생성형 AI 용도의 UI/UX

김태석 [1], Anh H. Vo[1], Marvin John Ignacio[1], Khuong G. T. Diep[1], 김용국 [1]
[1] 세종대학교 컴퓨터공학과 석사과정
[1] 세종대학교 컴퓨터공학과 박사과정
[1] 세종대학교 컴퓨터공학과 교수

22110617@sju.ac.kr, vohoanganh@sju.ac.kr, mjci@sju.ac.kr, khuongdiep@sju.ac.kr,
ykim@sejong.ac.kr

UI/UX for Generative AI

Tae-Seok Kim[1], Anh H. Vo[1], Marvin John Ignacio[1],
Khuong G. T. Diep[1], Yong-Guk Kim[1]
[1]Dept. of Computer Engineering, Sejong University

요 약

본 논문은 다양한 종류의 생성형 AI 용도의 UI/UX 중 텍스트 기반 UI/UX, 이미지 기반 UI/UX, 오디오 기반 UI/UX, 그리고 Multi-modal 을 기반으로 둔 UI/UX 와 같은 다양한 유형의 UI/UX 를 살펴보고 최신 기술을 활용한 미래전망에 대해 알아 보도록 한다. 현재 생성 모델은 다양한 산업 분야에서 광범위하고 다양한 응용 프로그램으로 사용되고 있으며, 최근 연구자와 실무자들로부터 상당한 관심을 받고 있다.생성형 AI 용도의 UI/UX 를 사용하면 생활에 편리해지며 시간과 돈이 매우 절약이 된다. 특히 사용자들이 편안하게 사용할 수 있는 생성형 AI 의 UI/UX 대한 연구방향에 대해 알아 보도록 한다.

1. 서론

UI 는 User Interface 이고 UX 는 User Experience 를 의미한다. 본 논문에서는 생성형 AI 모델이 소프트웨어 엔지니어, 제품 관리자, 사용자경험 디자이너를 UI/UX 면에서 어떻게 지원할 수 있는지에 대한 방법을 알아 보고자 한다. 이런 작업에서는 UI 의 설계와 구현이 동시에 변화해야 하며, 많은 종류의 UI/UX 속에서 최적의 선택을 해야만 한다. 많은 UI/UX 가 존재하는 이 시점에서 우리는 가끔 무엇을 사용해야 할지 고민에 빠질 수 있다. 따라서 그것에 대해 제대로 된 정의와 사용법이 필요하다. 생성형 AI 의 UI/UX 을 이용함으로써 시간을 절약하고, 여러 작업에 들어가는 비용을 아낄 수 있다. UI 를 생성하기 위해서는 여러 작업이 필요하다. 디자인, 코딩, 아이디어등의 작업이며, 그를 위해 기획자, 디자이너, 컴퓨터 엔지니어가 있어야 한다. 하지만 생성형 AI 를 이용하면 누구든 편리하고 간단하게 생성할 수 있다. 결론적으로 이 논문을 통하여 누구나 생성형 AI 의 종류와 발전가능성에 대해서 알 수 있게 하는 것이 목표이다. 대표적인 작업의 타임라인을 요약하면 다음과 같다. 전반적인 본 논문의 내용은 생성형 AI 를 위한 UI/UX 의 여러가지 개발 동향을 파악하고, 최근 진행 사항에 대하여 알아보고자 하며, 논문의 뒷 부분은 다음과 같이 구성되어 있다. 섹션 2 에서는 여러가지 종류의 UI/UX 설명과 사용 방법이 소개 될 것이며, 섹션 3 에서는 여러 기반의 UI/UX 의 미래방향 및 발전 가능성에 대해 논의 하도록 한다.

