Cangjie's Poetry: An Interactive Art Experience of a Semantic Human-Machine Reality

WEIDI ZHANG, University of California, Santa Barbara, Media Arts and Technology Program, Santa Barbara, CA 93106, USA

DONGHAO REN, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

GEORGE LEGRADY, University of California, Santa Barbara, Media Arts & Technology Graduate Program, Santa Barbara, CA 93106, USA

This paper describes the conceptual background, artificial intelligent system design, and visualization strategies of an interactive art experience: *Cangjie's Poetry*. This artwork provides a conceptual response to the human-machine reality in the context of language, symbols, and semantic meanings. In the *Cangjie's Poetry* art installation, the intelligent system (*Cangjie*) constantly observes surroundings through the lens of a camera, writes poetry using its symbolic system based on its interpretation, and explains the evolving poem in natural language to audiences in real time. Due to the global pandemic of COVID-19, multiple presentation formats of this work were developed, which include in-person installation, virtual installation, and a special edition with pre-rendered video. This work prioritizes ambiguity and tension between machinic vision and human perception, the actual and the virtual, past and present.

CCS Concepts: • **Computer systems organization** → Real-time system architecture; • **Human-centered computing** → Visualization; • **Computing methodologies** → Machine learning.

Additional Key Words and Phrases: Experimental Visualization, Intelligent System Design, Interactive Art

ACM Reference Format:

Weidi Zhang, Donghao Ren, and George Legrady. 2021. *Cangjie's Poetry: An Interactive Art Experience of a Semantic Human-Machine Reality*. *Proc. ACM Comput. Graph. Interact. Tech.* 4, 2, Article 19 (August 2021), 9 pages. <https://doi.org/10.1145/3465619>

Cangjie's Poetry starts with a conceptual question: If artificial intelligence can imitate human beings to create a symbolic system to communicate based on their understanding of the universe and begin interacting with us actively, how will this recontextualize and redefine the coexistence between humans and machines in this intertwined reality?

Inspired by an ancient Chinese legendary historian, Cangjie (c. 2650 BCE), who invented Chinese characters based on the characteristics of everything on the earth [Cangjie, 2013], we trained a neural network that we call *Cangjie* to learn the constructions and principles of over 9000 Chinese characters. After successful training, *Cangjie* can interpret images through the lens of Chinese characters and produce new symbols constructed by Chinese strokes. Those new symbols appear in the aesthetic of Chinese characters, but they are meaningless novel symbols. To create meanings for this symbolic system conceptually and aesthetically, we implement a pre-trained model (DenseCap)

Authors' addresses: Weidi Zhang, University of California, Santa Barbara, Media Arts and Technology Program, Santa Barbara, CA 93106, USA, weidizhang@ucsb.edu; Donghao Ren, University of California, Santa Barbara, Santa Barbara, CA 93106, USA, donghao.ren@gmail.com; George Legrady, University of California, Santa Barbara, Media Arts & Technology Graduate Program, Santa Barbara, CA 93106, USA, glegrady@ucsb.edu.

This work is licensed under a [Creative Commons Attribution International 4.0 License](https://creativecommons.org/licenses/by/4.0/).

© 2021 Copyright held by the owner/author(s).

2577-6193/2021/8-ART19

<https://doi.org/10.1145/3465619>

Fig. 1. A gallery mock-up view of the *Cangjie's Poetry* art installation. (left) The first projection: Live streaming imagery writes new symbols. (right) The second projection: Intelligent visual poetry describes surroundings, 2020. (©and Photo: Weidi Zhang)

to simultaneously generate descriptive sentences of subregions within images in natural language, making those novel symbols understandable [Johnson et al., 2016].

In the art installation, the *Cangjie* system captures the audiences and the gallery surroundings using a camera and transforms the real-world streaming into a cluster of ever-changing new symbols in real time. Those novel symbols intertwined with the captured imagery are then visualized algorithmically as an abstract pixelated landscape, which constantly moves and evolves. We project this semantic landscape on the wall in the exhibition space as the first projection. Meanwhile, *Cangjie* also generates descriptive sentences of surroundings based on its interpretation. Those sentences are designed as flowing poetry written in visual simulation with dynamic ink diffusion, assembled with real-time captured imagery fragments as the second projection in the space (Figure 1).

Since the beginning of the COVID-19 pandemic, three of this work's scheduled physical exhibitions were unfortunately reconfigured to be viewed online. To better accommodate the experience, we built an online gallery space with recorded visualizations on virtual screens. There is also a special edition of this work designed explicitly as a response to the pandemic. We will present this special edition in detail in section 4.

