Journal for Geometry and Graphics
Volume 27 (2023), No. 1, 69–79

An Analysis of Visual Interest Detection in 2D Game Concept Art

Tiancheng San, Yoshihisa Kanematsu, Koji Mikami

Tokyo University of Technology, Tokyo, Japan
d312000256@edu.teu.ac.jp, kanematsuyh@stf.teu.ac.jp, mikami@stf.teu.ac.jp

Abstract. This study proposes a scheme that demonstrates the pattern of color combinations in the visual and non-visual centers. To achieve this goal, we first have to clearly distinguish the visual and non-visual center areas in the game concept design. Up to now, the saliency studies have been similar to our study. However, after our tests, the saliency studies do not distinguish well between the visual and non-visual center areas in the concept art. Therefore we propose this research method from the artist's perspective. We first partition the image into small local regions, similar to bags of feature model, and then we give each local region a hash string to determine their similarity. Then we classify similar local regions. In the painting principle, similar or repeated regions are often non-visual centers. Finally, we check the variation of HSV of local regions to confirm the accuracy of the visual center. At the end of the study, we compared the performance of our study with two other major saliency methods in a sample of game concept designs. The results show that our method can better show the visual center of the game concept design diagram.

Key Words: visual interest, game concept art, color contrast, segmentation, 2D digital artwork

MSC 2020: 91E10

1 Introduction

The application area of our study is to identify the visual center of the game concept art. There are many types of analysis with pictures. For example, Arnheim proposed the relationship between visual center and shape, form, and balance as early as 1960 [2]. The research from Nishiyama analyzed the color harmony using a combination of the local region to evaluate the photo's aesthetic quality [12]. They mentioned the method of bags of color, which is inspired by the bags of the feature model. With LDA-dual, Ali Jahanian analyzed the connection between the color of the magazine cover and the theme of the magazine [7]. Recently, there have been some studies on game art such as Andreas Lambrantd's research attempt to improve the color

ISSN 1433-8157/$ 2.50 © 2023 Heldermann Verlag

selection in game art [9]. All these researches are looking for the relationship between objective pictures and subjective feedback from audiences. The objects of these studies are mostly photos, but there are differences between photos and concept artworks. John [14] pointed out that game art also has the unique feature that game art combines art and technology. Unlike pictures and photos, game concept design drawings include visual centers designed by game concept designers. This makes many previous studies hard to apply to game concept art. For example, there exist many studies of color combinations and websites that automatically generate color palettes. But these studies analyze the color combinations for the whole picture. So it is difficult to apply these conclusions in game concept art. We need to analyze the visual center and non-visual center areas separately. To achieve this, we first need to distinguish between visually centered and non-visually centered areas.

Game concept design is an indispensable part of game design. Game concept design is to set the tone for the game as a whole and summarize it into some intuitive visual symbolic information. It allows the next steps to unify the style. Also, game concept design involves specific design work, from nothing to creating specific characters, mechanics, scenes, and atmosphere. This study is focused on the game concept art of scenes. This type of design is mainly for the initial atmosphere and tone of the game.

The critical part is the visual center. First of all, the visual center refers to the one on the screen where the viewer will spend time and attention. A good game concept art designer will deliberately use various factors such as color, composition, object, and light to guide the viewer's attention in the pictures.

Therefore, visual interest is an indispensable factor in game art concepts, and it has always been considered during the design process. The visual center helps artists to generate rhythm and make the picture attractive. Previous research has presented saliency methods that can accurately detect primary objects in pictures. Most of these image studies are aimed at photos, pictures in magazines, or object recognition. For example, one research analyzed the color harmony using a combination of the local region to evaluate the photo's aesthetic quality [12]. They mentioned the method of bags of color, which is inspired by the bags of the feature model. These researchers are looking for the relationship between objective pictures and subjective feedback from audiences. The objects of these studies are mostly photos, but there are differences between photos and concept artworks. Concept art includes different styles, especially fantasy, and cartoon, using significantly different colors and object shapes in their artwork. Most of the photos describe the world objectively, but most game concept art is subjective designs. Our research also hopes to fill up the lack of research in digital artworks. Of course, with the advancement of science and technology, many systems are aimed at game concept design. For example, the website called artbreeder. This website can generate game concept art of different styles by inputting two pictures or more than two references. Users can change the attributes that they already set up to get the result in different art styles. These generators undoubtedly provide great help for the conceptual design of the game. For instance, game concept artists can generate an excellent draft for their starting. The Game designer can also quickly produce imaginary sketches to communicate effectively with the concept artist via this system. However, some result also has many limitations, and in many aspects, it does not follow the basic principles of artistic design. So, it is not available for game concept art. They leave some space for improvement in the generated results.

