ThoughtDiffusion: An Interactive Installation for Exploring Neuro-Art from EEG Data with Stable Diffusion Models

Janaka Rajapakse
Tainan National University of the Arts, No. 66, Daqi, Guantian Dist., Tainan, 720005, Taiwan
Email: janaka@mail.tnnua.edu.tw

Abstract

ThoughtDiffusion is an interactive installation that combines cognitive processes with the generative capabilities of AI to generate artistic images in real time. The installation system uses Stable Diffusion models and state-of-the-art neural decoding techniques that allow the mapping of brainwave patterns into coherent visual representations. The installation is based on a non-invasive commodity EEG headset that records users' brain signals fed into a stable diffusion model to output images corresponding to the intended mental state but unique to the participants. This installation uses a Kinect V2 sensor to capture users' body movements, which advance multimodal interaction significantly relating to relaxation, the state of being calm, and the state of attention.

Keywords: passive interaction, Neuto-art, Stable Diffusion, TouchDesigner, art interaction

1. Introduction

Neuroart is a growing field that combines art and neuroscience. It transforms neural activity into visual expressions, providing a unique perspective on the human mind. The emergence of the dissemination of thought represents a significant intersection between technology and creativity, nourishing an environment where artistic expression is transformed by ideas gleaned from neurological data. In addition, integrating neurotechnology into artistic practices generates discussions on authenticity and the property of creative expression. The question arises: To what extent does technology influence art production? While artists use brain-computer interface (BCI) technology to evoke specific artistic results, they engage in a crucial dialogue between intentionality and inadvertent expressions generated by cognitive processes [1], [2], [3], [4]. This duality can enrich artistic efforts, in which the interaction between brain activity and technological influence leads to innovative forms of creative expression [5].

In the scope of this study, the thought-of thought appears as a revolutionary installation that transcends traditional borders by exploiting electroencephalography (EEG) data to forge visual expressions in real time. This innovative company not only represents a convergence of technological and artistic disciplines but also facilitates an enriching interaction between users and works of art, ultimately improving their emotional experiences. In conjunction with EEG data, ThoughtDiffusion installation uses the capabilities of the Kinect-V2 sensor and the RGB camera to collect relevant user data, which improves the global interactive experience. The Kinect-V2 facilitates real-time movement monitoring, allowing users to engage with installation beyond simple passive observation [6]. The system captures the spatial dynamics of user movements, creating a reactive environment in which visual expressions dynamically correspond to cognitive states and physical interactions.

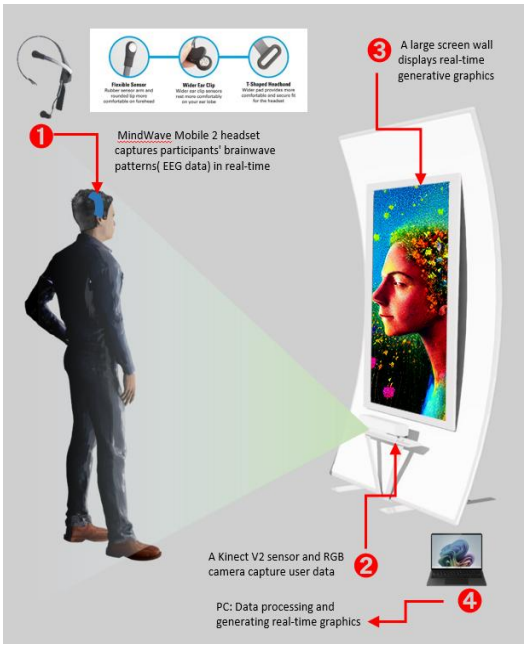


Fig. 1. ThoughtDiffusion Interactive Installation.

The evolution of diffusion models represents significant changes in the scenario of interactive art creation and was accompanied by the emergence of new applications that leverage its general approach [7]. As researchers and professionals navigate the interrelationships between diffusion methodologies and rapid strategies, the robust dialogue will undoubtedly promote future innovations in interactive art [8]. This installation's key aspect is its noninvasive approach. It uses a NeuroSky MindWave Mobile-2 EEG headset [9] to record users' brain signals, which are processed through the Stable Diffusion(SD) models to create unique visual expressions that correspond to the users' intended mental states. This installation proposes a scenario-

oriented prompt structure that the user can feed multiple prompts to interact passively to regenerate the thought of thought experience in real-time.

