Proceedings of the 4th International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2026.37324

Evaluation Framework for Emotional Experience in Visual Art Based on Psychometric Feedback

An Zhao

School of Humanities and Social Science, Xi'an Jiaotong University, Xi'an, China
zhaoan02529@outlook.com

Abstract. With the development of digital art exhibitions, intelligent display platforms, and interactive art systems, visual art experiences have gradually shifted from static viewing toward dynamic interaction based on user-state feedback. However, conventional art evaluation methods mainly rely on subjective questionnaires after viewing, which makes it difficult to capture emotional changes during the viewing process and support real-time visual optimization. To address this issue, an adaptive intervention evaluation framework for emotional experiences in visual art based on psychometric feedback is proposed. The framework integrates psychometric data, viewing behavior features, and facial behavior information to construct an emotional-state recognition model. Based on recognition results, visual factors including color, visual complexity, motion rhythm, and narrative prompts are dynamically adjusted. Sixty participants were recruited for experiments comparing static presentation and adaptive presentation conditions. The results show that the multimodal random forest model achieved superior emotional-state recognition performance, with an 85.00% accuracy in cognitive-load classification and a Macro-F1 score of 0.843. The adaptive presentation condition improved emotional valence, aesthetic engagement, immersion, and artwork understanding. This study provides a computational approach for emotional computing and personalized experience optimization in digital art environments, promoting the integration of artificial intelligence and visual art evaluation.

Keywords: Visual art, emotional experience, psychometric feedback, affective computing, adaptive intervention

1. Introduction

The emergence of digital technology has fostered a shift toward intelligent and interactive visual art presentations. Contemporary digital exhibitions and virtual art websites are using computational tools to improve audience involvement. However, emotional experiences in visual arts are very dynamic and personalized due to the fact that people's reactions depend both on artwork structure and visuals and people's psychological states. Traditional questionnaires used after the exhibition cannot properly capture the dynamic emotional changes occurring throughout the process [1]. The emergence of affective computing allows adopting new methods of evaluating artistic experiences via the combination of psychometric measures, behavior and physiological data. Computational systems can detect the emotional state of users and provide decision-making in intelligent interactive

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.37324

systems [2]. Contemporary researches have already studied the use of gaze pattern, facial expressions and physiological data for evaluating aesthetic experience; however, single source measures are sensitive to environment conditions and individual differences resulting in poor recognition and stability [3]. Thus, a comprehensive system that will allow for the integration of psychometric measures, multimodal recognition and adaptation is necessary in order to translate the subjective emotional experience into computational measure.

2. Literature review

2.1. Psychometric measurement of visual-art emotion

The investigation of the emotions in reaction to visual art usually adopts the models of valence and arousal to explain people's psychological responses to artworks. The tools such as Self-Assessment Manikin, Positive and Negative Affect Schedule, and aesthetic experience scales have been extensively used to assess pleasure, tension, and immersion, thus providing standard approaches for the assessment of artistic experience [4]. The past research also shows that color saturation, composition complexity, visual rhythm, and semantic information affect the emotions by way of perceptual processes and cognitive evaluations, which cause the generation of differential psychological feedback from different artistic elements [5]. Nevertheless, the traditional psychometric measures mainly concentrate on the overall assessment following the viewing session and cannot offer any explanation for the temporal change of emotion or the contribution of particular visual elements.

2.2. Computational recognition of emotional states

With the advent of affective computing, the study of the experience of visual arts has moved from subjective to computational analysis. Previous researches have created emotional recognition models based on eye movement paths, fixation time, facial gestures, physiological measures, and user behavior [6]. Multimodal fusion uses complementing data from various sources and increases the stability of the prediction in complicated settings, whereas machine learning finds associations between the behavior and psychological condition [7]. However, there still remain problems related to noise, individual differences, and lack of real-time feedback in emotional recognition in the context of visual arts. Computational measures are unable to describe the subjective experience of aesthetics. Psychometric measures need to be used as a reference point and combined with the behavioral measures to create an effective emotional recognition system.

2.3. Adaptive intervention in artistic experiences

The evolution of adaptive art systems is from static content recommendation to dynamic experience management in relation to the user's emotional state through the modification of visual stimuli and interaction approaches. The goal of such an approach is to make the viewer more involved and experience immersion [8]. Previous research shows that brightness, color, movement speed, visual complexity, and narrative prompts impact the emotions of people, which means that adjustment approaches may help create better experiences tailored to each user [9]. Unfortunately, the majority of the existing approaches focus on matching preferences but ignore the evaluation of changes in emotions post-intervention and preserving the artistic experience. A good adaptive approach should

DOI: 10.54254/2755-2721/2026.37324

include selecting appropriate modifications in relation to the current emotional state of the user and evaluating the outcomes of the intervention through experiments.

