BUILDING A COLOUR IMAGE DATABASE TO RECOMMEND ARCHITECTURAL COLOUR SCHEME USING CASE-BASED RETRIEVAL MECHANISM

WUN-BIN YANG
Graduate School of Computational Design, National Yunlin University of Science and Technology, Taiwan
g9234704@yuntech.edu.tw

AND

JI-HYUN LEE
Graduate School of Culture Technology
Korea Advanced Institute of Science and Technology, Korea
jihyunlee@kaist.ac.kr

Abstract. The purpose of this study is to develop a digitized Taiwanese colour image database for architectural colour scheme in Taiwan. This paper uses adjectives to present each colour's contribution to the colour image. The system uses the "colour difference formula" from the CIEDE2000 method, which calculates the difference between the two colour perceptions represented by the two given points. Using the "Group Nearest Neighbour" algorithm, the retrieval mechanism obtains a similarity measurement. This approach can help designers to know the meanings of colours and their associated colour images, which will help them develop the building image.

Keywords. Colour Scheme; Colour Image; Nearest Neighbour; Case-Based Retrieval.

1. Introduction

Since valuable research has been done by many Japanese investigators (Nagumo, 1999; Kobayashi, 1999; Oota and Kawahara, 1993), research on colour sensibility in Taiwan often uses Japanese colour image data. Chijiwa (Chijiiwa, 2002) explained, however, that people of different cultural backgrounds may have different colour sensibilities. Different Country also has different colour perception and colour image (Lai, 1999). Furthermore, all national standard colour value scales such as RGB, HSV, L*a*b, and L*u*v require much calculation time to convert among them. Therefore, building a digitized Taiwanese colour image database is necessary to enable the Taiwanese people to exchange and retrieve standard colour values more easily.

To analyze the colour identity of a case, a case-based retrieval mechanism is used. The mechanism needs to measure the similarity between the user's query and the existing cases. The colour information for each picture as a case can be calculated and compared with Taiwanese colour image data from the system database.

The purpose of this study is to develop a digitized Taiwanese colour image database for architectural colour scheme in Taiwan to define their colour identities using a case-based retrieval mechanism. This approach depends on the following assumptions: 1) the people of different country have different feelings about colour schemes; 2) in colour space, the distance representing colour similarity is equal to the similarity value of the colour image.

On the basis of these assumptions, the following steps were taken to achieve the goals of this research: first, Taiwanese colour image data were acquired using questionnaires and the data were stored in a colour image database; second, a colour space, which is the geometrical representation of colour perceptions for each colour image, was built and calculated using the CIEDE2000 colour difference formula; third, a distance measure for each colour of the colour space was calculated using the similarity measurement to retrieve the closest colour images for each case; and finally, the colour identity for each case was suggested by the system.

This paper develops a system prototype to prove these concepts, but the theories and principles involved in this prototype can be extended to other design domains.

2. Background

2.1. PERCEPTUAL AND COLOUR IMAGE

People's perceptual experience to objective subject, our mind would influence by colours in order to see different of them. It produces different kind of feeling and connection that arises psychological stimulation. We cannot make a conclusion that same colours must bring same feeling. Yet we can think it tends to such impression. (Tada and Kawahara, 1993). According to Itten (1970), the concept of colour harmony can be removed from the realm of subjective attitude and placed into that of an objective principle, and objectivity helps explain communal consensus (Kraut, 1992). Sensibility ergonomics can be used to understand the relationship between the factors associated with designer feelings and design elements. One of the most efficient ways of describing feelings about colour is through a set of adjectives called colour image (Lee and Qian, 2004).

Research in the colour sensibility field in Taiwan often uses Japanese colour image data, because many Japanese scientists have performed research in this area. In response, Lai (1999) and Lai (1997) have tried to construct a Taiwanese colour perceptual space and to establish a set of Taiwanese colour images using a restrictive-association questionnaire to obtain adjectives for colour images, with the aim of building a colour image database for Taiwan. However, the perceptual space in this work was a two-dimensional space. Therefore, the present study is advancing one step further by digitizing the colour image database in 3D perceptual space.

