Human Factors in Artificial Intelligence and Social Computing, Vol. 203, 2026, 128–140
https://doi.org/10.54941/ahfe1007325

Evaluating Public Art in Commercial Complexes: A Dual-Channel Emotion Recognition Framework Fusing Facial Micro-Expressions and Semantic Analysis

Jianing Hu[1], Xiayun He[2], and Lianmei Dong[1]

[1]School of Architecture and Applied Arts, Guangzhou Academy of Fine Arts, No.257, Changgang East Road, Guangzhou 510260, China
[2]Innovation School of Greater Bay Area, Guangzhou Academy of Fine Arts, Panyu District, Guangzhou 510006, China

ABSTRACT

The development of public art in commercial spaces increasingly emphasizes emotional experience, but traditional evaluation methods struggle to objectively capture the real-time dynamics and inherent complexity of user emotions. This study proposes a dual-channel emotion recognition framework integrating facial micro-expression analysis and semantic understanding. We define seven emotional categories and construct a dataset of facial expressions and speech data. Feature-level fusion between DeepFace-processed facial data and DeepSeek-R1-analyzed semantic data generates a unified recognition model. An empirical study analyzing five types of public art media with 2,100 speech transcripts and 1,200 facial images reveals that dual-channel fusion achieves 93.2% accuracy, significantly outperforming single-modal approaches. Interactive art generates the strongest emotional stimulation, platform-based art enables efficient communication through social attributes, and functional art provides unique emotional buffering. This research offers a quantifiable methodology for commercial public art evaluation.

Keywords: Public art, Emotion recognition, Micro-expression recognition, User environmental experience, Commercial architecture complex

INTRODUCTION

Contemporary commercial architecture is undergoing a fundamental transformation—from utilitarian containers of retail activity to curated environments mediating multi-sensory experiences. This shift aligns with the growing predominance of emotional consumption, where nearly 30% of young consumers prioritize affective engagement and self-referential meaning in commercial encounters (China Youth Consumption Research Center, 2024). Public art has consequently emerged as a critical medium for shaping atmospheric quality and eliciting targeted emotional responses, accounting for over 60% of spatial programming in leading commercial complexes (China Commercial Real Estate Association, 2023).

Received February 4, 2026; Revised April 23, 2026; Accepted May 13, 2026; Available online July 20, 2026

While sensory marketing theory establishes that artistic media elicit stronger emotional fluctuations than purely functional services (Krishna, 2012), the field lacks a coherent theoretical framework explaining how different artistic interventions trigger specific emotional mechanisms. Existing research remains predominantly output-focused, failing to deconstruct the internal "emotion-behavior" black box, which results in a disconnect between aesthetic investment and measurable consumer experience. This gap manifests in three critical areas:

1. inadequate micro-level, psychologically-grounded models explaining the cognitive-affective pathways through which artistic elements translate to consumer emotions;
2. limited integration of interdisciplinary findings, particularly from cognitive neuroscience, which reveals automatic, non-conscious emotional characteristics;
3. an over-reliance on correlating stimuli with behavioral outcomes, neglecting the dynamic process of emotional transition and regulation within individuals.

The methodological landscape for assessing emotional responses in commercial spaces presents substantial limitations. Traditional methods like questionnaires and interviews rely on subjective self-reporting, which cannot capture the immediacy and complexity of emotional states and is susceptible to recall bias (Tracy & Robins, 2008). While AI-assisted tools (e.g., computer vision) offer more objective analysis, uni-modal approaches are constrained by environmental interference and inherent modality limitations, failing to capture the full emotional spectrum. A primary challenge remains the lack of robust methods for multi-modal data fusion to create coherent emotional profiles, coupled with the poor ecological validity of many advanced methods when applied to real-world, in-situ commercial environments.

To address these gaps, this study develops and validates a dual-channel emotion recognition framework that integrates facial micro-expression analysis with speech semantic understanding. It aims to:

1. technically verify a robust methodology for quantifying public emotional responses to artistic interventions;
2. systematically profile and compare the emotional efficacy of five prevalent typologies of commercial public art;
3. derive evidence-based design strategies for aligning emotional impact with broader architectural and commercial objectives.

