Aesthetic Research on Intelligent Automation Design Combined with Virtual Reality Under the Background of Green Environmental Protection

Xiaotang Sun and Xiaoqi Liu(✉)

Beijing Institute of Fashion Techonology, Beijing, China
20180028@bift.edu.cn

Abstract. In recent years, with the great breakthrough of deep learning technology, the upsurge of artificial intelligence (AI) has been set off and flourished. Virtual reality (VR) technology based on AI has also made great progress. Remote interactive VR technology can effectively reduce carbon emissions and is of great significance to environmental protection. Therefore, this paper studies the aesthetics of VR of AI. AI and VR can help computers build technologies and environments that meet aesthetic needs, and then carry out intelligent and automated aesthetic design. Firstly, the basic situation of image aesthetic evaluation is introduced, including the task model, database and evaluation index of image aesthetic evaluation. Secondly, an image aesthetic evaluation algorithm that integrates visual saliency and compositional edge information is proposed. Finally, the current mainstream algorithm and the proposed algorithm are compared and tested. The experiment results show that the image aesthetic evaluation algorithm combining visual saliency and proposed composition edge information is superior to the current mainstream algorithms regarding the algorithm performance. The results have the reference significance for the aesthetic research of AI virtual technology design.

Keywords: Virtual technology · Intelligent automation aesthetic design · Image aesthetic evaluation algorithm · Visual saliency · Composition edge information · Artificial intelligence aesthetics

1 Introduction

With the advancement of science and technology, artificial intelligence (AI) has appeared in all aspects of lives, whether it is a smart phone, household appliances, or transport tools are updated with technology, continue to implant new intelligent programs, more and more comprehensive and humane [1, 2]. AI is a product created by human beings. The ultimate goal of its aesthetic production is to better understand human thinking and emotion itself, and to bring people a better life [3].

Virtual reality (VR) is based on computer technology, combined with related science and technology, to generate a digital environment that is highly similar to a certain

C. F. Peng et al. (Eds.): EIMT 2023, AHSSEH 8, pp. 757–768, 2023.
https://doi.org/10.2991/978-94-6463-192-0_97

range of real environments in terms of sight, hearing, touch, etc. [4, 5]. The emergence of the new medium of VR makes the VR environment with the characteristics of multi-dimensionality, interactivity, immediacy, and immersion come into being [6]. VR technology has been put into use in a number of areas, such as: three-dimensional virtual tourism, providing a new model of tourism education and training and technology, building the ecological environment of education, and using VR technology to build a three-dimensional scenic spot. These technologies allow people to stay at home and can feel face-to-face experience [7, 8]. It not only greatly facilitates people's lives, but also environmental protection plays an important role in reducing man-made environmental pollution [9].

In order to have a better visual experience, AI aesthetic is essential to be applied to VR technology, and the "aesthetic" problem, since its inception has been classified as a philosophical problem [10]. If looking at the "Turing test" in this way, it is true that current AI technology can distinguish between the works or styles of a particular author through its vast data support and data analysis [11]. Therefore, the problem here is that it does not deny that the computer has this kind of ability to match people, or even surpass people's 'intelligence' to 'aesthetic', but the computer does not have people's 'consciousness' and 'experience'. Therefore, the current image aesthetic evaluation research for deep learning methods is not sufficient [12]. Aesthetics is an indispensable part of human emotional life, especially in the era of big data. It is of great practical significance and broad application prospects to simulate human aesthetics through computers and make corresponding evaluations of image aesthetics [13]. Existing research on image aesthetics evaluation fails to fully consider the saliency and attention mechanism of the human brain in the process of perceiving and evaluating aesthetics, so it is difficult to achieve high performance [14].

To sum up, the research on image aesthetics evaluation starts from the two aspects of human visual saliency and the human brain's attention to image aesthetics. Based on fully considering the human brain's visual evaluation mechanism, a new method based on fusion vision is proposed. The image aesthetic evaluation algorithm of saliency and composition edge information is applied to 3D virtual scenes, so that the construction of 3D virtual scenes is more in line with human aesthetics. This work has important research significance for the research of intelligent automation aesthetic design combining VR technology with AI.

