Bull. Soc. Photogr. Imag. Japan. (2015) Vol. 25 No. 2: 19–24

Original Paper

Toward Image Emotion Space: Evaluation of Image Impression Shifts with Kansei Words

Kenji Kagitani[1], Deng Pei[2], Ei Matsuda[2], Tomoyasu Aizaki[1], Satoshi Nakamura[1], Toshihiro Suzuki[1], Michiyoshi Tachikawa[1], Naokazu Aoki[2], Hiroyuki Kobayashi[2]

Abstract: An image emotion space defining the relationship between the manipulation of a photographic image and the resulting change in the impression it produces was constructed by presenting original images and their manipulated versions in pairs to a panel of observers and asking them to identify the one giving the stronger impression for each of 59 emotion words. Picture-independent results could be obtained by comparing raw images of various pictures with those having various effects applied at random. Components governing the responses were extracted from the results by principal component analysis to construct a two-dimensional space with the components lightness and contrast and chroma as axes. Furthermore, it was shown that the differences in impressions can be clustered into 12 clusters by cluster analysis. A Kansei space model showing pathways to "preferred" was obtained by multiple linear regression analysis using "preferred" as the dependent variable and the representative words of their respective clusters as the independent variable. It consisted of two pathways, one for lightness information and one for color information.

Key words: image emotion space, effect, Kansei space model, pathways to "preferred"

1. Introduction

One of the main objectives of photograph retouching is to manipulate the photograph to provide a predetermined impression to the viewer. For example, in printing workflow, a request or instruction for photo retouching is transmitted using emotion words. Now that photo retouching applications for digital cameras and smartphones are widespread and we can share manipulated images, it is very important to know the change in the impression produced by an effect. It is of interest to construct an image emotion space[1)] defining the relationship between the manipulation of a photographic image and the resulting change in the impression it produces, not only from a practical point of view but also from a cognitive science point of view.

In this study, we presented original images and their manipulated versions in pairs to a panel of observers and asked them to identify the one giving the stronger impression for each of 59 emotion words. Components governing the responses were extracted from the results by statistical analysis to construct a two-dimensional space with the components as axes. Another goal was to construct a Kansei structure including a causal relationship between the word "preference" and the other emotion words.

An image includes not only the image quality but also a picture, and picture dependency must always be taken into consideration in subjective evaluations. Therefore, it would be ideal to understand the picture dependence of our impressions. This is difficult, however, because types of pictures are infinite in number and impossible to quantify. Bernat et al. measured physiological responses such as blinking and skin conductance as measures of affect for themes including erotica, adventure, victim, and threat[2)], but this is a rare case. In order to avoid picture-dependence, many image studies using the Kansei model[3)] have used a single color patch[4)] or a combination of two color patches[5)] as a stimulus. In contrast, in this study the aim was to obtain a picture-independent result by comparing raw images of various pictures with those having various effects applied at random and, further, to investigate the differences between the impressions created by the raw images and those created by the processed images.

2. Experiment

2.1 Stimuli

We obtained 60 photos depicting a variety of scenes and objects from the photo-sharing site stock.xchng (now called Freeimages)[6)], and we applied an image manipulation effect selected at random from 16 different types of Adobe Photoshop actions to each photo. Each effect (action) is a series of image manipulation procedures that combines the color tone and sharpness of change. Multiple ap-

Received: 24th, April 2015; Accepted: 9st, September 2015

[1]Ricoh Institute of Technology, Ricoh Co., Ltd.
16-1 Shineicho, Tsuzuki-ku, Yokohama 224-0035, Japan
[2]Graduate School of Advanced Integration Science, Chiba University
1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

