Statistical Analysis of Human Visual Impressions on Morphological Image Manipulation of Gray Scale Textures

Liang LI, Akira ASANO and Chie Muraki ASANO[1]

Graduate School of Engineering, Hiroshima University, Japan
[1]*Department of Lifestyle Design, Yasuda Women's University, Japan*

A method of evaluating human visual impressions of gray scale textures using morphological manipulation is proposed. To study the effects of textural features on human *Kansei*, we introduced a texture analysis method based on mathematical morphology. *Kansei* is a Japanese word for sensibility or emotion. *Kansei* engineering is an approach to connect human sensibility with engineering applications. The proposed method allows us to manipulate global and local properties of a texture separately. Variations of textures were generated by repetitively modifying arranged objects and configurations of the arrangements of original textures. The manipulated textures were presented to human respondents and the similarity of those textures based on human impressions was evaluated. Hierarchical clustering was applied to the similarity matrix generated from respondents' observations. The results of the human evaluation were compared with that of the objective similarity evaluation adopting six global textural features. The global features such as density, regularity, and directionality of the point configurations were shown to have significant effects on human visual impressions and identification of textures. In the case of a texture without significant characteristics in its point configuration, local features such as grain shape have an effect on visual impressions.

KEYWORDS: *Kansei*, visual perception, texture, texture manipulation, mathematical morphology

1. Introduction

Systems for evaluating human sensibility to texture have attracted much attention and interest. *Kansei* engineering is an approach to connect human sensibility with engineering applications. Studies of textural features corresponding to human *Kansei* provide a better outlook for texture analysis. Various methods of texture analysis, for example, the co-occurrence matrix method and the spatial frequency method, have been proposed.[1–4] Texture characteristic measurements are often employed for texture classification and discrimination. Julesz[5,6] conducted the earliest studies concerning texture statistics and human perception; however, Julesz and others[5–9] focused on the preattentive aspect of perception. In a typical preattentive experiment, after being presented with a pair of textures, the respondent is required to perform a discrimination task within a limited duration. The studies on early vision mechanisms mainly focused on whether two textures are different, rather than how great the difference is.

A number of studies have focused on exploring salient textural features used by the human visual system to judge texture similarities in attentive discrimination. Tamura *et al.*[10] performed a psychological experiment in which six basic textural features (coarseness, contrast, directionality, line-likeness, regularity, and roughness) were evaluated by human respondents. Human judgments and computational measures were also compared and their results indicate that strong correlations exist between certain features. Amadasun and King[11] carried out a similar study by exploring five textural features, namely, coarseness, contrast, busyness, complexity, and texture strength. Fujii *et al.*[12] presented a set of texture parameters based on autocorrelation function analysis for describing three salient perceptual properties of texture: contrast, coarseness, and regularity.

Rao and Lohse[13,14] constructed a three-dimensional feature space for texture classification based on human perception. The three features they identified were repetition, orientation, and complexity. Cho *et al.*[15] extended the dimensionality of the feature space to four: coarseness, regularity, contrast, and lightness. These studies suggest that the feature space for perceptual texture discrimination can be constructed using three or four salient textural features; however, how to explore the salient features and the statistical measurements of those features are still in question.

We introduce a novel method of analyzing human textural perception based on a morphological texture model called the "Primitive, Grain, and Point Configuration (PGPC)." This model regards a texture as an image composed of a regular or an irregular arrangement of objects that resemble each other[16] and are much smaller than the size of the image. The objects arranged in a texture are called grains, and they are considered to be derived from one or few typical objects called primitives. The point configuration of the grains is expressed by a morphological skeleton.[17–19] Compared with conventional texture models, the PGPC model discriminates and describes local and global features of a texture explicitly. The PGPC texture model has provided a method that enables separate modification of local and global features.

We applied binary texture manipulation based on the PGPC texture model to investigate human visual impressions of textures.[20] We found that not only the respective influence of a local structure and the entire structure of the texture, but also their mutual interactions were important for identifying of human visual impressions.

In this paper, we extend our investigation to the gray scale textures as they present natural textures better than binary textures. Moreover, they contain additional characteristics and provide flexible analysis methods. We modify three original textures from database[21] by separately controlling global and local parameters using the PGPC texture model. The manipulated textures are presented to human respondents. The similarities of the textures are evaluated by the impression of the respondents. As a new step, the results of human evaluation are analyzed by multivariate statistical analysis. We introduce hierarchical clustering to this research to achieve more precise and intuitive analysis than in our previous work.[20] The results of human evaluation are also compared with the results of objective evaluation adopting six global textural features, namely, skeleton density, isolated regions on the skeleton, grains of the largest size, uncovered regions, length ratio between minor and major axes of edge orientation histogram, and roughness of edge orientation histogram.

The experimental results show that global features, such as density, regularity, and directionality of point configurations have significant effects on human visual impressions. In the case of a texture without significant characteristics in its point configuration, local features such as the shape of the grains have some effect on visual impressions.

2. Mathematical Morphology and Morphological Skeleton

In the context of mathematical morphology, an image object is defined by a set. In the case of binary images, this set refers to the pixel positions included in the object. In the case of gray scale images, an image object is defined by an umbra set. If the pixel value distribution of an image object is denoted as $f(x)$, where $x \in \mathbb{Z}^2$ is a pixel position, its umbra $U[f(x)]$ is defined as

$$U[f(x)] = \{(x,t) \in \mathbb{Z}^3 | -\infty < t \leq f(x)\}. \tag{1}$$

Consequently, when we assume a "solid" whose support is the same as a gray scale image object and whose height at each pixel position is the same as the pixel value at this position, the umbra is equivalent to this solid and the whole volume below it within the support.

Another object, called a structuring element, is defined in the same manner. The structuring element is equivalent to the window of an image processing filter and is supposed to have a much smaller extent than the image object.