2 Materials and Methods

2.1 Intelligent Automation Design Principle in VR Environment

With the convenience and popularization of AI and VR technology, the process of design aesthetics is no longer limited to linear physical attributes, but exists with the appearance of VR technology as immaterial, digital, networked and virtualized. The scene system provided by VR technology will provide users with real-time visual, auditory, tactile and other direct sensory stimuli, and the perception management system will feed back the perception to users. After receiving the perceptual feedback, the human brain will make the second conversion and send the action instructions to the body again. This cycle can improve the working efficiency of the whole system. VR can be used to visualize and manipulate some conceptual designs or ideas, and to simulate the live effects of reality. The application of VR technology in aesthetic design has changed the original aesthetic experience mode, and the static and planar perception is developing towards a dynamic and three-dimensional situation. VR system is an open and diverse environment. In the intelligent virtual environment, aesthetic subjects can interact in real time in the VR environment, exchange data and information with the virtual environment, and can use real-time data variables fed back by users. They are in a dynamic environment, and thus get a dynamic aesthetic experience. As far as traditional aesthetic design is concerned, VR environment can embed the subject into the design in real time, and the explicit result of the design work becomes the result of the participation and creation of designers, equipment and users.

Intelligent design depends on computers and electronic devices for display and transmission. Taking the personal computer as an example, based on the openness and diversity of design aesthetic operating system, we design our own functional blocks within the framework of application programs. These functional blocks can be separated and supported among layers to form a framework with clear division of labor. In order to realize its functions, computers and intelligent automatic design equipment must rely on the above applications. And the operating systems of computers and mobile electronic devices are actually an application program. Each functional platform needs to design its own application program and man-machine interface to receive and feedback the user's operation instructions. Human-machine interface plays a vital role in interactive design, making it easy to use, easy to use, efficient and direct to complete information transmission, and achieving good results in visual communication.

2.2 Image Aesthetic Evaluation Task Analysis

Image aesthetic evaluation is generally divided into five-layer tasks, as shown in Fig. 1. The five-layer tasks are aesthetic classification, aesthetic evaluation, aesthetic distribution, aesthetic factors, and aesthetic description [15].

Aesthetic classification is to classify the images in the database according to the degree of aesthetics, generally divided into two categories: high and low. The classification problem has been widely studied in the field of computer vision. The aesthetic evaluation is to determine the score of the image according to the aesthetic standard by establishing a regression model. The aesthetic distribution is based on the fact that users tend to present a distribution state when scoring images. In the early days, the average of these scores was generally used as the aesthetic label of the image. Both aesthetic classification and aesthetic evaluation represent aesthetics in the form of scalars, and aesthetic distribution contains more valuable information, so a more accurate and intelligent way is needed for aesthetic representation. With the continuous breakthrough of image aesthetic evaluation research, the image aesthetic evaluation method of convolutional neural network has become a new idea for researchers to solve image aesthetics.

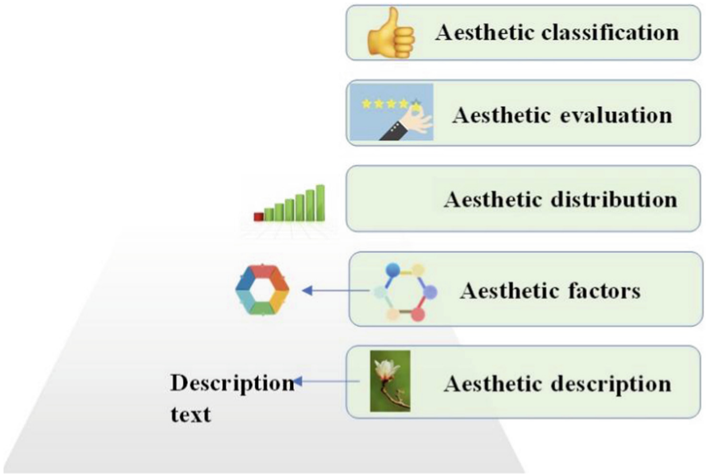

Fig. 1. Five-layer task model for image aesthetic evaluation

2.3 Image Aesthetics Evaluation Algorithm Fusing Visual Saliency and Compositional Edge Information

Visual saliency is an important part of the research on human attention mechanism, and the related research is based on the mechanism of human visual system [16]. Visual saliency generally means that when humans observe the actual scene or the scene in the image, they will always be attracted by a certain area or part of these scenes, so they will pay more attention to the area / part, and pay less or even no attention to the other parts of the scene. As shown in Fig. 2, visual saliency is a hot research topic in computer vision, and it has a wide range of applications in the fields of image compression, image recognition, image scaling, image inspection, image segmentation, and object recognition.