Our approach attempts to detect the visual center by mimicking the way the artist designed the game concept art. First, we divide the image into 16 equal local regions and assign a value to each local region by a hash string. Each hash string is based on the color of the local

region. Then we compare the local regions to find the specific local region with a big color variation. We verify the color variation by HSV color space. Finally, we can get the visual center in game concept art.

2 Related Works

There are many related or similar studies in recent research, including Saliency, image recognition, the study of the visual center in traditional artworks, composition, and aesthetics. The following will briefly overview these researches and describe the relationship and differences between their studies and our study.

2.1 Saliency

Saliency is an analysis of virtual objects in a series of pictures. In earlier studies, Saliency was based on pixel information to make objects stand out from their surroundings and thus get our attention [1]. It uses image processing technology and computer vision algorithms to locate the most significant area in the picture. The salient area refers to the eye-catching area or the more important area in the picture. The objects of this type of research are mainly photos. Saliency's research has many different methods and is constantly optimizing.

There are three main saliency detection algorithms in the saliency module of OpenCV:

Static saliency: This algorithm relies on image features and statistical information to locate the saliency area in the image.

Motion saliency: The target of this saliency method is videos or a series of frames. This algorithm tries to detect and track mainly moving objects as the saliency areas.

Objectness saliency: This type of saliency algorithm calculates proposal areas; these are considered saliency area.

Recent research on saliency has focused on some detail factors, such as the relationship between compositional balance and saliency [8], the research using color and depth attention rules to improve the saliency algorithm from D-RGB image [4], and re-design the programming framework to improve the accuracy of saliency [11].

The main difference between saliency's research and our research is the target samples. In saliency research, authors are looking for a way to detect the particle objects in the photos. It depends on the analysis of edge detection, pixel information, and other methods to confirm the weight function of the object in the photo. Especially in the most extensive dataset MSRA, most samples have a clear main object in the pictures. However, the game concept artwork is almost large scenes, and usually, there is no single object in the picture. Game concept art usually connects all the object in the painting. This work is made with the assumption that concept artists are focusing on an attention area rather than frameable objects. It refers to an area in which there can be one or more objects. The concept art design is a subjective creating process, but it also follows certain objective principles during the design process.

2.2 Object Recognition

Object recognition is part of the vital references for our research. There are some research on object recognition in images. Previous research [6] proposed that 16×16 would be the best segmentation for object recognition in images or photos. However, it was mentioned in the work of Wang et al [15] that the picture segmentation should depend on the target object.

Cutting with a smaller level will save the calculation cost and achieve similar results when the image is elementary, such as apples and oranges. A 4×4 subdivision is sufficient for larger objects to identify the object's content in the picture. Based on these studies, our research will adopt a 4×4 level segmentation to subdivide and analyze the image. In the future, we do have plans to continue to enrich the research on the results of different level segments.

2.3 Visual Center in Traditional Artwork

This study analyzed what factors specifically guide the viewer's gaze in traditional paintings by tracking the participants' gaze. Participants were asked to write down feedback after viewing the digital versions of eight artworks [10]. One example is in Figure 1.

