

실시간 음악 시각화를 적용한 생성형 AI 의상 디자인 연구

요 약

본 연구는 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인 연구에 관한 것이다. 연구의 목적은 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인의 가능성을 탐구하고, 이를 위한 방법론을 제시하는 데 있다. 연구의 배경은 실시간 음악 시각화와 생성형 AI 의상 디자인의 융합에 있다. 연구의 방법론은 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인을 위한 방법론을 제시하는 데 있다. 연구의 결과는 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인의 가능성을 탐구하고, 이를 위한 방법론을 제시하는 데 있다. 연구의 결론은 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인의 가능성을 탐구하고, 이를 위한 방법론을 제시하는 데 있다.

본 연구는 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인 연구에 관한 것이다. 연구의 목적은 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인의 가능성을 탐구하고, 이를 위한 방법론을 제시하는 데 있다. 연구의 배경은 실시간 음악 시각화와 생성형 AI 의상 디자인의 융합에 있다. 연구의 방법론은 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인을 위한 방법론을 제시하는 데 있다. 연구의 결과는 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인의 가능성을 탐구하고, 이를 위한 방법론을 제시하는 데 있다. 연구의 결론은 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인의 가능성을 탐구하고, 이를 위한 방법론을 제시하는 데 있다.

본 연구는 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인 연구에 관한 것이다. 연구의 목적은 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인의 가능성을 탐구하고, 이를 위한 방법론을 제시하는 데 있다. 연구의 배경은 실시간 음악 시각화와 생성형 AI 의상 디자인의 융합에 있다. 연구의 방법론은 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인을 위한 방법론을 제시하는 데 있다. 연구의 결과는 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인의 가능성을 탐구하고, 이를 위한 방법론을 제시하는 데 있다. 연구의 결론은 실시간 음악 시각화를 적용한 생성형 AI 의상 디자인의 가능성을 탐구하고, 이를 위한 방법론을 제시하는 데 있다.

I. 서론

최근 몇 년 동안, 인공지능(AI) 기술의 발전은 디자인 분야에 혁명적인 변화를 가져왔다. 특히, AI는 디자인 프로세스의 효율성을 높이고, 창의적인 아이디어를 생성하는 데 크게 기여하고 있다. 본 연구는 AI가 디자인에 어떻게 활용되는지를 탐구하고, 이를 통해 디자인의 질과 생산성을 어떻게 향상시킬 수 있는지를 분석한다.

본 연구의 주요 목적은 AI를 디자인에 적용하는 방법과 그 효과를 분석하는 것이다. 이를 위해, AI가 디자인에 활용되는 다양한 분야(예: 그래픽 디자인, 패션 디자인, 제품 디자인 등)를 살펴보고, 각 분야에서 AI가 어떻게 활용되는지를 분석한다. 또한, AI를 디자인에 적용하는 데 따른 장단점을 분석하고, 이를 통해 디자인의 질과 생산성을 어떻게 향상시킬 수 있는지를 분석한다.

본 연구의 주요 연구 질문은 다음과 같다:

- 1. AI가 디자인에 활용되는 방법과 그 효과는 무엇인가?
- 2. AI를 디자인에 적용하는 데 따른 장단점은 무엇인가?
- 3. AI를 디자인에 적용하여 디자인의 질과 생산성을 어떻게 향상시킬 수 있는가?

본 연구의 결과는 AI를 디자인에 활용하는 방법과 그 효과를 분석하고, 이를 통해 디자인의 질과 생산성을 어떻게 향상시킬 수 있는지를 분석하는 데 크게 기여할 것으로 기대된다.

II. 이론적 배경 및 선행 연구

본 연구는 AI가 디자인에 활용되는 방법과 그 효과를 분석하고, 이를 통해 디자인의 질과 생산성을 어떻게 향상시킬 수 있는지를 분석하는 데 크게 기여할 것으로 기대된다.

(Ramesh et al., 2022; Rombach et al., 2022), AI (Choi et al., 2023).

1. 음악 시각화

1) 음악 시각화의 개념과 발전

음악 시각화는 음악의 특성을 시각적으로 표현하는 것을 의미한다. 이는 음악의 리듬, 멜로디, 화성 등을 시각적으로 표현하여, 음악의 감정을 전달하고, 음악의 특성을 시각적으로 표현하는 것을 의미한다.

