



Visual Performance of Psychological Factors in Interior Design Under the Background of Artificial Intelligence

Yunkai Xu¹ and TianTian Yu^{2,3*}

¹School of Textile Apparel and Design, Changshu Institute of Technology, Suzhou, China, ²Wenzhou Polytechnic University, Wenzhou, China, ³School of Housing, University of Science, Penang, Malaysia

OPEN ACCESS

Edited by:

Deepak Kumar Jain,
Chongqing University of Posts and
Telecommunications, China

Reviewed by:

Zhiyong Jiang,
Guilin University of Aerospace
Technology, China
Jiansong Fang,
Guangdong Baiyun University, China

*Correspondence:

TianTian Yu
2021000528@wzpt.edu.cn

Specialty section:

This article was submitted to
Human-Media Interaction,
a section of the journal
Frontiers in Psychology

Received: 11 May 2022

Accepted: 20 June 2022

Published: 28 July 2022

Citation:

Xu Y and Yu T (2022) Visual
Performance of Psychological
Factors in Interior Design Under the
Background of Artificial Intelligence.
Front. Psychol. 13:941196.
doi: 10.3389/fpsyg.2022.941196

Sensation (the reflection of past experience in the mind) is the reflection of the brain on the individual attributes of objective things that directly act on the sense organs. Feeling is the most elementary cognitive process and the simplest psychological phenomenon. Vision is a kind of sense, and sense is produced by objective things acting on the sense organs. But at present, it is rare to analyze interior design exhibition from the perspective of visual psychology, an emerging science, as an interdisciplinary attempt, only in interior design research. Therefore, the study of sensory process should start from its external stimuli, in order to first understand how it acts on the sensory organs to produce sensory phenomena. This paper mainly studies the visual performance of psychological factors in interior design under the background of artificial intelligence. This paper proposes a K-means clustering algorithm and a localization algorithm fused with visual and inertial navigation. The distance thresholds corresponding to the SIFT feature descriptors of threshold T1, 128D, 96D, 64D, and 32D are 170, 160, 150, and 90, respectively. This verifies that the candidate image with the highest number of matching points is considered the best matching image.

Keywords: artificial intelligence, interior design, optical illusion, psychological factors, visual performance

INTRODUCTION

How to make the design form and design method of interior design closer to the audience's thinking from the perspective of human visual psychology, so that the content delivered by interior design can take into account the accuracy and artistry. According to different needs of people, sometimes the room needs to be warm and comfortable; sometimes it needs to be solemn and dignified; and sometimes it needs to be relaxed. When designers do interior design, they always try to make the room transparent, spacious, and bright. But the problem we face is that the plane of the room is fixed, the height of the room is also fixed, and the position and size of the windows in the room are also fixed. The hallucination image is not vivid and vivid, and it occurs in the patient's subjective space, such as the brain and the body. Hallucinations are not obtained through the sense organs. For example, if you hear a voice in your stomach, you can see a human figure in your mind without your own eyes. This all needs to be changed through the technique of optical illusion decoration.

In China, many rooms only face north. In today's crowded social housing conditions, there are many small units. The population density is high, and people have more and more items. Designers need to decorate the room to make it less crowded. Interior design needs to use optical illusions to cover up the shortcomings of the room, or learn from each other's strengths. Through comprehensive analysis, it can meet the different needs of today's people for indoors. According to the current national conditions, such as apartment type, space size, etc., are fixed, it is necessary to consider how to make the narrow space appear spacious, bright, low, and depressed room. As for how to increase the permeability of a relatively closed room, etc., it can be improved by means of optical illusions. Although the illusion has been discovered for a long time, its application in interior design is relatively new at home and abroad.

The innovation of this paper is that the perfect explanation makes use of optical illusion effects. It responds perfectly to the visual analysis of interior design, and utilizes advanced technologies and instruments for indoor positioning to perform experimental measurements to verify conjectures. In this paper, the K-means clustering algorithm is further improved and utilized. Interior design is closely related to people's visual senses and psychological characteristics. The human eye has a focus and sequence when viewing things. This visual psychological habit of people will guide our visual process and psychological thinking, so we are carrying out In interior design, text, graphics, and colors should be arranged according to people's visual and psychological habits, and at the same time to achieve the esthetic sense of formal art and people's emotional acceptance, so as to better transmit the design content.

RELATED WORK

Interior design is to create an indoor environment with reasonable functions, comfortable and beautiful, and meet people's material and spiritual needs according to the nature of use of the building, the environment, and the corresponding standards, using material technical means and architectural design principles. Logos in different regions may have different characteristics that affect the likes of the logo. Zhu et al. (2017) mainly discussed how logo design features affect consumer responses based on visual representation. To understand the extent to which indoor environmental factors influence the creation of an optimal rehabilitation environment, Mahmood and Tayib (2019) proposed that contemporary rehabilitation environments may be affected by the visual effects of indoor environmental quality. Celadyn (2018) discussed the issues that the modification of contemporary interior design strategies and methods should be introduced into the interior design model oriented by environmental responsibility. On the basis of previous research results, Zhang and Zheng (2017) explored the ideas and methods of realizing realistic indoor scenes based on the interactive real-time rendering of the open-source project MNRT. He used NVIDIA's powerful application engine, ray tracing engine Opti X, scene management engine SceniX, and PhysX engine to design the commercial design of interior design software. Aiming at the significant and abundant

color difference between regions in interior design, Fang et al. (2017) proposed a color transfer algorithm for interior design based on region matching based on topology information. He first introduced the segmentation of interior design images, and then proposed to calculate the topological information of each region to determine the matching relationship between regions. This, in turn, improves the accuracy of color transfer between images for enhanced visual effects. Celadyn (2017) analyzed the options for adapting an interior design education model to the principles of environmental sustainability in the creation of interior environments. This is because the current curriculum of most interior design schools is established according to the traditional design scheme. Ju and Yang (2021) research developed and applied a service-learning curriculum. The program combines university courses with housing and interior design in the local community. These studies have better explained the importance of the central psychological factors of interior design to visual performance, but they are still not comprehensive enough and not innovative enough.

THE VISUAL EXPRESSION ALGORITHM OF PSYCHOLOGICAL FACTORS IN INTERIOR DESIGN

K-means Clustering Algorithm

Artificial intelligence is the study of making computers to simulate certain thinking processes and intelligent behaviors (such as learning, reasoning, thinking, planning, etc.) and enables computers to achieve higher-level applications. The K-means algorithm is an unsupervised method that was formally mentioned in the early 1960s. The basic idea of the algorithm is to classify the m data types in the data set according to the preset threshold rules, and then divide them into p types. This satisfies the condition that the similarity between the same type of data after classification is high, while the similarity between different types of data is relatively low (Chang and Yang, 2020). It assumes that the p cluster centers of the dataset $w=(a_1, a_2, a_3, \dots, a_m)$ that need to be clustered are (b_1, b_2, \dots, b_m) . By default, it uses the Euclidean distance to measure the distance between two samples.

$$d(a_i, a_j) = (A_{i1} - A_{j1})^2 + (A_{i2} - A_{j2})^2 + \dots + (A_{im} - A_{jm})^2 \quad (1)$$

Among them, a_i, a_j represents two m -dimensional data objects (Caulfield, 2018).