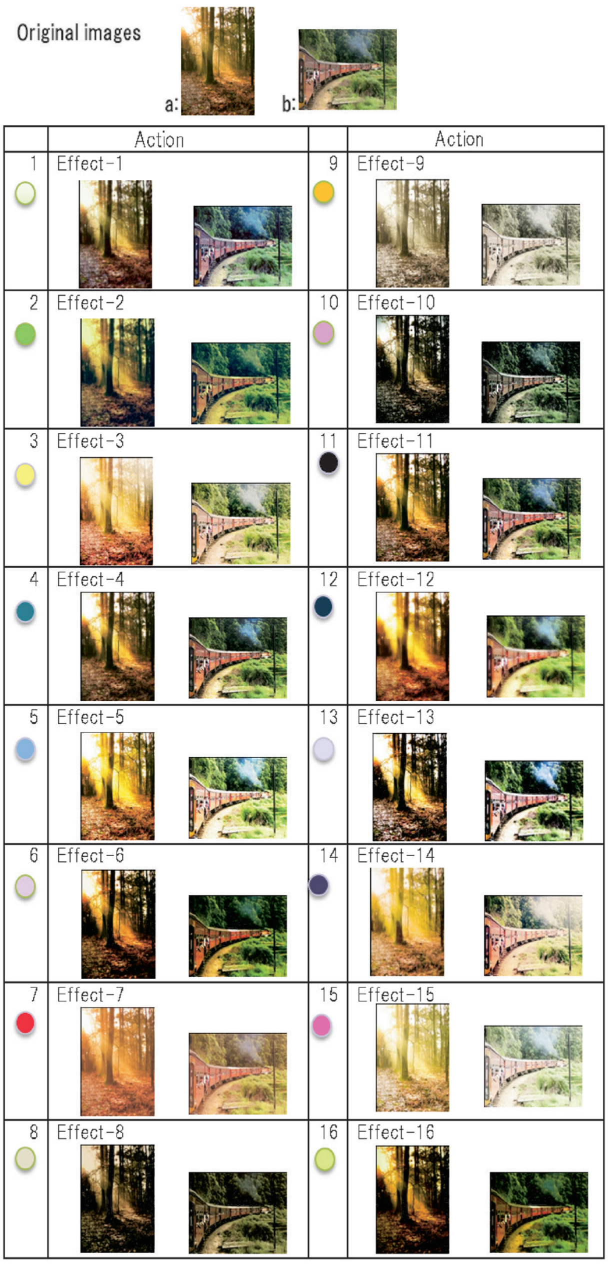

Fig. 1 Two original images and their versions after processing with each of the 16 actions.
The number and colored mark for each action will be used in the figures shown later; refer to this figure when viewing those figures.

plications of each action were allowed. The subjective evaluation experiment was performed using 60 pairs, each consisting of an original image and its processed version.

Further, in order to confirm the presence or absence of picture dependence, each of the sixteen effects was applied to two images of significantly different picture type, and the same subjective evaluation experiment was performed using these pairs. The original and processed images are shown in Fig.1.

2.2 Procedure

The panel of observers was shown an image pair consisting of the original and its processed image, arranged side by side or one above the other on the display, to evaluate them. The positional relationship of the pair was changed randomly every time to avoid prejudg-

Fig. 2 A sample pair (Image[9]).

Table 1 Emotion words (59 words)

natural, lonely, quiet, tasteful, real, dense, open, solemn, light, crisp, transparent, mild, brilliant, reasonable, clear, unstable, preferred, vital, hard, pure, cold, deep, cool and bitter, giving a moist feeling, giving a you-are-there feeling, giving a profound feeling, calm, rich, with texture, strange, sophisticated, swinging, well-modulated, unique, voluminous, smooth, smoky, pale, large-scaled, evil, sharp, impressive, peaceful, relaxed, three-dimensional, obscure, refreshing, magnificent, eerie, shining, healthy, evoking a feeling of the season, warm, garish, fresh, giving a sizzle feeling, giving an aerial feeling, simple and clear, dry

ment by the panel. The viewing time was not limited.

For each of 59 emotion words (Table 1), the panel was asked to select one of three choices: which image of the two presented was more appropriate for that emotion or (the third choice) neither of them. The adjectives were selected from a vast number of words and were chosen for their availability and their correlation with other words as determined through a preliminary experiment. Because the evaluation was performed using the emotion words in Japanese, Japanese ones are put with their English words.