2. Related Works

2.1. *Interactive Art and Diffusion Models*

Recent progress in artificial intelligence has generated a new paradigm in the kingdom of interactive art, in particular through the advent of diffusion models. These models, which create data distributions capable of producing complex and different results, have basically transformed artistic expression, improving multimodal interactions and deepening human involvement [10]. These probabilistic approaches allow the generation of high-dimensional outputs with unprecedented loyalty and varieties, which is particularly relevant in the artistic contexts in which the visual, auditory, and tactile modalities intersect. By facilitating the creation of works of art that can dynamically adapt to user interactions through different sensory ways, artists have the power to invent new forms of involvement [11]. The multi-sensory experience not only fascinates participants but also promotes a deeper emotional connection between the public and art, thus increasing the commitment [12]. In addition to redefining the artistic process, these models also push innovative paths for the involvement of the public. Interactive art that uses diffusion models can be experienced in various ways, inviting a broader audience to commit. This potential for customization allows individual experiences to differ according to the interactions, preferences, and unique answers of a public member.

Since the diffusion models generate interactive experiences of art and influence, the need for a critical discourse on the implications of the machine's involvement in creativity is presented [13]. Although these models can improve aspects of artistic expression, dialogue must also understand the complexities of how human artists navigate their roles within an increasingly automated landscape [14]. In conclusion, new diffusion models have started a transformative movement within interactive art, improving multimodal interactions and enriching human participation, simultaneously unlocking new realities of artistic expression and public involvement. While artists adopt these technologies, they create a more collaborative and engaging experience that redefines the boundaries of artistic creation. ThoughtDiffusion integrates technologies like EEG monitoring and depth sensors, facilitating the development of a creative landscape in which interactivity is deeply rooted in the interaction between human cognition and artistic representation.

2.2. *Prompt Engineering and Generative Expression in Art*

The evolution of rapid engineering in interactive art represents a critical intersection of technology, creativity, and artistic expression. As generative artificial intelligence tools become increasingly integrated into the creative process, the improvement of instructions and the integration of multiple instructions have emerged as vital components to improve artistic ideas [15]. This literature revision explores recent progress in this sector, underlining the meaning of evolving practices in rapid engineering. Rapid engineering, the practice of designing and refining inputs to arouse desired results from AI systems, has witnessed substantial developments that facilitate a deeper exploration of creative concepts. In addition, rapid design's role extends beyond simple refinement; It includes integrating multiple suggestions to improve creative processes. Recent investigations have examined the possibilities that support using different suggestions in artistic education [8].

Prompt-based rapid design practice promotes iterative processes that allow emerging students and artists to unlock their creative potential. Combining various instructive suggestions allows the creators to explore multiple dimensions of a singular artistic idea, thus promoting a richer and more varied creative output [15]. Mizrahi [17] underlines that the mastery of rapid engineering can accelerate the transition of an artist from a beginner to a competent creator, observing that the "*language of creativity*" plays a fundamental role in the generative process. By providing educational paintings focused on the art of creating engaging instructions, instructors can allow students to exploit the generative tools of artificial intelligence more effectively [16]. Artists can improve their ideation and expression through the iterative improvement of generative instructions and the integration of multiple suggestions. This aspect of artistic practice not only expands the creative potential of individuals but also underlines the educational importance of understanding and mastery of prompt engineering as a critical element of contemporary art production. Therefore, the continuous exploration in this domain benefits from both the artistic community and the broader field of creative education. ThoughtDiffusion installation provides a platform for redefining artistic scenarios through users' multiple prompts, allowing for adjustments to the artwork using their physical and cognitive inputs in real time; these models make the creative process more accessible to explore new forms of interaction with more human involvements.