The average algorithm for all samples is as follows.

$$Meandist(W) = \frac{2}{m(m-1)} \times \sum_{i \neq j}^m d(a_i, a_j) \quad (2)$$

The most commonly used objective function is the squared error criterion function, which is defined as.

$$Q = \sum_{l=1}^P \sum_{j \in m_l} \|a_j - c_l\|^2 \tag{3}$$

In the above formula, m_i represents the i th cluster set and c_i represents the center of the i th cluster. Euclidean distance is also called Euclidean distance. In n -dimensional space, the length of the shortest line is its Euclidean distance. It is a commonly used definition of distance, which is the true distance between two points in m -dimensional space. Q represents the sum of squares of the distances between all the data in the dataset and their respective cluster centers. The distance here refers to the Euclidean distance.

The specific steps of K-means clustering are as follows: (1) It first pre-determines the number of clusters k according to actual needs, and then selects k data from the given sample data. This selection is random and it is used as the initial clustering point. (2) It calculates the distances between all remaining data objects and the selected K cluster center points, and classifies them into the corresponding categories with the smallest similarity. (3) For each generated class, it calculates the average sum of all data in that class and takes it as the new cluster center point (Wallace, 2019). (4) It judges whether the iteration can be completed according to the requirements of the iteration completion. It assumes that the condition can be completed iteratively. The conditions for the end of the iteration of the K-means algorithm:

1. The cluster no longer changes in any way.
2. The difference between the objective function values of the two iterations before and after is lower than the pre-computed threshold.
3. It iterates more than the previously specified number of times. And as long as any one of the above three conditions is satisfied, the iteration can end. If either condition (1) or (2) is satisfied, the clustering is convergent. The K-means algorithm is really fast for big data processing, but its shortcomings are also obvious. The algorithm is very sensitive to the selected cluster center points; however, the results are indeed greatly affected by this. At first, it selects different initial random seed points, which may eventually lead to different clustering results, or even farther apart.

SVM Classifier

In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models related to related learning algorithms that can analyze data, identify patterns, and use for classification and regression analysis. Vector machines were formally proposed in the last century as a machine learning method for classification (Shubha and Meera, 2019). The core of the algorithm is to exploit the principle of risk minimization. It aims to find a hyperplane that separates all the different data and thus achieves the highest classification accuracy (Grobelny et al., 2021). It assumes that the set of samples to be trained is D , as shown in Formula (4). It contains n features, and the classification hyperplane is recorded as Formula (5).

$$R = \left\{ (x_i, y_i) \mid x_i \in P^d, y_i \in \{-1, 1\} \right\}_{i=1}^m \tag{4}$$

$$wx + t = 0 \tag{5}$$

For data that need to be classified, it is usually divided into two different cases according to whether it is linear or not.

For the first ideal case, the classification interval is usually set to $2/\|w\|$, and there are constraints shown in Formula (6) below (Jain et al., 2022).

$$y_i (wx_i + t) \geq 1, 1 \leq i \leq m \tag{6}$$

According to the constraints of Formula (6) above, the problem of constructing the optimal hyperplane is transformed into the solution problem shown in the following Formula (7):

$$\min \alpha(w) = \|w\|^2 / 2 \tag{7}$$

With the introduction of the Lagrangian operator, Formula (7) can be transformed into its dual problem, which can be obtained as:

$$\begin{cases} \max(\beta) = \sum_{i=1}^m \beta_i - \frac{1}{2} \sum_{i=1}^m \beta_i y_i y_j x_i x_j \\ s.t. \beta_i \geq 0, \sum_{i=1}^m \beta_i y_i = 0, 1 \leq i \leq m, 1 \leq j \leq m \end{cases} \tag{8}$$

Combining the above formulas, the optimal solutions of the hyperplanes such as Formulas (9) and (10) can be obtained.

$$w = \sum_{i=1}^m \beta_i y_i x_i \tag{9}$$

$$t = y - wx_i \tag{10}$$

Finally, the optimal classification function can be obtained by solving the Formula (11).

$$f(x) = \sum_{i=1, j=1}^m \beta_i \beta_j x_i x_j + t \tag{11}$$

In view of the fact that the vast majority of sample data is linearly inseparable, if the linear SVM classification method is also used at this time, it is likely to lead to serious wrong classification results (Horton, 2017). In view of this situation, researchers at home and abroad have improved the SVM30, that is, it introduced the kernel function. In the face of linear

TABLE 1 | Confusion matrix of scene classification results for bag-of-words model.

Bedroom	0.90	0.00	0.00	0.10	0.00
Boardroom	0.00	0.80	0.15	0.05	0.00
Corridor	0.05	0.15	0.75	0.00	0.00
Kitchen	0.00	0.10	0.15	0.00	0.00
Living room	0.20	0.10	0.00	0.80	0.00
Stairway	0.05	0.05	0.10	0.00	0.90

TABLE 2 | Confusion matrix of scene classification results for pyramid bag-of-words model.

Bedroom	1.00	0.00	0.10	0.10	0.05
Boardroom	0.00	0.70	0.15	0.05	0.00
Corridor	0.05	0.15	0.80	0.00	0.10
Kitchen	0.15	0.10	0.15	0.90	0.00
Living room	0.20	0.15	0.05	0.80	0.50
Stairway	0.05	0.05	0.10	0.00	0.90

inseparability, the way vector machines work is to first choose to complete the calculation in a low-dimensional space. It then uses the function of the kernel function to map the input feature space to a relatively high-dimensional spatial feature space, and finally realizes the high-dimensional feature space. The construction of the optimal separating hyperplane is done in the feature space. This thus completes the separation of non-linear data that is not easily separable on the plane itself.

$$\max(\beta) = \sum_{i=1}^m \beta_i - \frac{1}{2} \sum_{i=1, j=1}^m \beta_i \beta_j y_i y_j K(x_i x_j) \quad (12)$$

k is the introduced kernel function.

