Animation in the Age of AI: Creative Dialog With Algorithms

1ˢᵗ Celine Pham, 2ⁿᵈ Philipp Wintersberger, 3ʳᵈ Juergen Hagler

University of Applied Sciences Upper Austria, Hagenberg Campus
Hagenberg im Mühlkreis, Austria
{celine.pham}, {philipp.wintersberger}, {juergen.hagler}@fh-hagenberg.at

Abstract

LUCID is an experimental animated short film that explores the interplay of dance, music, and artificial intelligence (AI). Our objective was to examine the role and impact of style transfer technologies and other AI animation tools and the potential of creative human-AI collaboration in animation filmmaking. We investigated how these technologies change artistic processes, open new creative possibilities, and redefine the relationship between artists and their tools. A particular focus was placed on the creative Human-AI collaboration and on the interaction of technological innovation and artistic expression to gain a comprehensive understanding of the current and potential impact of AI-based style transfer and AI-generated images in the art world. Using the film *LUCID* as an example, we show how the creative collaboration process between artists and AI tools may become a central theme of the animated film.

Keywords

Computer animation, Animation, Art, Artificial intelligence, Machine learning algorithms, Human-AI cooperation, Conceptual art

Introduction

In the realm of digital art, Artificial Intelligence (AI) plays a transformative role, introducing various innovative methods to the art-making process. A central aspect of this transformation is the concept of generative image synthesis. Here, AI algorithms are utilized to create unique visual content, pushing beyond the traditional boundaries of art creation. Alongside this, style transfer emerges as another key method. Style Transfer is a technique in machine learning using deep neural networks that manipulates digital images or videos to adopt the appearance or visual style of another image [3, 4]. Ruder, Dosovitskiy, and Brox [12] extended this computer vision technique to video. Style transfer for videos applies the same principle as image style transfer but includes an additional consideration for the temporal consistency between frames.

This article introduces the animated short film *LUCID* as an example of how an artistic-creative design process between various artists and deep learning algorithms can emerge. In this case, it is a collaboration between animation artists, two dance performers, and a sound designer. *LUCID* goes beyond using AI tools to transfer a specific visual style

to moving images and follows a formal and self-reflexive approach [13, 16, 15], where the human-AI dialog becomes the central concept. The narrative and thematic concern is determined by expanding the aesthetic and technical parameters of style transfer, and its own construction is revealed for critical discourse.

The starting point of the dialogue was a first sound design draft that was interpreted by the dances. The recorded dance performance was the input for the artistic experimentation with AI tools. And vice versa, the video footage, the utilized prompts, and the output followed back into further sound design processing.

Over time, the film's construction and the creative dialogue between the artists become visible. Towards the end, the film shimmers along the boundary between the style-frame-output and the underlying life recording. The end credits show the prompt and the tools used.

Related Work

Style Transfer and AI-driven Image Generation

The intersection of AI and visual arts has undergone a significant transformation over recent decades, most notably in the realms of AI-driven image generation and style transfer. Initially considered as separate technological achievements, these two domains have evolved into a synergistic tool in the world of digital art. The roots of AI-driven image generation trace back to early experiments in computer graphics and algorithmic art in the 1960s [2]. However, a pivotal moment in this domain was marked by the advent of Convolutional Neural Networks (CNNs) [5] and especially the introduction of Generative Adversarial Networks (GANs) in 2014 [6]. These developments enabled the generation of complex visual content from scratch or based on textual descriptions [8]. In parallel, style transfer, facilitating the application of one image's style to another, has its origins in the advancements in computer vision and deep learning since 2015 [1]. Applications like *Prisma*[1], introduced in 2016, popularized this technology by allowing users to transform their own images into renowned artistic styles. The amalgamation of AI-driven image generation and style transfer opens a new spectrum of creative possibilities. These range from creating unique images through AI and subsequent stylization to hybrid creations that

---

[1]https://prisma-ai.com/

influence both content and style through AI. Advanced AI models integrating both techniques offer the ability to seamlessly and with a single input manipulate an image's content and style. The ongoing development and integration of AI-driven image generation and style transfer are turning points in digital art. These technologies not only expand the toolkit of artists and designers but also redefine the role of AI as a collaborative partner in the creative process [17]. The future of this synergy promises to continue blurring the boundaries between technology and artistic creativity, enabling new forms of aesthetic expression.