3. System Overview

The system of installation consists of two aspects: hardware and software, as shown in Fig. 2. As Hardware devices were used, both movement data from Kinect-V2 and brainwave data from the MindWave Mobile 2 EEG

Headset reach a multimodal interaction that enriches users' involvement. The Kinect-V2 RGB camera was used as a sensor to detect the user's skeleton data, depth, and silhouette. EEG headset was introduced to capture users' brainwave data in real-time; the device mechanism allows the system to record EEG frequency spectrum (alpha, beta, etc.) and proprietary *eSense* indicators for the level of *attention* and *relaxation* through EEG data. The development environment was Intel i9 PC on Windows 10 Pro system, with the NVIDIA GeForce RTX 2080Ti graphics card.

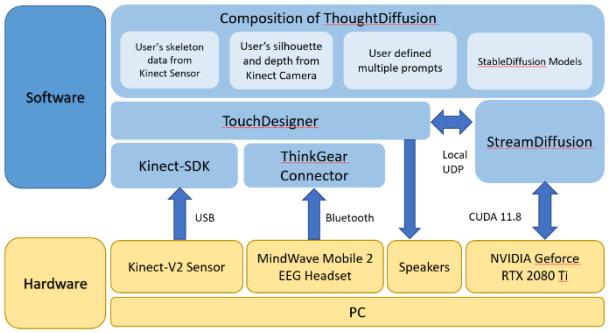


Fig. 2. System Overview

3.1. *Hardware*

The installation is equipped with a large screen wall to display real-time generated graphics, as shown in Fig. 3. The Kinect-V2 sensor is mounted under the large display, which can easily capture the user's full-body skeleton data. Kinect has two built-in sensors: an RGB color video camera and a depth sensor [6], which helps to capture the user's silhouette and depth for video-to-video generation in StreamDiffusion [19] model in real-time. Fig. 3 depicts the installation space. The x-axis corresponds to the horizontal direction across the sensor's field of view, the y-axis represents the vertical direction in the sensor's field of view, and the z-axis represents the normal of the sensor, which can detect depth information, indicating how far away the user is from the sensor.

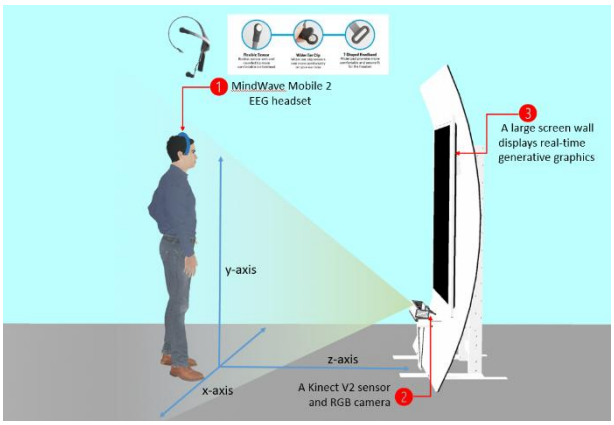


Fig. 3. Affected by pedestrians crossing

This installation mainly uses the functions of Kinect to perform the following three operations:

- Calculate the user's relative position by detecting values of the user's hip-joint coordinates, which control the step counts of the denoise schedule.
- Measure the displacement of the user's hip joint along the x-axis to control the weight of multiple prompts.
- The camera captures the user's silhouette for video-to-video interactive generation in the StreamDiffusion module.

While Kinect captures users' physical movements, Bluetooth connects the user-worn MindWave Mobile 2 EEG headset to the system, which allows the capture of brainwaves for passive interaction.

3.2. *Software*

The system was developed with TouchDesigner [18], a node-based visual programming language for the creation of real-time interactive media content. The Kinect SDK allows integration with TouchDesigner and its built-in channel operator, *Kinect CHOP*, which reads positional and skeletal tracking data. The built-in texture operator, *Kinect TOP*, captures data from the Kinect depth camera and RGB color camera, which is used to compute and isolate the user's silhouette.