In mathematical morphology, opening is a fundamental operation. In the case of the binary image and structuring element, the opening of an image object X with respect to a structuring element B, denoted X_B, has the property

$$X_B = (X \ominus \check{B}) \oplus B, \tag{2}$$

where $\ominus$ denotes Minkowski set subtraction, $\oplus$ denotes Minkowski set addition, and $\check{B}$ denotes the symmetrical set of B with respect to the origin.

Minkowski set subtraction and Minkowski set addition are defined as

$$X \ominus B = \bigcap_{b \in B} X_b, \tag{3}$$

$$X \oplus B = \bigcup_{b \in B} X_b, \tag{4}$$

where X_b indicates the translation of X by b, defined as

$$X_b = \{x + b, x \in X\}. \tag{5}$$

In the case of the gray scale image and structuring element, opening is similarly defined by replacing sets X and B with their umbra.

The skeleton $SK(X,B)$ is defined as

$$SK(X,B) = \bigcup_{n=0}^{N} SK_n(X,B), \tag{6}$$

$$SK_n(X,B) = (X \ominus n\check{B}) - (X \ominus n\check{B})_B, \tag{7}$$

where nB is n-times homothetic magnification of B defined as

$$nB = B \oplus C \oplus \ldots \oplus C \quad ((n-1) - times\ \ of\ \ \oplus), \tag{8}$$

$$0B = 0, \tag{9}$$

where C is another small structuring element. This definition differs from the usual one; however, we employ it to avoid the inconvenience that the difference between nB and $(n+1)B$ is too large for a large B in the original definition.

The gray scale image composed by assigning a pixel value n to the pixels in $SK_n(X,B)$ is called the medial axis transform. The original binary image is reconstructed by locating nB on every pixel of $SK_n(X,B)$ and calculating the union of all ns. Thus, $SK_n(X,B)$ is regarded as the grain location configurations of the texture X if we assume that B is the estimated primitive and nBs are the grains.

3. PGPC Texture Model and Texture Modification

The PGPC texture model represents a texture image X as

$$X = \bigcup_{n=0}^{N} B_n \oplus \Phi_n \tag{10}$$

for nonempty Φ_n, where B_n denotes a grain, and Φ_n is point configuration, that is, a set indicating pixel positions to locate the grain nB.

Here, we assume that $\{0B, 1B, \ldots, nB, \ldots\}$ are homothetic magnifications of a small object B as defined in eqs. (8) and (9), and that B_n in eq. (10) is equivalent to nB for each n. In this case, B is regarded as the primitive, n as the size of the magnification, and X_{nB} as the textural image composed of the arrangement of nB only.

In the primitive estimation scheme, the preferably large homothetic magnifications of an arbitrary candidate primitive that provides the simplest arrangement are considered to be optimum.[22)]

Once the primitive B is estimated, an estimation of Φ_n is obtained by the morphological skeleton employing B as the structuring element. Φ_n represents the location of grains of size n. An example of binary primitive and morphological skeletons of different size distributions is shown in Fig. 1 with the original texture. We can generate new textures from the original by changing three parameters: primitive, size distribution of grains, and skeleton.

4. Methods and Experiments

4.1 Texture Manipulations

We have applied the image processing method based on the PGPC texture model to describe a texture by local and global characteristics separately. We have considered that the estimated primitive or the grains of a texture show some local features, whereas the skeleton of a texture and the grain size distribution show some global features.[16)] To investigate and describe the effects of textural character-

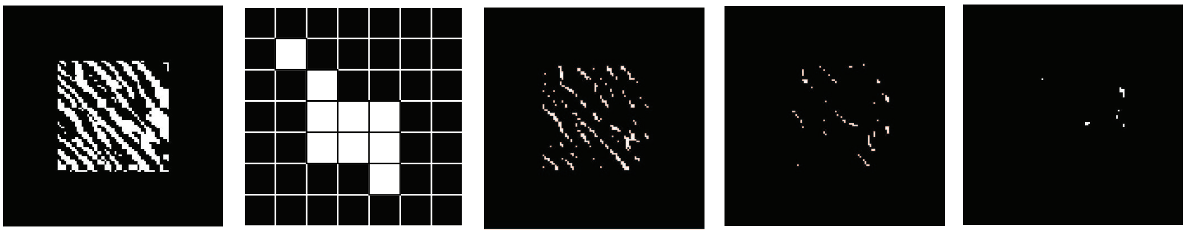

Fig. 1. Example image: (a) original texture, (b) estimated primitive, (c) skeleton Φ_1, (d) skeleton Φ_2, and (e) skeleton Φ_3.

Table I. Texture modifications.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Grains of the largest size	P	R	P	R	P	P	P	R
Grains of other sizes	P	P	R	R	P	P	P	R
Skeleton of the largest size	P	P	P	P	R	P	R	R
Skeletons of other sizes	P	P	P	P	P	R	R	R

Notations: P: Preserved, R: Replaced.

istics on human visual identification and impressions, the parameters of the texture in sample images were modified (Table I).