Current artistic positions

Since the introduction of *DALL-E*[2] in January 2021 and other user-friendly image generation tools, the field of AI-generated art has seen significant advancements and growing public interest. A short time later, the first animated films were made using this technique, attracting the audience's and the jury's attention at media art and animation festivals. The following presents a current selection of award-winning animation works demonstrating experimental approaches. At the beginning of 2022, the AI film *The Crow* (2022) [9] by Glenn Marshall won the Cannes Short Film Festival, and *Hysteresis* (2022) [14] by Robert Seidel was awarded a Special Mention at ITFS Stuttgart. In *The Crow*, digital artist Marshall used style transfer and other AI tools to transform a dancer into a crow.[3] Marshall gives an insight into his current approach in the follow-up project *Expanded Animation Trailer* [10], where he merges a dance performance and a particle system.

Hysteresis [14], the first AI film of Berlin-based artist Robert Seidel, goes a step further regarding creative human-AI collaboration. The film merges Seidel's artwork with a dance performance. Seidel edited the recorded video of the performance using AI tools; the output was then projected onto the dancer's body, and in a final step, these recordings were again processed by machine-learning tools. Based on the two animated short films *Dissolution* (2023) and *Cunabula* (2023), Sujin Kim illustrates how artists can use style transfer effectively to support a particular artistic concept [7]. Despite the artistic achievements of the AI tools, Kim points out some challenges regarding achieving a distinctive output. Extensive parameter adjustments are necessary to create a desired visual style due to the high unpredictability of the production. However, unpredictable results can stimulate the process if you embark on an experiment with an open outcome [17].

Even though AI tools have improved significantly in recent months, the introduced case studies point to an experimental human-AI collaboration. The artists embark on a joint search, a creative dialogue characterized by random and unpredictable outputs.

LUCID

The experimental animated short film *LUCID* [11] is an innovative dialogue between artistic expression and the advanced application of AI. Developed in 2023, a year in which countless AI tools have emerged, the film stood under the impact of integrating AI tools such as *Kaiber.ai*[4], *Runway Gen2*[5], *Midjourney*[6] and *Stable Diffusion XL*[7]. These technologies made it possible to set new limits in the quality and efficiency of image generation and thus both simplify and refine the process of artistic creation.

With *LUCID*, we explored the relationship between improvised dance and a carefully curated musical composition. The starting point was a dance choreography inspired and guided by a composition. The performance was recorded as a one-shot using a handheld camera. The image composition was closely aligned with the music, enabling a synergetic integration of the life action shot and AI-generated images (see Fig. 2).

The video material was processed using *Automatic1111*[8], a user-friendly interface for *Stable Diffusion* that generates images from textual descriptions. This interface simplifies the creation process with a graphical user interface. The initial step involved converting the video into an image sequence, from which keyframes were extracted. The *img2img* method with *ControlNet* was employed to process each of these keyframes. *ControlNet*[9] is a technique in *Stable Diffusion* that enables specific areas of a source image to be targeted and modified based on textual descriptions. This approach allows for the alteration of individual aspects of the image without requiring a complete redesign of the entire picture. The resulting images, including original elements and modifications prescribed by *ControlNet*, were defined as new keyframes for further processing with *EBSynth*[10]. *EBSynth* utilizes these keyframes to transfer the edited style to other images or videos. The stylistic differences between the original and the edited keyframe are analyzed, and these changes are applied to the target material to achieve a consistent style. For a more comprehensive understanding of the process, kindly refer to the corresponding figure in the documentation (see Fig. 3). The visual style evolves from an initially abstract form to a more precise visualization of the dancer, symbolizing the evolution of the relationship between the artists and the AI tools. Thus, the film not only illustrates the creative use of AI in art but also serves as a reflection on the role of technology in artistic practice.

Creative process

The film was created through a collaborative and multidisciplinary process, involving dynamic interaction and interpretation between artists and AI. The project began with a sound

---

Figure 1: Images from the dance performance video and final stills of *LUCID* [11].

designer's musical composition, which shaped the mood and choreographic direction of the film. The music acted as inspiration and a structural element for subsequent creative steps. The choreography, performed by two dancers, was captured on camera, with the music driving the interpretation of movement. This served as the basis for the visual component of the film. The next phase involved a dialogue with AI, which analyzed the music and video footage and extracted keywords and themes. This influenced the visual aesthetics and themes of the film. Based on the keywords, a prompt was created to guide the AI in generating images. The AI-generated images and sketches were then shared with the sound designer and digital artist, influencing the composition of the film and soundtrack. The animator and sound designer integrated the AI-generated text for a deeper integration of music and visual art. This process demonstrates how the fusion of human creativity and technological innovation can result in a synergistic artistic expression that utilizes AI technology's possibilities while honoring individual artists' contributions.

