Urban Symphony: An AI and Data-Driven Approach to Real-Time Animation for Public Digital Art

Rem RunGu Lin
Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
rlin408@connect.hkust-gz.edu.cn

Koo YongEn Ke
Funtheory Studio
Guangzhou, China
corrie@pineberry.cc

Kang Zhang*
Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
kzhangcma@ust.hk

ABSTRACT

Projection mapping is a form of urban public art that uses light and video to transform buildings and structures into dynamic canvases. However, producing high-quality projection mapping content with compelling storytelling requires extensive time and resources, as it involves integrating local culture, urban spatial understanding, and animation production. To address this challenge, this paper proposes a method that combines artistic co-creation with AI, audio-visualization, and data-visualization techniques. The authors present a case study: "Urban Symphony," an immersive public art installation that showcases our method and leverages AI and data-driven storytelling. This method fosters interdisciplinary research collaboration and explores the potential of projection mapping as a bridge between art, technology, and society. The paper describe the motivation, design, and production of the artwork, the outcomes of the performance, and the challenges and limitations of our method. The authors also suggest future directions for further improvement and exploration in this domain.

CCS CONCEPTS

• **Human-centered computing** → **Visualization**; • **Applied computing** → *Arts and humanities*; • **Computing methodologies** → Artificial intelligence.

KEYWORDS

Audio-visualization, data-visualization, co-creative AI, animation production, projection mapping, urban public digital art

ACM Reference Format:
Rem RunGu Lin, Koo YongEn Ke, and Kang Zhang*. 2023. Urban Symphony: An AI and Data-Driven Approach to Real-Time Animation for Public Digital Art. In *The 16th International Symposium on Visual Information Communication and Interaction (VINCI 2023), September 22–24, 2023, Guangzhou, China.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3615522.3615553

VINCI 2023, September 22–24, 2023, Guangzhou, China

ACM ISBN 979-8-4007-0751-3/23/09…$15.00
https://doi.org/10.1145/3615522.3615553

1 INTRODUCTION

1.1 Background

Projection mapping has emerged as an engaging form of urban public art, garnering widespread attention. It utilizes light and video to metamorphose architectural structures into dynamic canvases that pulse with life, enabling artistic expression and collaboration on an urban scale. By transforming cityscapes and nurturing community engagement, projection mapping adds a vibrant layer to the urban fabric.

However, crafting visually appealing and contextually appropriate 3D content for projection mapping poses significant challenges. The complexity of this creative process often demands substantial investment in terms of time and resources, as well as necessitates specialized skills.

In this paper, we address these challenges by proposing a method for the design and production of real-time animation for projection mapping through AI and data-driven animation. We leverage Stable Diffusion to streamline the concept design phase, and utilize Unreal Engine to facilitate the production of the animation. Within the Unreal Engine framework, we employ an audio spectrum to animate 3D objects, and harness local data visualization as a storytelling tool, thereby creating a more efficient production pipeline. By integrating AI, audio-visualization, and data-visualization into the artistic process, we seek to transcend the constraints of traditional techniques and push the boundaries of what projection mapping can achieve in the realm of urban public art.

1.2 Contributions

The main contributions of our research are:

- We present a method for the design and production of real-time animation for projection mapping by integrating artistic co-creative AI, audio-visualization, and data-visualization. This methodology offers a more efficient and cost-effective production of projection mapping content.
- We provide a case study, "Urban Symphony," exemplifying the creative potential of AI and data-driven storytelling in real-time animation for urban public art. This example illustrates how AI and data-driven techniques can be applied to urban public art, showcasing the potential of such approaches to enrich the urban landscape and engage the community in thought-provoking experiences.

Figure 1: "Urban Symphony" being displayed during the opening ceremony of BBUA 2023 (© Rungu Lin 2023).