We use 8-bit gray scale textures of 64×64 for the experiments,[21)] twenty-four textural images were constructed from three. The original textural images and the primitives used in the experiments are illustrated in Figs. 2 and 3, respectively. The modification results are shown in Figs. 4-6, where (a) only removed elements of grain size 0 (the residue) from the original image; (b), (c), and (d) respectively replaced the original primitive with the alternate, whereas the skeleton and size distribution of the original image were preserved; (e), (f), and (g) preserved the primitive, whereas (e) replaced the skeleton of the largest size with a random distribution while fixing the skeletons of smaller sizes, (f) replaced the skeletons of other sizes with random distributions while fixing the skeleton of the largest

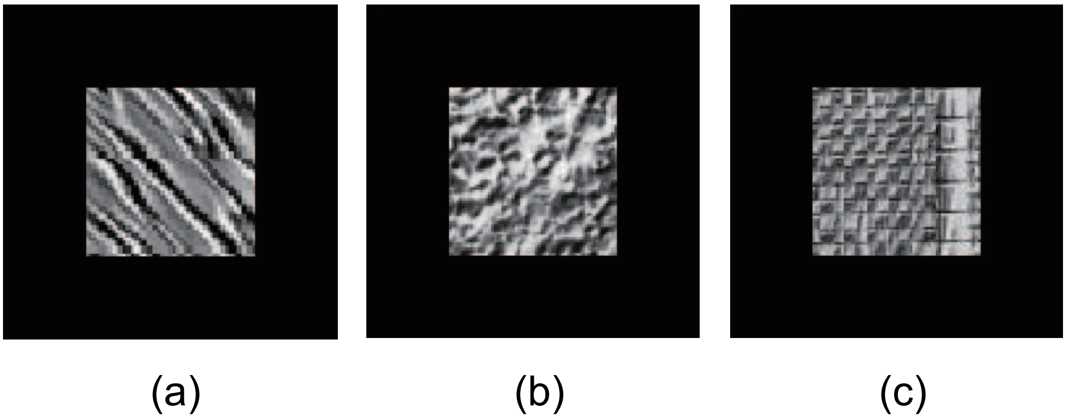

Fig. 2. Original textural images used in the experiments: (a) textural image “naname,” (b) textural image “rock,” (c) textural image “tatami.”

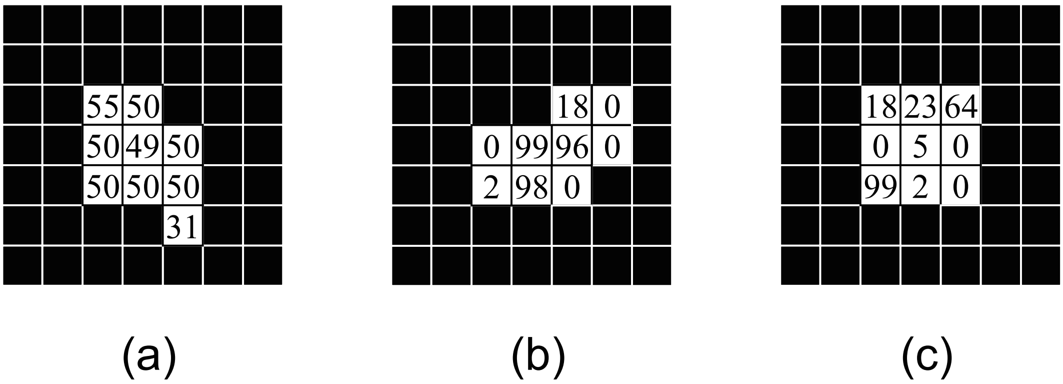

Fig. 3. Primitives used in the experiments: (a) primitive estimated from textural image naname, (b) rock, and (c) tatami.

size, (g) replaced the entire skeletons with random distributions while fixing the size distribution of grains; (h) altered both the primitive and the entire skeletons when fixed the size distribution of grains.

4.2 Visual Evaluation

The manipulated images were arranged randomly on a single screen and shown to human respondents. The screen of rearranged manipulated texture images used for the subjective visual experiment is shown in Fig. 7. The screen was displayed on a 15-inch 4:3 LCD monitor. Thirty-two respondents, who had no information about the original images, were asked to look at the screen for one minute before classifying these images into arbitrary groups while visually evaluating and discriminating the similarity levels of the textural images. The respondents could refer to the screen at any time.

To investigate the visual impressions of the respondents, results of the human evaluation were analyzed by applying hierarchical clustering. This is a general approach of cluster analysis, wherein the objects are classified into subsets or clusters, such that those within each cluster are more closely

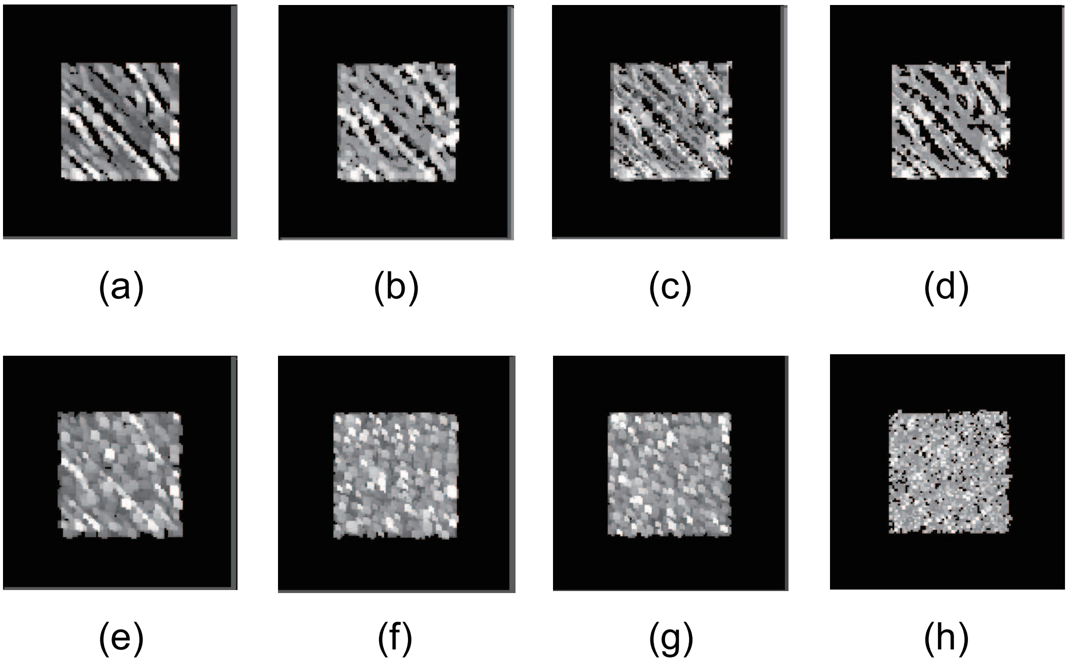

Fig. 4. Constructed textural images (original image: naname, alternate primitive: tatami).