Forty-three subjects in their 20s through 60s participated in this experiment, including students studying image technology and engineers working in an imaging-technology company.

Subjects in the experiment viewed the samples on their own liquid crystal display of 17 to 24 inches. The room illumination and the viewing distance were not specified.

2.3 Data analysis

Subject evaluations were scored by assigning -1 when the original image was deemed more appropriate for the emotion word, +1 when the manipulated one was chosen, and 0 when neither was selected as appropriate, and the data were statistically analyzed using the IBM SPSS Statistics (IBM Co.) software package[10].

The correlations between the score profiles for each emotion word for the 60 image pairs and the synonymity or antonymy between the words as represented by plus and minus, respectively, were reasonable for each word, using their dictionary meanings. This confirms the validity of the scoring method.

3. Results and Discussion

3.1 Principal component analysis (PCA)[11]

Five principal components were extracted from the experimental data by PCA. Contribution ratios of the first and the second principal component are 33.2 and 23.0, respectively. The cumulative contribution ratio of the two principle components exceeds 56%. The words of large component loading for the first principle component are "voluminous", "impressive", "magnificent", "pale", and "dense", and those for the second principle component are "cool and bitter", "quiet", "lonely", "calm", and "cold". It is interpreted that the respective components summarize "dignity" and "vividness".

The loadings for each principal component were obtained for each emotion word. The component score for each principal component was calculated for each image pair by adding loading for the principal component multiplied by average score of each subject together. As shown in Fig.3, plots for image pairs manipulated by the same effect (shown as plots of the same color) were grouped in a certain area in the scatter plot with axes of the first and the second principal component. Impressions produced by the images were classified roughly according to the effects; the results suggested that impressions were independent of the pictures in the images. This was confirmed by applying each effect to images of the same picture and plotting the results of PCA in the same space, as shown in Fig.4. The results for the two pictures fell in almost the same areas as in Fig.3 for each respective effect.

The changes produced by the actions were investigated by applying them to color patches of the Macbeth Color Checker. The results are shown in Fig.5 comparing with the color chart without the actions. Changes in lightness (L*), chroma (C*), and tone brought about by applying the actions are shown in Fig.6. The actions are grouped into classes by the results of principal component analysis. The L* values were obtained by averaging lightness of six grayscale color patches (White, Neutral 8, Neutral 6.5, Neutral 5, Neutral 3.5, and Black) and the C* values were obtained by averaging chroma of the primary colors patches (Red, Green, and Blue). Tone curves were obtained by plotting the L* values of the six grayscale colors patches against those of the original. They are shown in the same four groups with Fig.6 in Fig.7.

Actions located in negative score region of the first principle component increase L* and those located in positive score region decrease or do not change L*. Clear differences in C* were not ob-

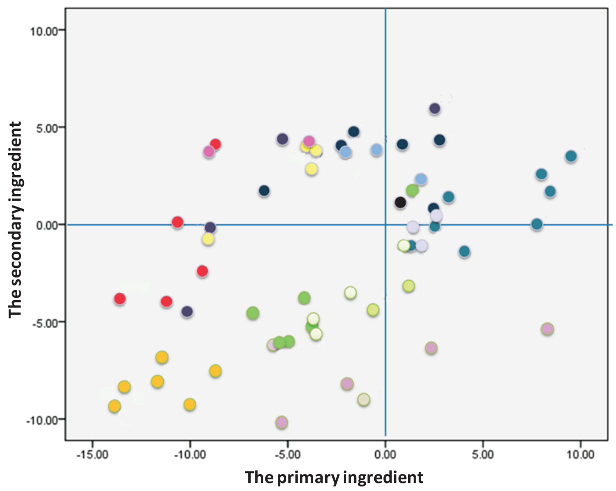

Fig. 3 A scatter plot of the primary and secondary ingredients. Color marks: See Fig.1.

