An Investigation of the Relationship between Visual Features and Kansei Perception for Scenery Images

An Investigation of the Relationship between Visual Features and Kansei Perception for Scenery Images

Yanyu Gao*, Zhiqiang Liu(M1), Takashi Uozumi **
* SVBL **Department of Computer Science and Systems Engineering, Muroran Institute of Technology

1. Introduction

Discovering and constructing the relationship between image's visual features and the subjectivity of human perception is an important problem in affective computing and Kansei engineering. However, since human's perception mechanism, not only to the whole image, but also to the basic visual components, such as color, texture, and spatial layout, has not been identified completely, this task still remains to be one of the most difficult issues.

In this paper, we try to solve this problem from the angles of psychology and statistics. We first select 5 pairs of adversative adjectives to describe human's feeling and impression to natural scenery images by questionnaire experiments and principal component analysis (PCA). These adjectives are also called as Kansei factors [1]. Then, multidimensional scaling (MDS) technique is utilized to obtain the semantic visual elements of each pair of Kansei factors. Using these semantic visual elements, we can connect human's subjective perception with image's visual features.

2. Kansei factors

In this paper, we focus on the natural scenery images and generate our Kansei factors by two small-scale questionnaire experiments.

In experiment I, we collected 35 pairs of adjectives from travel books and magazines, which were regarded as being able to comprehensively represent the feeling and emotion that viewers may have when they gaze at scenery images. Then, 20 subjects (18 male and 2 female) were asked to pick up appropriate adjectives to represent their feeling to 36 natural scenery images, with at least one word for each image. These images included different outdoor sceneries in different seasons and period. All subjects had normal color vision and were not familiar with these images. According to the questionnaire results, 16 pairs of adjectives with using frequency more than 35% were selected (as shown in Table 1).

Table 1. Sixteen pairs of commonly used adjectives and their frequencies

Affective words	Rate	Affective words	Rate
plain / gorgeous	53%	vivid / simple	40%
depressed / cheerful	62%	warm / cool	54%
modern /traditional	37%	peaceful / excited	46%
energetic / decadent	55%	crude / refined	38%
desolated / flourishing	42%	calm / agitated	35%
placid/ active	38%	bright / dark	66%
refreshing / heavy	51%	natural / artificial	37%
beautiful / horrible	60%	romantic / austere	51%

The experiment II, aiming at removing those synonymous or correlative adjectives words, was performed by another 20 volunteers (17 male and 3 female). Each participant was shown the same scenery images as those used in experiment 1 along with an answer sheet that contained a list of the 16 pairs of emotional words. For obviating the influence from display order of images to human's subjective evaluation, these images were shown in different order for different subjects and the order was determined randomly. The participants were asked to answer "how much do you feel to the image" by selecting a number between -5 and 5, where "0" represents not yet having the feeling, and "5" represents having the positive feeling with the strongest degree, while "-5" represents having the negative feeling with the strongest degree. In this way, we obtained 20 answer sheets with each sheet defined as a 36×16 matrix. Then, we adjusted rows of each subjective matrix to make these matrixes have the same row order and averaged them to be a single 36×16 matrix, where each image is represented as a 16-dimensional feature vector.

The average matrix can also be regarded as 16 Kansei feature vectors, with each vector representing the feature of a pair of Kansei adjectives and having 36 elements. For removing synonymous and correlative words, we use principal component analysis (PCA) to uncorrelate these features vectors. After expurgate those attribute-meaning adjectives, such as bright and dark, finally we got 5 pairs of adversative adjectives, which in our opinion are irrelevant and can describe outdoor sceneries from different angles:

$$K_{ij} = \begin{Bmatrix} \text{cool, warm} \\ \text{depressed, cheerful} \\ \text{placid, active} \\ \text{austere, romantic} \\ \text{plain, gorgeous} \end{Bmatrix}, \text{ with } i=1,\ldots,5, j=1, 2.$$

3. Multidimensional scaling to obtain semantic visual elements

Multidimensional scaling (MDS), which was developed primarily by psychometricians and statisticians, has been a popular technique for analyzing experimental data in the physical, biological, and behavioral sciences [2]. It concerns with arranging a set of points in a low-dimensional space, so that the distances between the points resemble the perceived dissimilarities between objects as closely as possible. This concept has been used in image analysis community to identify the perceptual dimensions of textures [3] and color pattern [4], and in the image retrieval to measure the similarity of images based on their color histograms [5].