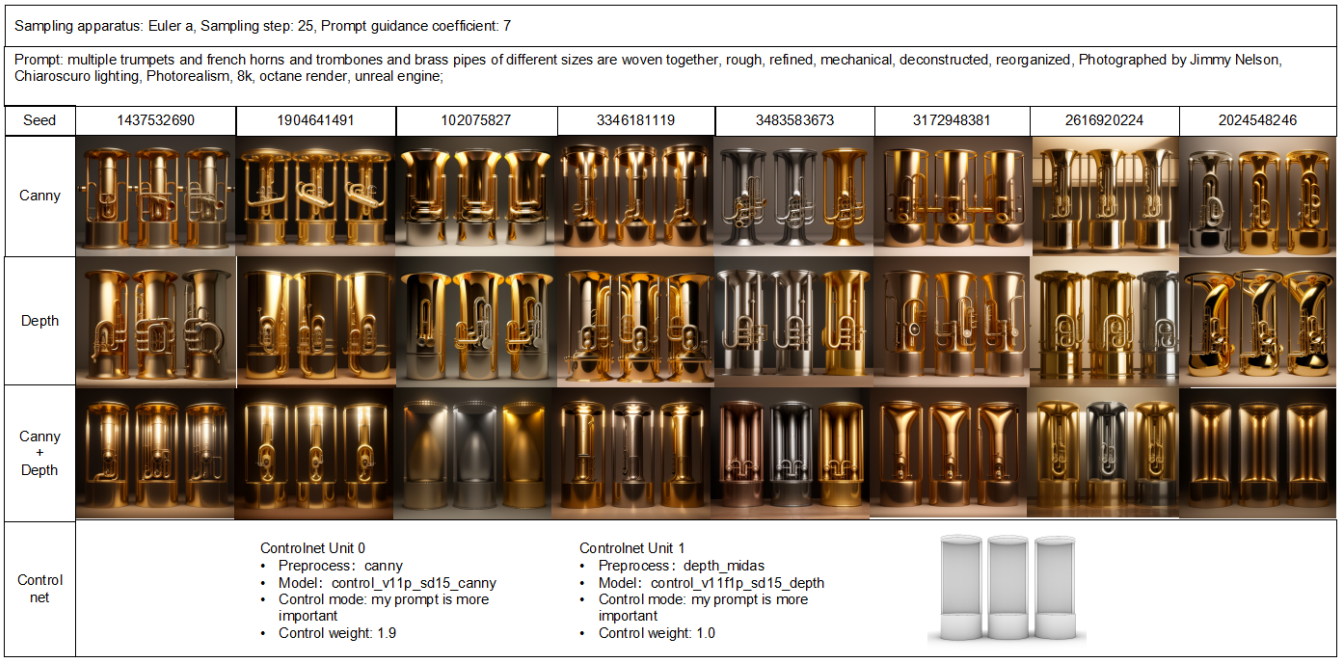

Figure 2: The results generated by different combination of ControlNet Units (© Rungu Lin 2023).

2 RELATED WORK

The related work for this research can be categorized into three main areas: AI-driven art, audio-visualization, and data-visualization. This section provides a brief overview of these areas

and highlights some key studies and projects that have influenced the development of the "Urban Symphony" project.

AI-driven Art: AI-driven art refers to any artwork, especially images and animations, created through the use of AI programs, such as genetic algorithms, machine learning, or text-to-image models [6][7]. This interdisciplinary approach has resulted in numerous innovative projects that challenge conventional artistic practices and explore the creative potential of AI. For example, Anadol's "Machine Hallucination" [4] uses machine learning algorithms to generate dynamic visualizations based on large-scale data sets, offering an immersive and thought-provoking experience. The use of AI in art has also been explored in the context of human-computer interaction, with researchers developing conceptual frameworks and methodologies for artistic human-computer collaboration [16]. Moreover, some recent projects have employed AI to create realistic images and art from natural language descriptions, such as DALL·E 2 [13], which is an AI system that can create realistic images and art from a description in natural language; and Stable Diffusion [2], which is a latent text-to-image diffusion model capable of generating photo-realistic images given any text input.The text-to-image AI tool shows great potential and is gaining increasing attention today. However, its application to projection mapping and the creation of an effective iteration between humans and AI remain unexplored.

Data-Visualization: Data-visualization is the process of representing data or information in graphical forms, such as charts, maps, or diagrams. Data-visualization can be used for various purposes, such as communicating insights, revealing patterns or trends, or supporting decision making. Data-visualization can be achieved by using different methods and techniques, such as selecting appropriate data types and formats, choosing suitable visual encodings and layouts, or applying interactive features or animations [14]. For example, Lindner [11] analyzed how data-visualization can be used to visualize urban change in modern New York by using historical photographs and maps. Another example is Lundman [12] who explored how site-specific video can be used as a method for participatory urban planning by engaging citizens in co-creating urban visions. Moreover, some projects have used data-visualization to create artistic expressions, such as Anadol's "Wind of Boston" [3], which is a data sculpture that visualizes wind patterns in Boston in LED screen. Nonetheless, the use of data-visualization as a storytelling tool within the scope of projection mapping in public art needs further exploration.

