Mining and Construction of User Experience Content: An Approach of Feature Analysis Based on Image

Di Wang[✉], Nan Liang, Jiaming Zhong, and Liqun Zhang

Design Management Institute, Shanghai Jiao Tong University, Shanghai, China
cheirsh_9@l63.com, zhanglliqun@gmail.com

Abstract. This research is a preliminary study for a professional museum of print ads during the period of Republic of China. We selected 571 pieces of print ads of the Republic of China as sample to analysis the feature of them to make the visitors of museum could more immersive feel the print ads of the Republic of China and experience the unique charm of imagery modeling and visual language. The main methods used in the research are image feature analysis and image feature quantization calculation, we extract and summarize the common elements and culture style feature based on the analysis of multidimensional design elements. The research results could provide effective guidance for the design of the professional museum, including the overall atmosphere of the museum, thematic construction and situation creation.

Keywords: Image analysis · Feature quantification · Multidimensional scaling analysis · Correspondence analysis

1 Introduction

The period of the Republic of China was important for the transformation of Chinese advertising from tradition to modern. It was not only the beginning of our modern advertising, but also the first development peak of our advertising. Advertising can be seen as the living fossil of the history. Through the study of advertising during the period of the Republic of China, our historical and cultural heritage, from different dimensions, we can understand the width and height the design practice can reach and feel the economic development status of the society in that period which is also called as the early stage of modern design in China.

Culture heritage is the essence extracted from the people's work and living after a long time, and the fusion of a variety of traditional life inherited by generations and closely related to the life of people. With society gradually paying attention to the protection of historical and cultural heritage, more and more libraries and museums concerning the intangible cultural heritage are built. The expression of culture heritage should combine the carrier itself with experience and feelings of people during visiting to make visitors not only have certain knowledge of the carrier but also get a quiet different interactive experience including emotional and cultural awareness and aesthetic, etc. through the interactive relationship between visitors and museum.

In order to abstract design elements which meet the demands of contemporary aesthetic and are in accord with the experience of modern aesthetic, we choose 571 pieces of print ads during the period of the Republic of China as samples to analyze their style and characteristics based on the requirement for culture heritage protection.

2   Design the Research

This paper will do research from the following aspects:

- Using a custom-developed software to identify the image feature, abstract the information of image feature significantly related to advertising sample and form the variables indicators.
- Basing on the technology of feature analysis, study the specific attributes in information architecture, visual representation for information, the aesthetic characteristics and construction method, technology and cultural elements, etc. and form the research variable indicators in the process of advertising samples labeling.
- Explore the image significant correlation between the objective elements and subjective cognitive elements and find the main features by analyzing the research variable indicators with statistical calculation method and technologies, such as cluster analysis, factor analysis, correspondence analysis, multidimensional scaling analysis, etc.
- Provide professional and effective guidance for the design of museum basing on the research result, including the overall atmosphere, thematic construction, situation creation, etc.

The samples used in this paper were popular with the economic zone taking Shanghai as the focus during the period of 1910-1940. Figure 1 shows that it contains pieces of print ads, such as posters, paper wrappings, cards, etc. In order to do research, the following pictures are all coded.

Fig. 1. 571 pieces of print ads (partial)

3   Research

3.1   Procedures

- Research on recognizing the image feature of advertising samples basing on calculation: Through chromatographic, color tune and singularity analysis, find the image feature with which sample variable indicators have a tight correlation, extract and label it.
- Image analysis, variable definition and labeling basing on feature analysis (do parallel with step 1): do feature analysis for all the visual advertising samples, make a label list and form nominal level, ordinal level and scale variables for the subsequent statistical analysis.
- Data mining basing on multidimensional indexes: basing on the data achieved from step 1 and 2, excavate the dominant or potential significant correlation in multiple dimensions and multi-facts, build the framework of influencing factors and extract the common and individual feature of samples.

3.2   Research on Recognizing the Image Feature of Advertising Samples Basing on Calculation

The image feature of advertising samples contains several different dimensions. In this paper, image feature refers to image color feature which includes theme tune and representative color.