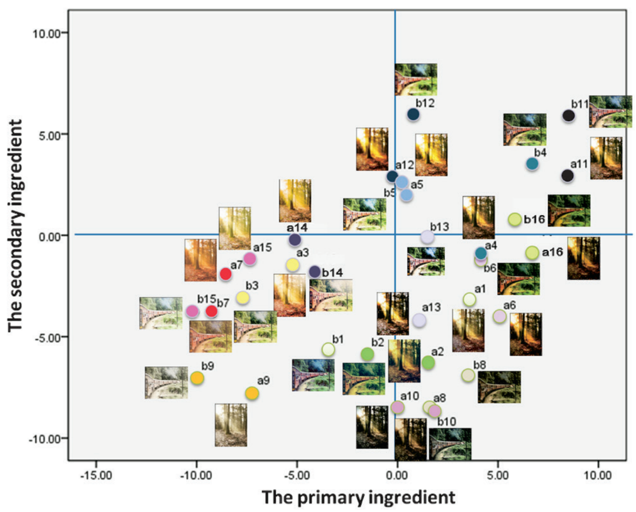

Fig. 4 Validation of the results shown in Fig.3 by applying each of the actions to an image. Color marks and numbers: See Fig.1.

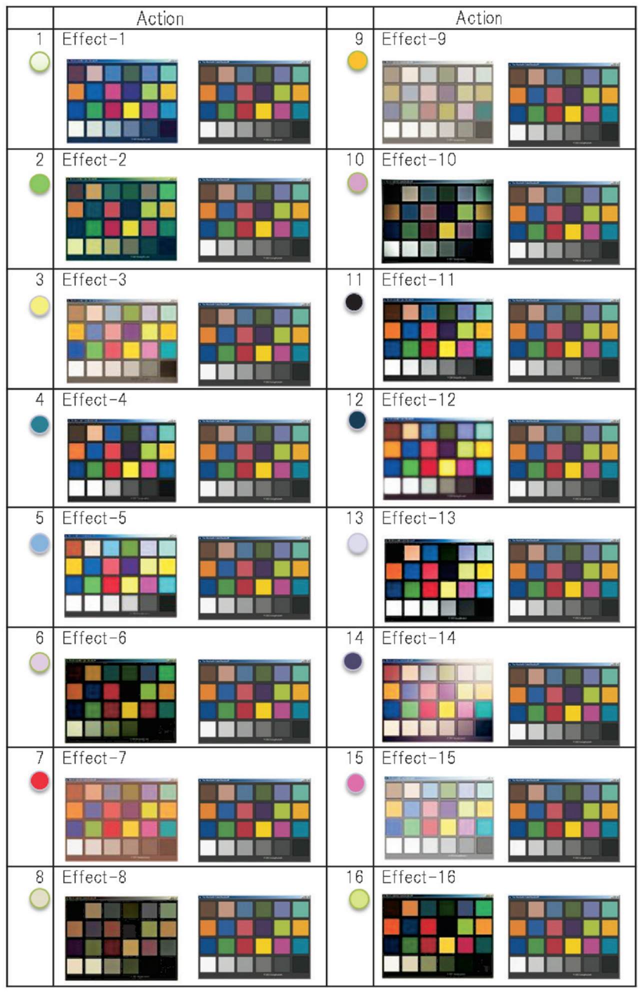

Fig. 5 Color patches of the Macbeth Color Checker. after processing with each of the 16 actions (left) compared with the original (right)

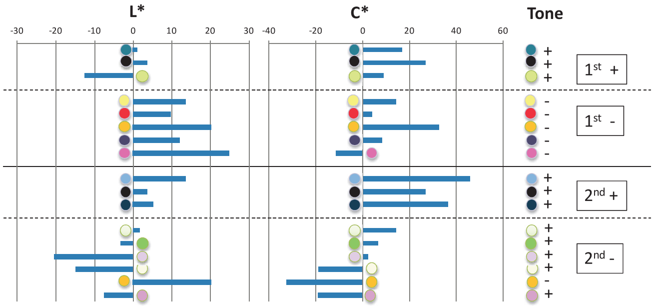

Fig. 6 Change in lightness(L*), chroma (C*), and tone brought about by applying the actions. The actions are grouped into classes by the results of principal component analysis. The marks + and – for tone indicate "become harder" and "become softer" by application of actions, respectively (see Fig.7).