Discussion and Conclusion

Interviews with artists

After the completion of the movie, we interviewed the involved artists. In the following, we summarize the discussion on the main key issues.

The animator, the project's initiator, and the only one working with AI acted as a link between the dancers and the sound designer. It was essential for the animator that the sound designer and dancers engage in an open-ended experiment. Since sound and visuals were closely intertwined and coordinated with each other, the sound designer also gained a deeper insight into working with AI. Like the animator, the sound designer had to work with the generated images, incorporating her interpretation of these visuals into the sound design. The project presented some challenges due to the vast array of possibilities offered by AI tools, which made it difficult to determine the outcome and direction. Additionally, new techniques were frequently introduced, and the rapid development over six months occasionally caused the process to be undervalued up to that point.

All the artists involved responded that they could contribute to the concept and felt they could incorporate their own language into the project. Ultimately, all the involved artists are interested in the potential of AI in animation filmmaking, envisioning it as a tool for enhanced communication and understanding of the interplay of dance movements, sound, and animation while posing questions on its implications for further artistic collaboration. Speaking about the integration of humans and technology in art, the artists acknowledge the physical limitations of the human body. However, they believe that human emotion and soul cannot be replaced by machines. Collaborations between humans and technology can open up new possibilities and create holistic artistic experiences.

Figure 2: Stills of the animated short film *LUCID* [11].

Artistic Human-AI collaboration

The emergence of AI and AI animation can be likened to the birth of other revolutionary technologies like film and computer animation, often marked by an avant-garde phase. AI art and animation are in a self-reflective period, with artistic discussions centering on their essence and structure. In the realm of AI, one predominant topic is collaboration with AI tools.

The experimental short *LUCID* explores human-AI collaboration and highlights its challenges, particularly in animation, dance, and sound. Artists engaging with AI must embrace an experimental phase, replete with uncertainties and iterative concepts. The many and sometimes random results of AI tools may hinder the artistic process due to the vast array of setting parameters associated with style transfer and AI image generation. There is a risk of oversimplification and arbitrariness in the artistic concept, and the presence of copyright issues, biases, and clichéd outputs in AI tools should be critically acknowledged.

To summarize, the case study shows that when interacting with AI tools, the artists are ultimately in charge of the artistic direction of this process but can still run the risk of consciously or unconsciously giving up some of the artistic direction.

References

[1] Dumoulin, V.; Shlens, J.; and Kudlur, M. 2016. Supercharging style transfer. `https://blog.research.google/2016/10/supercharging-style-transfer.html`. accessed January 12, 2024.

[2] Garcia, C. 2016. Harold cohen and aaron-a 40-year collaboration. *Computer History Museum Website*.

[3] Gatys, L. A.; Ecker, A. S.; and Bethge, M. 2015. A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576*.

[4] Gatys, L. A.; Ecker, A. S.; and Bethge, M. 2016a. Image style transfer using convolutional neural networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2414–2423.

[5] Gatys, L. A.; Ecker, A. S.; and Bethge, M. 2016b. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2414–2423.

[6] Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial nets. *Advances in neural information processing systems* 27.

[7] Kim, S. 2023. Reimagining animation making through style transfer. In *SIGGRAPH Asia 2023 Art Papers*. Association for Computing Machinery. 1–6.