The installed NeuroSky's Windows Developer Tools 3.2 provides ThinkGear Connector, which lets TouchDesigner transfer data from the MindWave Mobile 2 EEG headset through a *TCP/IP DAT*(data operator node) socket server (localhost with port 13854) [20]. Bluetooth automatically connects to the device whenever the system needs to read the data from the EEG headset. The recorded serial data is converted to a table (*TouchDesigner Table DAT*), which quickly reads the data into the separate channels for focus levels: *attention* and *relaxation*, and EEG data values: *delta*, *Theta*, *lowAlpha*, *highAlpha*, *lowBeta*, *highBeta*, *lowGamma*, *highGamma*.

The TouchDesigner adopted the submodule StreamDiffusion-TD [19], which is installed with Python 3.6, CUDA 11.8, and the Git desktop client. This submodule supports generating real-time generative AI content based on text2img and img2img. The measured levels of attention and relaxation in EEG data control the *seed* and *seed weight* of StreamDiffusion-TD in real-time. Using multiple Text COMP nodes, set up the multiple Prompt Boxes to read user-given prompts, which can be controlled by users' physical displacement along the x-axis determined by the Kinect CHOPs.

4. Results of the Interaction

The system promotes different modes of interaction.

1. ***Language of creativity***: The user can add their ideation and expression through multiple prompts, interact while moving along the horizontal direction(x-axis), pay attention to their input prompts, and iteratively improve generative prompts with multiple suggestions. Not only side-by-side comparison, but users can also move towards the display (along the z-axis, the normal of the Kinect sensor).

2. ***Pose for prompts***: Users (multiple users up to 3) can pose to the Kinect Camera while paying attention to the added input prompt. The system captures the user's poses for video-to-video interactive generation in the StreamDiffusion module. This scenario-oriented practice not only expands individuals' creative potential but also provides a more collaborative and engaging experience that redefines the boundaries of interactive creation.

Model: txt2img (sd-turbo)

z-distance from the display (denoise schedule)	Prompt: *Gandalf from Lord of the Rings, diffuse lighting, fantasy, intricate elegant highly, detailed lifelike photorealistic digital painting, art station.*	Prompt: *a perfect bonsai.*
0.5m (4 steps)		
1.5m (8 steps)		
2.0m (12 steps)		

Fig. 4. Creative expressions and physical engagement of the body (a sample Moving towards the display(along the z-axis, normal of the Kinect sensor).

4.1. *Creative Expressions*

The ThoughtDiffusion installation offers users an opportunity for an unprecedented transformation in the interaction between creative expression and physical engagement of the body. These modes of interaction, characterized by their multifaceted methods, support a complex interplay of language, user input, and bodily movements in iterative feedback, which collectively fosters richer and more dynamic creative processes. This exploration deepens these interaction ways, examining how they improve the user's experience and promote collaborative creativity. As shown in Fig. 4, language is a fundamental interaction element at the center of the involvement process. When the user moves closer to the display, the number of steps of the *denoise schedule* in StreamDiffusion gradually decreases. Incorporating physical body movements not only modifies the generative content on display but also enhances collective intelligence, as the intuitions lead to richer and more nuanced interactions that modify the prompt and passively adjust the body position to achieve creative results.

Model: img2img (sd-turbo)

Prompt: *A Boy with super powers and curly and brown hair, illustration concept art anime key visual trending pixiv fanbox by wlop and greg rutkowski and makoto shinkai and studio ghibli*

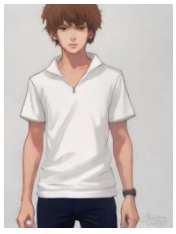
z-distance from (denoise schedule)	input	output
2.0m (12 steps)		
3.5m (16 steps)		

Fig. 5. Posing for the prompts (video-to-video interactive generation).

Fig. 5 illustrates how users can respond to a given prompt. The dynamic nature of creative processes requires a structure, such as *a method of acting or posing*, in which users can perfect their ideas. This iterative approach promotes an experimentation environment in which users are encouraged to take creative risks, knowing they can adjust their production in response to the evolving collaborative scene.