related to one another than to objects assigned to other clusters. The frequency of classifying each pair of textures into the same group by the visual impressions of respondents was adopted for clustering. The result of hierarchical clustering can be illustrated in a graphical representation called a dendrogram. Unlike other clustering methods, hierarchical clustering does not require prespecifying the number of clusters, which can be selected using the dendrogram after the clustering scheme. The hierarchical clustering dendrogram of the visual evaluation is shown in Fig. 8. Two typical clustered groups are circled. The textures in each group are considered to be highly related.

4.3 Image Feature Analysis

An objective evaluation employing textural image features was developed for investigating the most effective features of a texture and comparing with the subjective visual evaluation. The six image features employed in this study are described as follows.

Skeleton density ρ_{SK} is defined as

$$\rho_{SK} = \frac{n_{SK}}{n_D}, \tag{11}$$

where n_{SK} denotes the number of pixels on the binary skeleton, and n_D denotes the number of pixels in the definition domain of the 2-D texture plane.

Isolated regions on the skeleton ISO is defined as

$$ISO = SK - \gamma_C(SK), \tag{12}$$

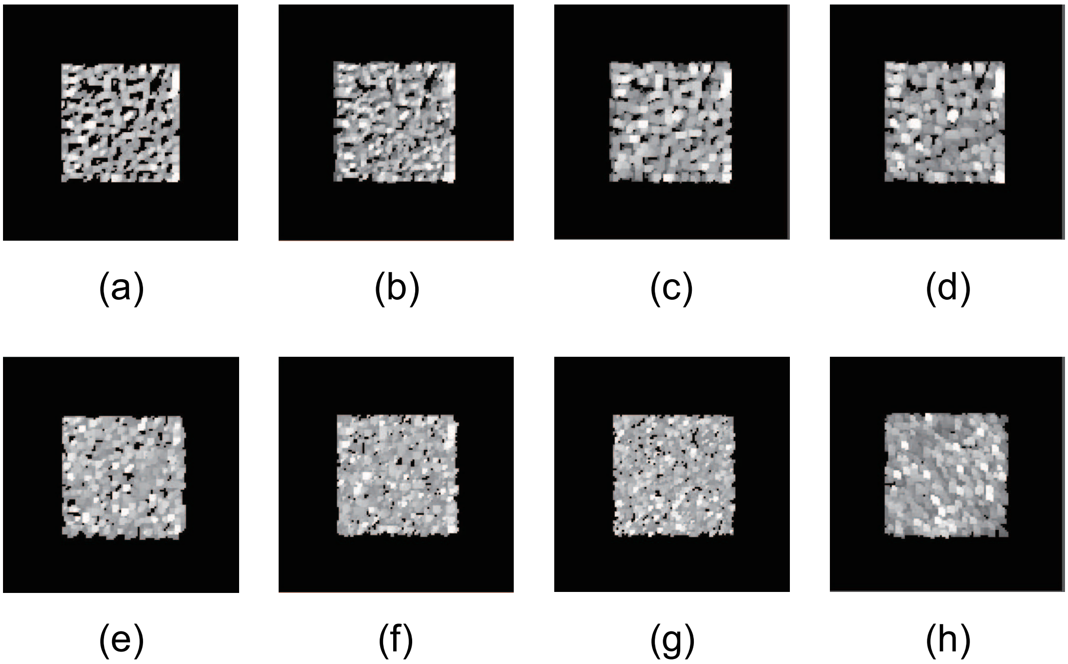

Fig. 5. Constructed textural images (original image: rock, alternate primitive: naname).

where SK denotes the binary skeleton of a textural image, and $\gamma_C(SK)$ denotes the opening of the skeleton by a four-pixel square structuring element C. This parameter can be defined in terms of the ratio between the number of pixels of the isolated regions and the number of pixels of the skeleton:

$$\rho_{ISO} = \frac{n_{ISO}}{n_{SK}}. \tag{13}$$

Grains of the largest size ρ_{Lar} is defined as

$$\rho_{Lar} = \frac{n_{Lar}}{n_T}, \tag{14}$$

where n_{Lar} denotes the number of grains of the largest size, and n_T denotes the total number of grains.

Uncovered regions ρ_U is defined as

$$\rho_U = \frac{n_U}{n_D}, \tag{15}$$

where n_U denotes the number of pixels in the uncovered background regions of the texture image, and n_D denotes the number of pixels in the definition domain of the 2-D plane.