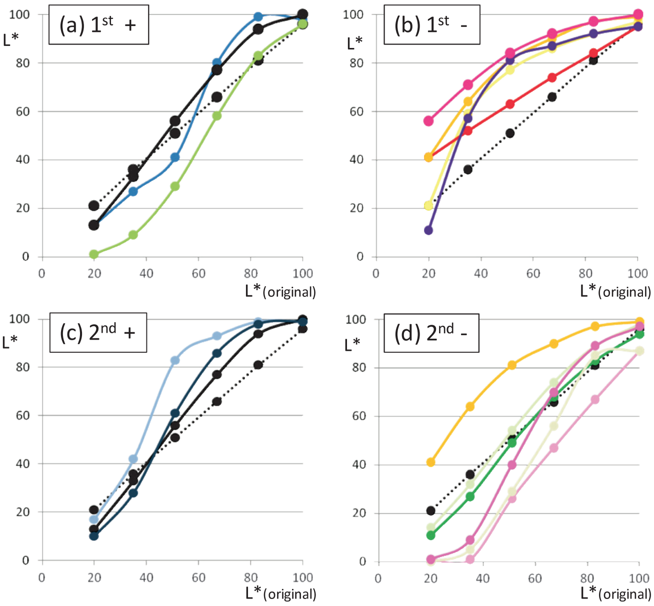

Fig. 7 Change of tone curve by application of the actions. The actions are grouped into classes by the results of principal component analysis as in Fig.6.

served between the two action groups. With regard to tone an obvious difference was observed that actions in positive score region make harder and those in negative score region make softer.

With regard to the second principle component no obvious difference in change of L* and tone was observed between the positive score group and the negative score group. On the other hand, C* of the positive score group increases much larger than the negative score group.

From all of these results it is suggested that the first principle component axis is for lightness and contrast and the second principle component axis is for chroma and that the combination of lightness with contrast and chroma control "dignity" and "vividness", respectively.

3.2 Multiple linear regression analysis (MLRA)[12] and covariance structure analysis (CSA)[13]

The emotion words were clustered according to similarities in the evaluation results, and we used MLRA to obtain a structural model for the emotional response that would show causal connections be-

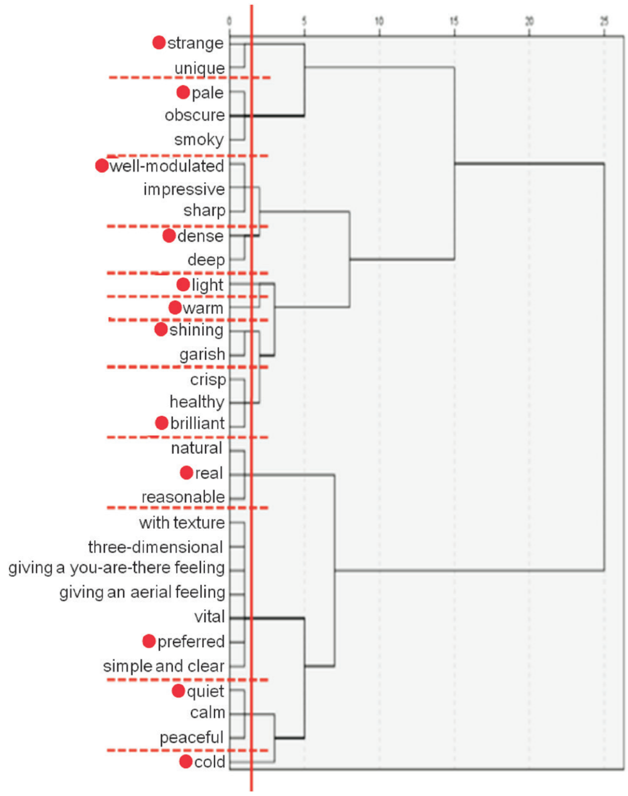

Fig. 8 Dendrogram obtained by cluster analysis (Ward's method).