가 Biner et al.(2024)
「TV(Participation TV)」(1963/1998) SonicDiffusion Xing et al.(2024) Seeing and
Hearing
가 , Huang et
(Paik, 1963/1998). Foote(1999) al.(2025) (MIR),
, Levin(2000)
<Table 1>
. Taylor et al.(2006)
, 가 , 가
(mapping)
, Perrotin and d’Alessandro(2014)

2) 음악 시각화와 의상 디자인

가 (cross-modal) . Kim and Kim(2014)
. Su and Ling(2023)
(Alexander McQueen) 2014 S/S(spring
and summer)
, Reymore and Lindsey(2025) , Kim and Kim(2016)
(Giorgio Armani) 2000 S/S

Table 1. 음악 시각화 선행 연구의 음악 요소·시각 요소·시각화 형식.

Paik (1963/1998)		, TV	
Foote (1999)	,	(matrix),	
Levin (2000)	,	,	
Taylor et al. (2006)		, 가	가 ,
Perrotin & d’Alessandro (2014)	,		
Su & Ling (2023)	, ,	, , , ,	TouchDesigner
Reymore & Lindsey (2025)	(timbre)		
Biner et al. (2024)			
Xing et al. (2024)	,	,	
Huang et al. (2025)	+	,	AI (MIR + LLM +)

가

,

,

2. 음악 분석

가 <S.T.A.Y.>

AI

, ,

(Akkerman, 2023; Richardson

•

‘ ‘

1. 연구 설계 및 절차

,

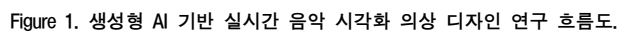


Table 2. 음악 구간별 주요 음악 요소 특성 분석.

(:)					
0:00 - 1:30	, , ,	, ,	, ,	, ,	, ,
1:30 - 2:50	, ,	가 , ,	, , 가	, ,	, ,
2:50 - 4:00	, , , ,	, ,	, , , ,	, ,	, ,
4:00 - 6:30	, , , ,	, ,	, ,	, ,	, ,

Figure 2. (Figure 2).

『Interstellar』 (Spectrogram) (Figure 2).

AI (OpenAI ChatGPT) (Elliot & Maier, 2014; Ou et al., 2004), GPT “Music Master: Song Analysis & Composition Insights” (Valdez & Mehrabian, 1994).

(0:00-1:30), (1:30-2:50), (2:50-4:00), (4:00-6:30)

<Table 2>

3. 무드보드 및 스타일 방향 설정

『Interstellar』 가 가 가 (moodboard) 가

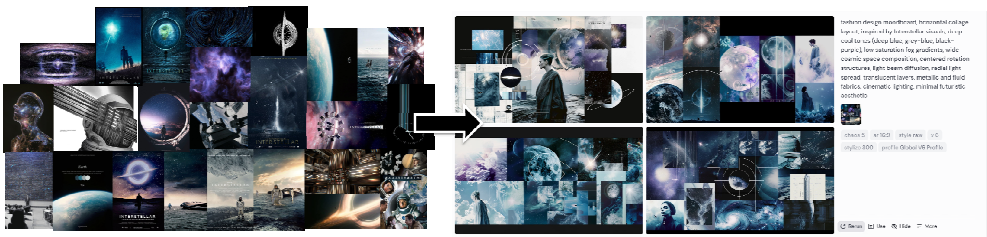


Figure 2. Midjourney를 활용한 무드보드 생성 과정.
From online sources and generated using Midjourney; composed by the author. (April 17, 2025)

Table 3. 음악 구간별 대표 색상 설정 기준.

(:)			Line MAT RGB(0-1)	
0:00-1:30	,	,	0.34, 0.58, 0.62	188, 0.45, 0.62
1:30-2:50	,	, 가	0.37, 0.35, 0.78	243, 0.55, 0.78
2:50-4:00	,	,	0.58, 0.22, 0.86	274, 0.74, 0.86
4:00-6:30	,	,	0.82, 0.94, 1.00	200, 0.18, 1.00

,	.			
.	Line MAT RGB	Radius X=2, Y=6, Divisions=50		
0-1	,			
	Line MAT			
RGB	.	audioAnalysis CHOP		
	(hue angle),	(low), (mid), (high),		
	.	(spectralCentroid), (RMS Power)		
<Table 3>	.	, (kick) (snare)		

4. TouchDesigner 기반 실시간 시각화 구현 Limit CHOP Filter CHOP

.		Line MAT RGB	.
,	.	,	,
.	.		
.	,	가	,
,	가	가	.
.			
Circle SOP	, Radius Divisions	<Table 4>	.