3. Experimental methods

3.1. Participants, materials, and data collection

The experiment recruited 60 participants as the research sample. The selection criteria included an age range of 18 to 35 years, normal or corrected-to-normal vision, and previous experience with digital artworks. Participants with visual perception disorders, neurological diseases, or inability to complete the full experimental procedure were excluded. The experimental materials consisted of 12 digital visual artworks. Three experts with backgrounds in art and design evaluated these artworks based on color representation, compositional complexity, visual rhythm, and information density. The artworks were then categorized into three complexity levels: low, medium, and high. Before the experiment, participants' demographic information, digital art exposure experience, artistic viewing frequency, and initial emotional states were collected. During the formal experiment, each artwork was displayed for 90 seconds. Self-Assessment Manikin (SAM) ratings of valence and arousal were collected at 30-second and 60-second intervals. Eye-tracking data, facial action data, and interaction behaviors were recorded simultaneously. Eye-tracking features included fixation duration, the number of visits to interest regions, and gaze dispersion. Facial data included the intensity of facial action units. Interaction data included mouse movement speed, pause frequency, and operation trajectories. All psychological, behavioral, and visual feedback data were synchronized through timestamps to construct a multimodal dataset for subsequent emotional-state recognition.

3.2. Emotional-state recognition model

Based on the multimodal data collected in Section 3.1, an emotional-state recognition model for visual art experience was developed. The model used psychometric measurements, eye-tracking features, facial action features, and interaction features as input variables. The outputs included valence state, arousal level, and cognitive-load category [10]. First, different modal features were standardized to reduce deviations caused by individual differences. Then, a weighted fusion strategy was applied to construct a unified emotional feature representation, as shown in Formula 1:

$$F=\sum_{i=1}^{N} w_i F_i \tag{1}$$

Where F represents the fused multimodal feature vector, F_i represents the i_{th} input feature, including psychometric features, eye-tracking features, facial action features, and interaction features, w_i represents the weight of each modality, and N represents the number of feature types.

The fused features were then input into a random forest classification model. Multiple decision trees were combined to identify the current emotional state, and the final category was determined through majority voting, as shown in Formula 2:

$$\hat{y}=\operatorname{argmax}_k \sum_{m=1}^{M} I\left(h_m(F)=k\right) \tag{2}$$

DOI: 10.54254/2755-2721/2026.37324

Where $\hat{y}$ represents the predicted emotional category, M represents the number of decision trees, $h_m(F)$ represents the classification result of the mth decision tree, and I represents the indicator function. Five-fold cross-validation was applied during model training. Accuracy, Macro-F1 score, mean squared error, and mean absolute error were used to evaluate model performance. The predicted emotional states were further used as the basis for adaptive intervention, connecting emotional recognition with visual adjustment.

3.3. Adaptive intervention and evaluation procedure

Based on the emotional-state evaluation in Section 3.2, the system developed rule-based adaptive interventions to artworks in the visual modality. If the model indicated low valence, the system decreased visual stimulation through lowering the color saturation and simplifying the local visual complexity in order to provide a more comfortable viewing experience. In case of a high arousal state, the system decreased the dynamic change to avoid excessive visual stimulation [11]. In case of a high cognitive load state, the system provided brief narrative hints to facilitate comprehension of the artwork. Only one visual dimension was intervened at once with three adjustment levels to avoid influencing the original artistic experience with multiple adjustments. After 30 seconds of intervention, the evaluation of the viewer's feedback was repeated. This evaluation included valence, aesthetic experience, immersion, artwork comprehension, cognitive load, and intention to continue viewing. The within-subject crossover design was used, when participants underwent both static and adaptive presentations. Statistical paired comparison between the two presentation modes was performed.