2.2. UNIFORM COLOUR SPACE

Colour models provide a conceptual framework for thinking about colour sequences by describing the ways in which colours can be defined. Specifically, a colour model specifies the basic components used to describe a colour. Components can be primary colours which are added or subtracted from each other, perceived qualities of the colour, proposed perceptual mechanisms (Rheingans, 1997). All colours which can be created by orthogonal three-space, is convenient to regard a colour as a point in this colour space (Joblove, 1978).

The colour space is a representation of a three-dimensional perceptual space. Wyszecki (1981) defined a geometrical representation of colour perceptions as points in three-dimensional space, in which the distance between any two points can be taken as a measure of the magnitude of the difference between the two colour perceptions as represented by the two given points.

In order to capture perceptual uniformity in a colour representation space, it is crucial to rely on the distance criterion which states that the distance D(c1, c2) between two colours c1 et c2 is correct if and only if the distance value is close to the difference perceived by the human eye (Paschos, 2001). Therefore, perceptual uniformity in a given colour space means that the perceptual similarity of two colours is measured by the distance between the two colour points (Sarifuddin and Missaoui, 2005).

Use the proper colour definitions which are more suited for human perceptions. L*C*H* space is selected because its definitions and measurements are suited for vision perception psychology (Wang and Yu 2005).

2.3. COLOUR DIFFERENCE FORMULA

Digitized colour allows easy translation from one colour mode to another. The colour space is a representation of a three-dimensional perceptual space. However, the uniformity values for CIE 1976 L*u*v*and CIE 1976 L*a*b* are still not perfect. For the reason, CIE develops the new colour difference formula CIEDE2000 based on CIE L*a*b*.

This colour difference formula often calls CIEDE2000 (Equation 1) or %E00 (Luo, Cui and Rigg, 2001). The CIEDE2000 is four steps to finish whole calculation as follow:

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T\left(\frac{\Delta C'}{k_C S_C}\right)\left(\frac{\Delta H'}{k_H S_H}\right)}, \quad (1)$$

2.4. MEASURING SIMILARITY FOR CASE-BASED RETRIEVAL

In case-based reasoning, the retrieval part finds previous cases that are similar to the current one. Retrieval identifies which cases have all or a subset of those features, and selection evaluates the retrieved cases (Maher, 1997). One of the retrieval issues is finding nearest neighbours (NN) (Roussopoulos et al. 1995) that can match effectively (Pearce, Goel et al. 1992). The challenge is to establish a retrieval mechanism to access relevant colour image cases from a digitized colour image database.

As an extension of the NN query, the group nearest neighbor (GNN) query (Papadias et al. 2004) has more than one query point, and its objective is to minimize the sum of distances from each resultant point to all query points. The difficulty is that query points may be distributed within the data space in arbitrary ways, creating a large search region (Li et al. 2005).

From the point of view of CPU requirements, the preferred methods are SPM and MBM based on GNN. The single-point method (SPM) (Papadias et al. 2004) processes a GNN query in a single traversal of the R-tree. SPM first determines the centroid q of Q, which is a point in the data space with the minimum value of dist(q,Q). Then a depth-first kNN search is performed with q as the query point (Figure 1).

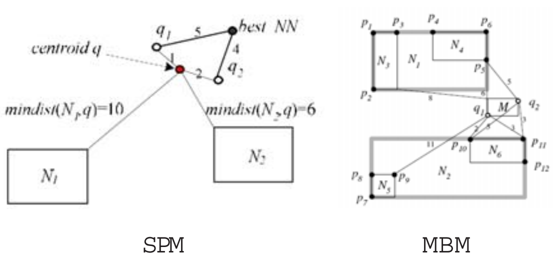


Figure 1. Pruning of nodes in SPM and Query processing of MBM (Papadias et al. 2004).