The so-called radial basis function is a scalar function that is symmetrical along the radial direction. It is usually defined as a monotonic function of the Euclidean distance between any point x in space and a certain center x_c . Gaussian radial basis kernel function:

$$k(x, y) = \exp\left(-\frac{|x - y|^2}{2\varepsilon}\right) \quad (13)$$

It assumes that the choice of parameter ε determines the width of the radial basis function. Therefore, its selection has a great influence on the final classification result. This kernel function can be said to be the most widely used among the multiple kernel functions. It will have a better effect regardless of the size of the sample. And it also has the advantage that it has fewer parameters than the polynomial kernel function. Therefore, it can consider the preferred Gaussian kernel function without knowing the specific case of the sample. According to pattern recognition theory, linearly inseparable patterns in low-dimensional space may be linearly separable by non-linear mapping to high-dimensional feature space, but if this technique is directly used for classification or regression in high-dimensional space, there are certain forms and parameters of the linear mapping function, the dimension of the feature space, etc., and

TABLE 3 | Classification results for all categories of images.

	The first time (%)	The second time (%)	The third time (%)	Average value (%)
Classic bag-of-words model	77.5	75.8	78.3	77.2
Pyramid bag-of-words model	82.5	80.8	83.3	82.2
Pyramid that only improves the K-means algorithm bag-of-words model	85.8	84.2	85.8	85.2
The improved pyramid bag-of-words model in this paper	90.8	91.7	90.8	91.1

the biggest obstacle is the “curse of dimensionality” that exists in the operation of the high-dimensional feature space. The use of kernel function technology can effectively solve such problems.

For indoor scene classification algorithms, we generally use the two indicators of accuracy and running time to judge (Ponde et al., 2017). The confusion matrix of scene classification results based on bag-of-words model is shown in **Table 1**. The confusion matrix of scene classification results based on the spatial pyramid bag-of-words model is shown in **Table 2**.

The classification of the two scenes, Kitchen and Boardroom, is easy to be confused, and classification and recognition errors are prone to occur. Its error rate can even reach 30% (Resch, 2019). The accuracy rate does not meet the high-standard scene classification requirements. The classification accuracy results of all scene classification algorithms are shown in **Table 3**. The running times of all scene classification algorithms are shown in **Table 4**. In order to ensure the reliability of the algorithm, we simulate and run each algorithm for three times in the simulation experiment, and finally take the average value.

From the data in **Table 3**, it can be seen that the accuracy of the scene classification algorithm based on the spatial pyramid bag-of-words model is 5% higher than that of the bag-of-words classification algorithm without the SPM model. Its scene classification rate can reach more than 80% (Leung and Zhang, 2019). Combined with the K-means clustering algorithm rule of the roulette method, it is also improved by 3% compared with the original bag-of-words model classification algorithm based on the spatial pyramid. In addition, it can be seen that there is room for improvement in this paper. The bag-of-words model of the pyramid is also effectively optimized by the classification algorithm to improve the defects of the K-means clustering algorithm. In general, the scene classification accuracy of the improved algorithm in this paper is indeed improved compared with the original scene classification algorithm based on the spatial pyramid model. This can verify that the improved algorithm has the highest scene classification accuracy.

It can be seen from the data in **Table 4** that the running time of the spatial pyramid bag-of-words model based on the K-means clustering algorithm combined with the roulette method differs by nearly 3 min from the running time of the algorithm. Overall, the scene classification algorithm improved by the bag-of-words model based on the spatial pyramid has the longest running time. It is almost twice as fast as the bag-of-words-based classification algorithm, but the overall time is different.

TABLE 4 | Temporal statistics for classification of all classes of images.

	The first time (min)	The second time (min)	The third time (min)	Average value (min)
Classic bag-of-words model	10.5	10.1	11.5	11.7
Pyramid bag-of-words model	14.1	13.2	15.1	14.1
Pyramid that only improves the K-means algorithm bag-of-words model	17.6	16.8	18.2	17.5
The improved pyramid bag-of-words model in this paper	19.4	18.2	20.2	19.3

Localization Algorithm Fusion of Vision and Inertial Navigation

The visual positioning and navigation technology is carried out on the basis of a series of work such as feature matching of images of a series of frames before and after. This results in that the accuracy of visual localization navigation depends to a large extent on whether the amount of features of the image is sufficiently accurate. When the features are scarce or there is a mismatch, the accuracy of the localization is prone to large errors (Jeamwathanachai, 2019). In order to improve this situation, in this chapter, a positioning algorithm combining visual and inertial navigation is proposed by taking advantage of the advantage of the inertial measurement unit with high positioning accuracy in a very short time. The basic principle of visual automatic positioning technology is to transmit the collected physical image to the PLC image-processing system through the CCD brought by the machine equipment, calculate the offset position and angle through the image processing positioning software, and then feed it back to the external platform for motion control. The position correction function is completed by the precision servo drive. The accuracy can be manually set, and if the accuracy exceeds the range, the system will report an alarm if it cannot be completed.

The inertial coordinate system ($G_i x_i y_i z_i$ is denoted as the i system) refers to a coordinate system that has no acceleration and rotation relative to the rest of the universe, and there is no absolute inertial coordinate system. When the motion carrier is limited to the motion near the earth, if the orbital motion of the earth is ignored, the geocentric inertial coordinate system can be used as an ideal approximation of the inertial coordinate system. Any point w near the earth can be represented by three-dimensional coordinates $P(x, y, z)$. Of course, any point on the earth's surface can also be represented by the latitude and longitude coordinate system. It assumes that the latitude and longitude coordinates of any point P are expressed as $P(\phi, L)$, and the conversion relationship between the latitude and longitude coordinate system to the earth coordinate system can be expressed by a formula.

$$\begin{cases} x = R_M \cos L \cos \phi \\ y = R_M \cos L \sin \phi \\ z = R_M (1 - \nu)^2 \sin L \end{cases} \quad (14)$$

The conversion relationship between the earth and the latitude and longitude coordinate system is shown in Formula (15).

$$\begin{cases} \phi = \arctan \frac{y}{x} \\ L = \arctan \left[\frac{1}{(1 - \nu)} \times \frac{z}{x^2 + y^2} \right] \end{cases} \quad (15)$$

R_M and ν are the radius of curvature and flattening of the Earth, respectively.