4. Results

4.1. Emotional-state recognition performance

The emotional-state recognition model established in Section 3.2 was evaluated using the test dataset. The results are shown in Table 1. The psychometric-feature model achieved basic emotional-state recognition but showed relatively high prediction errors due to the absence of continuous behavioral information. Its cognitive-load classification accuracy was only 72.50%, with a Macro-F1 score of 0.691. After incorporating eye-tracking and interaction features, model performance improved significantly. The cognitive-load classification accuracy increased to 80.83%, and the Macro-F1 score reached 0.785, indicating that behavioral information provided additional evidence for subjective emotional feedback. After further integrating psychometric measurements, eye-tracking features, facial actions, and interaction behaviors, the multimodal random forest model achieved the best performance. The mean squared error of valence prediction decreased to 0.512, the mean squared error of arousal prediction decreased to 0.586, and the cognitive-load classification accuracy reached 85.00% with a Macro-F1 score of 0.843. These results demonstrate that multimodal feature fusion effectively reduces emotional recognition errors and improves state estimation stability in visual art experiences.

DOI: 10.54254/2755-2721/2026.37324

Table 1. Performance comparison of emotional-state recognition models

Model	Valence Prediction MSE	Arousal Prediction MSE	Cognitive Load Accuracy/%	Macro-F1
Psychometric Feature Model	0.842	0.913	72.50	0.691
Psychometric + Behavioral Feature Model	0.621	0.724	80.83	0.785
Multimodal Random Forest Model	0.512	0.586	85.00	0.843

4.2. Adaptive intervention effectiveness

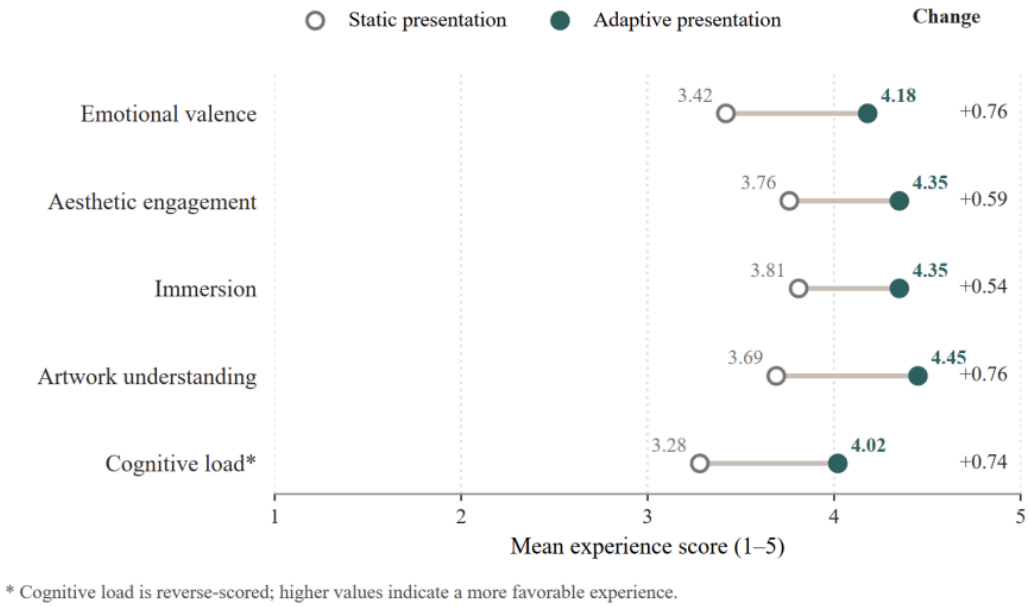


Figure 1. Comparison of visual art experience dimensions under static and adaptive presentation conditions

According to the method of the adaptive intervention introduced in Section 3.3, the differences between the experience of visual art through static presentation and adaptive presentation have been investigated. The findings have been represented in Figure 1. The adaptive presentation mode enhanced the general quality of visual art experience. As regards the five evaluation criteria, there was a significant improvement in the scores of emotional valence, aesthetic engagement, immersion, artwork understanding, and cognitive load, through reverse normalization. Emotional valence improved from 3.42 to 4.18; aesthetic engagement from 3.76 to 4.35; immersion from 3.81 to 4.35; artwork understanding from 3.69 to 4.45; and cognitive load from 3.28 to 4.02 after reverse scoring. It should be noted that the experience structure of adaptive presentation was relatively more balanced than that of static presentation in terms of all evaluation dimensions. Artwork understanding and immersion experienced especially notable improvements, which suggest that emotional real-time feedback helps enhance emotional and cognitive interactions with artworks. Figure 1 depicts the comparison of the experience dimensions of visual art through two presentation modes.

DOI: 10.54254/2755-2721/2026.37324

As shown in Table 2, paired-samples t-tests confirmed statistically significant improvements across all five evaluation dimensions under adaptive presentation, with all p-values below 0.001. Effect sizes ranged from medium to large, indicating that the observed differences were practically meaningful. Artwork understanding produced the largest standardized effect, followed by emotional valence. Aesthetic engagement, immersion, and cognitive-load regulation also improved consistently. These findings demonstrate that adaptive presentation strengthened both affective responses and cognitive engagement during visual art viewing.