Audio-Visualization: Audio-visualization is the process of transforming sound into visual forms, such as images, animations, or graphs. Audio-visualization can be used for various purposes, such as enhancing musical expression, facilitating sound analysis, or creating artistic experiences. Audio-visualization can be achieved by using different methods and techniques, such as mapping sound features to visual parameters, applying mathematical transformations or algorithms, or using machine learning models to learn associations between sound and image domains [9][1]. For example, Rubab et al. [15] proposed a novel method for exploring effective relationships between visual-audio channels in data visualization based on perceptual principles and user preferences. Another example is Specterr (2022), which is an online music visualizer that allows users to create stunning videos with sound-responsive elements and atmospheric 3D scenes. Additionally, some projects have focused on the use of urban soundscape data for artistic creation, such as URB [8], which collects and stores raw data from soundscapes analysis and provides them to artists for reinterpretation. Yet, the application of audio-driven techniques in animated storytelling to enhance audience immersion is an area that remains untapped.

This research builds upon the existing literature and projects in the fields of AI-driven art, audio-visualization, and data-visualization. By integrating these interdisciplinary approaches, the "Urban Symphony" project seeks to transform the animation production pipeline from a human-driven process to a human-AI-data co-driven approach, creating a captivating and immersive urban public art installation. The method developed in this research offers a comprehensive framework for interdisciplinary collaboration, enabling artists, technologists, and urban planners to work together in the creation of innovative public art installations that engage the public and enrich the urban landscape.

3 URBAN SYMPHONY

"Urban Symphony" is a real-time animated art installation displayed during the opening ceremony of the Shenzhen-Hong Kong Bi-City Biennale of Urbanism&Architecture 2023 (Figure 1). The exhibition canvas is an industrial heritage site in central Shenzhen, specifically using nine silos as the art medium. Ten projectors employing projection mapping technology illuminate the silo facades, meticulously aligning visuals to create a grand spectacle.

"Urban Symphony" is segmented into three chapters, each represented by geometric elements, city buildings, and brass instruments. These chapters progressively narrate Shenzhen's evolution - its early construction, urban development, and cultural prosperity. The geometric elements, abstract expressions derived from architectural structures and brass instruments, symbolize the core elements of the city's development. Iconic buildings selected represent the city's architectural achievements, and the brass instruments, deconstructed and reassembled, symbolize the cultural vibrancy of the city. In an immersive setting, "Urban Symphony" forms a semi-enclosed space, inviting audiences to engage both visually and auditorily. It encourages viewers to reflect on the city's evolution and their individual place within this vibrant metropolis.

By incorporating co-creative AI, audio-visualization, and data-visualization techniques, "Urban Symphony" epitomizes the creative potential of interdisciplinary collaboration in urban public art. This installation showcases an approach to real-time animation production, demonstrating the potential of human-AI-data co-driven processes in creating public art installations. It delivers a captivating, engaging experience that resonates with the community and contributes to the city's cultural heritage, adding a rich layer to the urban landscape.

4 4.CONCEPTUAL FRAMEWORK AND METHODOLOGY

4.1 Co-creative AI in concept design for projection mapping animation

In the development of "Urban Symphony," we harness the capabilities of co-creative AI, specifically the AIGC tool, Stable Diffusion, to

facilitate the conceptual design for projection mapping animation. This approach fosters the integration of AI-generated content, ensuring visual continuity and adaptability to the physical structure of the projected building - an essential requirement for projection mapping [10].

Our methodology begins by using the image of the section of the silos as the reference for ControlNet. In our trials, we experimented with a combination of two ControlNet Units: The first ControlNet Unit consisted of a canny edge detection preprocessing step, followed by the `Control_v11p_sd15_canny` model; the second ControlNet Unit incorporated a `depth_midas` preprocessing phase and the `control_v11f1p_sd15_depth` model (Figure 2). After comparing the results generated by each individual model and the combined models, we observed:

- The structure expression of the silos' sections was more accurately portrayed when using the first ControlNet Unit (canny) alone or in combination with the second ControlNet Unit (depth) than when using the second ControlNet Unit (depth) alone.
- The richness of the generated results was superior when using the first ControlNet Unit (canny) or the second ControlNet Unit (depth) alone, compared to the overlaid results of both models.