1 CONCEPTUAL BACKGROUND

Our inspiration comes from the Chinese folklore of a legendary historian: Cangjie (c. 2650 BCE). He is considered to be the person who invented Chinese characters, one of the world's oldest logographic systems. Cangjie struggled with creating characters until he saw a phoenix drop an object with a special hoofprint that he could not recognize. A hunter told him that it was a hoofprint of a Pixiu—a Chinese mythical hybrid creature whose hoofprint is unique. Cangjie realized that it would be a perfect prototype for a written system if he could create symbols that set apart everything on the earth [Wikipedia, 2021].

We artistically imagine the future language from this ancient truth of a logographic system. Suppose intelligent machines can create a symbolic system based on their observations of everything

on the earth like Cangjie did thousands of years ago. How will this influence our coexistence? Humans and machines are in constant conversations where humans start the dialogue by using programming languages compiled to binary digits that machines can interpret. However, today's intelligent machines not only are observers of the world but also make their decisions. Thus, we trained our intelligent machine (*Cangjie*) to actively communicate with human beings using its symbolic system. Meanwhile, real-time descriptions are delivered in natural language (English) to make this interactive experience go beyond a mere visual celebration, but more importantly, be a conceptual semantic experience that engages audiences.

2 RELEVANT WORK

The renowned contemporary artwork *Book from the Sky* by Chinese artist Xu Bing manipulates language in visual forms. He creates more than 4000 pseudo-Chinese characters; carves those fabricated characters in movable type; prints them in scrolls, books, and posters; and installs them as art installation work [Fraser and Li, 2020, p. 89]. This contemporary artwork questions the nature of language with meaningless meanings. Influenced by the artist's personal experience under the Chinese cultural revolution, *Book from the Sky* challenges the authority of the written text and the power of language [Fraser and Li, 2020, pp. 27–28]. In recent media art practices, several independent media artists, such as the artist group Ouchhh and Hong Kong-based artist Hanqian Zhang, use Chinese characters to train machine learning models and produce pseudo-Chinese characters [Cheung, 2020]. Although those artworks bring creative and innovative perspectives on pseudo-Chinese characters' generation to explore machine learning visuals, they only provide limited engaging experience for non-Chinese speakers.

In contrast, *Cangjie's Poetry* revisits the origins of Chinese characters as one of the oldest logographic systems by building associations between real-world imagery of the art installation scene with the real-time generation of the symbols. The *Cangjie* system elaborates on the reflection of the language system in the context of human-machine reality. *Cangjie's Poetry* makes the responsive art experience an intelligent and poetic conversation that engages audiences by providing interactive symbolic visualization in the aesthetic of ink-wash painting with AI-generated semantic information.

3 INTELLIGENT SYSTEM, EXPERIMENTAL VISUALIZATION, AND INSTALLATION DESIGN

Cangjie's Poetry emphasizes equal contributions from three perspectives: the machine's creative decisions, the author's artistic approaches, and audiences' real-time feedback. Thus, this work's methodology mainly consists of multi-modal intelligent system design, experimental visualization strategies, and human-computer interactions. The following subsections present a detailed discussion from concept to techniques of these three components.

3.1 Machine's Creation: Intelligent System Design

The intelligent system of *Cangjie's Poetry* (Figure 2) mainly consists of two parts: 1. a neural network that transforms the live streaming image data into a cluster of new symbols constructed by Chinese strokes in real time and 2. a pre-trained model (DenseCap) that both localizes and describes salient regions in the live streaming image using natural language.

In the first part of the system, we use unsupervised learning techniques to model Chinese character strokes and then use the learned model to create novel characters from images. We use the machine learning model: Bi-directional Generative Adversarial Networks (BiGAN) [Donahue et al., 2016]. This network is trained using an open-source Hanzi writer dataset, which contains vector stroke data for over 9000 Chinese characters. It is noticeable that we rendered partial characters by

Fig. 2. A system diagram of *Cangjie's Poetry* with a top view of the art installation on the left and system design on the right, 2020. (©Weidi Zhang)

Fig. 3. A trained neural network (*Cangjie*)'s real-time reconstruction of the live streaming (left) to the image of new symbols (right), 2020. (©and Photo: Weidi Zhang)

only taking four consecutive strokes from existing characters to avoid the model learning to create actual Chinese characters. Thus, the model only learns the basic composition of strokes. After successfully training the model, we get a pair of generator and encoder networks that reach a stable state—a Nash equilibrium. The network learns a low-dimensional latent representation of these images. When the *Cangjie* system processes the live streaming imagery of the installation view, including the audiences, the gallery space, and portions of the projection output, it passes video frames into the encoder network to produce latent representations of each frame. The generator network can then reconstruct characters from the latent representations. Since the neural network is trained on Chinese characters, it interprets the real-world images through the lens of its latent representation of Chinese characters and generates the symbols constructed by Chinese strokes based on the given latent features (Figure 3).