Edge detection can be seen as a method to extract visually salient edges and object boundaries from natural images [17, 18]. As shown in Fig. 3, images selected from the aesthetic image database, image saliency regions and edge detection result graphs obtained using the visual saliency detection algorithm and the edge detection algorithm are displayed.

The edge features of the image not only contain the content of the image but also reflect the image composition and other information, and this information belongs to the aesthetic attributes of the image. Therefore, it can be concluded that the image edge information can represent the aesthetic attributes to a certain extent, which is helpful for the deep neural network to improve. It can learn image aesthetic information well and improve the accuracy of aesthetic prediction. Based on this, an image aesthetic evaluation model by fusing image composition edge features and visual saliency features is proposed.

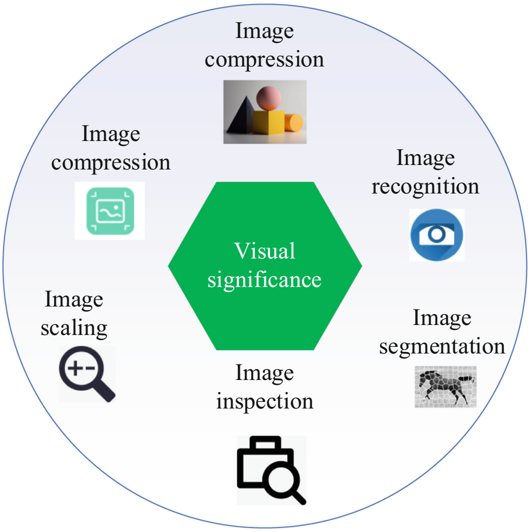

Fig. 2. Examples of visual saliency scenarios

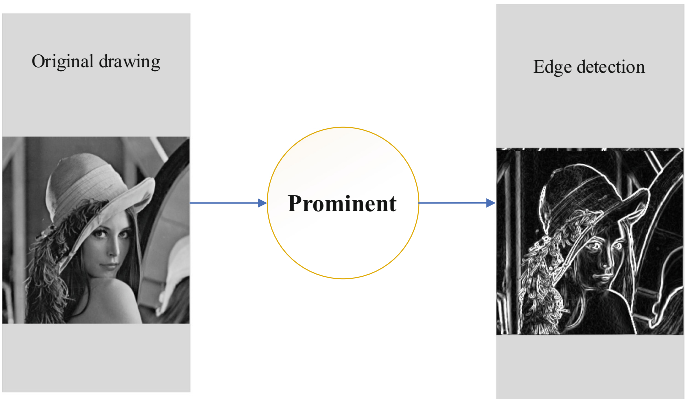

Fig. 3. Schematic diagram of edge detection

2.4 An Aesthetic Evaluation Model for Fusion of Visual Saliency and Compositional Edge Information

The image aesthetic evaluation algorithm based on the fusion of visual saliency and compositional edge information is composed of three parts: visual saliency detection module, compositional edge information feature module and deep feature learning module. Firstly, the saliency area map corresponding to the image in the aesthetic image database is obtained through the visual saliency detection algorithm. Second, the composition edge feature map of the image is obtained through the edge feature network model. Third, the depth feature of the image in the aesthetic image database is extracted through the deep neural network. The obtained saliency features, edge features and deep learning features