Table 2. Paired-samples t-test results for adaptive presentation effects

Evaluation Dimension	Mean Difference	95% CI	t	p	Cohen's d_z
Emotional Valence	0.76	[0.53, 0.99]	6.51	< 0.001	0.84
Aesthetic Engagement	0.59	[0.38, 0.80]	5.58	< 0.001	0.72
Immersion	0.54	[0.34, 0.75]	5.27	< 0.001	0.68
Artwork Understanding	0.76	[0.54, 0.98]	6.82	< 0.001	0.88
Cognitive-Load Regulation	0.74	[0.49, 1.00]	5.81	< 0.001	0.75

5. Discussion

Based on the above experiment results, we can see that multimodal feature fusion successfully enhances the performance of recognizing emotions in art experiences. By means of the complementary nature of psychometric feedback and viewing behaviors, the recognition of the complicated emotion transitions becomes more accurate. In addition, by means of the better performance of cognitive load classification accuracy, it shows that continuous behavioral information plays a very important role in the explanation of visual information processing. Moreover, based on the adaptive intervention results, it demonstrates that through the adjustment of the visual art, emotional response can be improved and enhance the understanding and immersion of art works. The above conclusions also indicate that the optimization of artistic experience needs to take affective response and cognition into consideration. Unlike the traditional preference-based recommendation system, the novel framework combines psychometric feedback into the visual adjustment and forms a closed loop from emotional recognition to experience optimization.

6. Conclusion

Visual art experience is characterized by high levels of subjectivity and variation, and thus poses limitations on the use of conventional evaluation approaches in the intelligent artistic setting. An approach for evaluation of adaptive intervention based on the feedback through psychological measurement and visual behavior interaction has been developed. The evaluation approach incorporates an entire process of recognition of emotional states, visual adaptation and evaluation of experience. Results from experimental verification prove the validity of the multi-modal recognition model and the fact that adaptive presentation enhances emotional experience, aesthetic involvement and understanding of artworks. The framework serves as an extendable methodology for digital exhibitions, intelligent art platforms and human-computer interactive arts.

DOI: 10.54254/2755-2721/2026.37324

References

[1] Schubert, E. (2024). The aesthetic emotion lexicon: A literature review of emotion words used by researchers to describe aesthetic experiences. *Empirical Studies of the Arts, 42*(1), 3–37.
[2] Nummenmaa, L., & Hari, R. (2023). Bodily feelings and aesthetic experience of art. *Cognition and Emotion, 37*(3), 515–528.
[3] Stamkou, E., et al. (2024). Emotional palette: A computational mapping of aesthetic experiences evoked by visual art. *Scientific Reports, 14*(1), 19932.
[4] Darda, K. M., & Chatterjee, A. (2023). The impact of contextual information on aesthetic engagement of artworks. *Scientific Reports, 13*(1), 4273.
[5] Salgues, S., et al. (2024). Self-reference and emotional reaction drive aesthetic judgment. *Scientific Reports, 14*(1), 19699.
[6] Al-Kire, R. L., et al. (2023). Original photographic art induces self-transcendent emotions. *Psychology of Aesthetics, Creativity, and the Arts*.
[7] Castellotti, S., et al. (2023). Psychophysiological and behavioral responses to descriptive labels in modern art museums. *PLoS One, 18*(5), e0284149.
[8] Casteau, S., & Smith, D. T. (2024). How does contextual information affect aesthetic appreciation and gaze behavior in figurative and abstract artwork? *Journal of Vision, 24*(12), 8–8.
[9] Kühnapfel, C., et al. (2024). How do we move in front of art? How does this relate to art experience? Linking movement, eye tracking, emotion, and evaluations in a gallery-like setting. *Empirical Studies of the Arts, 42*(1), 86–146.
[10] Castellotti, S., D'Agostino, O., & Del Viva, M. M. (2024). Effectiveness of labels in digital art experience: Psychophysiological and behavioral evidence. *Frontiers in Psychology, 15*, 1342667.
[11] Rodriguez-Boerwinkle, R. M., & Silvia, P. J. (2024). Visiting virtual museums: How personality and art-related individual differences shape visitor behavior in an online virtual gallery. *Empirical Studies of the Arts, 42*(2), 439–468.