By translating subjective aesthetic experience into quantifiable, actionable data, this research seeks to establish a new paradigm for emotion-aware evaluation and design in commercial architecture.

THEORETICAL FRAMEWORK: FROM SENSORY MARKETING TO A DUAL-CHANNEL EMOTION RECOGNITION MODEL

Theoretical Evolution: From Sensory Marketing to Multi-Modal Emotion Evaluation

The evaluation of user satisfaction in commercial environments is undergoing a theoretical shift from a uni-dimensional focus on functionality to a multi-sensory, emotion-centric paradigm. Grounded in sensory marketing theory, which posits that multi-sensory stimuli significantly influence consumer perception and behavior (Zhong, Wang, and Yang 2016; Krishna, 2012; Hultén, 2011), this evolution recognizes artistic media as pivotal "transmitters" of spatial emotion. Such media utilize visual and auditory cues—such as color palettes, materiality, dynamic lighting, and ambient soundscapes—to evoke subconscious emotional resonance and shape behavioral intentions (Lee & Lee, 2014; Bruner, 1990; Chen, 2024; Lehrl et al., 2007; Zhang et al., 2017; Spangenberg et al., 2006).

To empirically identify which sensory channels dominate public perception in commercial art contexts, this study conducted a web scraping and semantic analysis of 663 social media posts. The analysis revealed that visual (48.2%) and auditory (38.9%) stimuli collectively account for 87.1% of sensory influences on consumer emotion (see Figure 1). This empirical finding substantiates the prioritization of visual and auditory modalities in constructing the emotion recognition framework for this study. Therefore, this study focuses specifically on these two dominant sensory channels to further investigate their impact on consumer emotional perception.

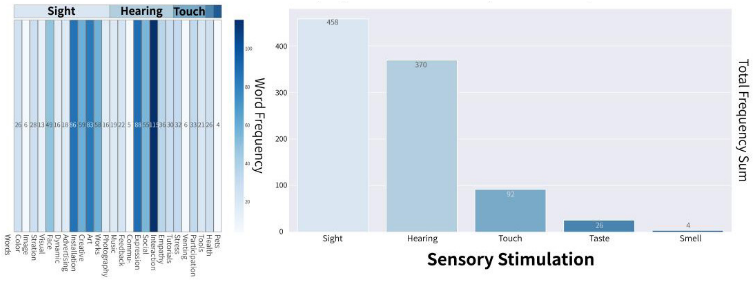

Figure 1: Distribution and comparison of online word frequency under different sensory stimuli in commercial spaces.

Technological Pathways in Emotion Recognition: From Uni-Modal to Bimodal Fusion

Emotion recognition technology has advanced considerably, with deep learning enabling accurate affect detection through various modalities. Visual emotion recognition, primarily via facial expression analysis using convolutional neural networks (CNNs), achieves high accuracy by decoding emotion-related facial landmarks (Naidoo et al., 2022; Kwon & Yu, 2024). In parallel, auditory channel recognition—encompassing speech emotion recognition (Sahoo et al., 2021; Wu et al., 2021; Song et al., 2020) and textual semantic analysis (Wang et al., 2020; Mahima et al., 2021)—interprets affective states from vocal features and linguistic content.

However, uni-modal approaches exhibit inherent limitations: facial analysis can be affected by environmental noise or expressive suppression, while semantic understanding struggles with linguistic ambiguity, cultural nuance, or intentional emotional masking (Wang et al., 2020; Sahoo et al., 2021). A comparative analysis of these technological paths highlights their complementary strengths: micro-expression recognition captures rapid, subconscious affective leakage with high objectivity (e.g., DeepFace achieves 97.53% accuracy), whereas semantic understanding decodes contextualized, self-reported emotional content (e.g., DeepSeek-R1 attains 89.7% fine-grained classification accuracy). This complementarity motivates a bimodal fusion strategy to overcome the constraints of single-modality systems.