Table 4. 음악 시각화 요소와 TouchDesigner 수치 변화 매핑 기준.

		Line MAT / RGB	<Table 3>	
	kick	Switch TOP / Index	0 1	
	snare	Twist SOP / Strength	0 1	가
	spectral Centroid	Noise SOP / Period	-1-1	
	low (20-250 Hz)	Noise SOP / Amplitude	0-1	가
	mid (250-2000 Hz)	Noise SOP / Offset	0-1	가
	high (2000-20000 Hz)	Particle SOP / Time Inc	0-1	가
	RMS Power	Circle SOP / Anchor U	0-1	가
	RMS Power	Transform SOP / Uniform Scale	0-1	

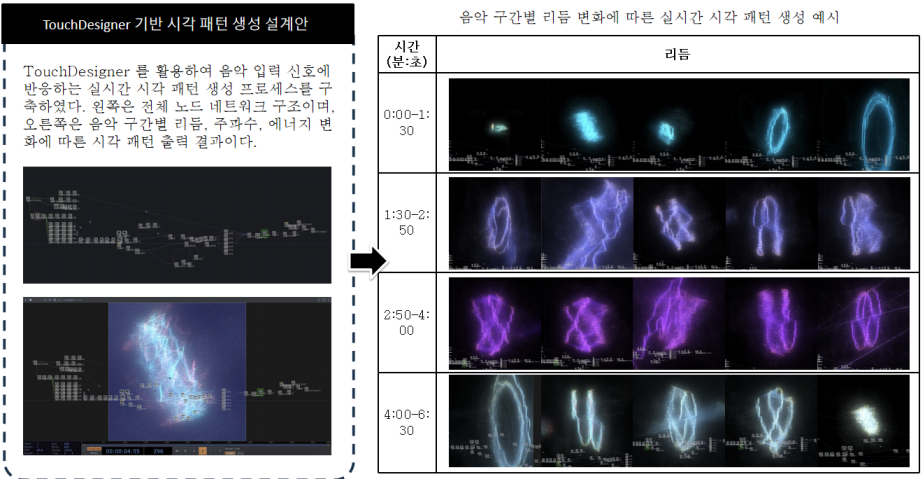


Figure 3. TouchDesigner를 활용한 음악 반응형 시각 패턴 생성.
Made by the author. (April 10, 2025)

(Figure 3).

AI

<Table 5>

0:00-1:30	， ， ，	high-collar wrap coat, layered front structure, minimal futuristic silhouette, vertical line
1:30-2:50	， 가, ，	sleeveless hooded A-line dress, smooth panel composition, enveloping upper form, soft expansion
2:50-4:00	， ， ，	long A-line dress, precise panel layout, architectural silhouette, minimal vertical seam lines
4:00-6:30	， ， ，	sculptural wrap dress, X-shaped panel structure, geometric silhouette, overlapping front panels

IV. 연구 결과 및 논의

1. 생성형 AI 기반 의상 이미지 생성

Midjourney V7, raw mode, 2:3, seed 1234, 20, 80, <Table 6>

Table 6. 음악 구간별 의상 이미지 생성 프롬프트 및 결과.

	prompt	
0:00-1:30	full-body front view fashion model wearing a high-collar structured wrap coat with layered construction, made from thick matte white cotton fabric, smooth matte surface, no pattern, no wrinkles, paired with black ankle boots, light gray background, soft studio lighting, futuristic minimalist style	 Figure 4. 도입부 의상 이미지. Screenshot by the author. (April 12, 2025)
1:30-2:50	full-body front view fashion model wearing a sleeveless hooded A-line dress with structured silhouette and smooth panel construction, made from thick matte white cotton fabric, smooth matte surface, no pattern, no wrinkles, paired with black ankle boots, light gray background, soft studio lighting, futuristic minimalist style	 Figure 5. 전개부 의상 이미지. Screenshot by the author. (April 12, 2025)
2:50-4:00	full-body front view fashion model wearing a long A-line dress with precise panel construction and architectural silhouette, made from thick matte white cotton fabric, smooth matte surface, no pattern, no wrinkles, paired with black ankle boots, light gray background, soft studio lighting, futuristic minimalist style	 Figure 6. 절정부 의상 이미지. Screenshot by the author. (April 12, 2025)

	prompt	
4:00-6:30	full-body front view fashion model wearing a sculptural wrap dress with X-shaped panel construction and geometric silhouette, made from thick matte white cotton fabric, smooth matte surface, no pattern, no wrinkles, paired with black ankle boots, light gray background, soft studio lighting, futuristic minimalist style	 Figure 7. 완화부 의상 이미지. Screenshot by the author. (April 12, 2025)