In this paper, all samples are JPG format data files and their color models are RGB. In the past, some color reduction research for the monuments establishing color reduction algorithm with the help of implements' real color. It made the precedent in color reduction for the monuments. But in this paper, the real objects can't be found to apply the reduction algorithm because the real objects in the picture of samples are abundant, and the samples span 30 years. Because of the limit and in order to protect the initial data, this paper will analyze directly the samples by RGB color model.

The research tools for recognizing the image feature in this paper are developed by our research team with JAVA language, including color spatial distribution analysis tool, theme tune extraction tool and theme tune similarity analysis tool.

Color spatial distribution analysis tool. Color spatial distribution analysis tool generates color points according to pixels in samples (Fig. 2), forming a color space, clusters the points and reduces their dimension by vector computation and principal component analysis to help researchers analyze the color feature of samples.

Theme tune extraction tool. Theme tune extraction tool does color feature calculation of pixel, obtains the constituent of global color tune separately (Fig. 3), then extracts component of the theme tune which can describe 80 % tune in samples (Fig. 4), does batch processing and analysis and generates a theme tune table of every sample for the following multidimensional scaling analysis among samples.
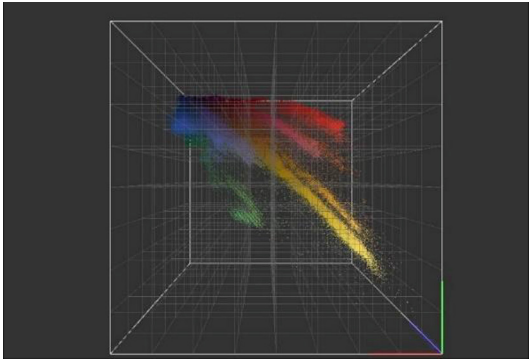
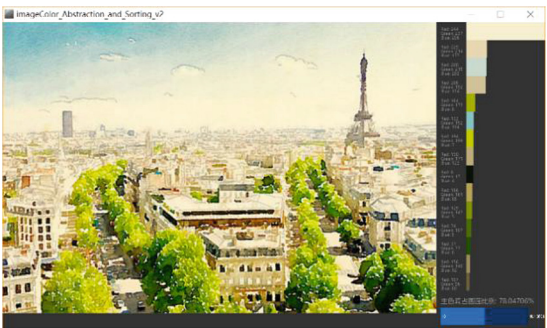**Fig. 2.** Color spatial distribution analysis tool colorSpace_3D_v1

Fig. 3. ImageColor_Abstraction_and_Sorting_v2

Fig. 4. ImageColor_Abstraction_and_Sorting_v3

Theme tune similarity analysis tool. Theme tune similarity analysis tool (Fig. 5) calculates the similarity of the theme tune between any two and generates a table with. csv format. The table will be used for the following multidimensional scaling analysis and statistical calculation of MDS.