Multiplication of quaternions is not commutative. Quaternions are, unambiguously, non-commutative extensions of complex numbers. If the set of quaternions is considered as a multi-dimensional real number space, the quaternion represents a four-dimensional space, which is a two-dimensional space relative to complex numbers. In the calculation of the change of position of inertial navigation and positioning by considering the idea of representing the rotation matrix through the quaternion, the core part is the quaternion differential formula. It now combines some inherent properties of quaternions to analyze its differential formula as follows:

It is assumed in advance that the coordinate system $GXYZ$ at time d has a rotation relative to the original coordinate system $G_i x_i y_i z_i$. It assumes that this rotation is G_1 , then formula Q_1^{-1} can be obtained.

$$R_i(d) = Q_1 R(d) Q_1^{-1} \quad (16)$$

In Formula (16), Q_1 is a quaternion representing rotation, and we get the form of rotation operator through Q_1 , that is, $Q_1 Q_1^{-1}$. For the short-time variation range of Δd , due to the angular velocity w of the coordinate system $Oxyz$, the relative positions of the above two coordinate systems also change immediately. At this point, we see that the coordinate system $G_i x_i y_i z_i$ of $Oxyz$ has changed to Q_2 for the previously fixed coordinate system. At this time, the formula can be obtained.

$$R_i(d + \Delta d) = Q_2 R(d + \Delta d) Q_2^{-1} \quad (17)$$

For a brief time range of Δd , the motion changes and the rotational form is now represented by $Q_1^{-1} Q_2$. Since we assume that the value of Δd is very small, the change of the angular rate S existing for the coordinate system is regarded as a relatively fixed constant value. It can then obtain the angular displacement of the dynamic system shown in the formula. At that time, it was mainly text-based image retrieval technology. Use text description to describe the characteristics of the image, such as the author, age, genre, size, etc., of the painting. After the 1990s, the image retrieval technology that analyzes and retrieves the content semantics of the image, such as the color, texture, and layout of the image, has emerged, that is, the content-based image retrieval technology. The improved bag-of-words model algorithm is its core algorithm.

$$\Delta \phi = |S| \Delta d \quad (18)$$

In Formula (18), the value of s is reflected by the modulo $|s|$ of s , but the direction of $\Delta 1$ still depends on the direction of s . It assumes a unit vector $\eta = \frac{s}{|s|}$ to obtain Formula (19):

$$Q_1^{-1}Q_2 = \cos \frac{|s|\Delta d}{2} + \eta \sin \frac{|s|\Delta d}{2} \quad (19)$$

From Formula (19), Q_2 can be obtained after simple calculation, and Q_2 can be expressed by Formula (20).

$$Q_2 = Q_1 \left(\cos \frac{|s|\Delta d}{2} + \eta \sin \frac{|s|\Delta d}{2} \right) \quad (20)$$

It finally performs the derivation work for the quaternion Q , and the derivative of Q can be easily obtained as shown in Formula (21).

$$Q'(t) = \lim_{\Delta d \rightarrow 0} \frac{\Delta d^{2\alpha} - \Delta d^{2\beta}}{\Delta d} = \lim_{\Delta d \rightarrow 0} \Delta d^{\alpha} \left(\cos \frac{|s|\Delta d}{2} + \eta \sin \frac{|s|\Delta d}{2} - 1 \right) \quad (21)$$

It firstly studies the theoretical part of inertial navigation. Aiming at the problem of the cumulative error of visual positioning, it proposes a visual positioning algorithm fused with inertial navigation by taking advantage of the high accuracy of inertial navigation in a short time. The core of the algorithm is to use the rotation matrix R and translation vector T of the IMU solution result as the initial value of the camera motion. It sets a threshold and removes the matching points that obviously do not satisfy the motion model between adjacent frames of the camera as outliers. Through the simulation experiments of two indoor scenes, it is compared and analyzed that the visual positioning algorithm of the integrated inertial navigation proposed in this paper has higher positioning accuracy and better effect than the visual positioning algorithm.

EXPERIMENTAL ANALYSIS ON THE VISUAL PERFORMANCE OF DESIGN BY PSYCHOLOGICAL FACTORS IN INTERIOR DESIGN

Experimental Method

In this paper, an efficient vision-based indoor localization system is proposed. It uses the bag-of-words model algorithm in the image retrieval technology to extract and retrieve the image information requested for positioning, and obtain the most similar images and positions, so as to realize the positioning. SIFT, that is, scale-invariant feature transformation, is a description used in the field of image processing. This description has scale invariance and can detect key points in the image. It is a local feature descriptor. The system combines SIFT feature extraction, bag-of-words model, and location matching to search and retrieve indoor scene images. Users capture and send query images and get location feedback through their smartphones. This paper proposes a system to build a bag-of-words model based on SIFT features to obtain the visual vocabulary of indoor scenes. It establishes an inverted file index for the image data set of the indoor scene, and then performs location matching. In the image retrieval stage, the first step is to retrieve images with the same image by matching

the visual vocabulary of the query image. The second step is to perform position matching. It calculates the similarity between the candidate images retrieved in the first step and the query image and ranks them. Then, it votes for the position of the top three images. If the vote is passed, the position will be returned, and if the vote is not passed, the position matching algorithm will be carried out one by one. The position matching algorithm combines the homography constraints between SIFT feature points, which ensures the accuracy of position matching. In order to ensure the simplicity of the user's operation, the user's client only requires the user to shoot and send the scene in which he is located, and the server returns the client's user's location through retrieval and matching. The main system framework and process of this paper are shown in **Figure 1**.

The Experimental Environment of This Experiment

Pentium(R) dual-core CPU, main frequency 3.20 GHz, memory 2 GB, operating system Windows 7, and development environment is MATLABR2008a. In order to verify the performance of the improved SIFT feature descriptor, a simple matching algorithm based on kd-tree is used. It uses an improved SIFT algorithm, which extracts SIFT feature points and SIFT feature descriptors from the training atlas and stores them in kd-tree. kd-tree is widely used in SIFT feature point matching, which can quickly and accurately find the nearest neighbors of query points. The

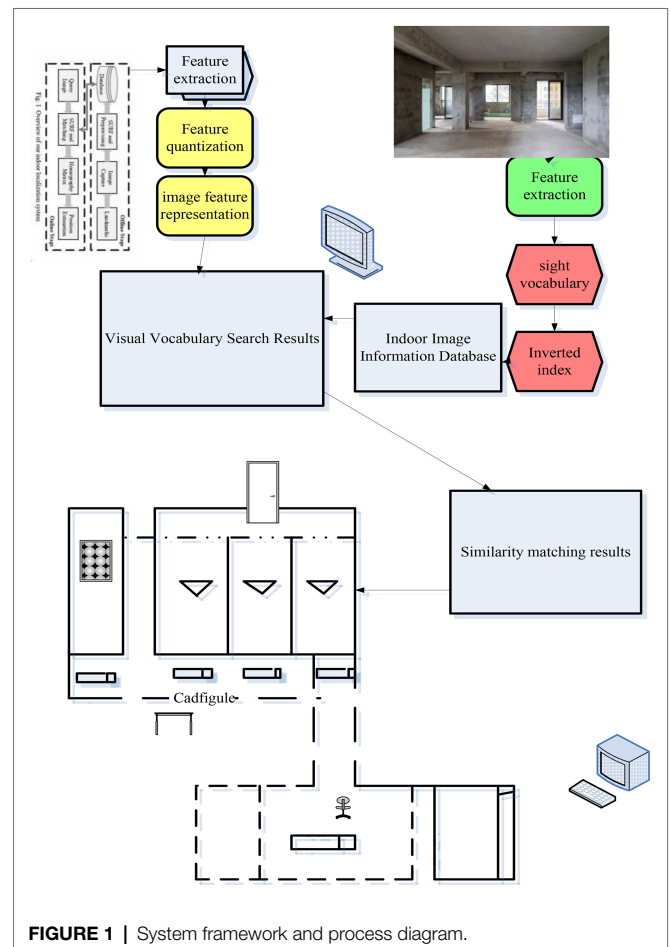


FIGURE 1 | System framework and process diagram.