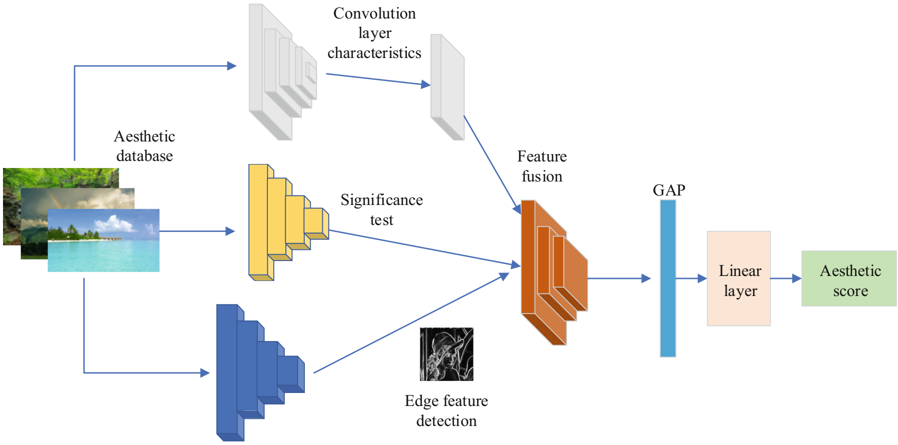

Fig. 4. Aesthetic evaluation algorithm model for fusion of visual saliency and compositional edge information

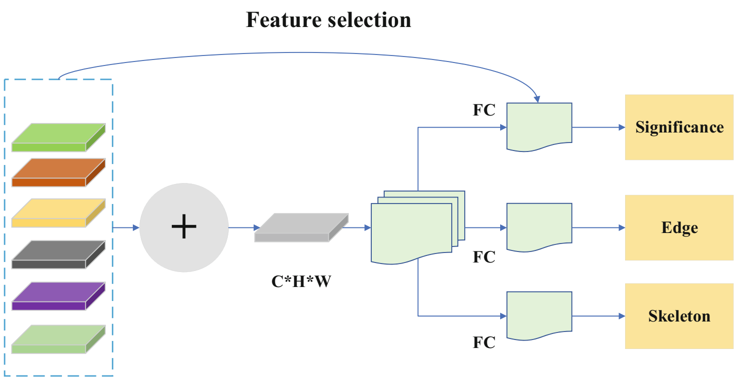

Fig. 5. DFI algorithm dynamic feature integration module diagram

are subjected to feature fusion operation, and the fused features are used to predict the final image aesthetic score. The algorithm frame diagram is shown in Fig. 4.

Through the joint efforts of many researchers, the current research on image saliency detection has been relatively mature, and some existing saliency detection algorithms can detect the target area of the image well. The visual saliency detection module refers to a new saliency detection algorithm Deep/Dynamic Flow Inspection (DFI) proposed by Liu et al. (2020) of Nankai University. It is a dynamic feature integration algorithm that can be used to simultaneously detect salient objects, edges, and skeletons of images [19]. Figure 5 presents the DFI algorithm dynamic feature integration module.

2.5 Image Aesthetics Database

The research on image aesthetic evaluation has only started in the past ten years, and related research based on deep learning algorithms has always relied on a data-driven

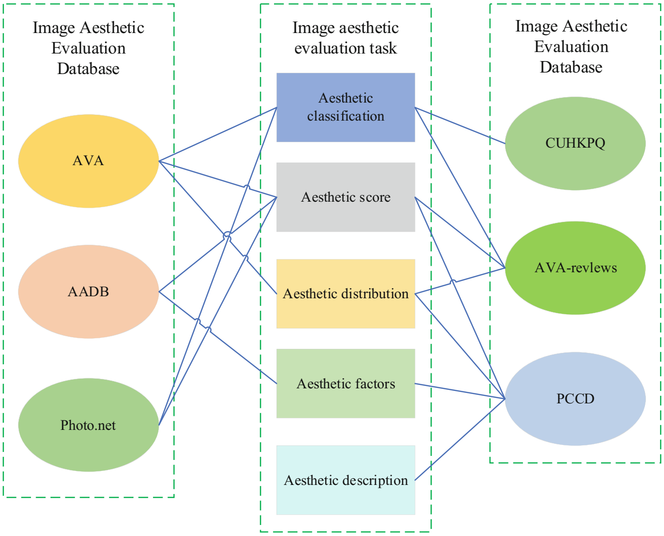

Fig. 6. Introduction to the public aesthetic evaluation database

approach. The construction of image aesthetic evaluation dataset is generally to obtain the required aesthetic labels through manual subjective evaluation. In terms of manual subjective evaluation, it can be divided into manual scoring in laboratory environment, online image website scoring and download collection, and crowdsourcing evaluation [20]. The data sets A Large-Scale Database for Aesthetic Visual Analysis (AVA) and Aesthetic with Attributes Database (AADB) formed by online image website scoring download collection and crowdsourcing evaluation are the most abundant samples. Therefore, the experiment is mainly based on two public image aesthetics databases, AVA, and AADB, which are shown in Fig. 6.