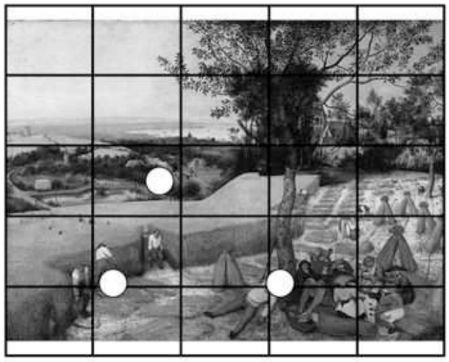

Figure 1: An example of visual interest research [10]

The research attempt to analyze the way that audiences obtain information. They get the result by using the processing model and visual aesthetics principle. According to this model, the perceptual-cognitive processing of artistic stimuli starts with the rapid generation of essential responses, and researchers are looking for the connection between aesthetic methods and audience responses. This research is more closely related to our research than saliency. To a certain extent, we are also looking for the connection between design principles and audiences' attention. Nevertheless, our methods and module are based on the artist's way of thinking. Because the painting principles have already set the ways to connect with the audience. One of the critical processes is guiding the audience's attention. So, the final goals of the two studies are quite similar, but the starting points and methods are entirely different.

2.4 Composition

Whether in saliency, the algorithms used are not assuming that the images have been created using artistic principles. But some research noticed that. The study of Chen et al. [3] includes the basic composition principle and saliency algorithm. It mentions that composition is one factor that can affect people's attention. The composition has always been an unavoidable problem, such as triangle composition, the golden ratio, and the rule of third. Composition is indeed one of the ways to guide the sight, but artists do not only use composition to guide the audience's sight. In the module of our research, we include the influence of some composition factors. At the same time, we combined the other art principles to improve the accuracy and reliability of our result.

3 Method

As mentioned in Sharp's paper[14], Game art is the crystallization of art and technology. The principles of this research are based on basic principles of art painting, Here the author briefly introduces what principle the game concept designer uses to guide the audience's attention:

The size of the color area. A large area of the same color, such as the blue sky, is often not the visual interest. On the contrary, a small area that has intense colors is more attractive to people's eyes. Different ways of composition. For example, the well-known triangular composition, symmetrical composition, and the rule of third. These composition methods are also critical considerations when designing visual interest.

The perception of color can be identified through various color spaces such as HSV, Labspace, CMY, RGB, and more. For our purposes, we opted for HSV due to its similarities with the traditional approach to color in painting. In traditional painting, the initial colors available to us are high purity, high brightness base colors, which are blended in different proportions to produce a variety of colors. In the process of color mixing, contrasting colors are used to reduce the purity and brightness, black and white are used to adjust the brightness and reduce the purity, and the addition of the initial single color increases the purity and brightness. These basic color grading methods respond to HSV's color space approach. Our algorithm is based on the HSV color space approach and aims to analyze the color and visual center in game concept art from a designer's perspective. Also, a number of studies have suggested that HSV is closer to human perception of color. For example, Zhang's[16] research mentions the difference of colorspace in color palette research. They mention the comparison of five color models: RGB, HSV, LAB, YIQ, and Opponent. The RGB color model is the fastest to use, but it is also the least accurate, whereas the HSV color model, which is more accurate[16]. And Cui [5] also mentions that compared to other color spaces, the HSV color space accurately represents the human perception of color by demonstrating a high degree of consistency with the subjective understanding of color by human eyes. This is also the reason why game designers use the HSV panel in Photoshop to select colors when drawing game concept art.

As shown in Figure 2, our method tries to restore the three most important attributes in the game concept design. Based on the above factors, we first use composition to establish the visual interest of the entire picture. Then we use large and small areas of color to enhance the rhythm of the picture and enhance the contrast between the visual center and the non-visual center.

At the same time, the proportion of colors, saturation, lightness, and hue also vary extremely in two different areas. This also means that we can follow this thought to find the visual interest of the picture. Based on these principles, we designed this module to achieve the functions we need. The module consists of four steps to imitate the design principle.

3.1 Image Segmentation

Based on previous studies [15], in image processing algorithms, researchers generally segment the pictures and then analyze and classify each local region to get a conclusion. This method inspires our study. The study [15] mentioned that scenes could be detected within a level of 4×4 segmentation, achieving an accuracy rate of object recognition is 91.4%. 4×4 segmentation can ensure the algorithm's efficiency and keep high accuracy simultaneously. So we adopted a 4×4 segmentation in this experiment.