4.2. *Level of Attention and Relaxation*

Neurological feedback on attention and relaxation levels is another important aspect of this installation's complexity interaction. While the levels of attention and relaxation cannot be directly associated with the meaning

of seed and seed weight parameters of the StreamDiffusionTD model, our approach attempted to map and trigger their values and intuitively demonstrate the capability of EEG signals to reconstruct dynamic visual perceptions when the user keeps the attention on the given prompt. As shown in Fig. 6, different levels of attention can significantly modulate to generate real-time output using the Stable Diffusion model *txt2img (sd15)* for a given prompt: "*An elegant girl in an urban outfit, cute fine face, round eyes, digital painting, glasses, blonde hair, fashion magazine*." The different levels of relaxation experimented with the Stable diffusion model *img2img (sd15)* for a given prompt: "*A boy in a village outfit, cute fine face, round eyes, digital painting, face paint, curly hair, art magazine*."(Fig. 7). As a new form of human-computer interaction, the ability to effortlessly adjust users' physical postures to achieve desired outcomes in real-time has become a major focus in training attention and relaxation practices, which can also be considered art therapy.

Model: txt2img (sd15)

Prompt: *An elegant girl in an urban outfit, cute fine face, round eyes, digital painting, glasses, blonde hair, fashion magazine*

Attention: 50 Attention: 60 Attention: 65 Attention: 80

Fig. 6. Different levels of attention.

Model: img2img (sd15)

Prompt: *A boy in a village outfit, cute fine face, round eyes, digital painting, face paint, curly hair, art magazine.*

Relaxation: 60 Relaxation: 70 Relaxation:78 Relaxation: 85

Fig. 7. Different levels of relaxation.

5. Conclusion

As users engage in this vibrant form of human-computer interaction, user gestures, and neuro-feedback mechanisms generate new content and co-evolving their creative identities. The interaction of distinct prospects and experiences occurs in the form of collaborative creation richer in the sum of its parts. The installation acts as a co-creator, offering instructions and possibilities that use users to think in addition to their initial ideas and hypotheses. This coevolution of thought and creativity cultivates a sense of collective property of creative production, thus improving the intrinsic motivation to collaborate further.

In summary, this installation system's multifaceted interaction method significantly enriches creative expression and collaboration. By exploiting the dynamics of language, user poses, and iterative neurological feedback, this system creates a rich space for innovative ideas and collective creativity. Through the development of this installation, it is clear that the evolution of technology continues to transform the nature of creative collaboration, allowing users to interact with each other and their creative processes in profoundly transformative ways.

References

1. E. M. Holz, L. Botrel L, and A. Ku¨bler, "Independent home use of Brain Painting improves the quality of life of two artists in the locked-in state diagnosed with amyotrophic lateral sclerosis." Brain-Computer Interfaces, Vol. 2, no. 2–3, pp.117–34, 2015.
2. C-T. Lin, R.P.C.J. Rajapakse, and Y. Tokuyama, "NeuroPhyllotaxis: An Interactive Application for Generative Art Based on EEG Data" Proc. of the International Conference on Artificial Life and Robotics 2021, pp. 724-727, 2021, https://doi.org/10.5954/ICAROB.2021.OS20-1
3. R.P.C.J. Rajapakse, and Y. Tokuyama, Y. (2021, January 21–24). Thoughtmix: Interactive watercolor generation and mixing based on EEG data. Proc. of the International Conference on Artificial Life and Robotics 2021, pp. 728-731, https://doi.org/10.5954/ICAROB.2021.OS20-3
4. C-T. Lin, R.P.C.J. Rajapakse, and Y. Tokuyama, "Development of EEG Data-driven Generative Art Application for Real-time and Dynamic Interaction" Journal of Robotics, Networking and Artificial Life, vol. 8, no. 2, pp. 117-121, -727, 2021, https://doi.org/10.2991/jrnal.k.210713.010
5. K. Yamashiro, N. Matsumoto, and Y. Ikegaya Y, "Diffusion model-based image generation from rat brain activity." PLOS ONE, vol.19, no. 9,: doi: https://doi.org/10.1371/journal.pone.0309709
6. "Kinect for Windows", https://learn.microsoft.com/zh-tw/windows/apps/design/devices/kinect-for-windows
7. R. Po, W. Yifan, V. Golyanik, K. Aberman, J. T. Barron, A. Bermano, E. Chan, T. Dekel, A. Holynski, A. Kanazawa, C.K. Liu, L. Liu, B. Mildenhall, M. Nießner, B. Ommer, C. Theobalt, P. Wonka, G. Wetzstein, "State of the Art on Diffusion Models for Visual Computing," Computer Graphics Forum, Vol. 43, No. 2, : e15063, 2024, https://doi.org/10.1111/cgf.15063
8. J. Hutson, and P. Cotroneo, "Generative AI Tools in Art Education: Exploring Prompt Engineering and Iterative Processes for Enhanced Creativity," Metaverse, vol. 4, no. 1, pp. 1-14, 2023, https://doi.org/10.54517/m.v4i1.2164
9. NeuroSky MindWave Mobile-2 EEG Headset https://store.neurosky.com/pages/mindwave
10. V. Boutin, T. Fel, L. Singhal, R. Mukherji, A. Nagaraj, J. Colin, and T. Serre, "Diffusion Models as Artists: Are we Closing the Gap between Humans and Machines?" https://doi.org/10.48550/arXiv.2301.11722
11. Georgios Marentakis, David Pirrò, and Marian Weger. 2017. Creative Evaluation. In Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17). Association for Computing Machinery, New York, NY, USA, 853–864. https://doi.org/10.1145/3064663.3064710