2. 다양한 종류의 UI/UX

2-1.TexT 기반 시스템의 UI/UX

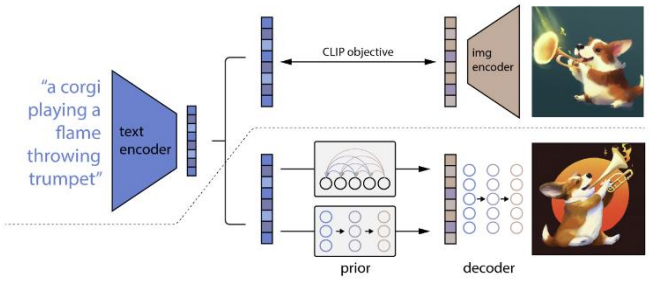
텍스트 기반의 UI/UX 에는 현재 많은 종류가 있다. 그 중 대표적인 UI/UX 는 ChatGPT 와 Microsoft 에서 사용중인 Bing 이 있다 두 모델은 텍스트를 기반으로 새로운 텍스트 생성하며, 두 모델은 GPT3.5[1], GPT4[2]를 각각 사용한다. Chat GPT 모델은 Open-AI 에서 개발한 자연어 처리 기반의 대화형 인공지능 모델이며, GPT3.5 를

사용하여 대화, 질문/응답, 자동 텍스트 생성 등의 응용 프로그램에서 사용된다. Bing 은 GPT4 를 이용하며, 웹 상에서 정보 검색과 결과를 제공하며, 다양한 정보와 연계가 가능하고 대부분의 정보가 웹에서 가져온 최신 정보이다. 둘 다 GPT 시리즈를 사용했으며, GPT 는 Transformer[3]의 Decoder 를 사용하여, 인공지능이 사람과 같이 대화를 하며, 대화를 통해 사용자가 원하는 결론 및 결과를 제공한다.

(그림 1) Chat GPT VS Bing.

2-2. 영상기반 시스템의 UI/UX

영상 기반의 UI/UX 는 텍스트를 이미지로 변환하거나 이미지를 새로운 이미지로 변환시켜 이미지의 품질 개선 등으로 응용이 가능하다. 여기에는 VIT 을 기본으로 여러가지 모델을 추가해서 GAN[4], VAE[5], Diffusion[6]등을 이용해서 이미지변환이 가능하다. 대표적으로는 Open AI 의 DALL-E2[7]가 있다. DALL-E 는 DALL-E 시작으로 DALL-E2, DALL-E3 가 출시됐다. 가장 최신 텍스트-이미지생성인 DALL-E2 는 텍스트-이미지를 생성하는 모델이며, 텍스트를 입력하면 출력으로 이미지 생성이 가능하다. Open.AI 에서 글자를 입력시키면 이미지가 생성된다. CLIP[8]과 Prior 과 Decoder 를 이용해서 이미지를 생성한다. 점차적으로 이미지기반의 UI/UX 또한 기술발전으로 여러 종류가 생겨났고 앞으로도 점진적으로 기술발전이 예상되며, 여러가지 방면에 응용이 될 것으로 예상된다.



(그림 2) DALL-E2 모델 개요.

DALL-E2 는 점선의 윗부분을 CLIP 학습 과정을 묘사하고 점선 아래에는 Text- to- Image 생성 과정을 묘사하고 있으며, CLIP 의 Text Encoder 에서 나온 Text Embedding 을 prior 에 대입하여 Image Embedding 을 생성한다. 앞에 생성된 Image Embedding 과 캡션을 이용해서 Decoder 로 최종 이미지를 생성 한다.