	A	B	C	D	E	F	G	H	I	J	K
1	SampleCode0	SampleCode1	SampleCode2	SampleCode3	SampleCode4	SampleCode5	SampleCode6	SampleCode7	SampleCode8	SampleCode9	SampleCode10
2	0	29.223278	66.70082	123.85879	175.84084	26.627054	85.90111	155.18376	87.36704	66.71581	131.1259
3	29.223278	0	73.02739	124.34227	175.6303	53.99074	93.27915	150.43936	81.89628	82.78285	132.23463
4	66.70082	73.02739	0	58.034473	109.676796	59.04236	20.639767	92.52567	35.91657	30.757113	64.96922
5	123.85879	124.34227	58.034473	0	52.009613	116.64476	43.52011	38.06573	46.58326	76.66811	8.5440035
6	175.84084	175.6303	109.676796	52.009613	0	167.53806	93.13968	34.43835	95.3782	124.52711	44.877613
7	26.627054	53.99074	59.04236	116.64476	167.53806	0	74.60563	151.50908	88.588936	48.270073	123.146255
8	85.90111	93.27915	20.639767	43.52011	93.13968	74.60563	0	80.70316	41.689327	33.526108	49.487373
9	155.18376	150.43936	92.52567	38.06573	34.43835	151.50908	80.70316	0	69.70653	113.77609	35.52464
10	87.36704	81.89628	35.91657	46.58326	95.3782	88.588936	41.689327	69.70653	0	66.19668	54.08327
11	66.71581	82.78285	30.757113	76.66811	124.52711	48.270073	33.526108	113.77609	66.19668	0	82.29824
12	131.1259	132.23463	64.96922	8.5440035	44.877613	123.146255	49.487373	35.52464	54.08327	82.29824	0
13	62.26556	67.65353	6	62	113.84639	56.124863	25.65151	95.524864	36.249138	31.780497	69.260376
14	74.094536	77.92304	9.433981	49.96999	101.8332	67.96323	18.788294	83.7138	27.910572	38.79433	57.043842
15	95.32051	98.81296	28.86174	29.546574	80.88263	87.19518	16.27882	65.81793	32.04684	49.487373	36.138622
16	27.147743	27	84.0119	138.61458	190.23407	48.02083	102.81051	167.3828	100.6479	84.20214	146.5162
17	45.188496	43.312817	31.38471	81.981705	133.69368	50.783855	51.662365	110.79711	42.272923	51.351727	89.61027
18	46.67976	50.48762	26.1916	79.20859	130.916	47.644516	46.561787	110.059074	42.73172	45.738388	86.28442
19	95.885345	91.10434	39.153545	36.619667	85.24084	95.262794	39.255573	59.766212	12.124355	67.32756	44.56456
20	62.072536	68.03676	8.485281	63.118935	114.75626	55.118053	25.573423	96.87621	39.217342	29.086079	70.434364
21	89.721794	96.68506	23.853722	39.610603	89.050545	78.60661	4.582576	77.071396	41.12177	37.589893	45.321075
22	83.86298	80.454956	28.248894	44.474712	95.189285	83.03011	33.196384	71.44928	11.224972	57.253822	52.564247
23	101.0396	95.921844	44.11349	34.263683	81.27115	100.5286	43.011627	55.072678	14.764823	72.34639	41.70132
24	106.00943	100.86625	53.094257	42.461746	83.96428	106.47535	53.600372	58.532043	23.43075	82.29824	47.307507
25	28.809721	50.675438	44.82187	102.610916	153.85057	16.093477	61.81424	136.96715	73.28711	40.657104	109.15127

Fig. 5. The output result of Themetune similarity analysis tool

3.3 Image Analysis, Variable Definition and Labeling Basing on Feature Analysis

The main task of this paper is to form a structural system of variables to describe samples. After labeling the system, form a normal and ordinal description for variables to support the following descriptive statistics and correspondence analysis and excavate the common and different feature among samples.

This paper raises several variables for description basing on the overall analysis of the 571 pairs of advertising and labels the variables. The variables are classified into 6 categories.

- Content label (theme, product, product type, place of production, consumer orientation, brand, etc.);
- Composition label (diagonal, hub-and-spoke, zigzag, revolving shape, etc.);
- Expression label(direct display, emphasize on feature, contrast and foil, reasonably exaggerate, see big things through small, associative thinking, rich in humor, metaphorical transfer, use wonderful feeling to foil the product, suspense, idol, imitation, mystical method, multi pictures, etc.);
- Visual expression label (oil painting style, printing style, watercolor style, photo style, line drawing style, ink style, meticulous style, etc.);
- Image theme label (person, animal, objects, scene, etc.);
- cultural label (the traditional oriental, the west);

3.4 Measure Data Matrix

The data matrix is a contingence table used for management of samples, the corresponding labels and value. All the value is combined by image feature and artificial label. After the necessary standardization, all the data will be Imported into the SPSS software to do the following descriptive statistics and advanced statistical calculation (Table 1).

Table 1. Sample measure data matrix sheet

Sam-ple No	Smaple	con-tent tags	Composi-tion tag	Expres-sion tag	Vi-sion tag	Them e tag	Fea-ture tag
label	Image						
001							
002							
003							
...							
571							

3.5   The Statistical Calculation and Analysis Basing on Multidimensional Indexes

In this paper, qualitative research and quantitative research are the main method to analyze the information and data. The data required from the process of feature analysis are mainly nominal data, for example, the expression methods used in the sample. It offers data for analyzing the variable distribution in all samples.