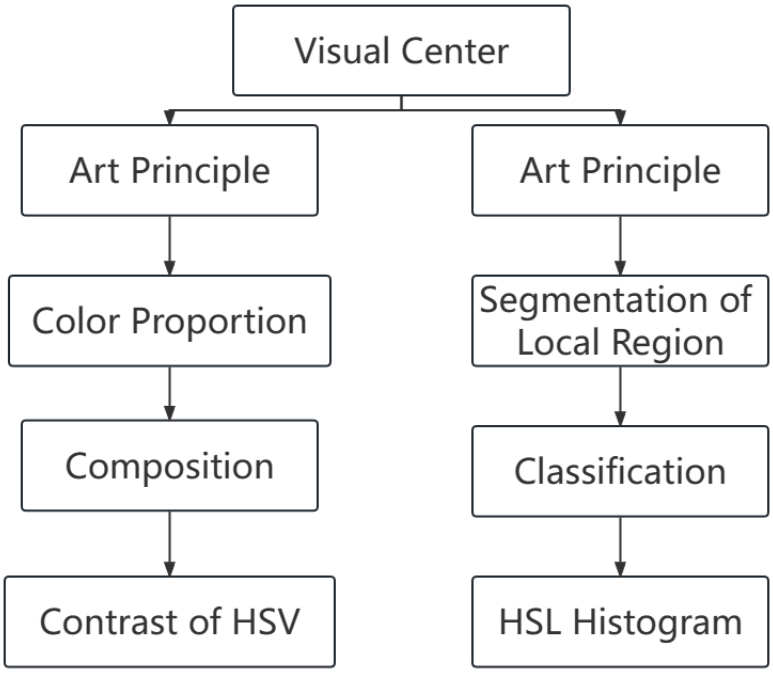

Figure 2: The framework shows how we match design rules and algorithms

Figure 3: A sample of segmentation

3.2 Perceptual Hash Algorithm and Classification

The perceptual hash algorithm is used to test the similarity of two local regions. The function of this algorithm is to generate a "fingerprint" string for each image, and then compare the "fingerprints" of different images [13]. There are three main types of hash algorithms: Ahash. Phash and Dhash. Among them, Dhash has high accuracy and speed. So we are using Dhash, which follows the step below to test the similarity of our samples. The first step reduces the size of the local region to 8 pixels × 8 pixels. This step tries to remove the details and retain essential information such as structure and shading. At the same time, we can abandon differences in local region size. Second, we conduct graying processing in order to optimize the speed of the entire module. Third, comparing the color intensity of the previous pixel and the second pixel makes them true (1) or false (0) depending on the differences in intensity. So, we can get a string with 64 values to describe the information in the local region. In the last, calculate the similarity of the two local regions. We get the 64-bit hash value of each local region by comparing it in pairs. Then record them separately to determine the similarity between regions. Finally, we make sure that each local region has been compared with other

local regions.

Based on the hash value obtained in the previous step, we arrange the hash value, which is the similarity of the local region, from high to low. We start with the two local regions with the highest similarity and put them into the dictionary. If the next local region matches the previous local region, put them in the same dictionary. If the following local region does not match any existing dictionary, create a new dictionary. Next, we conduct this operation until all local regions are classified.

3.3 Detect Visual Interest

Finally, we determine the visual interest areas based on the local region's frequency in the dictionary. In Figure 4, the horizontal coordinate is the value of HSV, from 0 to 255, and the vertical coordinate is the number of pixel points. In Figure 4, the peaks of HSV are concentrated in one point, which means that the colors in this local region are very close. In Figure 5, the peaks of the HSV are scattered in several points, which means that multiple colors is used in this local region.