basic idea is to divide the search space hierarchically without overlapping. It searches the training atlas for the k closest SIFT features to the SIFT features in the test images. When the distance between the closest feature point searched and the query feature point is within a certain threshold, it records the feature point as a matching feature point. When most of the features in an image are matched with the query image, the image is regarded as a matching image. The algorithm flow is shown in **Figure 2**:

In the experiment $k=1$, so searching through kd-tree will return a nearest neighbor feature point. It is at a distance d from the query image feature. It only accepts this match if $d < T1$, otherwise it rejects. This paper observes the similarities and differences between various SIFT feature descriptors. In the experiment, SIFT features with thresholds T1, 128D, 96D, 64D, and 32D were determined for SIFT feature descriptors of different dimensions in image matching. The distance thresholds corresponding to the descriptors are 170, 160, 150, and 90 (Pian et al., 2017). The candidate image with the largest number of matching points is considered as the best matching image.

The matching results of the rotated images are shown in **Figure 3**. The results show that the SIFT feature descriptors of 32D, 64D, 96D, and 128D all have good orientation invariance.

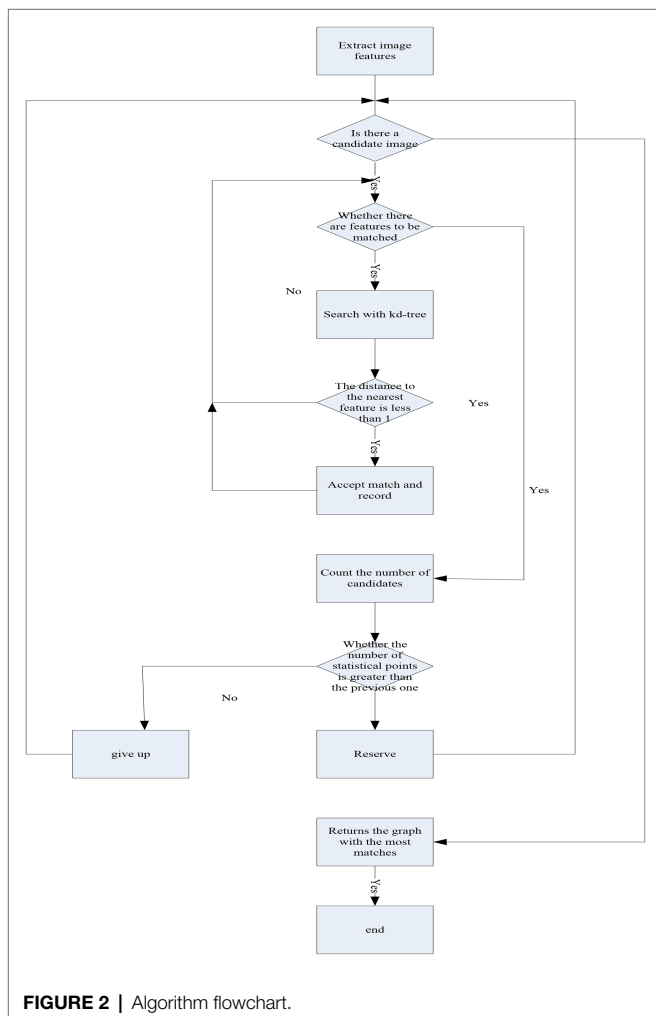


FIGURE 2 | Algorithm flowchart.

Experiments With Illumination Invariance

To test the illumination invariance of the SIFT feature descriptor, it adds or subtracts a portion of the pixel offset from the same 500 images, such as increasing the red, green, and blue intensity of each pixel at the same time. As shown in **Figure 4**, the offset is generally between 0 and 255. The pixel offsets tested are 50, 70, 100, 120, -30, -50, -70, -100, and -120.

Noise Invariance Experiment

To test the noise invariance of the SIFT feature descriptor, the experiment uses the imnoise function in MATLAB to add noise to the image. It uses three kinds of noise: Gaussian noise, salt and pepper noise, and speckle noise. Before adding noise, the pixels of each test image should be normalized first. Here, we use white Gaussian noise of $\phi^2 = 0.1$, salt and pepper noise of 15% and 30%, and a mean of 0. The experimental results of speckle noise with a variance of 0.04 are shown in **Figure 5**. Experiments show that the matching accuracy of SIFT feature descriptor will be significantly reduced in the environment of salt and pepper noise. This situation is more obvious for small-dimensional SIFT descriptors.

As shown in **Figure 6**, the matching performance of SIFT feature descriptors of all scales is not affected in the case of image blur. When $\sigma=20$, the test image is too blurred, which reduces the matching performance of 32D SIFT feature descriptors. In the case of high ambiguity, large-dimensional SIFT feature descriptors are less affected (Katsanos and Vamvatsikos, 2017).

Scale Transformation

To test the scale invariance, 500 objects were selected from the image dataset in the experiment. Each object has two images of different scales, with a total of 1,000 images. All the selected object images have scale transformation, and some