Theoretical Evaluation Framework and Research Workflow Design of Dual-Channel

Integrating the theoretical underpinnings of sensory marketing with insights from comparative emotion recognition technologies, this study proposes an "Audio-Visual Bimodal → Emotion → Behavioral Intention" theoretical model (see Figure 2). The framework posits that synchronized capture and computational fusion of facial micro-expressions (visual) and speech-derived semantics (auditory) yield a more robust and ecologically valid emotional assessment. The resulting emotion indices serve as predictors for behavioral intentions—such as purchase likelihood, dwell time, and overall satisfaction—enabling data-driven optimization of public art interventions in commercial settings.

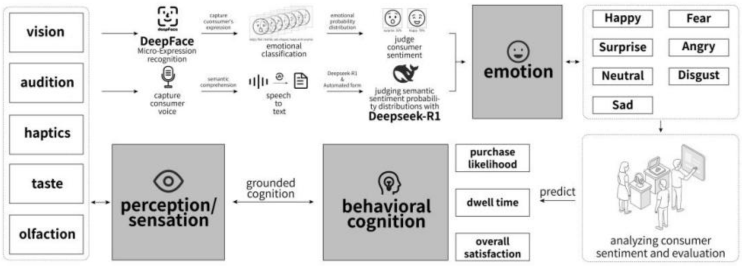

Figure 2: "Audio-Visual Bimodal – Emotion – Behavioral Intention" model.

RESEARCH DESIGN AND COMPUTATIONAL METHODOLOGY: A DUAL-CHANNEL FRAMEWORK FOR EMOTIONAL EFFICACY EVALUATION

To empirically validate the proposed "Audio-Visual Bimodal – Emotion – Behavioral Intention" model, this study developed and executed a three-phase research protocol and the overall technical architecture is depicted in the diagram below (Figure 3):

1. a real-time and on-site-based technical validation of the dual-channel emotion recognition testing;
2. an in-situ field data collection within a representative commercial complex;
3. the application of a computational model for affective analysis.

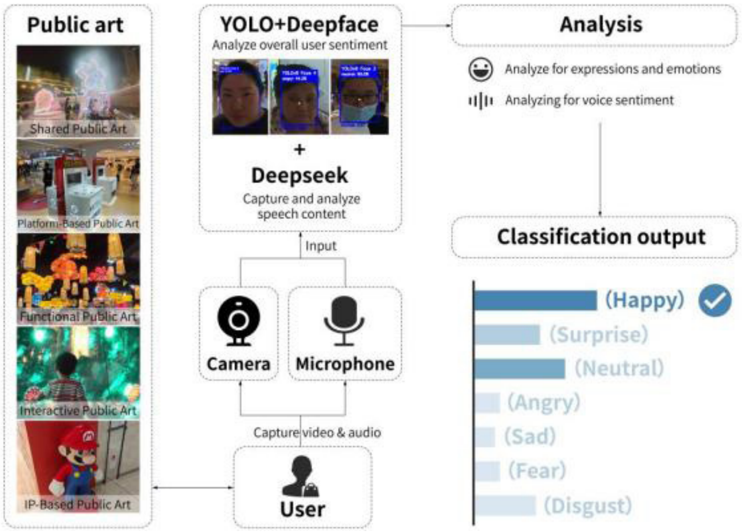

Figure 3: Dual-channel emotion recognition technology architecture.

Technical Validation and Pre-Testing of the Dual-Channel Framework

Prior to field deployment, the integrated performance of the visual and auditory emotion recognition channels was validated in a controlled setting. This pre-testing phase confirmed the operational stability and accuracy of the core computational models.

1. Visual Channel Processing: The visual pipeline employed the DeepFace framework, a hybrid model integrating architectures such as VGG-Face and FaceNet, for micro-expression recognition. It utilized Deep Convolutional Neural Networks (CNNs) to analyze facial muscle movements. The YOLO object detection model was applied for real-time face detection and tracking within the commercial environment. This setup enabled the translation of facial variations into discrete emotional states (e.g., joy, surprise, anger) and the subsequent generation of spatiotemporal emotion heat-maps. Laboratory validation of the DeepFace model demonstrated a micro-expression recognition accuracy of 97.53%, confirming its capability to capture subtle emotional shifts (Figure 4).
2. Auditory Channel Processing: The auditory pipeline was built upon the DeepSeek-R1 model for semantic emotion recognition. This model leverages a Multi-task Mixture of Experts (MoE) architecture and a consumer psychology semantic corpus to perform fine-grained sentiment analysis on textual data. An automated robot workflow was implemented to simulate real-time data acquisition which could identify different person's speaking, handling speech-to-text transcription and its immediate conversion into emotion scores, categorized as positive or negative sentiment. This pre-test ensured the reliability and processing speed of the auditory channel.