SPM and MBM perform a single query with some pruning heuristics. SPM is a modified single depth-first NN search, with the centroid of Q being the query point, while MBM considers the MBR (Minimum Bounding Rectangle) of Q and can proceed in either a depth-first or a best-first manner (Papadias et al. 2004).

3. Methodology

3.1. ENVIRONMENT SETTING FOR EXPERIMENT

This experiment was conducted for the design students and designers trained by colour theory. Each of the 150 subjects was given a questionnaire to fill out. Through one-to-one way and the same equipment environment, then get the result about colour image. The equipment is monitor (Viewsonic); colour gamut is sRGB colour space profile and interior environment.

3.2. ESTABLISHMENT OF COLOUR IMAGE DATA

This paper uses Lai's (1999) 44 adjectives from Taiwanese colour image data. Besides, the perceptual uniform colour space (LCH), corresponding closely to the perceptual attributes of the human visual system. Therefore, this paper use colour samplings base on the equidistant value of L*C*H* and five simple colours, are shown in Table 1.

TABLE 1. Equidistant value of L*C*H*.

L	30	50	70
C	20	50	80
H	0,30,60, 90,120,150,180,210,240,270,300,330		

Therefore, it can get the 116 colour of the L*C*H* uniform colour space to be presented in the colour image questionnaire. By processing the restrictive-association questionnaire, this paper examined each colour included in each adjective.

The results for the adjective "classic", which received the most votes, are shown in Figure 2.

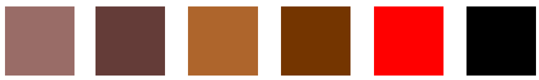


Figure 2. Representative "classic" colour images.

After this survey, it turned out that six colours exhibit agreement greater than 70% as acceptable options.

3.3. CONVERT EACH COLOUR OF EACH COLOUR IMAGE TO CIEDE2000 FORMAT

Digitized colour can be easily translated from one colour space to another. The perceptual similarity of two colours can be measured by the distance between the two colour points (Sarifuddin and Missaoui, 2005).

For this reason, this system calculates colour differences based on the new CIEDE2000 formula to get the L*, C*, H*.

3.4. MEASURE SIMILARITY FOR CASE BASE RETRIEVAL

After calculating the values in colour space for the image, the system needs to search for a matching colour image. This study uses SPM in GNN as the search approach. The first step is to calculate the centroid q, using the SPM method, from the coordinate locations in colour space. The centroid will be used as a search benchmark for the sample to be retrieved.

3.4.1. Improving the SPM in GNN Calculation of the Centroid of the Colour Image

The SPM in GNN searching approach uses a two-dimensional interpretation concept. However, the colour space is three-dimensional. Therefore, this method must transform the two-dimensional value into a three-dimensional one. Furthermore, the coordinates of the colour image occupy an irregular shape in colour space. This means that the mass center of the general form cannot be used as a basis for the calculations. Instead, the irregular shape is translated into a bounding three-dimensional rectangular solid according to the relative values of the coordinates, using the following procedure: first, obtain the bounding rectangular solid from the maximum colour value of L*, C*, H* obtained earlier; second, fit Equation 2 to the maximum value, mapping C* to x, H* to y, L* to z, which yields the centroid q as follow:
(http://mathworld.wolfram.com/GeometricCentroid.html)

$$\bar{x} = \frac{\iiint x\,\rho(x,y,z)\,dV}{M} \quad \bar{y} = \frac{\iiint y\,\rho(x,y,z)\,dV}{M} \quad \bar{z} = \frac{\iiint z\,\rho(x,y,z)\,dV}{M} \quad (2)$$

Each colour image is represented by a group N, based on adjectives for colour images in the questionnaire, so there are groups in the system. Each of the groups N1 to N.. indicates a different colour image. Using the CIEDE 2000 formula, a distance is calculated for each colour image (N1 to N..). Therefore, with reference to the bounding rectangular solid and the centroid q, the system can eliminate unrelated colour images and select only the nearest colour image based on the threshold value.