12. Claudia Rébola Winegarden, Nicholas Komor, and Scott M. Gilliland. 2010. Silent mutations: physical-digital interactions in spaces. In Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction (TEI '10). Associationn for Computing Machinery, New York, NY, USA, 277–278.
13. B. Wang, Q. Chen, and Z. Wang, "Diffusion-Based Visual Art Creation: A Survey and New Perspectives," arXiv:2408.12128, 2024. 4, 2024.
14. H. Cao, T. Cheng, Z. Gao, Y. Xu, G. Chen, P-A. Heng and Z. Li. Stan, "A Survey on Generative Diffusion Models." IEEE Transactions on Knowledge and Data Engineering, vol. 36, pp. 2814-2830, 2022, https://doi.org/10.48550/arXiv.2209.02646
15. J. Hutson and B. Robertson, "Exploring the educational potential of AI generative art in 3D design fundamentals: A case study on prompt engineering and creative workflows," Global Journal of HUMAN-SOCIAL SCIENCE: A Arts & Humanities-Psychology, Vol. 23, No. 2,2023,https://socialscienceresearch.org/index.php/GJHSS/article/view/103669/37415
16. J. Oppenlaender, R. Linder, and J. Silvennoinen, "Prompting AI art: An investigation into the creative skill of prompt engineering," International Journal of Human-Computer Interaction, pp. 1-23, November 2024. https://doi.org/10.1080/10447318.2024.2431761
17. G. Mizrahi and D. Serfaty. "Unlocking the secrets of prompt engineering: Master the art of creative language generation to accelerate your journey from novice to pro," Packt Publishing; 1st edition, January 12, 2024.
18. TouchDesigner, https://derivative.ca/, Accessed on 18th December 2024.
19. K. Kodaira, C. Xu, T. Hazama, T. Yoshimoto, K. Ohno, S. Mitsuhori, S. Sugano, H.g Cho, Z. Liu, and K. Keutzer, "Streamdiffusion: A pipeline-level solution for real-time interactive generation." arXiv preprint arXiv:2312.12491 (2023).
20. NeuroSky's Windows Developer Tools 3.2 and ThinkGear Connector, https://store.neurosky.com/products/pc-developer-tools

Authors Introduction

Dr. Janaka Rajapakse

R.P.C. Janaka Rajapakse is an Associate Professor at the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan. He received his Ph.D. in Knowledge System Science from the Japan Advanced Institute of Science and Technology in 2008. His research interests include computer animation, motion capture, VR/AR/MR, haptic interfaces, AI, computer graphics, and Kansei Engineering. He is a member of the Motion Capture Society, The Society for Art and Science, ASIAGRAPH, and SIG-Design Creativity.