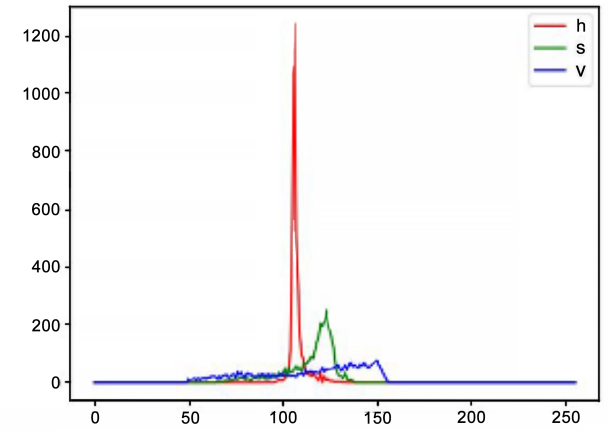

Figure 4: HSV changes in non-visual interest

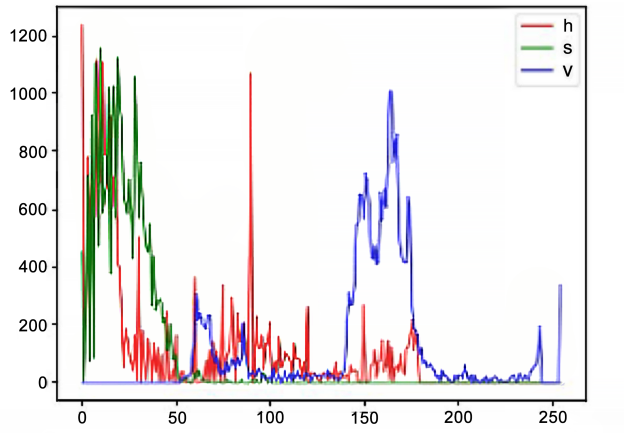

Figure 5: HSV changes in visual interest

The dictionary with the least local region is the target area. If multiple results are detected, based on the rules of composition in art principle, the visual center is generally not designed in the four corners of the picture. Assuming that the image segmentation level is $m \times m$ when multiple results appear, our module will prioritize excluding the dictionaries that contain local regions below: 0, $m-1$, $m(m-1)$, m^2-1. For example: In our study, we adopted a 4×4 sample cut. Therefore, our module will first exclude the four local regions of 0, 3, 12, and 15 at the four corners with multiple results. Then, we use the HSV histogram to confirm the magnitude of the color change in the local region. The non-visual interest area has no big color change, as shown in Figure 4. However, as shown in Figure 5, there will be drastic hue, saturation, and value changes in the local visual interest region. By comparing the two pictures, we can clearly distinguish the region of visual interest and the region of non-visual interest.

4 Evaluation

Saliency's research employs the MSRA sample set, which typically comprises images containing distinct subject objects, serving as a means to assess the efficacy of saliency methodologies through the accuracy of subject-object recognition. However, in the game concept art, images may feature intricate objects, rendering the saliency approach unsuitable for evaluating accuracy. Therefore, an alternative experimental approach was implemented. The methodology for this experiment is as follows:

To compensate for the dearth of available data sets pertaining to game concept art, we procured a total of 100 design drawings from Artstation (`https://www.artstation.com`), a well-regarded community for exchanging exemplary game concept designs. In order to ensure the quality of samples. The samples are also arranged according to all artists' evaluations from highest to lowest. The selected sample should contain story-telling in the concept art. The samples we extracted in this study are based on the game concept art with storytelling and do not specify particular themes and styles.

Figure 6: The example of experiment

We invited 50 individuals who were tasked with identifying the visual center of each image. Utilizing a simple interactive program devised by our team via python, the participants were able to designate the visual center of each image via a click. These responses were then aggregated into a single picture. As shown in Figure 6, the areas in which the dots appear most concentrated, indicate the visual centers. Ultimately, after establishing the visual centers of all 100 images, we compared the visual centers generated by our algorithm, the visual center areas resulting from the two saliency methodologies, and the visual center areas as identified by the participants.

We incorporated two saliency methods into our sample analysis. Despite the unavailability of open-source code for all saliency methods, we identified three methods that had been referenced in related work. However, due to the motion saliency intended usage in video

analysis, the motion saliency method was deemed unsuitable for our study. Consequently, we implemented the objectness saliency method and the static saliency method on our samples, as shown in Figure 7.