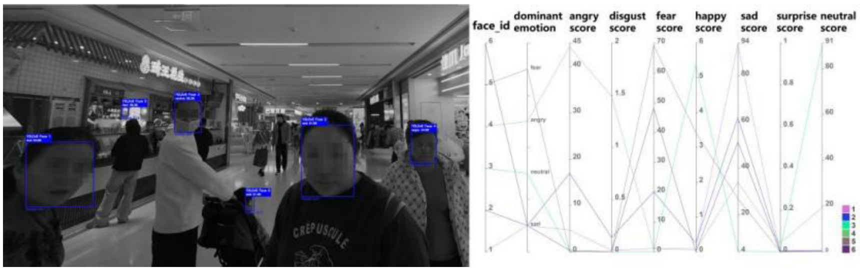

Figure 4: Real-time emotion detection data.

In-Situ Data Collection: Site Selection and Protocol

The field study was conducted at the Wushang Dream Times Mall in Wuhan, China, selected for its status as one of the world's largest single-structure commercial complexes (gross floor area: 800,000 m^2; average daily footfall: 120,000 visitors). Its high density and diversity of public art interventions—featuring 21 permanent installations, including a Team Lab digital art zone—provided an ideal context for a comparative study of emotional efficacy.

Artistic media in commercial public spaces represent not merely the display of public artworks with physical attributes, but rather a systematically constructed intervention model that integrates platform-based, interactive, functional, and IP-driven strategies to address diverse users' sensory stimulation and functional demands. Based on prevailing curatorial strategies, public art within the mall was categorized into five distinct typologies to facilitate systematic analysis:

1. Shared Public Art: Static visual artworks (e.g., sculpture, painting, digital projection) integrated into public areas for cultural enrichment and aesthetic ambiance.
2. Platform-based Public Art: Artistically curated events and programs (e.g., exhibitions, salons, workshops) designed to foster social exchange and cultural engagement.
3. Functional Public Art: Artistically treated utilitarian facilities (e.g., wayfinding systems, landscape furniture, ambient lighting) that merge aesthetic appeal with practical use to elevate the overall spatial quality.
4. Interactive Public Art: Installations employing mechanical, digital, or multimedia interfaces to necessitate active physical or behavioral participation from users. This category is to enhance immersive experiential engagement.
5. IP-based Public Art: Thematic decorations and installations derived from communicable intellectual properties, aimed at attracting specific demographic groups and generating cultural resonance.

Data acquisition was conducted during varied times across weekdays and weekends to ensure a representative sample of visitor interactions. Using action cameras positioned strategically around the five types of public art zones, non-intrusive video and audio data were collected. The final dataset comprised 2,100 valid speech transcripts and 1,200 non-intrusive facial images (Figure 5). All data streams were synchronized with timestamps accurate to within 10 milliseconds using the Network Time Protocol (NTP). Strict adherence to ethical standards was maintained: in compliance with China's Personal Information Protection Law, all collected visual data underwent facial blurring post-processing, and all data were stored in an encrypted format to protect individual privacy.

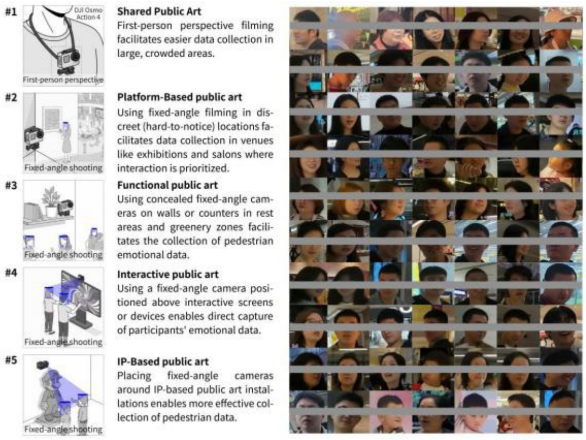

Figure 5: Data acquisition methods and equipment placement.