Owing to the large number of variables and there are more specific variables and value in six categories, this paper adapts correspondence analysis and multidimensional scaling analysis to dig the feature of samples basing on basic descriptive statistics.

4   Analysis

4.1   Result Analysis of Recognizing the Image Feature

The following are the results basing on the similar color space positioning analysis and calculation of samples (Figs. 6, 7 and 8).

Figure 8 shows that it is clearly seen that the position of all samples in two dimensional spaces are in accordance with the rule of two dimensions:

- In the horizontal dimension, the theme tune shows the transition from black and white color (left) to high saturation color (right).
- In the vertical dimension, samples show the transition from cool tone (down) to warm tone (up).
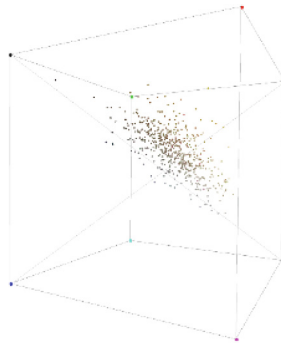
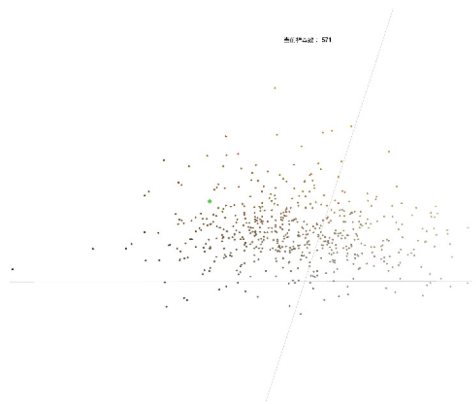**Fig. 6.** Position of the themetune of all samples

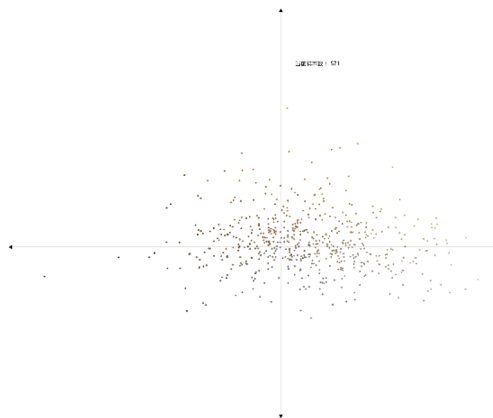**Fig. 7.** Position of the themetune rotated of all samples

Fig. 8. Position of the themetune of all samples whose dimension is reduced

- The distribution of samples in the horizontal direction is wider than in the vertical direction, i.e. saturation is more representative than color temperature for the difference of the theme tune.
- The number of samples gathering in the center of the picture is bigger than that in the surrounding area, i.e. the theme tune of the most samples are similar. It reflects the relative uniformity of the theme tune.

The following is the result of multidimensional scaling analysis of MDS using similarity analysis tool of theme tune (Figs. 9 and 10).
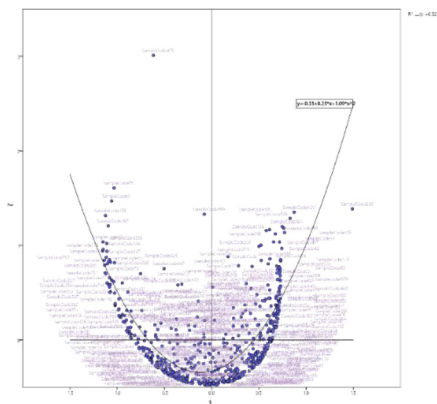

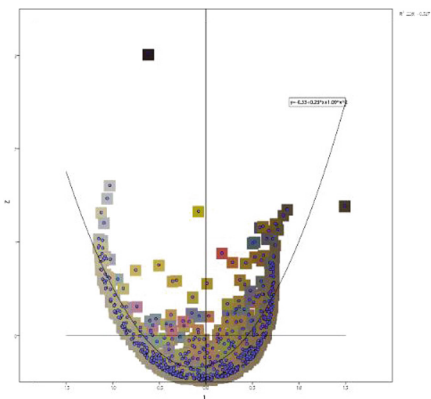**Fig. 9.** Position of tune similarity MDS