Length ratio between minor and major axes of edge orientation histogram — Texture anisotropy is evaluated by a kind of histogram, called an edge orientation histogram, indicating how frequently edges oriented in each direction appear. Figure 9 illustrates some typical textures and their edge orientation histograms. The graph in the center of each histogram is considered a connected region. An ellipse whose quadratic moment is the same as that of the connected region is determined for each region, and its minor and major axes lengths are calculated. The length ratio between the minor and

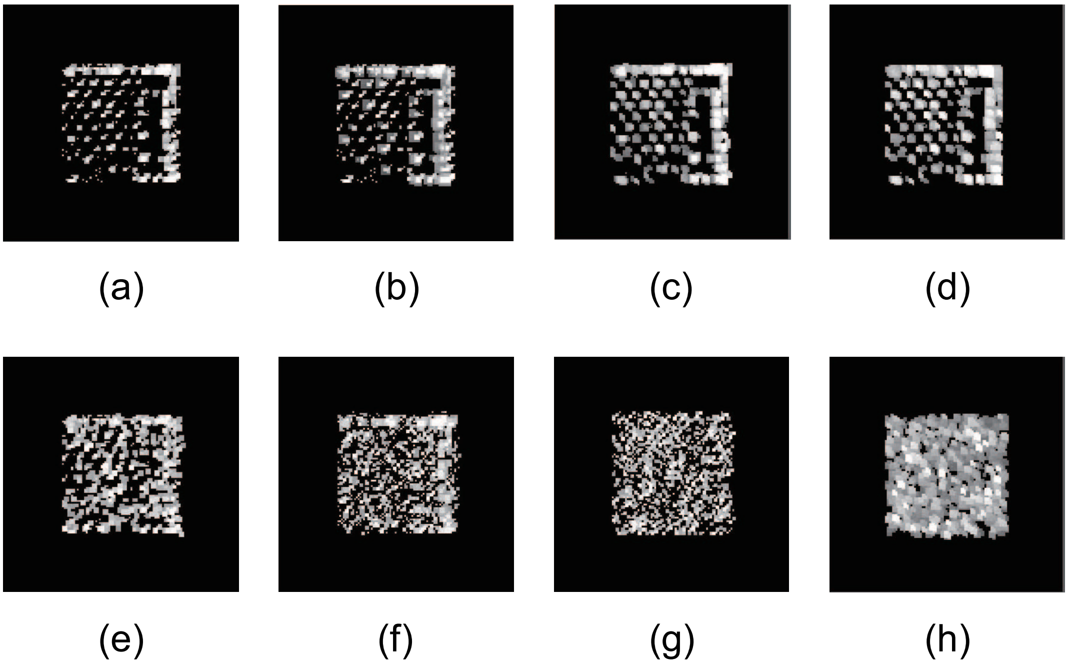

Fig. 6. Constructed textural images (original image: tatami, alternate primitive: naname).

major axes ρ_{Len} is defined as

$$\rho_{Len} = \frac{L_{min}}{L_{max}}, \tag{16}$$

where L_{min} denotes the minor axis length, and L_{max} denotes the major axis length. This parameter indicates the strength of the direction information in a textural image.

Roughness of edge orientation histogram ρ_R is defined as

$$\rho_R = \frac{AreaB}{AreaR}, \tag{17}$$

where $AreaB$ denotes the area between the connected region and the ellipse, and $AreaR$ denotes the area of the connected region.

The image features described above were extracted from twenty-four texture images. For applying clustering analysis, the data of the image features were standardized by the function

$$X_i^{STD} = \frac{X_i - Xmean}{\sigma}, \tag{18}$$

where X_i^{STD} denotes the standardized score of X_i, X_i denotes each value of the image feature for standardization in dataset X, $Xmean$ denotes the mean value of dataset X, and σ denotes the standard deviation of dataset X.

These standardized scores of the six textural image features were analyzed by applying hierarchical clustering. The result is shown in Fig. 10. For comparison with the results of the visual evaluation, revelant groups are circled in Fig. 10.

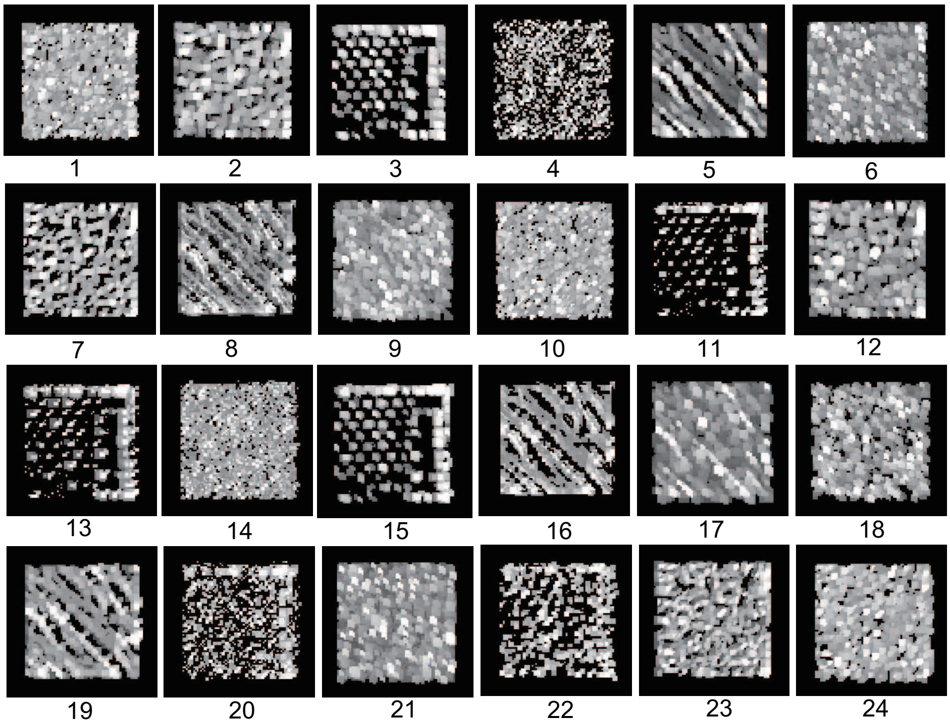

Fig. 7. Screen of randomly arranged manipulated textural images for the subjective visual experiment.

5. Discussion

The results show that the density, regularity, and directionality of the point configurations have significant effects on human visual impressions. For a texture without significant density, regularity, or clear orientation information in its point configuration, the primitive has an effect on the visual impressions.