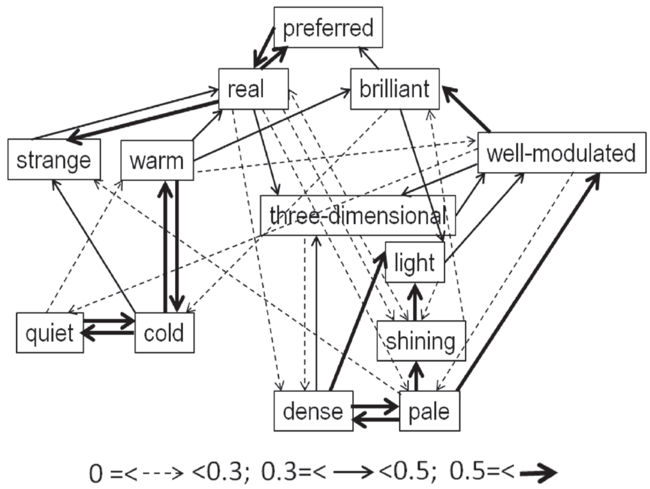

Fig. 9 A model representing two-way causal connections between 13 words (the representative words of the 12 clusters plus the word "preferred").
Independent variables are those at arrow tails, and dependent variables are those at arrowheads. Magnitudes of standard partial regression coefficients are represented by the thicknesses of the arrows.

tween the clusters. First, we investigated clustering methods[14] appropriate for improving the fit of models obtained by MLRA and selection methods for identifying a word representative of each cluster. It was found that dendrogram clustering according to distance and choosing the word with the largest dispersion as the representative one of the cluster were effective. In addition, performing the analysis after omitting words with a high frequency of "neither" resulted in an improved model fit, the best result being obtained at 50% of the frequency. In this case, 28 words were omitted, and 31

Table 2 Goodness-of-fit indices

Acronym	Explanation	Value	Critical values and polarity
CMNI (Minimum of discrepancy function)	chi-square value	47.892	CMIN/DF = 1.29 (Acceptable fit: 1~2)
DF (Degrees of freedom)		37	
P (Probability of CMIN)		0.108	Good fit: >0.1
GFI (Goodness-of-fit index)	proportion of variance that is accounted for by the estimated population covariance	0.900	Good fit: ~1
AGFI (Adjusted goodness-of-fit index)	GFI adjusted for degrees of freedom, with more saturated models reducing fit. Thus, more parsimonious models are preferred, while complicated models are penalized.	0.755	
RMSEA (Root-mean-square error of approximation)	measure of how well the model, with unknown but optimally chosen parameter estimates, would fit the population's covariance matrix	0.071	Good fit: <0.1

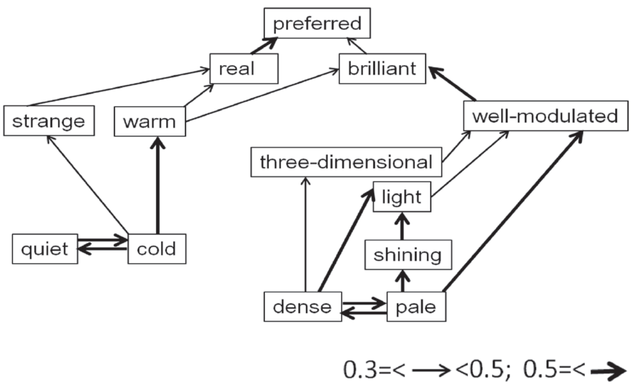

Fig. 10 A model representing causal connections.
Arrows with standard partial regression coefficients greater than 0.3 and pointing toward "preferred" are retained from Fig.7.

1 2 3

Fig. 11 Images that gave "strange" at high rates.
1[15]: With Action 12 applied; 2[16]: With Action 12 applied; 3[9]: With Action 9 applied.

words were clustered into 12 clusters (Fig.8). We executed stepwise MLRA for 13 words (the 12 words having the largest dispersion of each cluster plus the word "preferred") and constructed a model of high goodness of fit including two-way causal connections by using every word as the dependent variable and all other words as a candidate independent variable. Next, we used CSA to construct a model including all 13 words, and a model of high goodness of fit was obtained (Fig.9). Indices of model fit are listed in Table 2. The red solid line is the standard for clustering, and the broken lines are the cluster borders. The words with a red sphere are the representative words of their respective clusters.