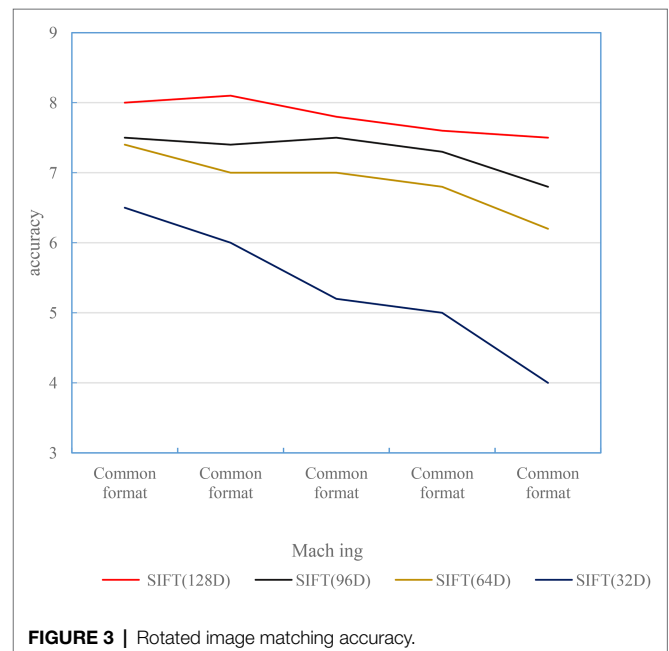
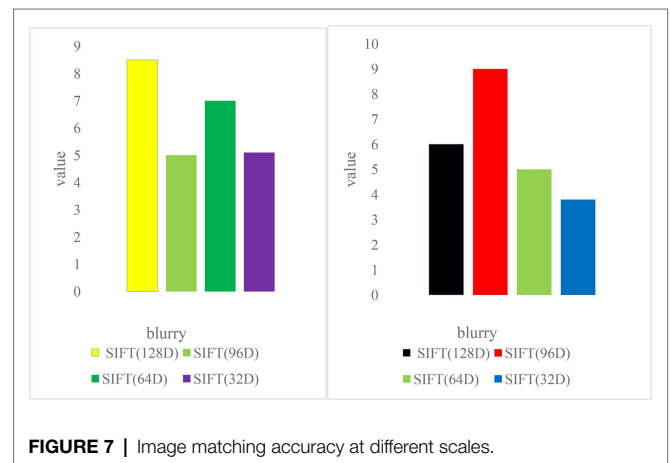
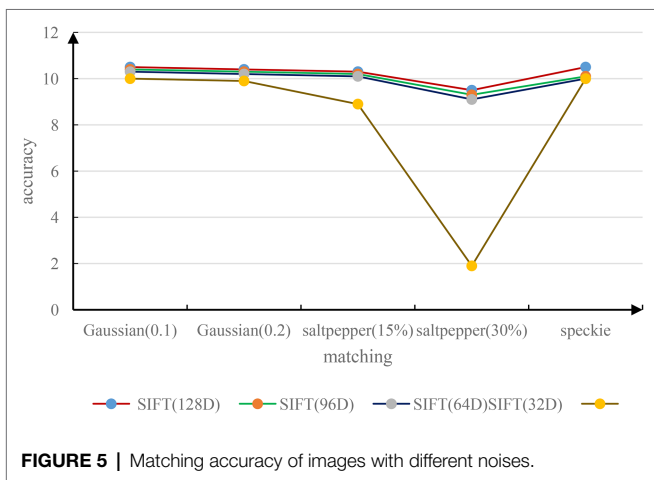
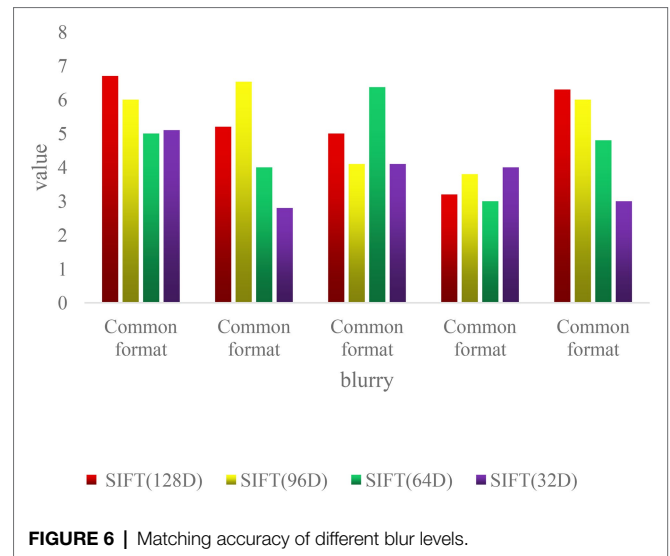
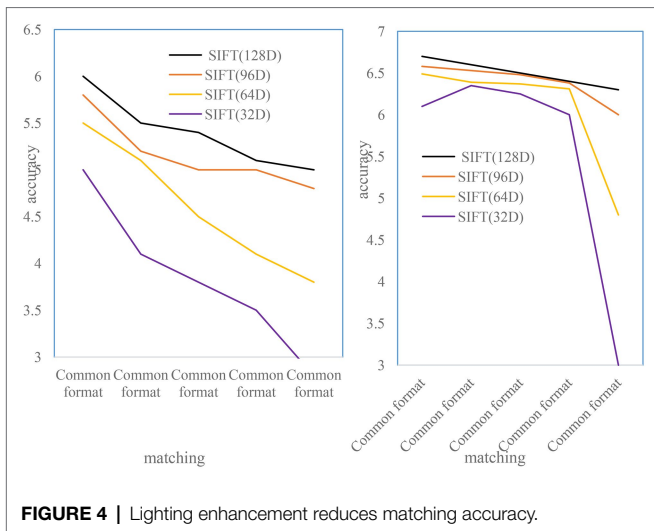


FIGURE 3 | Rotated image matching accuracy.



images also have perspective transformation. One image of each object is used for testing and the other is used for training. As shown in **Figure 7**, the 128D, 96D, and 64D SIFT feature descriptors have good matching performance in the case of scale transformation, and the 32D SIFT feature has poor matching performance. Moreover, the SIFT matching accuracy of 96D and 64D is even higher than that of the 128D SIFT feature descriptor. This is because the improved 96DSIFT and 64DSIFT make the matching center of gravity more concentrated in the vicinity of the feature points by calculating the gradient value weights of different regions.

Perspective Transformation

To test perspective invariance, images of 500 objects were selected from the image dataset. Each object has two images from different perspectives, with a total of 1,000 images. The perspective transformation of each object is different. The first image of the object is used for testing and the other is used for training. The experimental results are shown in **Figure 7**. The matching performance of all SIFT feature descriptors under perspective transformation is worse than that of other cases,

and the matching accuracy of 64D SIFT feature descriptors is higher than others (Oteda et al., 2021).

Experimental Results and Analysis

Through the operation of the above four different algorithms, the experimental results that can be obtained are mainly the position of the returned query image and whether the position matching algorithm is carried out in a query. The experimental results are shown in **Figure 8**. It will then analyze the results of the experimental summary data.

Matching Accuracy Analysis of the Four Algorithms

Each algorithm was run 15 times for cross-validation. The average matching accuracy is shown in **Figure 9**, and the system SD is shown in **Figure 10**. The results show that the homography-based location confirmation algorithm performs the best in terms of matching accuracy, and also has very high stability (low SD). The matching accuracy of the location matching algorithm based on the color histogram is more stable than that of the location confirmation algorithm based on the SIFT distance. This is because of the addition of color information.