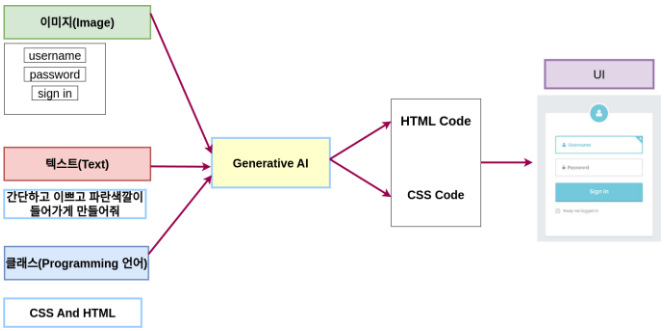
2-3. Audio 기반 시스템의 UI/UX

오디오 기반의 사용자 인터페이스 (UI) 및 사용자 경험 (UX)은 음성 및 오디오 입력 및 출력을 사용하기도 하고, 기기 또는 애플리케이션을 제어하고 상호 작용하는 방법을 의미한다. 이러한 종류의 UI/UX 는 음성 인식 기술과 음성 출력 기술을 기반으로 한다. 첫 번째로 Jukebox[9]에 대해서 설명한다. 장르, 아티스트, 가사를 입력 받아서 새로운 음악 샘플을 출력 가능하며, 다양한 음악과 노래 스타일을 생성하고 이전에 없었던 새로운 가사를 생성하기도 한다. Jukebox 는 VQ-VAE 모델을 이용하여, 새로운 장르의 오디오 생성이 가능하다. 두 번째로 스마트 스피커로 음성인식 기술을 사용하여, 명령을 내리고 여러가지 정보를 가지고 다양한 작업을 할 수 있다. 대표적으로 Google Home[10]이라는 UI/UX 이다. 다음은 음성 비서(Voice Assistants)이다. 이러한 음성 비서는 스마트폰, 태블릿, 스마트 스피커에 다양하게 사용이 가능하며, 음성을 통해 일상적인 작업을 수행하며, 정보를 요청할 수 있다. 음성 비서로는 Amazon Alexa, Google Assistant[11]등이 있다.

2-4. Multi-modal 기반의 UI/UX

Multi-modal UI/UX 는 사용자 경험을 향상시키기 위해 여러 종류의 입력과 출력 모드를 결합하는 방식이다. 이는 다양한 디바이스와 상황에서 사용자가 서로 다른 방식으로 상호 작용할 수 있도록 지원한다.[12].[13]. 주로 음성, 텍스트,

이미지 등 다양한 입력 및 출력 모드를 통합하여 사용자에게 최적의 경험을 제공하는 것을 목표로 한다. 이 아이디어는 Multi-model UI/UX 로 이미지와 텍스트 그리고 여러가지 컴퓨터 언어를 입력 하면 자동으로 코딩을 해주는 것이다. 코드라는 벽이 높기 때문에 코드를 익혀서 UI/UX 을 생성하기 위해서는 많은 시간과 경험이 필요하나 원하는 UI/UX 을 만들 수 있다.



(그림 3) Multi-modal 기반 UI/UX

그래서 원하는 프로그램을 생성할 수 있는 기능으로 원하는 이미지와 원하는 요구 사항을 입력하고 원하는 컴퓨터 언어를 입력 시키면 보다 쉽게 원하는 코드를 생성할 수 있다. 이 기술로 코드 미숙자들도 프로젝트에 참여가 가능하며, 비전문가들도 이것을 응용해서 이미지와 텍스트, 컴퓨터 언어를 입력 시켜서 생성형 AI 에 대입하여 코드를 자동으로 생성이 가능하다. 여기서 영상은 아이디어를 제공하며, 텍스트는 코드의 스타일을 제공함으로 코드의 색깔이나 방식을 제공하고, 컴퓨터 언어는 다양한 언어를 제공함으로써 사용처에 따라 유연하게 사용 가능하다. 여기서 코드는 바로 영상과 같이 UI/UX 로 이용하여 표시가 가능합니다.