2. 의상 이미지와 동적 패턴 결합

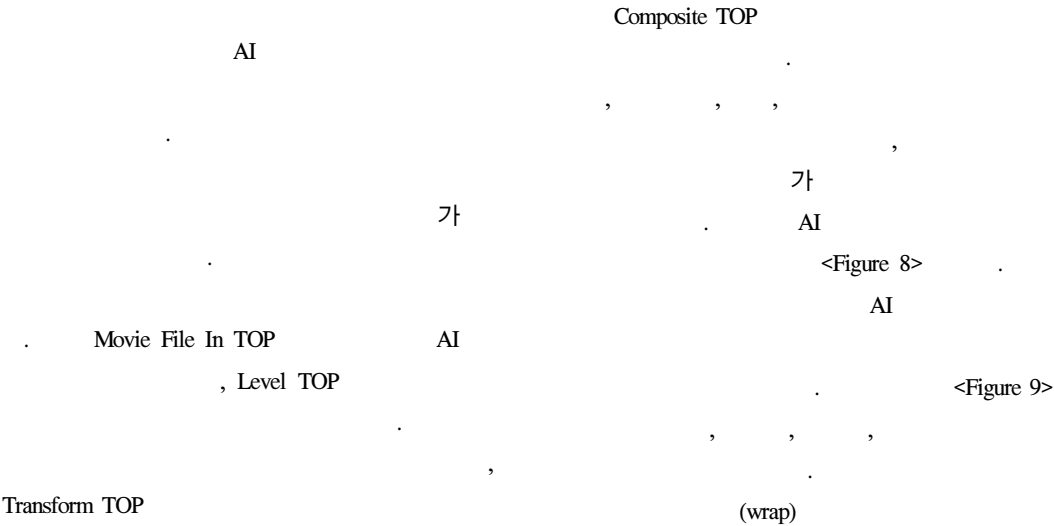


Figure 8. 음악 시각화와 디지털 의상 이미지의 결합.
Made by the author. (April 20, 2025)



Figure 9. 음악 구간별 시각 패턴이 적용된 디지털 의상 이미지 결과.
Made by the author. (April 20, 2025)

가 .833-.899 .80 가
(Table 7).
가 가 M=4.35
A M=4.62
가 , M=4.55
M=4.70,
M=4.80
AI
3. 전문가 평가 가
-AI M=4.18,
가 M=4.05 ,
-AI
가 가 . 가
가 가 ,
M=4.80,
10 M=4.40,
가 -AI , M=4.00 가
가 ,
가 4 16 가
5 (Likert scale)
Cronbach's α .967, M=3.50 가

Table 7. 전문가 평가 차원별 신뢰도 및 기술통계.

가	Cronbach's α	(M)	(SD)
-AI	.868	4.18	0.46
	.899	4.55	0.44
	.896	4.62	0.43
가	.833	4.05	0.48
	.967	4.35	0.44

가 가 ,

가 . , , , 가 .

가 .

가 .

가 .

가 가 M=4.30, ,

. . M=4.10, 가 가 가

. 가 M=4.00 .

. ,

, 3D 가

V. 결 론

M=3.80

. AI

가 .

, , 『Interstellar』 <S.T.A.Y.>

, AI ,

가 가 , ,

, 가

. .

. ,

가 (3 , , ,

), (2), ,

(2), (2) ,

가 . ,

, , RGB .

, ,

. (3), ,

(3), 가 (2), ,

(2) . 가 가 ,

M=4.35 ,

가 . AI

가 , - -

가 .

AI , , 가

가 . , ,

가

AI

<S.T.A.Y.>

가

가 .

가 .

가 .

가 가

가가 가 .