Fig. 10. Position of tune similarity MDS (including color tags) (Color figure online)

Figure 9 shows that it can be easily seen that similarity distribution of the theme tune of all samples shows significant rule. We can achieve the binary regression equation by using twice curve-fitting method.

$$Y = 0.33 + 0.25 * X + 1.09 * X2, \ R2 = 0.527$$

In order to make the analysis result easier to find the distribution rule, researchers give more information to Fig. 9, add color code of theme tune to every sample (Fig. 10).

As can be seen in the picture, although the attribute and content of all samples are different, the feature of theme tune shows obvious rule. Color brightness changes gradually from bright to dark from left to right, but in the vertical dimension, no clear rule can be found. In addition, saturation in center section enveloped by U-shaped curve is higher than that around.

4.2   Result Analysis of Separate Identification of Label Attributes

The following are analysis result of position picture from different variable dimension.

Analysis about tone richness and analysis from other variable dimensions

Analysis about composition way and tone richness. Figure 11 shows that there is no clear relationship between composition way and tone richness in all samples and there is no clear rule that can explain the distribution of composition way in all samples. In addition, it is rare that repeatedly arranged patterning method, diagonal composition method and symmetrical composition method are widely used in the samples.
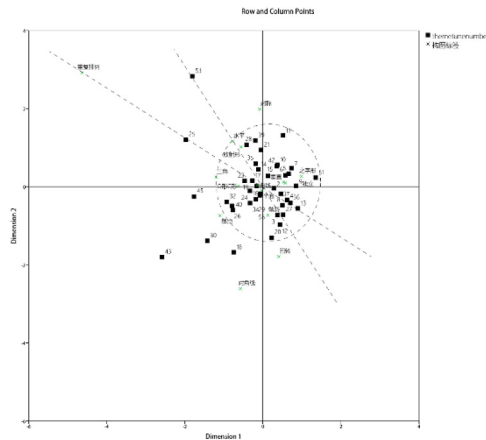

Fig. 11. Analysis about composition way and tone richness

Analysis about expression way and tone richness (Fig. 12). From the result, most of the expression ways can be easily seen in the 571 samples, but several ways, including text announcement method, suspense arrangement act, humorous imitation method, see big things through small ones, etc., are less used than other ways.
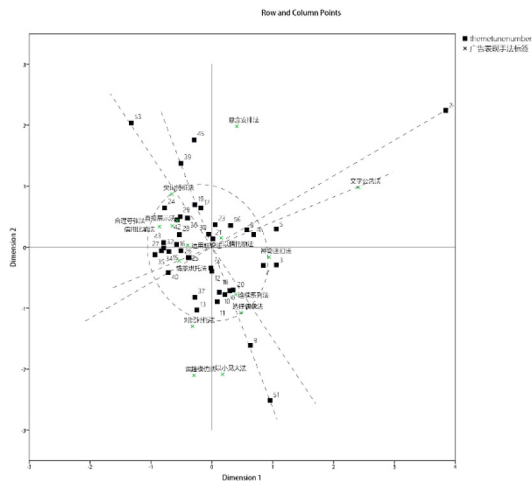
Fig. 12. Analysis about expression way and tone richness

Analysis about visual expression way and tone richness (Fig. 13). From the result, calligraphic style is relatively rare, so are photographic style and painting style, but the latter are used more than the former, the expression of the most samples are similar.

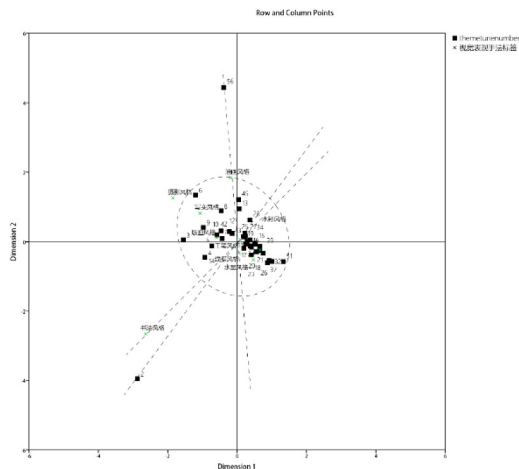**Fig. 13.** Analysis about visual expression way and tone richness

Analysis about culture label and tone richness (Fig. 14). In this picture, there is no clear relationship between culture label and tone richness in all samples. In this dimension, it is unique that historical figures and folk customs are used. A lower proportion of appearance may be the reason.