The colour images for colour image and adjacent regions can be represented as shown in Figure 3.

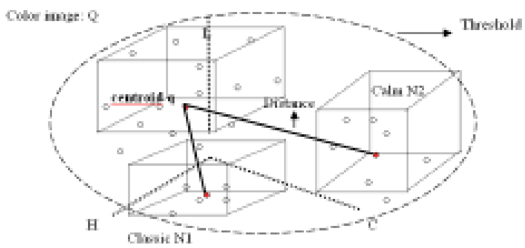
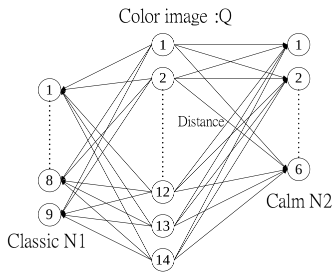


Figure 3. Colour space for each of the three groups.

3.4.2. Analysis Relation For Nearest Colour Image

After the system obtains the nearest colour images N1 and N2, distance values are calculated from each point of the sample's colour image to each point of another nearest colour image. Then the nearest point to each point of the sample's colour image Q is identified. Finally, by determining the nearest colour image with the largest number of identified points, the system can identify the nearest colour image.

Finally, this research will take the example form buildings, product, poster are the cases to verification. Besides, evaluate the result compare between system and user by questionnaire.



Q	Nearest Point	Nearest Color Image
1	N1-01	N1-Classic
2	N1-01	
3	N1-02	
4	N1-02	
5	N1-03	
6	N1-04	
7	N2-03	
8	N2-03	
9	N2-03	
10	N2-02	
11	N1-03	
12	N1-05	

Figure 4. Analytical relationship for nearest colour image.

4. Implementation

The system obtains a solution in the form of a retrieved colour image or an adaptation of a retrieved image, and then retains the new colour image to the case base.

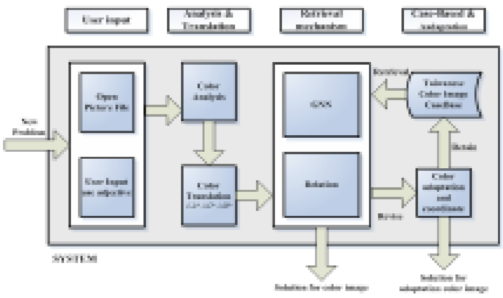


Figure 5. System architecture and module.

There are two ways to obtain colour image information. One is use a colour adjective to search, and then retrieve colour image information for each colour. Supposing the user wants only to select a colour image to meet his needs, in this case the user can directly select the desired colour image using colour matching. In the second approach, the user can immediately analyze the number of colours and the composition ratios in a picture by means of architectural pictures already in the system. When the user wants to refer to additional similar colour images, he can click the “Show similar colour image” button to retrieve additional colour images for colour matching.

The system uses the GNN method and threshold values to match the colour information entered by the user with the colour image database and to analyze the colour information in the colour image. By switching to another page in the software, the user can select the required colour images. In addition, a three-dimensional image is displayed in the right-hand portion of the screen to show the relationship between each colour image and other similar colour images. In this way, the user can conveniently observe the difference between individual colour groups. It is also possible to adjust the value of the threshold to observe more database images. Moreover, the software will display distances in colour space according to the results of the CIEDE2000 calculations, which can be used as reference data as well. Finally, designers needing colour images of building can use this system to access Taiwanese colour image data.

5. Conclusion

In the colour scheme stage, designers traditionally tend to carry out their colour scheme based on general stereotypes and design experience. Therefore, developing a Colour image system which is capable of suggestion.