References

- Akkerman, R. (2023). *Film music as a guide in complex temporal narratives: An analysis of the narrative function of Hans Zimmer's film music in Inception (2010) and Interstellar (2014)*. Unpublished master's thesis, Uppsala University, Uppsala.
- Baek, E., Haines, S., Fares, O. H., Huang, Z., Hong, Y., & Lee, S. H. M. (2022). Defining digital fashion: Reshaping the field via a systematic review. *Computers in Human Behavior*, 137, 107407. doi:10.1016/j.chb.2022.107407
- Bertola, P., & Teunissen, J. (2018). Fashion 4.0: Innovating fashion industry through digital transformation. *Research Journal of Textile and Apparel*, 22(4), 352-369. doi:10.1108/RJTA-03-2018-0023
- Biner, B. C., Sofian, F. M., Karakaş, U. B., Ceylan, D., Erdem, E., & Erdem, A. (2024). *SonicDiffusion: Audio-driven image generation and editing with pretrained diffusion models*. arXiv. <https://doi.org/10.48550/arXiv.2405.00878>
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). *A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT*. arXiv. <https://doi.org/10.48550/arXiv.2303.04226>
- Choi, W., Jang, S., Kim, H. Y., Lee, Y., Lee, S., Lee, H., & Park, S. (2023). Developing an AI-based automated fashion design system: Reflecting the work process of fashion designers. *Fashion and Textiles*, 10(1), 39. doi:10.1186/s40691-023-00360-w
- Chui, Y. T., & Taylor, G. (2012). Interactive fashion featuring of sound in fashion design. *Research Journal of Textile and Apparel*, 16(4), 1-13. doi:10.1108/RJTA-16-04-2012-B001
- Delgado, M., Llopart, M., Sarabia, E., Taboada, S., Vierge, P., Vilariño, F., Moya Kohler, J., Grimberg Golijov, J., & Bilkis, M. (2024). *Music-triggered fashion design: From songs to the metaverse*. arXiv. <https://doi.org/10.48550/arXiv.2410.04921>
- Dhariwal, P., & Nichol, A. (2021). *Diffusion models beat GANs on image synthesis*. arXiv. <https://doi.org/10.48550/arXiv.2105.05233>
- Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology*, 65, 95-120. doi:10.1146/annurev-psych-010213-115035
- Foote, J. (1999, October). Visualizing music and audio using self-similarity. *Proceedings of the Seventh ACM International Conference on Multimedia, Orlando*, 77-80. doi:10.1145/319463.319472
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144. doi:10.1145/3422622
- Ho, J., Jain, A., & Abbeel, P. (2020). *Denosing diffusion*