In the experiments, more than ninety percent of the respondents classified textures 5, 8, 16, and 19 in Fig. 7 into the same group and more than eighty percent classi-fied textures 3, 11, 13, and 15 in Fig. 7 into the same group.

Textures 5, 8, 16, and 19 are circled in Fig. 8 as Group 1, and textures 3, 11, 13, and 15 as Group 2(a). The textures of Groups 1 and 2(a) are illustrated in Fig. 11. Textures 5, 8, 16, and 19 share the same point configuration but have different primitives. The visual impressions varied very little because the original point configuration preserved much of the orientation information. A similar situation occurred with textures 3, 11, 13, and 15. In this case, the density and regularity of the point configuration dominated the visual impressions.

While identifying a texture whose point configuration featured unclear characteristics, respondents paid more attention to the impression of the primitive. For example, more than eighty percent of the respondents tendentiously classified textures 6, 9, and 21 in Fig. 7 into the same group. Although these were generated from different original textures, they share the same primitive. As a counterexample, less than thirty percent of the respondents classified textures 2, 7, 12, and 23 in Fig. 7 into the same

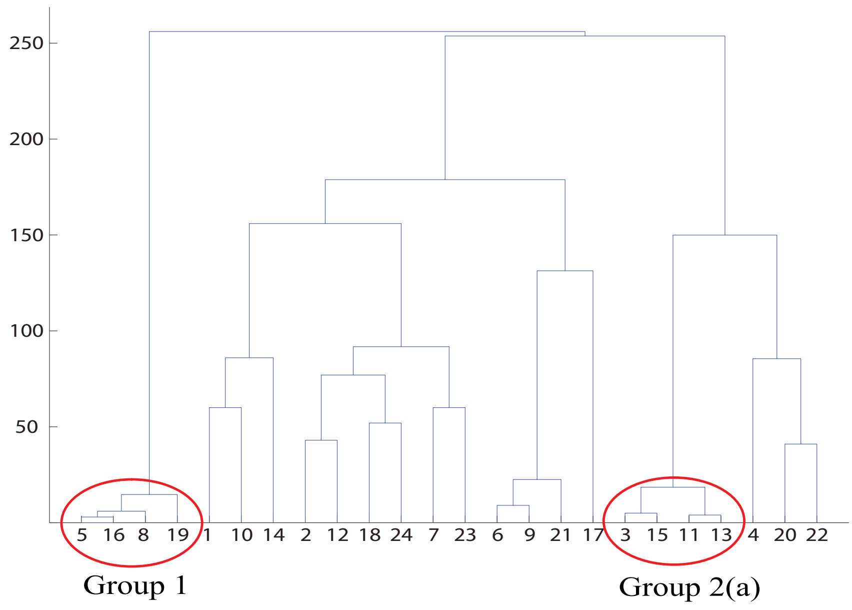

Fig. 8. Hierarchical clustering dendrogram of visual evaluation.

group, even though those textures share the same point configuration.

Textures with clear global density, regularity, and orientation information, such as textures 5, 8, 16, and 19 in Fig. 7, were clustered into the same group in the objective evaluation adopting statistical image features as in the results of the visual evaluation. Textures 3, 11, 13, 15, and 22 are circled in Fig. 10 as Group 2(b). The textures in Group 2(b) are almost the same as those in Group 2(a) except for texture 22. Texture 22 has similar density, directionality, and grain size to those of textures 3, 11, 13, and 15. However, in current image feature analysis we have not adopted regularity as one of the image features. This shows the importance of regularity in human texture perception. If we add regularity to the image feature analysis, Group 2(b) may have the same clustering result as Group 2(a).

The clustering results of image feature analysis are different from those of visual evaluation for the textures without significant global characteristics indicating that local textural features have become the discrimination criteria. More accurate statistical classification can be achieved by adopting relevant local textural features.

The results in the gray scale case of this investigation generally support those in the binary case of our previous work.[20] However, the comparison with image features analysis, such as texture anisotropy, can be achieved only in the gray scale case.

Compared with the supervised similarity judgment experiments used by many researchers,[10–12] the grouping task in our research is an unsupervised procedure similar to Rao's experiment.[13] In a typical supervised experiment, the respondent is given a brief explanation of the basic concept of textural features in the feature set before the respondent rates the strength of each feature. These supervised experiments may cause the respondents to focus only on the explained features, while their