This paper has contribution as follow: 1.) The approach used here leads to a system reflecting Taiwanese preferences for colour images. The capabilities of the system have been evaluated experimentally with users, and the results have shown that this system can suggest appropriate similar colour images that match user preferences; 2.) Presented a digitized Taiwanese colour image database. 3.) System which uses the GNN retrieval mechanism to address the computer science methodology can be applied into design domain.

Therefore, the system can be used to increase the efficiency of defining colour schemes and preparing designs. Not only can it help to enhance the performance of colour image retrieval, but it can also provide a larger library of colour image cases for designers to use as a reference.

References

Chijiiwa, H.: 2002, Inconceivable psychology and colour, Xin-chau, Taiwan.

Itten, J.: 1970, The Elements of Colour, Van Nostrand Reinhold, New York.

Joblove, G.H.: 1978, Greenberg D. Colour spaces for computer graphics. ACM SIGGRAPH Computer Graphics 12(3) 20-25.

Kobayashi, S.: 1999, Colourist: A Practical Handbook for Personal and Professional Use, Oxford University Press, USA.

Kraut, R.: 1992, The objectivity of colour and the colour of objectivity. Philosophical Studies ; 68: 265-287.

Lai, C.C.: 1997, Colour psychology of design, Culture of visual communication, Taipei.

Lai, J.B.: 1999, Construction of Taiwanese Colour Perceptual Space. MS Thesis, Graduate school of design, NTUST, Taiwan. (in Chinese)

Lee, J.H. and Qian, W.: 2004, Colour Your Feeling. Recent Advances in Design & Decision Support Systems in Architecture and Urban Planning, Dordrecht: Kluwer Academic Publishers, 113-125.

Li, H., Lu, H., Huang, B., Huang, Z.: 2005, Two ellipse-based pruning methods for group nearest neighbor queries, Proceedings of the 13th annual ACM international workshop on Geographic information systems, Bremen, Germany.

Luo, M.R., Cui, G. and Rigg, B.: 2001, The development of the CIE 2000 colour difference formula. Colour Res. 26: 340-350.

Maher, M.L.: 1997, Gomez de Silva Garza A. Case-Based Reasoning in Design. IEEE Exper 12(2): 34–42.

Nagumo, H.: 1999, Colour image chart, Graphic-sha, Japan.

Oota, A. and Kawahara, E.:1993, Colour and matching colours, New-image, Taipei.

Papadias, D., Shen, Q., Tao, Y., Mouratidis, K.: 2004, Group nearest neighbor queries, Proceedings of the 20th International Conference on Data Engineering, 301–312.

Paschos, G.: 2001, Perceptually uniform colour spaces for colour texture analysis: An exeprimental evaluation. IEEE Trans, on Image Processing 10(6): 932–937.

Pearce, M., Goel, A.K., Kolodner, J.L., Zimring, C., Sentosa, L. and Billington, R.: 1992, Case-Based Design Support: A Case Study in Architectural Design. IEEE Expert 7(5): 14-20.

Rheingans, P.: 1997, Colour Perception and Applications, SIGGRAPH 97 Course #33.

Roussopoulos, N., Kelley, S., Vincent, F.: 1995, Nearest neighbor queries, Proceedings of the ACM SIGMOD international conference on Management of data, San Jose, California, United States, 71-79.

Sarifuddin, M. and Missaoui, R.: 2005, A New Perceptually Uniform Colour Space and Associated Similarity Measure for Content-Based Image and Video Retrieval. Multimedia Information Retrieval Workshop, Salvador, Brazil.

Tada, A. and Kawahara, E.: 1993, Colour and Matching, New Image, Taipei. (in Chinese)

Wang W.N., Yu Y.L.: 2005, Image emotional semantic query based on colour semantic description, Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21.

Wyszecki, G.: 1981, Uniform colour spaces, Golden Jubilee of Colour in the CIE, Bradford, The Society of Dyers and Colourists, 52–77.