Based on these findings, we opted for using just one ControlNet Unit—canny edge detection preprocessing with the `Control_v11p_sd15_canny` model—for concept design generation. Post-generation, numerous images are evaluated by the author. Selected images are collaged and adjusted, then re-input into the system (as Img2) for subsequent iterations. This process is iteratively executed until a satisfactory final concept image (ImgN) is realized (Figure 3).

Figure 3: The approach to select and collage images for subsequent iterations (© Rungu Lin 2023).

This co-creative AI procedure involves repetitive cycles of human input, AI generation, and human evaluation, allowing the artist to focus on overarching creative decisions, while the AI system proffers novel visual concepts and combinations (Figure 4). Furthermore, this methodology promotes an efficient and adaptive design process, substantially reducing the time and effort spent on manual adjustments and refining visual elements (Figure 4).

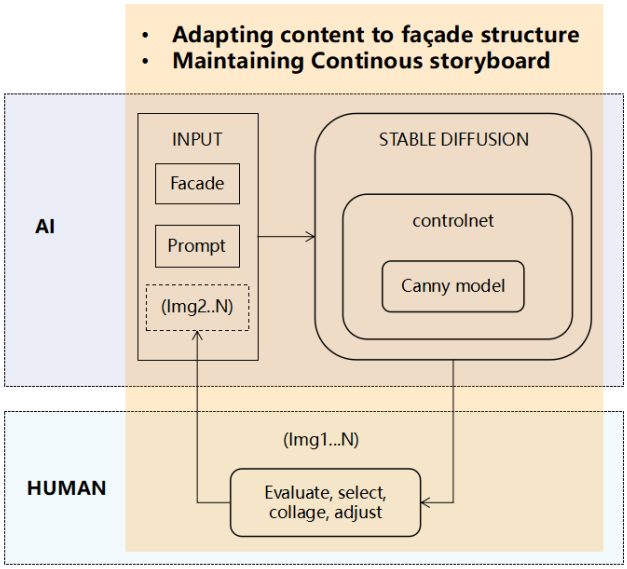

Figure 4: The approach of co-creative AI in concept design for projection mapping animation (© Rungu Lin 2023).

4.2 Data-visualization for animation production

In "Urban Symphony," data-visualization techniques are employed to tell the story of Shenzhen's urban development through the capabilities of Unreal Engine's particle system. By combining data-driven visualizations with artistic expression, the animation provides an engaging and informative representation of the city's history and economic progress. This approach allows the audience to better understand and appreciate the dynamic nature of Shenzhen's urban transformation.

Shenzhen's urban development history is characterized by the rapid rise of skyscrapers, symbolizing the city's remarkable progression. The city's transformation from a small fishing village in the 1970s to a bustling metropolis today is a testament to its dramatic growth. Following China's economic reforms, Shenzhen was designated as the country's first Special Economic Zone in 1980, spurring a period of unprecedented urbanization and economic development. The city soon emerged as a global hub for technology, finance, and innovation. We carefully selected 9 significant landmark buildings constructed between 1984 and 2020 to represent this astonishing transformation [17]. These landmarks were arranged horizontally, appearing in chronological order to emphasize the passage of time and showcase the city's evolving skyline (Figure 5).

The height of each building was scaled proportionally to its actual height, maintaining an accurate visual representation of their relative sizes. This attention to detail ensures that the audience can grasp the true scale of Shenzhen's urban growth and appreciate the architectural achievements that define the city's landscape.

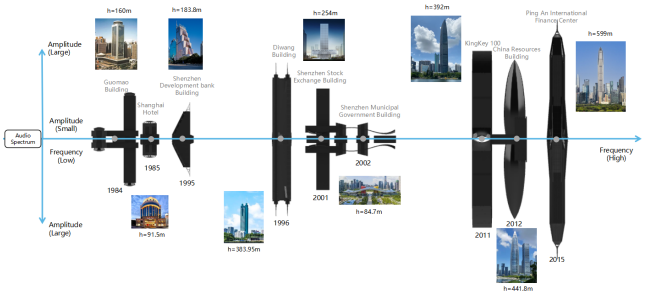

Figure 5: The landmarks of Shenzhen being arranged in chronological order and remapped according to audio spectrum (© Rungu Lin 2023).