Besides generating new symbols with a BiGAN, the *Cangjie* system also produces sentences to describe salient regions of captured live streaming of the surroundings in the installation, attaching meanings to the new symbols. We utilize pre-trained fully convolutional localization networks for dense captioning: DenseCap [Johnson et al., 2016]. This network is capable of both locating and

describing the salient regions in images in natural language (the pre-trained model we implement to produce descriptions in English) [Johnson et al., 2016]. We utilize DenseCap to determine sentences' arrangement and generate multiple descriptive sentences in the image's multiple regions. This model also outputs a confidence score to value how much the descriptive sentence corresponds to the ground truth. We purposefully select the lowest confidence outputs (the most inaccurate result) and the highest confident outputs (the most accurate result) and combine them as interactive intelligent poetry that produces metaphors with ambiguity and unexpectedness.

3.2 Artistic Approach: Experimental Visualization for Artistic Goal

Instead of using the machine's outputs as presentations directly, we value the collective voices from artists, intelligent machines, and audiences. We further develop the *Cangjie* system's output data with multiple experimental visualization strategies to accomplish our artistic goal.

Our visualization aesthetic resonates with twentieth-century Western abstract expressionism paintings and Eastern ink wash paintings. Abstract expressionist Willem de Kooning painted his black-and-white series by organizing the complex array of forms and shapes within a spatial structure using expressive brushwork to combine figuration with abstraction [de Kooning, 1948]. In the Eastern art world, Chinese ink wash paintings originated in the Tang dynasty (618–907 AD), using varying ink density, brush movement, and ink flow to control the tonality and shading and capture the spirit of the subject [Da-Wei, 2012]. Our development of machine-generated data has been inspired by the techniques and aesthetics of both Eastern and Western paintings.

We have designed the first interactive visualization of live streaming texture pixels to consistently write new symbols in an ink-flowing aesthetic (Figure 4). First, the image data of new symbols generated by the *Cangjie* system is manipulated and processed with image processing techniques, such as image differencing and alpha compositing, filter design, and data transformation. Then we use shaders written in OpenGL shading language (GLSL) to relocate the pixels from live streaming texture to a position determined by the image data of new symbols generated by *Cangjie*. The RGBA channels of live streaming image data control the density, point size, moving directions, velocity, and rotation of the pixels. Thus, we achieve the result of a monochrome pixelated fluid landscape of surroundings that evolves and transforms in the gesture of writing new symbols.

In the second visualization of the descriptive sentences generated by *Cangjie*, we combine the visualization of text with pictorial fragments of live streaming together to present a visual poetic assemblage in an ever-changing composition (Figure 5). Inspired by concrete poetry in the twentieth century, which arranges linguistic elements topographically in the space to convey meanings [Encyclopedia Britannica, 2019], we utilized the DenseCap model to determine the arrangement of sentences in the multiple regions of the visualization. We have designed the texts of the inaccurate results and the accurate results in the handwriting style with ink effects. The letters' size, spacing, and curvature are determined by the confidence score given by the DenseCap model. The most accurate sentences are in relatively small size with less spacing between the letters. In contrast, the inaccurate ones are relatively eye-catching with larger sizes, more spacing, much more organic shapes, and flowing animation. Portions of the live streaming are shown as pictorial fragments collaging with texts. The positions of the fragments are determined by the selected DenseCap model's detected regions of interest [Johnson et al., 2016]. The forms of the pictorial components are composed with the digital painted textures to create the quality of the ink wash painting. We design this interactive visualization using alpha channel compositing, image filter design, and GLSL. Following a number of image processing techniques and creative programming, the second visualization's outcome is an evolving collage of real-time video fragments with machine-generated texts in ink wash painting style.

Fig. 4. The first visualization generated by the *Cangjie* system: *Cangjie* writes poetry using pixels of live streaming textures, 2020. (©and Photo: Weidi Zhang)

Fig. 5. The second visualization generated by the *Cangjie* system: *Cangjie* describes surroundings as a visual form of poetry. (©and Photo: Weidi Zhang)

3.3 Interactivity: Exhibition Installation Design and Interactivity

Two visualizations are projected on two parallel screens. The screens' ideal size is 2.5 meters (height) by 7 meters (width) to engage the audiences in an immersive installation. We can vary the size of the screens depending on the exhibition space. A camera is set between the two screens to capture audiences and the environment (Figure 6). There are dimmed spotlights to illuminate the environment.