Evaluation Method: Variable Definition and Affective Computation of Dual-Channel Model

The collected bimodal data were processed through a unified computational model to evaluate the emotional efficacy of public art. Emotions were classified into three overarching categories: Positive (happiness, surprise), Neutral, and Negative (sadness, fear, anger, disgust).

The core of the analysis relies on a standardized sentiment analysis formula (Liu, 2015) to compute a composite emotional score (Sen), which quantifies emotional polarity on a normalized scale from −1 to +1. The formula (formula (1)) is defined as:

$$Sen = \frac{Pos - Neg}{Pos + Neg + Neu} \tag{1}$$

For the visual channel, emotional intensity was quantified as follows: Positive Intensity (Pos) was calculated as the sum of the probability scores for 'happiness' and 'surprise'; Negative Intensity (Neg) as the sum for 'anger', 'sadness', 'fear', and 'disgust'; and Neutral Intensity (Neu) directly as the probability score for 'neutral'.

For the auditory channel, the semantic emotion value—an internal metric of sentiment polarity derived from transcribed speech—was used directly.

This multi-modal scoring framework enabled the direct comparison and subsequent fusion of affective data from facial micro-expressions and speech semantics, yielding a robust, quantifiable measure of emotional efficacy for each art typology (Figure 6).

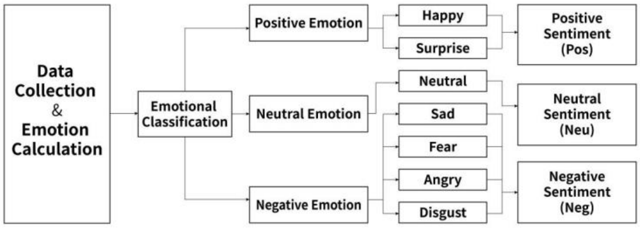

Figure 6: Real-time emotion detection data.

RESULTS AND DISCUSSION: EMOTIONAL EFFICACY AND DESIGN IMPLICATIONS

This chapter conducts a systematic analysis of the emotional efficacy of five categories of public art in commercial spaces, utilizing a dual-channel emotion recognition framework applied to 2,100 speech transcripts and 1,200 facial images collected at Wushang Dream Times Mall in Wuhan. Data acquisition spanned both weekdays and weekends across varying time periods to ensure the representative of the sample.

Comparative Analysis of Emotional Efficacy Across Art Types

By aggregating and analyzing emotional data from five categories of public art spaces, the overall emotional efficacy performance of different art types was determined (Table 1).

Table 1: Comparative analysis of emotional efficacy across different commercial public art media.

Commercial Public Art Category	Sample Size (N)	Positive Emotion Percentage (%)	Neutral Emotion Percentage (%)	Negative Emotion Percentage (%)	Comprehensive Emotion Score (Sen)	Average Dwell Time (Min)
Shared Public Art	412	52.3	35.4	12.3	0.41	8.5
Platform-Based Public Art	385	48.6	38.2	13.2	0.36	15.2
Functional Public Art	398	45.2	42.5	12.3	0.33	6.3
Interactive Public Art	476	68.4	24.3	7.3	0.62	18.7
IP-Based Public Art	429	61.7	29.6	8.7	0.54	12.4

Data analysis reveals a hierarchical distribution in the emotional efficacy of the five public art types (Table 2). Interactive art significantly outperforms others, with a comprehensive emotion score (Sen) of 0.62 and a positive emotion proportion of 68.4%, demonstrating the potency of participatory experiences. Platform-based art follows (Sen = 0.54, 61.7% positive), its social attributes reinforcing emotional transmission. Its lower neutral emotion proportion (29.6%) indicates stronger emotional polarization.

A non-linear relationship exists between dwell time and emotional intensity. Interactive art, despite highest positive affect, shows a moderate average dwell time (18.7 min), aligning with the peak-end rule where emotional "peaks" define the experience more than duration. Functional art exhibits a distinct pattern, with a significantly higher neutral emotion proportion (42.5%), acting as an emotional buffer or regulator in stimulating environments.