A complex model including bidirectional causal connections was obtained as a significant one. However, pathways toward "preferred" emerged when connections with standard partial regression coefficients smaller than 0.3 were removed, leaving arrows pointing toward "preferred" as shown in Fig.10. That is, the pathways that start with words related to lightness, such as "dense" and "pale", go through "light" or "three-dimensional", "well-modulated" and "brilliant" and reach "preferred". This suggests that from the viewpoint of image quality, the enhancement of tone and sharpness achieved by a change in lightness results in "preferred". On the other hand, the pathways that start with "calm" and "cold" go through "warm" or "strange", and "real", and reach "preferred". We can connect "cold" and "warm" with cool and warm coloration and interpret this pathway as one for color information. Images that gave "strange" at high rates are of very high or very low chroma (Fig.11), suggesting that this pathway is also one of color information.

4. Conclusion

Differences in impressions produced by the existence or nonexistence of various effects were evaluated by their aptness for 59 emotion words. A two-dimensional Kansei space having lightness and contrast axes and chroma axes was constructed from the result by PCA. It was verified that this result is picture-independent. Furthermore, it was shown that the differences in impressions can be clustered into 12 clusters.

A Kansei space model showing pathways to "preferred" was obtained by MLRA using "preferred" as the dependent variable. It consists of two pathways, one for lightness information and one for color information. It is very interesting that these two pathways are analogous to those of our visual system from retina to brain, which are the "Where" system and the "What" system18). The former carries information about lightness and is used for motion perception, depth perception, spatial organization, and figure/ground segregation, and the latter carries information about color and is used for object recognition, face recognition, and color perception.

References

1) B. Oicherman, C. Staelin, Kansei Imaging: First Steps, HPL-2009-12, 2009
2) E. Bernat, C. J. Patrick, S. D. Benning, A. Tellegan, Effects of Picture Content and Intensity on Affective Physiological Response, Psychophysiology, 43(1): 93(2006)
3) L. C. Ou, M. R. Luo, A. Woodcock, A. Wright, A Study of Colour Emotion and Colour Preference. Part III: Colour Preference Modelling, Color Research and Application, 29(5), 381(2004)

4) L. C. Ou, M. R. Luo, A. Woodcock, A. Wright, A Study of Colour Emotion and Colour Preference. Part I: Colour Emotions for Single Colours, Color Research and Application, 29(3), 232(2004)
5) L. C. Ou, M. R. Luo, A. Woodcock, A. Wright, A Study of Colour Emotion and Colour Preference. Part II: Colour Emotions for Two-Colour Combinations, Color Research and Application, 29(4), 292(2004)
6) http://www.sxc.hu/
7) M. Richert, http://www.sxc.hu/photo/1330703
8) Sheepy, http://www.sxc.hu/photo/225852
9) T. Gulick, http://www.sxc.hu/photo/36013
10) S. Landau, B. S. Everitt, A Handbook of Statistical Analyses using SPSS, Chapman & Hall/CRC, 2004
11) I. T. Jolliffe, Principal Component Analysis, 2nd Ed., Springer, 2002
12) J. Higgins, The Radical Statistician, 5th Ed., Chap. 4, The Management Advantage, 2005
13) J. L. Arbuckle, Amos 5.0 Update to the AMOS User's Guide (S. 77-85), Chicago: Small Waters Corp., 2003
14) B. S. Everitt, S. Landau, M. Leese, D. Stahl, Cluster Analysis, 5th Ed., Wiley, 2011
15) A. Kovacs, http://www.sxc.hu/photo/98822
16) M. Alberto, http://www.sxc.hu/photo/1349619
17) M. Livingstone, Vision and Art, Chap. 4, Abrams, 2002