Fig. 14. Analysis about culture label and tone richness

Analysis from other dimensions

Overall, 571 samples show the following common characteristics:

- Textile products are rich in the expression way and 'scenario contrast' and 'imaginations' are most used. Cigarettes, drugs and cosmetic products prefer 'idols' followed by use 'wonderful feeling to foil the product', 'suspense arrangement', 'humorous imitation'. Festive supplies use 'direct display' more. Daily utensils use 'contrast' more.
- Textile products are also rich in culture label and it use 'love', 'nature', 'story', etc. Stationery commodity prefers 'history'. Apparel fabrics use 'idols' more. Daily utensils usually use 'the west'. Tobacco and festival activities use 'a combination of Chinese and Western elements' widely.
- In consumer orientation and choice for culture label, cultural groups prefer 'history' label. 'A combination of Chinese and Western elements' and 'Stylish in business' labels are more used in the female consumer group. Smokers use 'idols' more. Professionals use 'story' more.

5   Conclusion

The research results can offer an effective guidance to the design of Professional museum for print ads mentioned above during the period of the republic of China, including the overall atmosphere, thematic construction, and situation construction of the museum.

The samples dominant tone presents obvious regularity in visual cognition. On the scatter diagrams, the tone of the horizontal dimension is clear changed gradually from light to dark. There is no striking rule on the vertical dimensions. The overall museum

atmosphere should be consistent with the samples tone, and keeping responding relation in light-shade & cold-warm.

The sample theme should reflect its own cultural label, such as by telling historical stories, reappearing tales of legendia. The characteristics of the time should be fully embodied by the help of the special idols from the Republic of China, fashionable product and arisen combination elements of Chinese and Western.

The construction of the Museum should pay attention to create story atmosphere, such as stories based on story background, story connection, suspense, and passion things, which makes the viewers be personally on the scene of the stories of the ads. It reflects a unique expression to the ads of Republic of China.

Further more, we can adapt this visual-based approach to design elements research in cultural heritage study, that could help us to evaluate visual products, and to evaluate the effect of visual cognition. This approach has a very high important significance in cultural archeology and design industry research as well as other fields.

References

1. Xie, H., Li, Q., Mao, X., et al.: Mining latent user community for tag-based and content-based search in social media. Comput. J. **57**(9), 1415–1430 (2014)
2. He, W.: Improving user experience with case-based reasoning systems using text mining and Web 2.0. Exp. Syst. Appl. **40**(2), 500–507 (2013)
3. Moscato, V., Picariello, A., Rinaldi, A.M.: Towards a user based recommendation strategy for digital ecosystems. Knowl. Based Syst. **37**, 165–175 (2013)
4. Qi, G.J., Aggarwal, C., Tian, Q., et al.: Exploring context and content links in social media: a latent space method. IEEE Trans. Pattern Anal. Mach. Intell. **34**(5), 850–862 (2012)
5. Shim, H., Lee, S.: Multi-channel electromyography pattern classification using deep belief networks for enhanced user experience. J. Cent. S. Univ. **22**(5), 1801–1808 (2015)
6. Martins, D.S., Oliveira, L.S., Pimentel, M.G.C.: Designing the user experience in iTV-based interactive learning objects. In: Proceedings of the 28th ACM International Conference on Design of Communication, pp. 243–250. ACM (2010)
7. Lei, T., Liu, X., Wu, L., Chen, T., Wang, Y., Xiong, L., Wei, S.: The impact of natural utilization of traditional Chinese cultural elements on the user experience in mobile interaction design. In: Rau, P. (ed.) CCD 2015. LNCS, vol. 9181, pp. 46–56. Springer, Heidelberg (2015)