As a prominent player in China's reform and opening up, Shenzhen's remarkable economic advancement serves as a testimony to its national significance. In this context, the author has procured official data reflecting Shenzhen's urban GDP from 1987 to 2022, directly from the Shenzhen Municipal Bureau of Statistics website http://tjj.sz.gov.cn/. This data forms the basis for driving the particle generation rate within the animation.

A unique approach is adopted where the city's GDP determines the quantity of particles generated, thus translating complex economic data into a visually engaging narrative. Concurrently, the particle velocity mirrors the gradient of GDP growth – the steeper the slope, the higher the velocity of the particles (Figure 6). This creative interpretation and representation of data offer an engaging mechanism to visually narrate the city's economic evolution, highlighting the tangible impact of its growth on the urban landscape. Through this innovative visualization technique, "Urban Symphony" provides a compelling perspective on Shenzhen's economic development, adding depth to its storytelling.

By integrating data-visualization techniques in the animation production pipeline, "Urban Symphony" offers a visually captivating representation of Shenzhen's development, demonstrating the potential of interdisciplinary collaboration in the creation of meaningful public art installations. This approach not only enriches the urban space but also fosters a deeper understanding of the city's history and development, ultimately contributing to the cultivation of a more engaged and informed community.

4.3 Audio-visualization for animation production

In "Urban Symphony," the implementation of audio-visualization techniques forms a critical component, enabling the creation of an immersive experience. By allowing the animation to synchronize with the auditory component, these techniques enhance the overall perception and reception of the installation, forging a powerful bond between the audience and the artwork [5]. The effective application of audio-visualization contributes significantly to the visual impact of the installation.

The initial phase of the performance deploys audio-driven animation, in which the scale of diverse abstract objects corresponds directly to the audio amplitude across various frequencies. For example, the high-frequency range corresponds to the scale of triangular geometrical bodies, mid-frequency range associates with circles, and low-frequency range with squares. Sliced and layered, these objects are mapped in order to the spectrum, yielding a dynamic visual representation of sound. This technique, relying on sound input to generate motion and scale, results in captivating visuals that underline the intricate interplay between auditory and visual elements.

During the second phase of the performance, landmark buildings of Shenzhen are arranged linearly, mirrored along the x-axis, and based on their construction timeline. The audio spectrum, a visual representation of the sound frequency and amplitude, is utilized to generate sound-responsive visuals. The buildings are mapped to the audio spectrum along the x-axis corresponding to their specific sound frequency, and their vertical displacement along the y-axis is governed by the music amplitude at distinct frequencies. This audio-reactive visualization fosters a multisensory experience, underscoring Shenzhen's historical development and its key role in China's urbanization.

The final phase of the performance centers on the motion of brass instruments. The amplitude of the music at various frequencies controls the movements of their pistons and tuning slides. In addition, the instrument bells expand and contract according to the brass section's frequency, reinforcing the synchronization of audio and visual elements. This stage underscores the artistic potential of uniting traditional musical components with cutting-edge animation techniques, offering a captivating experience for spectators (Figure 7).

By integrating audio-visualization techniques into the animation production process, "Urban Symphony" delivers an engaging, immersive audience experience. It thereby illustrates the potential for innovative artistic expression in public digital art installations.

5 CONCLUSIONS AND FUTURE WORK

"Urban Symphony" exemplifies the integration of co-creative AI, audio-visualization, and data-visualization into the animation production pipeline. This approach reshapes the conventional, human-centric process into a harmonious blend of human creativity, AI capabilities, and data insights. The outcome is an intriguing narrative that offers a glimpse into Shenzhen's urban evolution and its significance within the context of China's policy of reform and opening up. This multidisciplinary endeavor embellishes the urban landscape and invites the community to partake in an enriching experience, hinting at potential future directions for public art installations.

It is important to note, however, that the efficacy of this approach is yet to be empirically verified. For this reason, future research is proposed to involve comprehensive quantitative user studies among animation designers and audiences.

Moreover, we propose to enhance the co-creative AI process by incorporating real-time audience interaction. This could potentially elevate the immersive experience and deepen the connection between the public and the artwork.