2.6 Image Aesthetics Evaluation Algorithm Performance Evaluation Index

The evaluation index of image aesthetic evaluation is an important basis for measuring the performance of the algorithm. Different aesthetic evaluation tasks have different evaluation indicators. The evaluation indicators of aesthetic classification and aesthetic score are introduced below.

(1) One of the aesthetic classification evaluation indicators is the accuracy rate (ACC), and the calculation equation is as Eq. (1):

$$ACC = \frac{TP + TN}{P + N} \tag{1}$$

TP represents the number of images classified correctly by positive samples. TN represents the number of images classified correctly by negative samples. P represents the number of images in positive samples. N represents the number of images in negative samples. The value range of ACC is in the [0,1] interval, and the result is usually expressed as a percentage (%). The higher the value of ACC, the higher the classification accuracy, that is, the better the classification performance of the algorithm.

(2) The aesthetic evaluation task is based on the label in the database as the score value. Therefore, the aesthetic evaluation task is actually a regression task. It has three metrics: Spearman Rankorder Correlation Coefficient (SRCC), Pearson Linear Correlation Coefficient (PLCC), and Root Mean Square Error (RMSE). SRCC is used to measure the monotonicity of the aesthetic evaluation algorithm, and PLCC and RMSE are used to measure the accuracy of the aesthetic evaluation algorithm.

The equation for calculating SRCC is:

$$SRCC = 1 - \frac{\delta \sum_{I=1}^{N} d_i^2}{N(N^2 - 1)} \tag{2}$$

Among them, d is the difference between the score corresponding to the label of the ith image and the order of the predicted score of the algorithm. N is the number of image samples in the test set. The value range of SRCC is generally in the [0, 1] interval. The larger the value of SRCC is, the higher the monotonicity is, that is, the algorithm has superior performance.

The calculation equation of PLCC is:

$$PLCC = \frac{\sum_{i=1}^{N}(s_i - s)(p_i - p)}{\sqrt{\sum_{i=1}^{N}(s_i - s)^2}\sqrt{\sum_{i=1}^{N}(p_i - p)}} \tag{3}$$

Among them, s_i is the score value corresponding to the label of the i th image, s is the average value of s, N is the number of image samples in the test set, and p_i is the score value predicted by the algorithm for the i th image. The value range of PLCC is generally in the [0,1] interval. The larger the value of PLCC, the higher the accuracy, which means the algorithm has superior performance.

The equation for calculating RMSE is:

$$RMSE = \sqrt{(\frac{1}{N}\sum_{i=1}^{N} s_i - p_i)} \tag{4}$$

Among them, s_i s is the score value corresponding to the label of the ith image, p_i is the score value predicted by the algorithm of the ith image, and N is the number of image samples in the test set. The value range of RMSE is in the [0,1] interval. The smaller the value of RMSE, the closer the predicted score value of the algorithm is to the score value of the label in the database, that is, the better the performance of the algorithm.

2.7 Experiments Settings

The algorithm is tested on Windows 10 operating system platform, Intel (R) Core (TM) i5 2.5 GHz CPU, 8 G memory. Two image aesthetics databases, AADB and AVA, are used to test the performance of the proposed aesthetic evaluation algorithm. Two popular deep network models, Residual Network (ResNet) and EfficientNet, are used as the backbone network. Both backbone networks use the model trained on ImageNet as their initialization. The Euclidean Loss and Mean Absolute Error Loss (MAE) functions are used to jointly guide the network training, which can make the network training process more stable. The calculation equation is as follows (5):

$$Loss = \frac{1}{N}\sum_{i=1}^{N}\|s_i - p_i\|_2^2 + \frac{1}{N}\sum_{i=1}^{N}|s_i - p_i| \tag{5}$$

Among them, *s* represents the score value corresponding to the label of the ith virtual scene image, and *p* represents the score value predicted by the algorithm for the ith image.