- probabilistic models. arXiv. <https://doi.org/10.48550/arXiv.2006.11239>
- Huang, J., Weber, C. J., & Rothe, S. (2025, June). An AI-driven music visualization system for generating meaningful audio-responsive visuals in real-time. *Proceedings of the ACM International Conference on Interactive Media Experiences, Niterói*, 258-274. doi:10.1145/3706370.3727869
- Jin, Y., Yoon, J., Self, J., & Lee, K. (2024). Understanding fashion designers' behavior using generative AI for early-stage concept ideation and revision. *Archives of Design Research*, 37(3), 25-45. doi:10.15187/adr.2024.07.37.3.25
- Kim, J. Y., Oh, K. W., & Jung, H. J. (2015). Creating the idea of textile print pattern design using the visual expression of popular music. *The Korean Fashion and Textile Research Journal*, 17(4), 524-540. doi:10.5805/SFTI.2015.17.4.524
- Kim, S. S., & Kim, Y. S. (2014). Music visualization expression in modern fashion: Focus on the application of Mondrian's paintings to Alexander McQueen's 2014 spring ready-to-wear collection. *The Korean Fashion and Textile Research Journal*, 16(1), 55-65. doi:10.5805/SFTI.2014.16.1.55
- Kim, S. S., & Kim, Y. S. (2016). Fashion design expression of Wassily Kandinsky's paintings: Focus on music visualization. *Journal of the Korean Society of Costume*, 66(5), 16-32. doi:10.7233/jksc.2016.66.5.016
- Levin, G. (2000). *Painterly interfaces for audiovisual performance*. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge.
- Lima, H. B., Santos, C. G. R. D., & Meiguins, B. S. (2021). A survey of music visualization techniques. *ACM Computing Surveys*, 54(7), 1-29. doi:10.1145/3461835
- Mantri, K. S. I., & Sasikumar, N. (2023). *Interactive fashion content generation using LLMs and latent diffusion models*. arXiv. <https://doi.org/10.48550/arXiv.2306.05182>
- Ou, L.-C., Luo, M. R., Woodcock, A., & Wright, A. (2004). A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Research & Application*, 29(3), 232-240. doi:10.1002/col.20010
- Paik, N. J. (1963/1998). Participation TV. *Nam June Paik Art Center*. Retrieved April 16, 2026, from <http://njpart.ggcf.kr/collections/252>
- Pan, X., Tewari, A., Leimkühler, T., Liu, L., Meka, A., & Theobalt, C. (2023, August). Drag your GAN: Interactive point-based manipulation on the generative image manifold. *Proceeding of the ACM SIGGRAPH Conference, Los Angeles*, 1-11. doi:10.1145/3588432.3591500
- Perrotin, O., & d'Alessandro, C. (2014, June). Visualizing gestures in the control of a digital musical instrument. *Proceeding of the 14th international conference on new interfaces for musical expression, London*, 605-608.
- Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). *Hierarchical text-conditional image generation with CLIP latents*. arXiv. <https://doi.org/10.48550/arXiv.2204.06125>
- Reymore, L., & Lindsey, D. T. (2025). Color and tone color: Audiovisual crossmodal correspondences with musical instrument timbre. *Frontiers in Psychology*, 15, 1520131. doi:10.3389/fpsyg.2024.1520131
- Richardson, J., Pääkkölä, A.-E., & Qvick, S. (2021). Sensing time and space through the soundtracks of *Interstellar* and *Arrival*. In C. Cenciarelli (Ed.), *The Oxford handbook of cinematic listening* (pp. 385-406). Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780190853617.013.24
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022, June). High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans*, 10674-10685. doi:10.1109/CVPR52688.2022.01042
- Singh, A. K., & Patras, I. (2024). *FashionSD-X: Multimodal fashion garment synthesis using latent diffusion*. arXiv. <https://doi.org/10.48550/arXiv.2404.18591>
- Su, W., & Ling, J. (2023). Real-time visualization design of electronic music based on immersion theory. *Human-Centered Design and User Experience*, 114, 102-113. doi:10.54941/ahfe1004227
- Taylor, R., Boulanger, P., & Torres, D. (2006, April). Real-time music visualization using responsive imagery. *Proceedings of the 8th International Conference on Virtual Reality, Laval*, 62-69.
- Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology: General*, 123(4), 394-409. doi:10.1037/0096-3445.123.4.394
- Xing, Y., He, Y., Tian, Z., Wang, X., & Chen, Q. (2024). *Seeing and hearing: Open-domain visual-audio generation with diffusion latent aligners*. arXiv. <https://doi.org/10.48550/arXiv.2402.17723>

A Study on Generative AI-Based Fashion Design with Real-Time Music Visualization

Yusheng Yu • Jinsook Chae • Sookjin Kim⁺

Doctoral course, Department of Fashion Design, Sejong University, Fashion Design Research Institute
Concurrent Professor, Department of Fashion Design, Sejong University, Fashion Design Research Institute
Professor, Department of Fashion Design, Sejong University, Fashion Design Research Institute⁺

Abstract

Digital fashion is evolving from static garment imagery toward temporal and responsive visual expression, yet approaches to integrating dynamic musical information with garment images have received limited attention. This study proposes a method combining generative AI with real-time music visualization and examines its applicability. The soundtrack *S.T.A.Y.* from *Interstellar* was divided into four sections: introduction, development, climax, and relaxation, and analyzed in terms of emotional flow, rhythm, acoustic density, frequency, and energy. These characteristics were mapped to color, luminous effects, form transformation, and pattern density, and then implemented as music-responsive patterns in TouchDesigner. The results were translated into Midjourney prompts to generate garment images for each section, which were subsequently composited with TouchDesigner patterns. Ten experts evaluated music-AI garment connectivity, dynamic visual expressiveness, technological integration and innovation, and design quality and applicability using a 5-point Likert scale. Musical changes were reflected through variations in color, light, form expansion and contraction, and pattern density. The overall mean score was $M=4.35$, with technological integration and innovation and dynamic visual expressiveness receiving the highest ratings. This study presents an integrated process linking music analysis, visualization, AI image generation, compositing, and expert evaluation. Future research should extend this approach to broader genres, garment types, user evaluation, physical production, and exhibition applications.

Key words : music visualization, generative AI, digital garment design, TouchDesigner, Midjourney