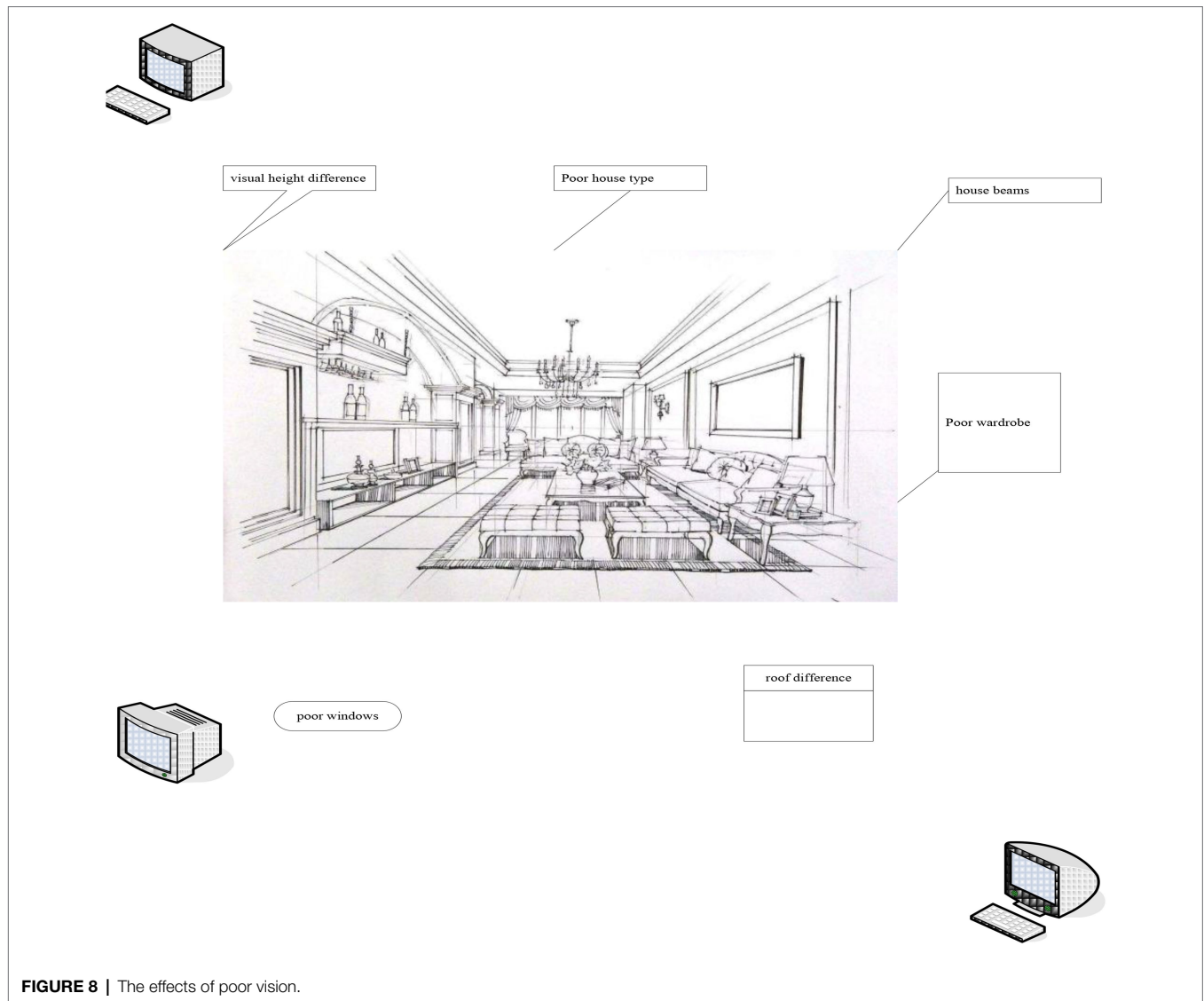


FIGURE 8 | The effects of poor vision.

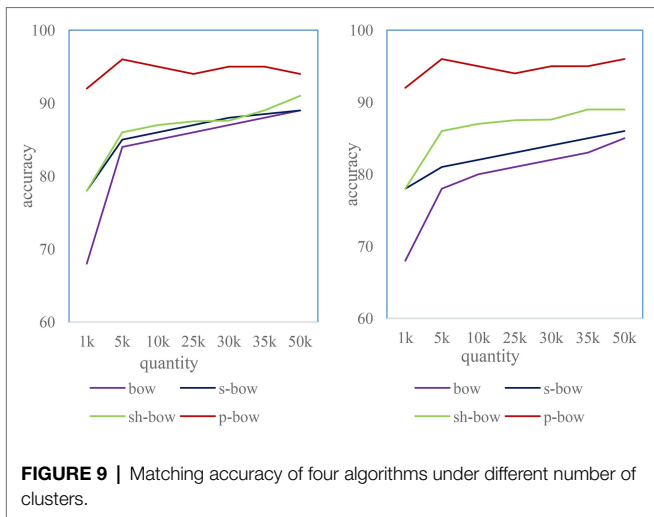
In order to further analyze the performance of the homography-based position matching algorithm, we choose to consider only the matching accuracy under non-ideal conditions in the following analysis. It serves as a metric in the following two aspects (Simpal, 2020).

1. Unable to match ratio (NoDecisionRate, R.R): the number of query images cannot get the best match through the position matching algorithm.
2. Correct matching ratio (Correctacceptancerate, C.A): it is the correct matching ratio after passing through the position matching algorithm.

DISCUSSION

In this paper, the use of image retrieval techniques to implement an indoor positioning system is introduced. It experimentally analyzes the effects of different locations from

the perspective of visual capture. However, because the vision-based indoor positioning system needs to collect and preprocess the indoor scene images in advance, the image quality of the image database and the richness of the image database will affect the accuracy of position matching. Large-scale datasets require more algorithmic improvements to improve retrieval efficiency (Deng, 2018). First, it combines optical illusion with various information for positioning, using GPS and WiFi positioning technology to combine with the vision-based indoor positioning system proposed in this paper. This enables the localization of more indoor scenes (Huang et al., 2017; Norford, 2018). For example, it uses GPS to first locate the location of the user. When the user puts forward a positioning requirement, the system can first retrieve the image data set of the user's location through the GPS positioning information, and then perform detailed positioning. This expands the influence of visual effects on positioning techniques in interior design.