Negative emotions remain relatively low across all types (7.3%–13.2%). Interactive art has the lowest proportion (7.3%), indicating effective negative experience mitigation, while IP-based art has the highest (13.2%), potentially due to cognitive bias or fatigue. A decision matrix using emotional efficacy (Sen) and resource investment (dwell time) classifies interactive art as high efficacy–high investment (ideal for flagship spaces), platform-based art as high efficacy–medium investment (optimal cost-effectiveness), shared art as medium efficacy–low investment (basic atmosphere), and functional art as a targeted functional complement.

Furthermore, validation confirms the synergistic effect of dual-channel fusion, achieving a recognition accuracy of 93.2%, superior to visual-only (82.4%) or auditory-only (85.7%) modalities. This fusion notably reduced the misjudgment rate for complex emotions from 15.95% to 6.8%.

Table 2: Accuracy comparison between single-modal and dual-modal emotion recognition.

Recognition Mode	Recognition Accuracy (%)	Emotion Consistency (%)	Misclassification Rate (%)	F1-Score
Visual Channel Only	82.4	78.6	17.6	0.81
Auditory Channel Only	85.7	81.3	14.3	0.84
Dual-Channel Fusion	93.2	91.5	6.8	0.92

To more intuitively demonstrate the classification performance of each model, the following presents confusion matrices based on single facial expressions, single speech recognition, and dual-channel fusion (Figure 7). These charts clearly illustrate the performance of different models in emotion classification, particularly highlighting the dual-channel fusion model's significant advantages in accuracy and complex emotion recognition.

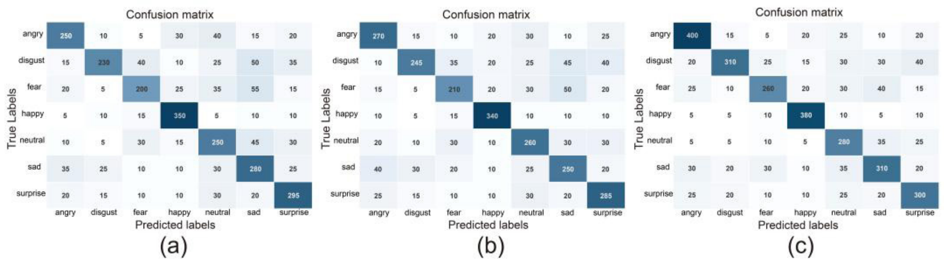

Figure 7: Combined confusion matrices for the emotion recognition framework ((a) Visual unimodal channel; (b) Auditory unimodal channel; (c) Dual-channel fusion).

Synergistic Effect of Dual-Channel Data Fusion

Through the seven basic emotions (happy, surprise, neutral, sadness, anger, disgust, fear) identified by the DeepFace model, we analyzed the specific emotional characteristics elicited by different types of public art (Figure 8), further investigating the differential impacts of various commercial public art forms on consumers' emotional composition.

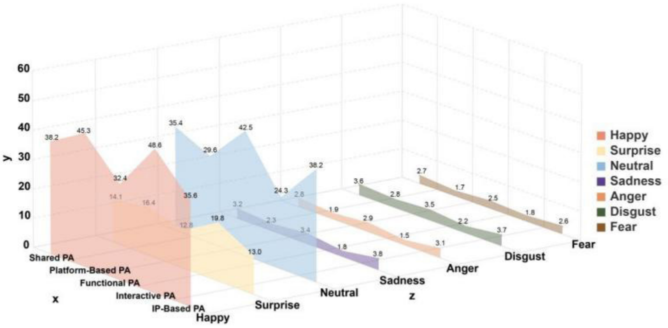

Figure 8: Fine-grained emotional distribution across different commercial public art.

In the dimension of positive emotions, interactive public art performed most prominently, with a joy proportion of 48.6%, significantly higher than the other four art types ($F = 12.34$, $p < 0.001$), supporting the view that participatory experiences enhance positive affect. Meanwhile, interactive art also ranked highest in surprise (19.8%), indicating its unique advantage in creating novel experiences and exceeding consumer expectations. Platform-based public art followed closely in joy (45.3%), reflecting the role of social interaction in facilitating positive emotions.