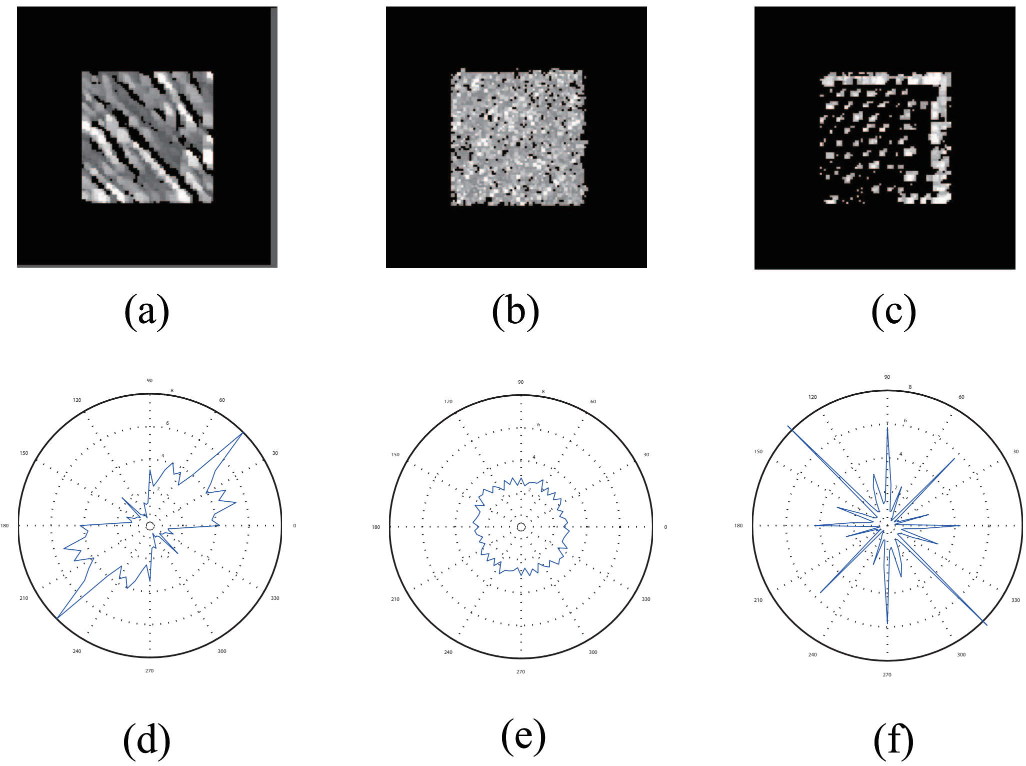

Fig. 9. Typical textures and their edge orientation histograms.

relationship and more salient textural features may be ignored in the rating tasks. The unsupervised experiment allows the respondents to evaluate and determine the similarities of the textures using their own knowledge. In this procedure, all the effective textural features are evaluated by the human visual system. However, in unsupervised experiments, it is difficult to investigate the effect produced by specific textural features separately. Further analysis of supervised and unsupervised experiments will be carried out.

Furthermore, visual impressions of a texture vary under different viewing distances or illumination conditions. Different conditions of perception may change the mutual interactions of global and local features in a texture. This assumption suggests that a computational texture model considering the parameters of those conditions is required. Investigating the relationship between human visual impressions and conditions of perception is one of our future projects.

6. Conclusions

A method of evaluating human visual impressions of gray scale textures created by morphological manipulation has been proposed. We are concerned with which features of a texture produce the most significant effects on human visual impressions. However, we have not employed conventional textural feature measurements such as Fourier analysis or the co-occurrence matrix method. We have introduced a novel texture analysis method that considers textural features in two categories: global and local. The proposed method also enables separate modification of local and global features of a

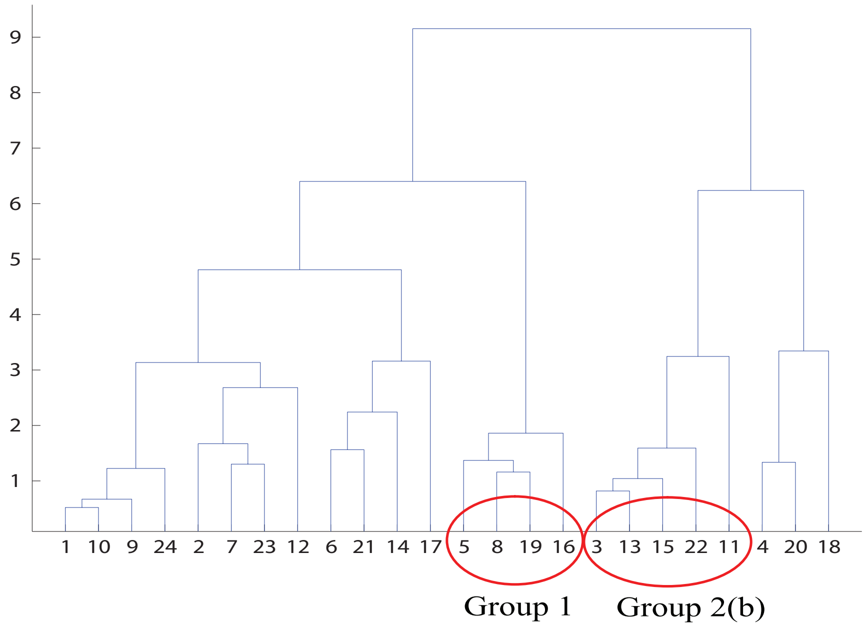

Fig. 10. Hierarchical clustering dendrogram of evaluation based on image features.

textural image. The initial manipulations made the experimental textures not only limited to existing natural textures but also specifically modified textures. This systematic manipulation enabled us to track the variations in human impressions by controlling the degree of texture modifications.

In this paper, features and characteristics of gray scale textures have been modified by applying morphological texture manipulations in which the primitives and skeletons of the textures have been manipulated systematically. The manipulated textures have been used as experimental textures in human evaluation of similarities. The results of the human evaluation have been analyzed by hierarchical clustering and compared with the results of evaluation based on six global image features, namely, skeleton density, isolated regions on the skeleton, grains of the largest size, uncovered regions, length ratio between minor and major axes of edge orientation histogram, and roughness of edge orientation histogram. The subjective human visual and objective image feature evaluations had similar classification results for textures with significant global features. The experimental results showed that global features of a texture have significant effects on human visual impressions. In the case of a texture without significant characteristics in its point configuration, local features have some effect on human visual impressions.

The types of textures used for visual evaluations were limited in this research because it was difficult for the respondents to classify multifarious textures at the same time. Further experiments on different types of textures will be carried out.