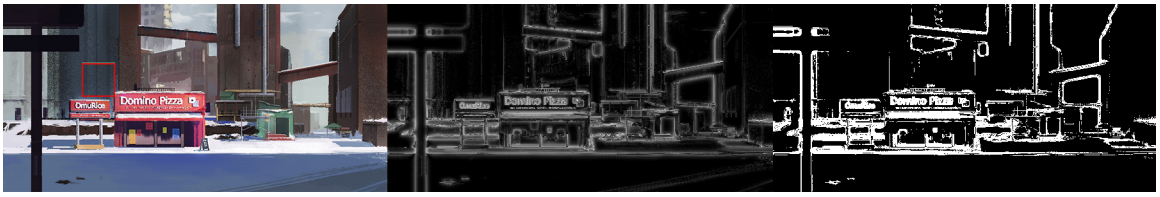

Figure 7: The result from object saliency method and static saliency method

The objectness saliency method directly displays the marked box that encapsulates the visual center area. Meanwhile, the static saliency method demarcates the area of significant pixel variation via white color.

5 Result

A comparison was conducted among the results obtained from the experiment of evaluation. The findings indicate that, out of 100 samples, the objectness saliency method was able to accurately detect 18 samples, while the static saliency method correctly identified 46 samples. In contrast, our methodology achieved a significantly higher level of accuracy, precisely detecting 67 samples.

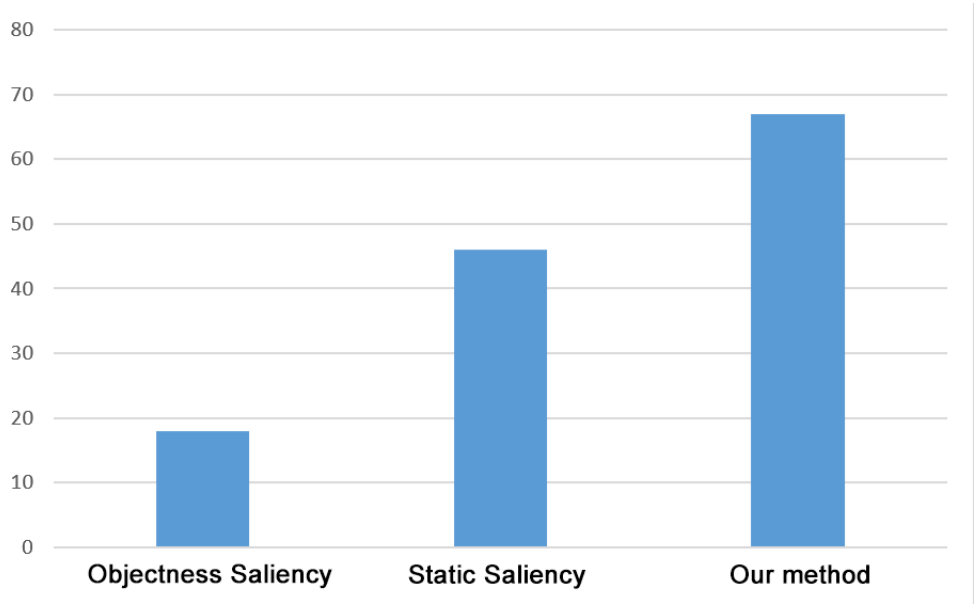

Figure 8: The result of different Saliency methods

Our findings indicate that our method can get better results in detecting the visual center area within game concept art. Because there are many complex elements in the game concept art, the objectness saliency method cannot detect the visual center area well. Static saliency mainly determines the areas with pixel variations. The static saliency method has better performance than the objectness saliency method on game concept art. However, in certain

instances of game concept art, the entire image exhibits pixel variations, leading to the static saliency method marking the entire picture with white, which fails to highlight the visual center area. Our proposed methodology demonstrates an improvement of 45.65% over the static saliency method.

6 Conclusion And Future Work

This study attempts to create a module of visual attention by starting from the artist's creative thinking and targeting game concept art. The module includes the core painting principle, and the module is used to simulate the painting process to detect the visual interest in the picture. At the same time, it paves the way for future research in the game design and painting domain.