3 Results and Discussion

In order to better explore the fusion algorithm, and explore the contribution of different features in the algorithm, the algorithm performance test is carried out for the fusion features of salient features, edge features and salient features, and edge features. The detection results are shown in Fig. 7:

Figure 7 shows that ACC and SRCC are not more than 60% after using significant features and edge detection features alone, but ACC and SRCC are 82.9% and 71.2% respectively after the two features are used together. ACC and SRCC data change significantly before and after the fusion algorithm is used. Therefore, the fusion of significant features and edge detection features has a significant enhancement effect on the performance of the algorithm.

The performance of the proposed algorithm and the existing mainstream algorithms are compared. The algorithms used for comparison include AVA (handerafted featuresl), ResNet101, and ResNet101 (improved version). The indicators of algorithm performance measurement include PLCC, SRCC and RMSE. The algorithm performance test results are shown in Fig. 8.

Figure 8 shows that the performance of the existing mainstream algorithm and the algorithm proposed is better than the three comparison mainstream algorithms in terms of PLCC, SRCC, and RMSE. The algorithm proposed is 70.9%, 70.4%, and 1.5% in PLCC, SRCC and RMSE, respectively. It can be concluded that the algorithm has strong applicability and practicality.

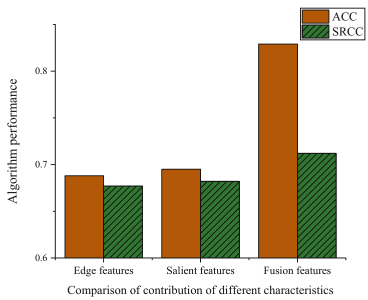

Fig. 7. Algorithm performance verification for different features

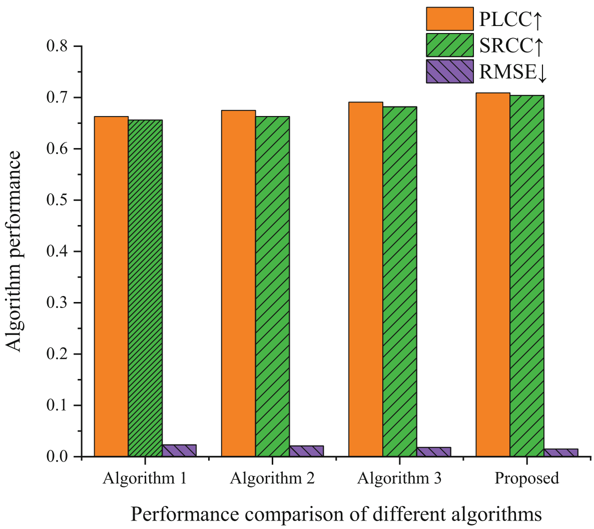

Fig. 8. Performance comparison with mainstream algorithms

4 Conclusion

Nowadays, society comprehensively advocates green environmental protection. In this context, in order to better construct the virtual scene and make the design of the virtual scene more in line with people's aesthetic, based on the basic situation of green environmental protection and image aesthetic evaluation, this paper proposes an aesthetic image evaluation algorithm combining visual salient features and edge features, and compares it with the current mainstream algorithm. AI and VR can help computers build technologies and environments that meet aesthetic needs, and then carry out intelligent and automated aesthetic design. According to the principle of intelligent automatic design in VR environment, an intelligent automatic virtual design aesthetic operating system platform is provided. Through analysis, the salient features and edge features of the image can well represent the aesthetic feeling of the image. Therefore, the saliency features and edge features of the image are extracted respectively to assist the backbone deep network to learn the image aesthetics. Experiments show that the accuracy of image aesthetic representation after the fusion of the two features reaches 82.9%, which is better than the current mainstream image aesthetics algorithm. Therefore, the image aesthetic evaluation algorithm proposed can effectively improve the efficiency and accuracy of image aesthetic representation. However, the work still has shortcomings. Visual saliency features need to rely on existing visual saliency detection algorithms, which will also be the focus of future research work.