The distribution of neutral emotions revealed a function-oriented pattern. Functional public art had the highest proportion of neutral emotions (42.5%), significantly exceeding other types. From an environmental psychology perspective, functional art serves a psychological regulatory role by providing an emotional buffer in high-stimulation commercial environments. In contrast, interactive art had the lowest proportion of neutral emotions (24.3%), further confirming its strong emotional polarization effect.

It is noteworthy that negative emotions remained at overall low levels but showed notable internal variation. The total proportion of negative emotions for all art types was controlled within 10%, indicating the overall effectiveness of commercial public art in emotional management. Among them, platform-based public art showed the lowest values across all four negative emotions, suggesting that social attributes may serve as a buffer against negative affect. In comparison, IP-based art had a relatively higher proportion of negative emotions, possibly due to expectation gaps arising from cognitive discrepancies related to IP. These findings provide important empirical evidence for emotional risk management in commercial spaces.

Research Discussions: Towards a Strategic Framework for Emotional Design

The Participation-Co-Creation Mechanism in Interactive Art

The study reveals that interactive public art generates the most substantial emotional impact, evidenced by its highest composite emotion score (Sen = 0.62) and positive emotion proportion (68.4%). This efficacy stems from its participation-co-creation mechanism, which transforms the audience into active contributors through responsive installations. It also recorded the lowest neutral emotion proportion (24.3%), indicating a polarization of affect. The relatively extended dwell time (18.7 $\pm$ 5.2 min) aligns with the peak-end rule, suggesting emotional intensity is achieved through brief, high-arousal peaks rather than prolonged exposure.

Functional Art as Emotional Buffer Zones

Functional public art—including artistic lighting and wayfinding systems—acts as an emotional buffer in high-stimulation commercial settings. It recorded the highest proportion of neutral emotional states (42.5%), which reflects a deliberate ambient regulation mechanism. This design strategy aligns with the concept of soft fascination from Attention Restoration Theory, offering cognitive restoration through perceptual stabilization rather than expressive intensification.

Re-Examining the Relationship Between Dwell Time and Emotional Intensity

A non-linear relationship is observed between dwell time and emotional intensity. Interactive art achieved the highest emotional score but a shorter average dwell time (18.7 min), whereas platform-based art sustained longer engagement with moderate scores. This suggests emotional impact operates through temporal compression and rapid emotional punctum in interactive works, while platform art employs experiential drip-feeding. Design strategies should therefore align art typology with specific commercial objectives: interactive art for immediate impact at decision points, and platform-based art for areas intended for prolonged social exploration.

Research Limitations

This study has several limitations. Data were collected from a single commercial context, limiting generalizability across cultures and environments. The tripartite emotion classification lacks granularity for complex affective states. Moreover, the analysis is perceptual and does not establish causal links between emotional responses and consumer behavior.

Future work should validate the framework in diverse cultural and spatial settings. Integrating finer-grained emotion models, such as facial action coding, would improve affective interpretation. Longitudinal tracking of behavioral metrics, like purchase decisions, could clarify causal relationships and better inform emotion-aware spatial design.

CONCLUSION

This study established a dual channel sentiment computing framework, achieving a recognition accuracy of 93.2%, significantly better than the unimodal method. Through large-scale empirical analysis, we have identified different emotional efficacy characteristics among five types of art: interactive art performs well through participation in co-creation mechanisms (Sen = 0.62), platform based art utilizes social identity for participation, and functional art serves as a key emotional buffer. It is crucial that we reveal the non-linear relationship between emotional intensity and dwell time, challenging traditional design assumptions. This evidence supports the paradigm shift of public art towards temporal sentiment analysis, enabling strategic planning of art interventions to integrate public values, commercial goals, and spatial contexts, thereby advancing emotional perception design in commercial environments.

ACKNOWLEDGMENT

We wish to show our gratitude to the project Research on Co-creative Participation and Public Education in Urban Public Art of the Guangdong-Hong Kong-Macao Greater Bay Area, funded by the Department of Social Sciences, Ministry of Education, People's Republic of China under grant number 24YJC760040. We also sincerely appreciate all the participants who contributed to this study.