Additionally, the integration of various types of urban data, such as environmental and social media data, could offer fresh perspectives on city development and create innovative visual narratives. These alternative data sources could potentially provide unique

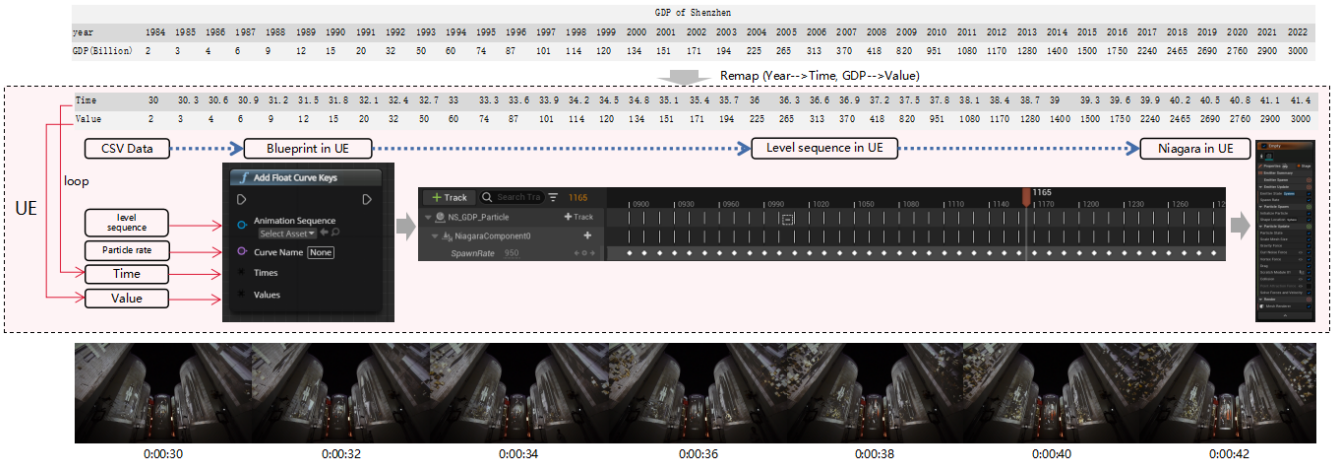

Figure 6: The approach to use to GDP data to drive the particle animation (© Rungu Lin 2023).