References

1. Cui, M., & Zhang, D.Y., "Artificial intelligence and computational pathology," *Laboratory Investigation*, *101*(4), 412-422 (2021).

2. Gunasekeran, D. V., Tseng, R. M. W. W., Tham, Y. C., & Wong, T. Y., "Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence," telehealth and related technologies. NPJ digital medicine, 4(1), 1-6 (2021).
3. Zeba, G., Dabić, M., Čičak, M., Daim, T., & Yalcin, H., "Technology mining: Artificial intelligence in manufacturing." Technological Forecasting and Social Change, 171, 120971 (2021).
4. Emmelkamp, P. M., & Meyerbröker, K., "Virtual reality therapy in mental health." Annual Review of Clinical Psychology, 17, 495-519 (2021).
5. Xi, N., & Hamari, J., "Shopping in virtual reality: A literature review and future agenda." Journal of Business Research, 134, 37-58 (2021).
6. Jang, J., Ko, Y., Shin, W. S., & Han, I. "Augmented reality and virtual reality for learning: An examination using an extended technology acceptance model." IEEE Access, 9, 6798-6809 (2021).
7. Riches, S., Azevedo, L., Bird, L., Pisani, S., & Valmaggia, L., "Virtual reality relaxation for the general population: a systematic review." Social psychiatry and psychiatric epidemiology, 56(10), 1707-1727 (2021).
8. Fracaro, S. G., Chan, P., Gallagher, T., Tehreem, Y., Toyoda, R., Bernaerts, K., & Wilk, M., "Towards design guidelines for virtual reality training for the chemical industry." Education for Chemical Engineers, 36, 12-23 (2021).
9. Baghaei, N., Chitale, V., Hlasnik, A., Stemmet, L., Liang, H. N., & Porter, R., "Virtual reality for supporting the treatment of depression and anxiety: Scoping review." JMIR Mental Health, 8(9), e29681 (2021).
10. Obwegeser, D., Timofte, R., Mayer, C., Eliades, T., Bornstein, M. M., Schätzle, M. A., & Patcas, R., "Using artificial intelligence to determine the influence of dental aesthetics on facial attractiveness in comparison to other facial modifications." European Journal of Orthodontics. 5(1), 17-27 (2022).
11. Feng, T. A. O., "An Overview on Sci-technological Aesthetics in the Age of Artificial Intelligence." Philosophy, 11(7), 527-532 (2021).
12. Wang, Y. "Lin Fengmian's Art Series and Aesthetic Education Research for Artificial Intelligence Aesthetics." Frontiers in Art Research, 4(4).1-15 (2022).
13. Barale, A., "Who inspires who? Aesthetics in front of AI art." Philosophical Inquiries, 9(2), 195-220 (2021).
14. Greenfield, R., & Cao, S., "The edge of life-as-we-know-it: Aesthetics of decay within artificial life and art." Technoetic Arts, 19(1-2), 185-201 (2021)
15. Zhang, J., Miao, Y., & Yu, J., "A comprehensive survey on computational aesthetic evaluation of visual art images: Metrics and challenges." IEEE Access, 9, 77164-77187. (2021).
16. Liu, J. J., Hou, Q., Cheng, M. M., "Dynamic feature integration for simultaneous detection of salient object, edge and skeleton." IEEE Transactions on Image Processing, 29(1), 8652-8667 (2020).
17. Liu, X., Chen, P., Ding, X., Liu, A., Li, P., Sun, C., & Guan, H., "A narrative review of cortical visual prosthesis systems: the latest progress and significance of nanotechnology for the future." Annals of Translational Medicine, 10(12), 22-28 (2022).
18. Versaci, M., & Morabito, F. C., "Image edge detection: A new approach based on fuzzy entropy and fuzzy divergence." International Journal of Fuzzy Systems, 23(4), 918-936. (2021).
19. Pham, L. T., Oksum, E., Le, D. V., Ferreira, F. J., & Le, S. T., "Edge detection of potential field sources using the softsign function." Geocarto International, 1–14 (2021).
20. Takimoto, H., Omori, F., & Kanagawa, A., "Image aesthetics assessment based on multi-stream CNN architecture and saliency features." Applied Artificial Intelligence, 35(1), 25-40 (2021).