However, this research also has certain limitations. First, different degrees of segmentation will lead to different results. In our preliminary test, too high and too low-level segmentation will lead to inaccurate results. Secondly, objects in the artworks, like characters, animals, machinery, and monsters, are inevitably involved in game concept design. They also cause different weights for visual interest detection. If they happen to be segmented during the segmentation stage, they will have relatively low weight during analysis and cause misdetection. Finally, the dark art style has few HSV changes in painting, so it is not easy to detect the result with this method.

In the future research plan, this research first tries to include most of the types and themes from a broad perspective. Future research will gradually subdivide the classification of research samples. We will divide future plans into two primary directions. The first is to continue to optimize the module and try to improve the accuracy and precision. The second is to analyze the differences between the colors of the visual interest area and the non-visual center area. We are continuously looking for the relationship between the color usage pattern and the visual center.

7 Acknowledgment

This work was supported by the Tokyo University of Technology (TUT), the School of media science, and the Hubei Provincial Library.

References

[1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk: *Frequency-tuned salient region detection.* In *2009 IEEE conference on computer vision and pattern recognition*, 1597–1604. IEEE, 2009.

[2] R. Arnheim: *Art and visual perception.* Stockholms Universitet, Institutionen för Konstvetenskap, 1960.

[3] J. Chen, H. Zhao, Y. Han, and X. Cao: *Visual saliency detection based on photographic composition.* In *Proceedings of the Fifth International Conference on Internet Multimedia Computing and Service*, 13–16. 2013.

[4] Y. Cheng, H. Fu, X. Wei, J. Xiao, and X. Cao: *Depth enhanced saliency detection method*. In *Proceedings of international conference on internet multimedia computing and service*, 23–27. 2014.

[5] J. Cui: *Image Style Migration Algorithm based on HSV Color Model*. In *2022 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, 111–114. IEEE, 2022.

[6] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al.: *An image is worth 16x16 words: Transformers for image recognition at scale*, 2020. arXiv: 2010.11929.

[7] A. Jahanian, S. Keshvari, S. Vishwanathan, and J. P. Allebach: *Colors–Messengers of Concepts: Visual Design Mining for Learning Color Semantics*. ACM Trans. Comput. Hum. Interact. **24**(1), 1–39, 2017.

[8] B. Kandemir, Z. Zhou, J. Li, and J. Z. Wang: *Beyond saliency: Assessing visual balance with high-level cues*. In *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, 26–34. 2017.

[9] A. Lambrant, F. L. Luro, and V. Sundstedt: *Avatar preference selection in game design based on color theory*. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, 15–18. 2015.

[10] P. Locher, E. A. Krupinski, C. Mello-Thoms, and C. F. Nodine: *Visual interest in pictorial art during an aesthetic experience*. Spat. Vision **21**(1), 55, 2008.

[11] D. E. Martinez-Rodriguez, M. A. Contreras-Cruz, U. H. Hernandez-Belmonte, S. Bereg, and V. Ayala-Ramirez: *Saliency improvement through genetic programming*. In *Proceedings of the 3rd International Workshop on Interactive and Spatial Computing*, 29–38. 2018.

[12] M. Nishiyama, T. Okabe, I. Sato, and Y. Sato: *Aesthetic quality classification of photographs based on color harmony*. In *CVPR 2011*, 33–40. IEEE, 2011.

[13] X.-m. Niu and Y.-h. Jiao: *An overview of perceptual hashing*. Acta Electron. Sin. **36**(7), 1405–1411, 2008.

[14] J. Sharp: *A curiously short history of game art*. In *Proceedings of the International Conference on the Foundations of Digital Games*, 26–32. 2012.

[15] Y. Wang, R. Huang, S. Song, Z. Huang, and G. Huang: *Not All Images are Worth 16x16 Words: Dynamic Vision Transformers with Adaptive Sequence Length*, 2021. arXiv: 2105.15075.

[16] M. Zhang, G. Qiu, N. Alechina, and S. Atkinson: *A Comparison of Five HSV Color Selection Interfaces for Mobile Painting Search*. In *Human-Computer Interaction–INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings, Part II 15*, 265–273. Springer, 2015.

Received August 16, 2022; final form June 13, 2023.