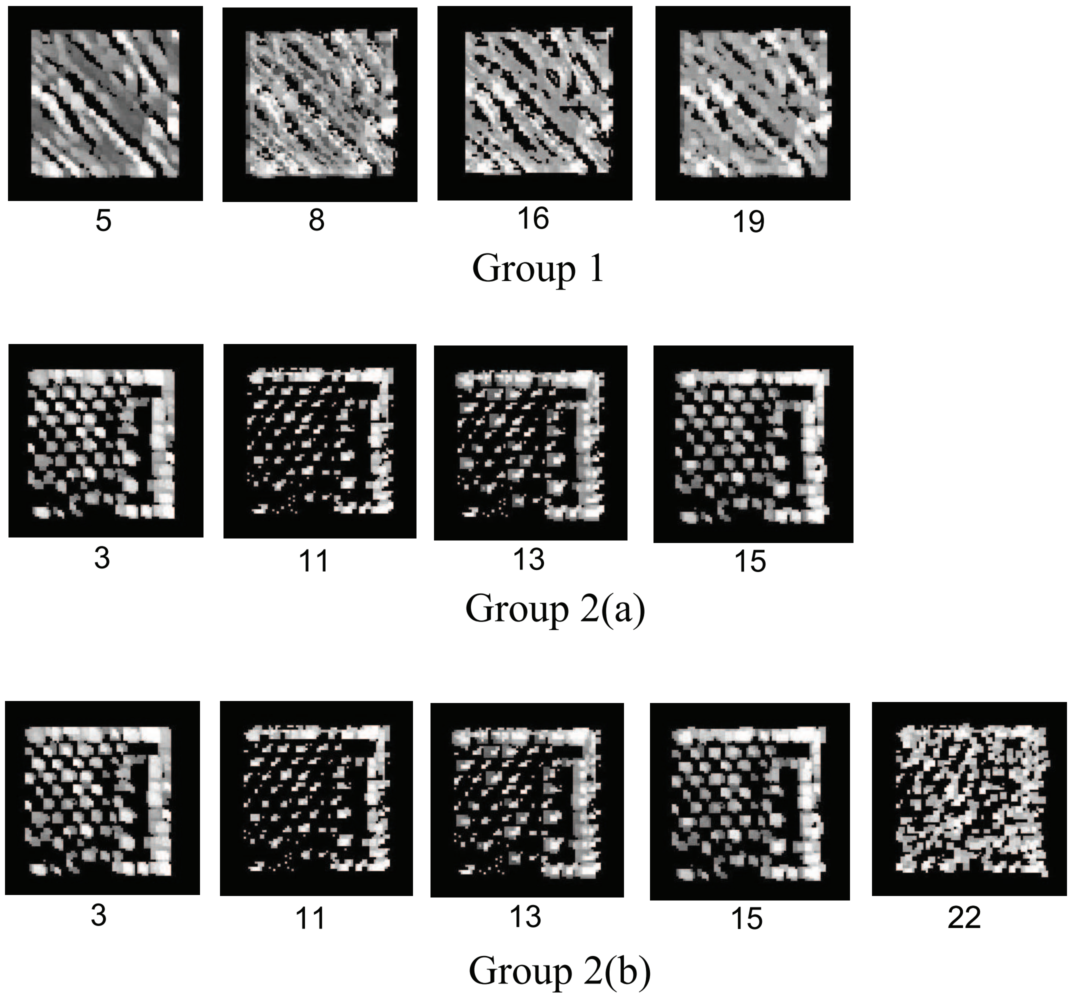

Fig. 11. Typical clustered groups in visual evaluation and evaluation based on image features.

Acknowledgment

This work was partially supported by a Grant-in-Aid for Scientific Research (C), *KAKENHI* (20560357).

The authors thank the respondents who participated in the experiments.

References

1) M. Petrou and P. Garicia-Sevilla: *Image Processing: Deal with Texture* (John Wiley & Sons Ltd, 2006).
2) T. Ojala and M. Pietikainen: in *Texture classification*, ed. R. B. Fisher (CVonline - Compendium of Computer Vision, 2001) Available: http://www.dai.ed.ac.uk/CVonline.
3) J. M. H. du Buf: Signal Processing **21** (1990) 221.
4) T. Chang and C. C. J. Kuo: IEEE Trans. Img. Proc. **2** (1990) 429.
5) B. Julesz: IRE Trans. Inf. Theory **8** (1962) 84 .
6) B. Julesz: Nature **290** (1981) 1619.
7) J. Malik and P. Perona: J. Opt. Soc. Am. A **7** (1990) No.5, 923.
8) A. Treisman and S. Gormican: Psychol. Rev. **95** (1988) No.1, 15.
9) B. J. Balas: Vision Res. **46** (2006) 299.
10) H. Tamura, S. Mori, and T. Yamawaki: IEEE Trans. Syst. Man Cybern. **SMC-8** (1978) 460.
11) M. Amadasun and R. King: IEEE Trans. Syst. Man Cybern. **19** (1989) 1264.
12) K. Fujii, S. Sugi, and Y. Ando: Psychol. Res. **67** (2003) 197.
13) A. R. Rao: CVGIP: Graphical Models and Image Processing **5** (1993) No.3, 218.
14) A. R. Rao and G. L. Lohse: Vision Res. **36** (1996) 1649.
15) R. Y. Cho, V. Yang, and P. E. Hallett: Perception and Psychophysics **62** (2000) No.4, 735.
16) A. Asano, T. Ohkubo, M. Muneyasu, and T. Hina-moto: Proc. 13th Scandinavian Conference on Image Analysis, 2003, p. 178.
17) J. Serra, ed.: *Image Analysis and Mathematical Morphology Volume 2. Technical Advances* (Academic Press, 1988) p. 217.
18) P. Soille: *Morphological Image Analysis* (Springer, Berlin, 2003) 2nd ed., p. 158.
19) H. J. A. M. Heijmans: *Morphological Image Operators* (Academic Press, 1994) p. 319.
20) C. Muraki, A. Asano, and T. Fujimoto: Proc. International Conference on Kansei Engineering and Emotion Research, 2007, F-8.
21) P. Brodatz: *Textures: A Photographic Album for Artists and Designers* (Dover, New York, 1966).
22) A. Fujiki, A. Asano, and M. Muneyasu: Proc. 2006 International Symposium on Nonlinear Theory and its Applications, 2006, p. 711.