Fig. 6. The mock-up view of the *Cangjie's Poetry* art installation, 2020. (©and Photo: Weidi Zhang)

In the exhibition space, the audiences can be captured by the camera as live streaming in real time processed by the *Cangjie* system, which dynamically transforms the projections on the two screens. On the left screen, the audiences see how the imagery of them is transforming into new symbols and abstract landscapes in real time. When audiences move in front of the camera, their body movements can immediately change the symbols the *Cangjie* system produces and the ink flowing gestures, including pixels' moving directions, density, velocity, and tonality. The second screen on the right side features the visual poetry of the imagery generated by the system, where the descriptive sentence changes every 9 seconds. The audiences can see themselves and their surrounding environments in texture fragments that are dynamically changing as part of the visual poetry. The descriptive sentences float fluidly between the textures. We design the *Cangjie* system to generate new texts for visual poetry every 9 seconds, making it possible for the audiences to read the descriptive sentences. Audiences are engaged in the visualizations of new symbols and intelligent poetry by building semantic connections between reality and abstraction with empathy.

4 SPECIAL EDITION: COLLECTIVE MEMORIES AS INPUT SOURCES

As an artistic response to the COVID-19 pandemic and social distancing regulations, we also created a special edition of *Cangjie's Poetry*. In this version, instead of using real-time live streaming of surroundings in the art installation as a video feed for the *Cangjie* system, we collect and curate submissions of daily footages worldwide. We edit the submitted footage from Canada, China, the United Kingdom, the United States, Japan, Korea, and Egypt. *Cangjie* interprets the edited footage and generates the visualization based on its interpretation. We render the animation generated by *Cangjie* and design the sound using the technique of granular synth.

This special edition was planned to be projected on two large screens in an immersive gallery as part of a new media art festival in San Francisco. Unfortunately, the physical exhibition was canceled and moved to an online web-based gallery due to curfew restrictions. To accommodate the online format, we then redesigned the special edition from two separate projections on parallel screens to one animation that combined both visualizations as a new entity (Figure 7). The visualization's new format can be interpreted as an unfolding scroll of evolving Eastern calligraphy or an opened poetry book with changing thoughts generated by machines and humans.

Fig. 7. Special edition of *Cangjie's Poetry*. A screenshot of the animation combined with two visualizations is generated by the *Cangjie* system by interpreting the video footage submitted worldwide, 2020. (©and Photo: Weidi Zhang)

Compared to the original, dynamically interactive version, this special edition of visualization is a pre-rendered audiovisual experience. However, it travels and extends the vision of *Cangjie* to different places by collecting the footage worldwide. The content of those collective memories is more diverse than art installation environments, which generate visualizations with varying, engaging, and meaningful results. Meanwhile, the personal footage of daily life builds a new level of intimacy between humans and humans, and humans and machines.

5 CONCLUSION

Cangjie's Poetry transforms the machine's interpretation of surroundings into an interactive semantic experience. It constructs a conceptual and futuristic communication between humans and machines. We consider intelligent machines' creative decisions as part of collective voices with artists and audiences, which compose a dynamically evolving assemblage. This work is not only a conceptual response to the tension and fragility in the coexistence of humans and machines but also an artistic expression of a future language that reflects on ancient truths, a way to evoke enchantment in this era of artificial intelligence. Like the legendary historian Cangjie did nearly 5000 years ago, the *Cangjie* system continuously writes poetry with humans collaboratively as long as the real world exists.

ACKNOWLEDGMENTS

We would like to thank Prof. Marko Peljhan and Prof. Curtis Roads for their help and support, as well as the Experimental Visualization Lab and Media Arts and Technology program (University of California, Santa Barbara).

REFERENCES

- Cangjie: Chinese God of Writing, Literature, Chinese Characters. 2013. *Ancient Chinese Mythology*. Retrieved from <http://ancientchinesemythology.blogspot.com/2013/08/cangjie-chinese-god-of-writing.html>
- Chris Cheung. 2020. 封—墨池 I No Longer Write—Mochiji. *Vimeo*, uploaded by Chris Cheung, April 2020. www.vimeo.com/406793526
- Kwo Da-Wei. 2012. *Chinese Brushwork in Calligraphy and Painting: Its History, Aesthetics, and Techniques*. Courier Corporation, 2012.
- Willem de Kooning. 1948. *Painting*. MoMA collection. www.moma.org/collection/works/79242
- Jeff Donahue, Philipp Krähenbühl, and Trevor Darrell. 2016. Adversarial feature learning. arXiv preprint. Retrieved from <https://arxiv.org/abs/1605.09782>
- Encyclopedia Britannica (Eds.) 2019. Concrete poetry. Encyclopedia Britannica. <https://www.britannica.com/art/concrete-poetry>

- Sarah E. Fraser and Yu-Chieh Li (Eds.). 2020. *Xu Bing: Beyond the Book from the Sky*. Springer Nature.
- Justin Johnson, Andrej Karpathy, and Li Fei-Fei. 2016. DenseCap: fully convolutional localization networks for dense captioning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, 4565–4574.
- Wikipedia. 2021. Cangjie (2021, June 12). Retrieved from Wikipedia: the Free Encyclopedia. <https://en.wikipedia.org/wiki/Cangjie>