CONCLUSION

People live in a three-dimensional world, but the images perceived by the human eye are indeed two-dimensional. People are exposed to the shape, size, distance, texture, and other information of objects since childhood. Then we analyze it through the brain, because people’s understanding of space is not innate. Our brains need to touch, feel, and analyze space to fully understand this three-dimensional world. At present, indoor visual positioning technology has not appeared in relatively excellent actual product systems, and most of them are only laboratory research. Considering some practical problems encountered in the research process of this paper, the following research work will conduct more detailed research and improvement in the following directions. The development of deep learning in recent years seems to have become a research trend. Deep learning has shown good performance in all aspects, so in the next research process, more in-depth research on

intelligent algorithms is required. It needs to learn from the research method of neural network algorithm to improve the scene classification and recognition rate. In terms of fusion algorithm, Kalman filter technology is used in this paper. Compared with other advanced filtering technologies, it lacks comparability, and the experimental environment is a relatively simple indoor environment. Therefore, it requires more complex experiments in follow-up work. In addition, we can consider trying to combine better filtering techniques to fuse the data and remove impurities. In the actual research process of this topic, due to the limited theoretical knowledge and the lack of practical experience, there are also some places that contradict the traditional knowledge structure and inertial thinking. The research results of the paper still need to be continuously enriched and improved. The theoretical breadth and depth of thinking are still lacking, and the shortcomings will be continuously improved and comprehended in future studies and work.

AUTHOR’S NOTE

YX was born in Jiangsu, China, in 1981. She received the Master degree from Tianjin Polytechnic University, China. Now, she studies in the School of Economics and Management, NanJing University of Aeronautics and Astronautics. Her field of study is environmental design.

E-mail: yunkai@nuaa.edu.cn.

TY was born in WenZhou, Zhejiang province, China, in 1987. She received the Master degree from Kunming University of MFA, China. Now, she works in the School of architectural engineering, Wenzhou Polytechnic. Her research interests include intelligent interior design for the elderly. Research on interior design culture.

E-mail: 2021000528@wzpt.edu.cn.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

YX: writing and static analysis of data. TY: guiding the research directions and ideas. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Ministry of Education Humanities and Social Sciences Research project “project approval number 20YJCZH218” under healthy Chinese perspective integrated into the intelligent elderly housing space planning design research.

REFERENCES

- Caulfield, J. (2018). "AI arrives," in *Building Design and Construction*. Vol. 59, 32–35.
- Celadyn, M. (2017). Environmental sustainability considerations in an interior design curriculum. *World Trans. Eng. Technol. Educ.* 15, 317–322.
- Celadyn, M. (2018). The use of visuals in environmentally responsible interior design. *World Trans. Eng. Technol. Educ.* 16, 212–213.
- Chang, A., and Yang, X. (2020). Research on the development of intelligent consulting model under the background of big data. *TPPC* 2, 486–495. doi: 10.35534/tpcc.0208037
- Deng, T. S. (2018). Visual representation of popular Science Publications of global warming. *J. Educ. Media Libr. Sci.* 55, 171–225. doi: 10.6120/JoEMLS.201807_55(2).0013.RS.BM
- Fang, L., Wang, J., and Lu, G. (2017). Color transfer algorithm for interior design through region matching based on topological information. *JCADCG* 29, 1044–1051.
- Grobelyny, J., Michalski, R., and Weber, G. W. (2021). Modeling human thinking about similarities by neuromatrices in the perspective of fuzzy logic. *Neural Comput. & Applic.* 33, 5843–5867. doi: 10.1007/s00521-020-05363-y
- Horton, A. (2017). The horse as a symbol of patriotism in the United States. *Int. J. Vis. Des.* 11, 1–16. doi: 10.18848/2325-1581/CGP/v11i04/1-16
- Huang, Y., Zhang, J. S., and Wang, J. D. (2017). Comparative study of visual performance of two types of rotationally asymmetric segmental-refractive multifocal intraocular lens. *Ophthalmol. China* 26, 248–252. doi: 10.13281/j.cnki.issn.1004-4469.2017.04.008
- Jain, D. K., Boyapati, P., Venkatesh, J., and Prakash, M. (2022). An intelligent cognitive-inspired computing with big data analytics framework for sentiment analysis and classification. *Inf. Process. Manag.* 59:102758. doi: 10.1016/j.ipm.2021.102758
- Jeamwathanachai, W. (2019). Spatial representation framework for indoor navigation by people with visual impairment. *J. Enabling Technol.* 13, 212–227. doi: 10.1108/JET-12-2018-0068
- Ju, S., and Yang, J. S. (2021). Development and application of housing and interior design courses work for the promotion of service-learning in home economics education. *Fam. Environ. Res.* 59, 99–112. doi: 10.6115/fer.2021.008
- Katsanos, E. I., and Vamvatsikos, D. (2017). Yield frequency spectra and seismic design of code-compatible RC structures: an illustrative example. *Earthq. Eng. Struct. Dyn.* 46, 1727–1745. doi: 10.1002/eqe.2877
- Leung, C., and Zhang, Y. (2019). An HSV-based visual analytic system for data science on music and beyond. *IJACDT* 8, 68–83. doi: 10.4018/IJACDT.2019010105
- Mahmood, F. J., and Tayib, A. Y. (2019). The role of patients' psychological comfort in optimizing indoor healing environments: a case study of the indoor environments of recently built hospitals in Sulaimani City, Kurdistan, Iraq. *HERD* 13, 68–82. doi: 10.1177/1937586719894549
- Norford, L. (2018). Energy accounts: architectural representations of energy, climate, and the future. *Technology Architecture Design* 2, 114–115. doi: 10.1080/24751448.2018.1420970
- Oteda, P. M., Fabellon, A. K., and Toquero, C. M. D. (2021). Expository text structure: a strategic visual representation in enhancing the reading comprehension skills of the learners. *Musamus J. Prim. Educ.* 3, 104–114. doi: 10.35724/musjpe.v3i2.3339
- Pian, Z., Shi, T., and Yuan, D. (2017). Application of hierarchical visual perception in target recognition. *JCADCG* 29, 1093–1102.
- Ponde, P. S., Shirwaikar, S. C., and Kharat, V. S. (2017). Knowledge representation of security design pattern landscape using formal concept analysis. *J. Eng. Sci. Technol.* 12, 2513–2527.
- Resch, G. (2019). Accounting for visual Bias in tangible data design. *Digital Culture and Society* 5, 43–60. doi: 10.14361/dcs-2019-0104
- Shubha, S., and Meera, B. N. (2019). The effect of visual representation in developing conceptual understanding of magnetic force. *JFDL* 3, 82–87. doi: 10.1007/s41686-019-00031-4
- Simpal, M. T. (2020). Use of visual representation and peer-assisted approach in developing students' mental models in solving physics problems. *HERJ* 10, 84–95.
- Wallace, J. (2019). Ethics and inscription in social robot design. A visual ethnography. *Paladyn* 10, 66–76. doi: 10.1515/pjbr-2019-0003
- Zhang, H., and Zheng, H. (2017). Research on interior design based on virtual reality technology. *Tech. Bull.* 55, 380–385.
- Zhu, Z., Cao, H., and Li, B. (2017). Research on logo design and evaluation of youth education brands based on visual representation. *J. Prod. Brand. Manag.* 26, 722–733. doi: 10.1108/JPBM-08-2016-1287

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xu and Yu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.