Since interpreting the dimensions of MDS configuration can lead to an understanding of the processes underlying the perceived nearness of entities, in this section we apply MDS technique to discover the relationship between image's visual feature and its perception effect. We first build 5 distance matrixes for a set of representative images from the angles of the five pairs of Kansei factors respectively. Then, MDS technique is adopted to plot these images in 5 maps with the distribution rule—"images that are perceived to be similar to each other in one Kansei viewpoint are placed near each other, while perceived to be different from each other are placed far from each other". By seeking visual commonness of those adjacent images, we may discover the relevant visual elements for each pair of Kansei factors.

3.1 Subjective evaluation and dissimilarity measurement

Although it is easy for human to evaluate the dissimilarities between images according to their visual contents, it is difficult for them to give similarity evaluation from one Kansei viewpoint. Here we propose an indirect measurement method. First, a small-scale subjective experiment was performed for obtaining the quantified Kansei measurement to 24 carefully selected images (as shown in Fig.1), in which 40 volunteers (35 male and 5 female) participated. All of these participants had full color perception and non-technical background. Their ages ranged from 20 to 40. Those representative images consisted of different color compositions, spatial frequencies and arrangements, and included broad contents, such as mountain, river, seashore, desert, and clouds.

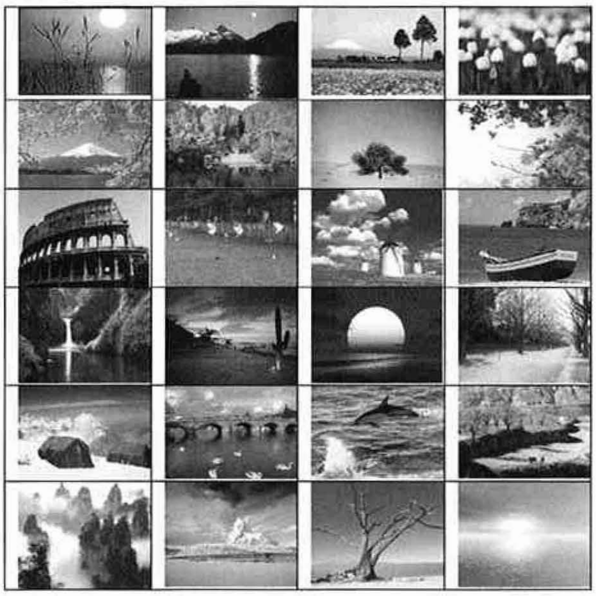


Fig.1. Twenty-four representative images

Participants were asked to watch these images and give their subjective evaluation to them from the angles that those Kansei factors represent using semantic differential (SD) method (as shown in Fig.2). In addition, participants were asked to explain their evaluation, which would be used later to interpret the MDS configurations.

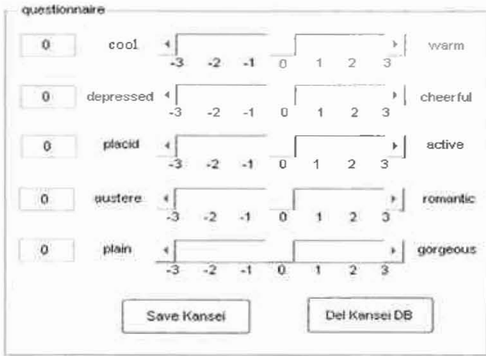


Fig.2. SD method for subjective evaluation

The evaluation of each subject constituted a 24×5 matrix (namely evaluation matrix), where each column represented the Kansei evaluations of all images from one viewpoint. According to column information of the evaluation matrix, we can calculate the dissimilarities of all 276 possible images pairs. With the consideration that the distance between contrary evaluations is much larger than that of same direction evaluations, we propose to use (1) to calculate the normalized Kansei distances:

$$\delta_{i,j} = \begin{cases} abs(K_{k,i} - K_{k,j})/10, & K_{k,i} \cdot K_{k,j} \geq 0 \\ (4 + abs(K_{k,i} - K_{k,j}))/10, & K_{k,i} \cdot K_{k,j} < 0 \end{cases} \quad (1)$$

where k =1,...5; i, j = 1,...,24. In this way, for each subject and from each Kansei viewpoint, we obtained a 24×24 dissimilarity matrix, where zero represents "the same feeling in the Kansei feature" and 1 represents "the furthest feeling in the Kansei angle". By averaging the dissimilarity matrixes got from all subjects, we obtained 5 mean dissimilarity evaluation matrixes Δ corresponding to the 5 pairs of Kansei factors. These dissimilarity evaluation matrixes are all symmetric matrixes with nonnegative elements and zeroes on the diagonal.

3.2 Classical multidimensional scaling

There are many kinds of MDS according to different objective function and optimization algorithms. Among them, classical MDS (CMDS) is a simple, global, noniterative technique, which only analyzes one kind of distance measurement.

The central concept of CMDS is that the distance d_{ij} between points in an n-dimensional configuration space X will have the strongest correspondence to the component δ_{ij} of the mean dissimilarity evaluation matrix Δ [4]. The degree of correspondence is usually measured by Kruskal Stress function:

$$stress\ (\Delta, X, f) = \sqrt{\frac{\sum_i \sum_j [f(\delta_{ij}) - d_{ij}]^2}{\sum_i \sum_j f(\delta_{ij})^2}} \quad (2)$$

where d_{ij} refers to the Euclidean distance between points i and j on the configuration space X, δ_{ij} refers to the component of similarity evaluation matrix Δ. The function of the input values $f(\delta_{ij})$ is a weakly monotonic transformation, which is usually computed via the monotone or isotonic regression, such as $d=f(\delta)=a\delta+b$, where a and b are constant for a given configuration.

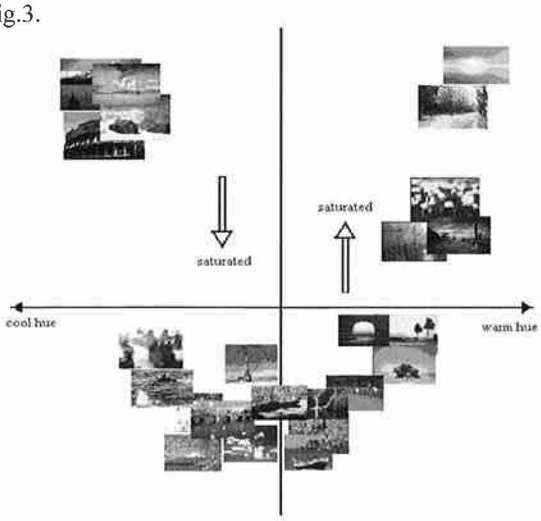
When the *stress* value gets to the minimum, the correspondence between points in n-dimensional configuration space X and components of the mean dissimilarity evaluation matrix Δ would be the best. Therefore, our task is to find the best transformation function f and configuration space X to make the *stress* minimum. The calculation procedure of CMDS can be summarized as follows: 1) Choose the initial configuration X_0 in L-dimensional space; 2) Find the best transformation function f for the initial configuration X_0 by isotonic regression; 3) Search for the best configuration X for the points in the subjective proximity evaluation space by the steepest descent method; 4) Calculate the stress function by equation (2); 5) Increase dimensionality L to L+1 and return to step 1) until further increase in the number of dimensions does not bring a reduction to the value of stress function.

Theoretically, those representative images would be clustered into 7 regions in the configuration space with one region representing an evaluated score. Therefore, we just need to calculate 2-D and 3-D CMDS configuration. The stress values of each pair of Kansei factors are shown in Table 2.