[8] Lawto, G. 2023. Cnn vs. gan: How are they different? `https://www.techtarget.com/searchenterpriseai/feature/`
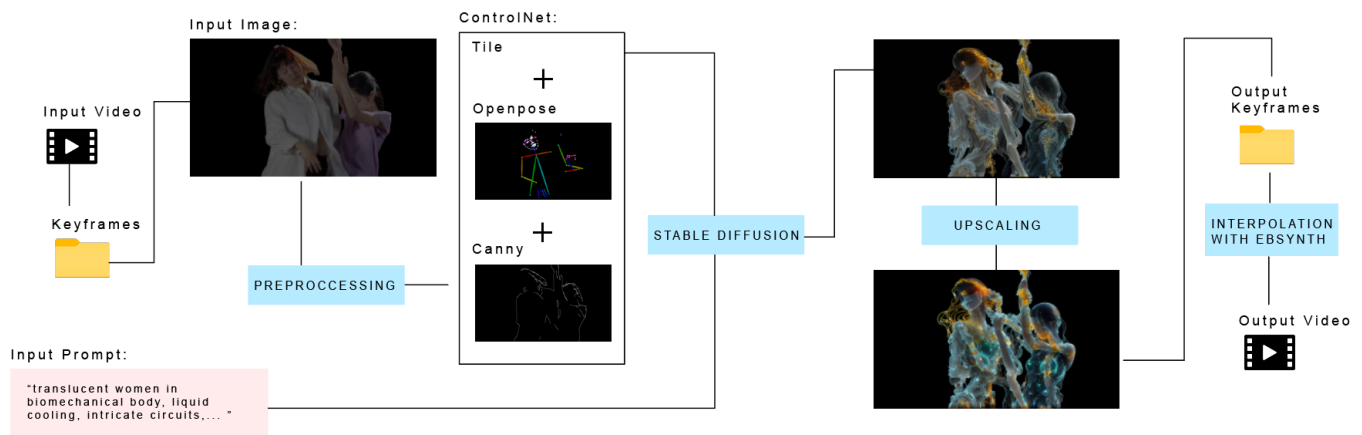
Figure 3: Workflow for image generation in *Stable Diffusion LUCID* [11]: 1. Input; 2. Creative Human-AI-Dialogue, 3. Output

CNN-vs-GAN-How-are-they-different. accessed January 12, 2024.

[9] Marshall, G. 2022. The crow. https://archive.aec.at/prix/showmode/70040/. accessed January 12, 2024.

[10] Marshall, G. 2023. Making of expanded animation 2023 trailer. https://www.youtube.com/watch?v=Px8xCDh6jQ0&t=84s. accessed January 12, 2024.

[11] Pham, C. 2023. Lucid. https://vimeo.com/873477157?share=copy. Music & sound design: Sara Trawoeger; dance performance: Naho Takeda & Shuy Wang, accessed January 12, 2024.

[12] Ruder, M.; Dosovitskiy, A.; and Brox, T. 2016. Artistic style transfer for videos. In *Pattern Recognition: 38th German Conference, GCPR 2016, Hannover, Germany, September 12-15, 2016, Proceedings 38*, 26–36. Springer.

[13] Russett, R., and Starr, C. 1976. *Experimental animation: an illustrated anthology*. Da Capo Press.

[14] Seidel, R. 2022. Hysteresis. https://robertseidel.com/hysteresis/. accessed January 12, 2024.

[15] Smith, V., and Hamlyn, N. 2018. *Experimental and expanded animation: New perspectives and practices*. Springer.

[16] Wells, P. 2002. *Animation: Genre and authorship*, volume 13. Wallflower Press.

[17] Wintersberger, P., and Hagler, J. 2023. Broken telephone: A tale of two ais in the loop. *10th International Conference on Illustration and Animation* 725–731.

Authors Biographies

Celine Pham , born in Linz, Austria, began her career with an education in information technology and worked as a software developer, where she developed her interest in technology. Alongside her technical background, she always maintained a strong interest in creative projects and analog art, which nurtured her passion for blending technology with art. This duality of interest eventually led her to pursue studies in Digital Arts, which she is currently undertaking in her 6th semester, following her attendance at a Technical College. Her unique technical expertise and artistic inclination enable her to explore innovative pathways in digital art.

Philipp Wintersberger is a Professor of Interactive Systems at the University of Applied Sciences Upper Austria (Campus Hagenberg) and a lecturer at TU Wien. He leads an interdisciplinary team of scientists on FWF, FFG, and industry-funded research projects focusing on human-machine cooperation in safety-critical AI-based systems. He has (co)authored various works published at major journals and conferences (such as ACM CHI, IUI, AutomotiveUI, or Human Factors), and his contributions have won several awards. Further, he is a member of the ACM AutomotiveUI steering committee and has contributed to HCI conferences in various roles in the past.

Juergen Hagler is an academic researcher and curator working at the interface of animation, game, and media art. He studied art education, experimental visual design, and cultural studies at the University for Art and Design Linz, Austria. Currently, he is a Professor of Computer Animation and Media Studies and the head of studies of the bachelor's and master's program Digital Arts at the University of Applied Sciences Upper Austria, Hagenberg Campus. Since 2014 he has been the co-head of the research group Playful Interactive Environments. He has been involved in the activities of Ars Electronica since 1997 in a series of different functions. Since 2017 he has been the director of the Ars Electronica Animation Festival and initiator and organizer of the Expanded Animation Symposium.