, San Ramon, CA. Her research interests include wavelets, multirate signal processing, data compression, and signal processing for communications. She is a coauthor of the book (with M. Vetterli) *Wavelets and Subband Coding* (Englewood Cliffs, NJ: Prentice-Hall, 1995). She is on the Editorial Boards of the *Journal of Applied and Computational Harmonic Analysis*, *Journal of Fourier Analysis and Applications*, and *Signal Procesing Magazine*.

Dr. Kovačević received the Belgrade October Prize, the highest Belgrade prize for student scientific achievements awarded for the Engineering Diploma Thesis in October 1986, and the E.I. Jury Award at Columbia University for outstanding achievement as a graduate student in the areas of systems, communication, or signal processing. She served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and as a Guest Co-Editor (with I. Daubechies) of the Special Issue on Wavelets of the PROCEEDINGS OF THE IEEE. She is on the IMDSP Technical Committee of the IEEE Signal Processing Society and was a General Co-Chair (with J. Allebach) of the Ninth Workshop on Image and Multidimensional Signal Processing.



Darren Kall was born in Paterson, NJ, in 1959. He received the M.A. degree in cognitive and perceptual experimental psychology from Dartmouth College, Hanover, NH, in 1984.

He has worked for Bell Laboratories, Lucent Technologies, Holmdel, NJ, since 1988, where he is presently Distinguished Member of Technical Staff. His research and design interests include the psychophysical categorization of perceived audio impairments, perceptually adaptive user interfaces, and designing technology for persons with disabilities. He holds nine U.S. patents and has three in progress.

Mr. Kall is a member of the HFES, ASA, and AES.

Robert J. Safranek (S'78–M'80–SM'92) was born in Manitowoc, WI, in 1958. He received the B.S.E.E. degree in 1980, the M.S.E.E. degree in 1982, and the Ph.D. degree in 1986, all from Purdue University, West Lafayette, IN. While at Purdue, he worked in the areas of parallel processing, speech coding, computer vision, and robotics.

Since 1986, he has been with Bell Laboratories, Lucent Technologies, Murray Hill, NJ. His research interests include human and machine perception, digital video, networked multimedia, film, and video production, and hardware/software systems for signal processing. He has worked on a variety problems in the area of image and video processing. He was a member of the team that developed the AT&T/Zenith submission to the United States HDTV standardization process. He has also been active in investigating and promoting the use of human perception in signal processing systems. Currently, he is working in the areas of networked multimedia and multimedia information retrieval. He holds 17 U.S. and numerous foreign patents and has published extensively.

Dr. Safranek was a awarded the IEEE Donald G. Fink prize for his paper "Signal compression based on models of human compression" (co-authored with J. D. Johnston and N. S. Jayant) in 1995. He is a member of Eta Kappa Nu, SMPTE, and the IEEE Image and Multidimensional Signal Processing Technical Committee.

S. Kicha Ganapathy received the B.Tech. degree in electrical engineering from the Indian Institute of Technology (I.I.T.), Bombay, in 1967, the M.Tech. degree in electrical engineering from I.I.T., Kanpur, in 1969, and the Ph.D. degree in computer science from Stanford University, Stanford, CA, in 1976, in the area of computer vision.

He was on the faculty at the University of Michigan, Ann Arbor, from 1976 to 1982 and conducted research in the areas of computer vision, computer graphics, and robotics. Since 1982, he has been with Bell Laboratories, Lucent Technologies, Murray Hill, NJ, where he became a Research Department Head in 1985. In that capacity, he co-directed work that put Bell Laboratories in the forefront of robotics. These include robot ping-pong player, telerobotics, and the pioneering work in silicon micromachines. Subsequently, he initiated and led the team that created virtual reality-based video games, an example of which can be seen at the AT&T Pavilion, Epcot Center, Orlando, FL. Currently, he is a Member of Technical Staff, Multimedia Communications Research Laboratory, and is engaged in the area of multimedia information retrieval based on psychophysics of human perception. In particular, he is leading an R&D effort that is applying these research ideas to the emerging field of electronic retailing and related businesses in electronic commerce.