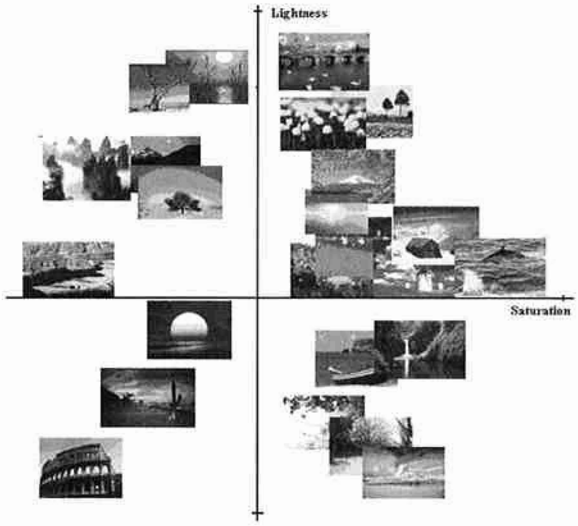
Table 2. The stress values for 2-D and 3-D CMDS configurations of each pair of Kansei factors

Kansei factors	Stress	
	2-D	3-D
cool vs warm	0.6266	0.3709
depressed vs. cheerful	0.6192	0.3075
placid vs. active	0.6159	0.3578
austere vs. romantic	0.5784	0.2994
plain vs. gorgeous	0.6387	0.3064

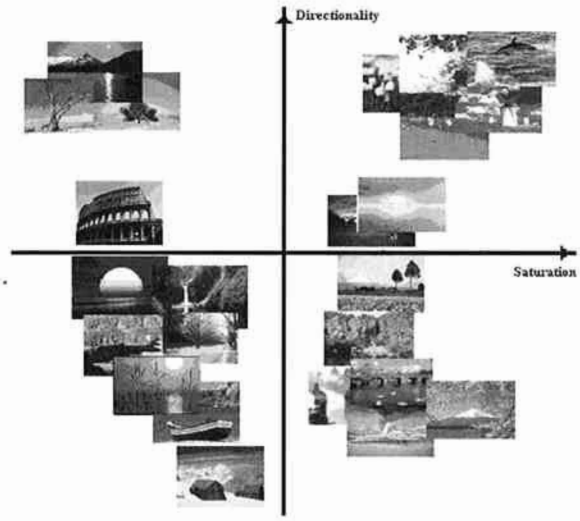
The 2-D configuration of each pair of Kansei factors and corresponding dimension definition are shown in Fig.3.