REFERENCES

Chen, C. (2024). "Research on the Emotionalized Installation Design of Urban Commercial Spaces." Journal of China-ASEAN Studies, 5 (1): 18–28.

China Commercial Real Estate Association. (2023). Commercial Space Artistic Renovation White Paper. Beijing: China Commercial Real Estate Association.

China Youth Consumer Research Center. (2024). China Youth Consumption Trend Report. Beijing: China Youth Consumer Research Center.

Hultén, B. (2011). "Sensory Marketing: The Multi-sensory Brand-experience Concept." European Business Review, 23 (3): 256–273.

Krishna, A. (2012). "An Integrative Review of Sensory Marketing: Engaging the Senses to Affect Perception, Judgment and behavior." Journal of Consumer Psychology, 22 (3): 332–351.

Kwon, D.H., Yu, J.M. (2024). "Real-time Multi-CNN-based Emotion Recognition System for Evaluating Museum Visitors' Satisfaction." ACM Journal on Computing and Cultural Heritage, 17 (1): 1–18.

Lee, H.Y., Lee, W.H. (2014). "A Study on Interactive Media Art to Apply Emotion Recognition." International Journal of Multimedia & Ubiquitous Engineering, 9 (12): 431–442.

Lehrl, S., Gerstmeyer, K., Jacob, J.H., Frieling, H., Riederer, A.J., Riederer, K., Weiss, M., Bleich, C., Kornhuber, J., Bleich, S. (2007). "Blue Light Improves Cognitive Performance." Journal of Neural Transmission, 114: 457–460.

Liu, B. (2015). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge: Cambridge University Press.

Mahima, M.A., Patel, N.C., Ravichandran, S., Bhatt, B., Chaudhari, K. (2021). "A Text-based Hybrid Approach for Multiple Emotion Detection Using Contextual and Semantic Analysis." In: 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 1–6. IEEE.

Naidoo, S.W., Naicker, N., Patel, S.S., Chowdhury, S., Maharaj, K. (2022). "Computer Vision: The Effectiveness of Deep Learning for Emotion Detection in Marketing Campaigns." International Journal of Advanced Computer Science and Applications, 13 (5).

Sahoo, K.K., Dutta, I., Ijaz, M.F., Woźniak, M., Damaševičius, R. (2021). "TLEFuzzyNet: Fuzzy Rank-based Ensemble of Transfer Learning Models for Emotion Recognition from Human Speeches." IEEE Access, 9: 166518–166530.

Song, P., Zheng, W., Yu, Y., Ou, S. (2020). "Speech Emotion Recognition Based on Robust Discriminative Sparse Regression." IEEE Transactions on Cognitive and Developmental Systems, 13 (2): 343–353.

Spangenberg, E.R., Sprott, D.E., Grohmann, B., Tracy, D.L. (2006). "Gender-congruent Ambient Scent Influences on Approach and Avoidance Behaviors in a Retail Store." Journal of Business Research, 59 (12): 1281–1287.

Tracy, J.L., Robins, R.W. (2008). "The Automaticity of Emotion Recognition." Emotion, 8 (1): 81–95.

Wang, X., Kou, L., Sugumaran, V., Luo, X., Zhang, H. (2020). "Emotion Correlation Mining through Deep Learning Models on Natural Language Text." IEEE Transactions on Cybernetics, 51 (9): 4400–4413.

Wu, X., Cao, Y., Lu, H., Zhang, S., Wu, Z. (2021). "Speech Emotion Recognition Using Sequential Capsule Networks." IEEE/ACM Transactions on Audio, Speech, and Language Processing, 29: 3280–3291.

Zhang, Q., Lai, T., Yang, Y., Chen, L. (2017). "Multisensory Interaction of Consumers: Performance, Formation Mechanism and Research Prospects." Foreign Economics and Management, 39 (7): 80–90.

Zhong, K., Wang, H., Yang, C. (2016). "A Review and Prospect of Sensory Marketing Research." Foreign Economics and Management, 38 (5): 69–85.