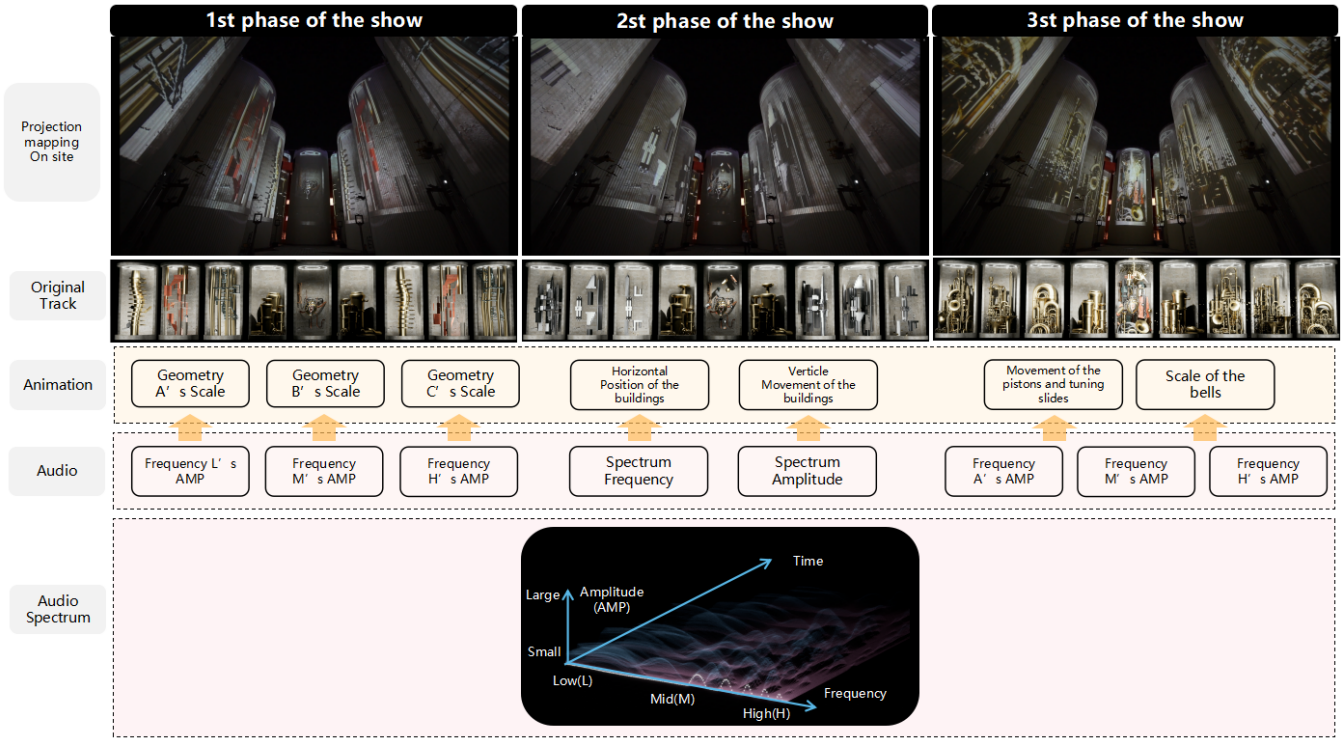

Figure 7: The approach to use the data from audio spectrum, amplitude and frequency to drive the animation (© Rungu Lin 2023).

insights into the social, cultural, and ecological aspects of urban life, thereby broadening the array of stories that can be conveyed through public art installations.

While "Urban Symphony" presents a novel approach to creating urban public art, we recognize the field of AI-driven art is constantly evolving. The potential for new forms of artistic expression and public engagement remains vast. Researchers, artists, and urban planners can collaborate, pushing the boundaries of what is possible, to craft meaningful experiences that resonate with diverse communities and contribute to the ongoing conversation surrounding urban development and transformation. However, further empirical investigations and user studies are needed to substantiate and refine these innovative methods.

REFERENCES

[1] Matthew Joseph Adiletta and Oliver Thomas. 2020. An artistic visualization of music modeling a synesthetic experience. *arXiv preprint arXiv:2012.08034* (2020).

[2] Stability AI. 2022. *Stable Diffusion.* https://stability.ai/stable-diffusion

[3] Refik Anadol. 2017. *Wind of Boston: Data Paintings.* https://refikanadol.com/works/wind-of-boston-data-paintings/

[4] Refik Anadol. 2019. *Machine Hallucination.* https://refikanadol.com/works/machine-hallucination/

[5] Matthew N Bain. 2008. *Real time music visualization: A study in the visual extension of music.* Ph. D. Dissertation. The Ohio State University.

[6] Margaret A Boden. 2009. Computer models of creativity. *AI Magazine* 30, 3 (2009), 23–23.

[7] Marcus Du Sautoy. 2019. The creativity code. In *The Creativity Code.* Harvard University Press.

[8] José Alberto Gomes, Nuno Peixoto de Pinho, Filipe Lopes, Gustavo Costa, Rui Dias, Diogo Tudela, and Álvaro Barbosa. 2014. Capture and transformation of urban soundscape data for artistic creation. *Journal of Science and Technology of the Arts* 6, 1 (2014), 97–109.

[9] Thomas Hermann, Andy Hunt, and John G Neuhoff. 2011. *The sonification handbook.* Vol. 1. Logos Verlag Berlin.

[10] Katerina Korola. 2014. Probing Light: Projection Mapping, Architectural Surface, and the Politics of Luminous Abstraction. *Intermédialités* 24 (2014).

[11] Christoph Lindner. 2013. After-Images of the Highrise City: Visualizing Urban Change in Modern New York. *The Journal of American Culture* 36, 2 (2013), 75–87.

[12] Riina Lundman. 2016. Bringing planning to the streets: using site-specific video as a method for participatory urban planning. *Planning Theory & Practice* 17, 4 (2016), 601–617.

[13] OpenAI. 2022. *DALL·E 2.* https://openai.com/product/dall-e-2

[14] Xuedi Qin, Yuyu Luo, Nan Tang, and Guoliang Li. 2020. Making data visualization more efficient and effective: a survey. *The VLDB Journal* 29 (2020), 93–117.

[15] Sadia Rubab, Lingyun Yu, Junxiu Tang, and Yingcai Wu. 2023. Exploring Effective Relationships Between Visual-Audio Channels in Data Visualization. *Journal of Visualization* (2023), 1–20.

[16] Uwe Seifert and Jin Hyun Kim. 2008. Towards a conceptual framework and an empirical methodology in research on artistic human-computer and human-robot interaction. In *Human Computer Interaction.* IntechOpen.

[17] Cong Sun and Charlie QL Xue. 2020. Shennan Road and the modernization of Shenzhen architecture. *Frontiers of Architectural Research* 9, 2 (2020), 437–449.