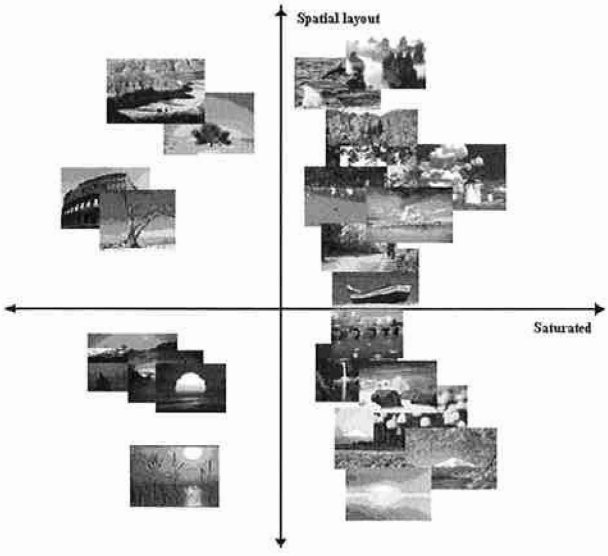
(a) cool vs. warm



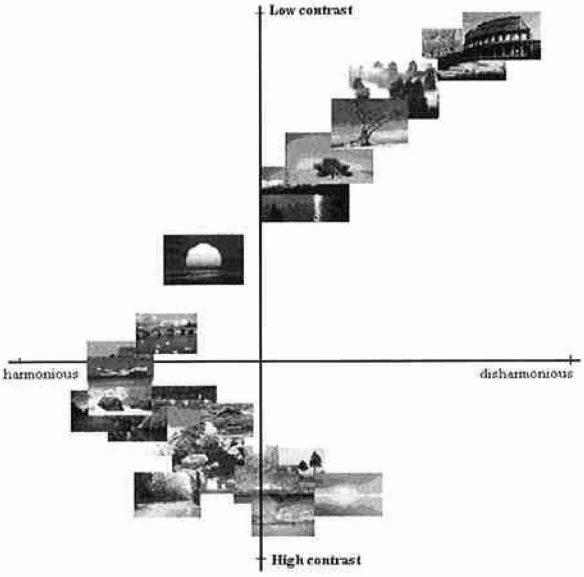
(b) depressed vs. cheerful



(c) placid vs. active



(d) austere vs. romantic



(e) plain vs. gorgeous

Fig.3. 2-D CMDS configuration for each pair of Kansei factors

3.3 Dimensions definition of CMDS configuration space

Once the 2-D and 3-D CMDS configuration of each pair of Kansei factors is obtained, we are left with the task of interpreting and describing the meaning of each dimension. Here we utilize the explanations from participants to accomplish this task. The dimension meanings of the five pairs of Kansei factors in 3-D space are shown in Table 3.

Table 3. Dimension meaning of the five pairs of Kansei factors

Kansei		D	Dimension meaning
1	Cool *vs.* Warm	1	**The hue of representative color** yellow undertone colors—warm blue undertone colors—cool
		2	**The saturation of representative color** saturated color—warm unsaturated color—cool
		3	**The luminance of representative color** dark colors—warm; light colors—cool
2	Depressed *vs.* Cheerful	1	**The saturation of representative color** saturated color— cheerful unsaturated color— depressed
		2	**The luminance of representative color** light colors— cheerful dark colors — depressed
		3	**Color scheme** harmonious color scheme— cheerful disharmonious scheme— depressed
3	Placid *vs.* Active	1	**The saturation of representative color** saturated color— active unsaturated color— placid
		2	**Directionality** isotropic texture— placid direction texture — active
		3	**Color scheme** hue disparity between 30° and 45° color scheme— active same hue color scheme—placid
4	Austere *vs.* Romantic	1	**The saturation of representative color** saturated color— romantic unsaturated color— austere
		2	**Spatial layout** quadrate layout— austere circle layout— romantic
		3	**Overall coarseness** coarse texture— austere fine texture— romantic
5	Plain *vs.* Gorgeous	1	**Color scheme** complementary color scheme—gorgeous adjacent color scheme—plain
		2	**Overall contrast** high contrast —gorgeous low contrast— plain
		3	**The saturation of representative color** saturated color— gorgeous unsaturated color— plain

These dimension definitions can be regarded as the semantic visual elements of the five pairs of Kansei factors, which bridge the gap between visual features and Kansei understanding.

4. Discussion and conclusions

In this paper, we propose to apply CMDS technique to investigate the relationship between image's visual features and the subjectivity of human perception. However, we only deduce 3 semantic visual elements for each pair of Kansei factors, which is an ideal assumption. Sometimes, these semantic elements are not enough to reflect the relationship between human's Kansei and images' visual features. Therefore, more semantic visual elements need to be discovered and investigated.

Moreover, the subjects that took part in our experiments are limited not only in quantity but also in age, career, etc. Therefore, large-scale questionnaires are necessary for obtaining objective results.

References

[1] J. A. Black, K. Kahol, P. Tripathi, P. Kuchi, and S. Panchanathan. Indexing natural images for retrieval based on query keywords, SPIE Proceedings of the Human Vision and Electronic Imaging Conference, vol.5292, pp.363-375, CA, 2004.

[2] T. F. Cox and M. A. A. Cox. Multidimensional Scaling, Chapman and Hall, London, 1994.

[3] A. R. Rao and G.L. Lohse. Identifying high level features of texture perception, CVGIP: Graphical Models and Image Proc., vol.55, no.3, pp.218-233, 1993.

[4] A. Mojsilovic, J. Kovacevic, D. Kall, R. J. Safranek, S. K. Ganapathy. The vocabulary and grammar of color patterns, IEEE Trans. on Image Processing, 9(3): 417-431, 2000.

[5] J. MacCuish, A. McPherson, J. Barros, P. Kelly, Interactive layout mechanisms for image database retrieval. Proc. SPIE, vol. 2656, pp.104-114, 1996.