3. 미래 방향 및 발전가능성

현재는 하나의 UI/UX 에서 하나의 베이스 또는 두 개가 합쳐진 Multi-model 의 UI/UX 을 이용 중이지만 미래의 UI/UX 에서는 기술의 발전과 사용자의 요구에 맞추어 계속 발전될 것이다. 사용자의 행동을 이해하고 예측하여 개인화된 경험을 제공하며, 작업의 자동화와 증강 현실 및 가상 현실에 새로운 UI/UX 을 사용한 새로운 기술로 현실세계와 가상세계를 통합한 환경에서 상호작용이 가능하며, 다양한 분야에서 혁신적인 경험을 제공할 수 있다. 또한 생성형 AI UI/UX 를 통해 DNA 생성도 가능 할 것으로 예상이 된다.[15].[16]. 바이오닉 기술과 생성형 AI 를 이용해 DNA 서열을 생성하고 원하는 속성을 갖추도록 할 수 있을 것이다. 한편 GAN, Diffusion model 을 이용한 합성 DNA 서열 생성 등 다양한 응용이 가능할 것이다. 기존 DNA 에 새로운 DNA 를 생성형 AI 에 대입하고, 새로운 DNA 를 생성 및 변경하여, 병해충에 관련된 문제를 해결하거나, 곡식이나 과일의 맛을 더 좋게 하여, 새로운 품종으로 생성이 가능 할 것이다.[17].[18].[19] 이런 작업들은 새로운 UI/UX 로 만들어 져서 누구나 편히 사용할 수 있을 것이다.

4. 결론

이 논문의 목표는 다양한 생성형 AI 용도의 UI/UX 의 사용과 방법에 대해서 소개하고 미래의 발전과 발전 가능성에 대해서 예측하는 것을 목표로 한다. 다양한 시각과 방법과 혁신적인 기법, 새로운 방향과 발전을 분석하고 요약하여 생성형 AI 의 UI/UX 의 개념을 구축 하고자 한다. 생성형 AI 을 위한 UI/UX 는 AI 를 사용하는 경험을 최적화하고 사용자와 AI 간의 상호 작용을 자연스럽게 만들어야 한다. 또한 생성형 AI 를 위한 UI/UX 은 사용자와 AI 간의 협력과 상호 작용을 중요하게 생각해야 한다. AI 의 능력을 향상 시키고 사용자의 요구에 부응하기 위해 지속적인 개선과 혁신이 필요하다. 결국은 사용자가 AI 을 쉽게 이해하고 활용할 수 있는 사용자 경험을 제공하는 것이 핵심이기 때문이다.

참고문헌

[1] Ye, Junjie, Xuanting Chen, Nuo Xu, Can Zu, Zekai Shao, Shichun Liu, Yuhan Cui, Zeyang Zhou, Chao Gong, Yang Shen, Jie Zhou, Siming Chen, Tao Gui, Qi Zhang and Xuanjing Huang. “A Comprehensive Capability Analysis of GPT-3 and GPT-3.5 Series Models.” ArXiv abs/2303.10420 (2023): n. pag.

[2] OpenAI. “GPT-4 Technical Report.” ArXiv abs/2303.08774 (2023): n. pag.

[3] Vaswani, Ashish, Noam M. Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin. “Attention is All you Need.” NIPS (2017).

[4] Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C. Courville and Yoshua Bengio. “Generative Adversarial Networks.” arXiv: Machine Learning (2014): n. pag.

[5] Kingma, Diederik P. and Max Welling. “Auto-Encoding Variational Bayes.” CoRR abs/1312.6114 (2013): n. pag.

[6] Dhariwal, Prafulla and Alex Nichol. “Diffusion Models Beat GANs on Image Synthesis.” ArXiv abs/2105.05233

(2021): n. pag.
[7] Ramesh, Aditya, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen and Ilya Sutskever. “Zero-Shot Text-to-Image Generation.” ArXiv abs/2102.12092 (2021): n. pag.
[8] Radford, Alec, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger and Ilya Sutskever. “Learning Transferable Visual Models From Natural Language Supervision.” International Conference on Machine Learning (2021).
[9] Dhariwal, Prafulla, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford and Ilya Sutskever. “Jukebox: A Generative Model for Music.” ArXiv abs/2005.00341 (2020): n. pag.
[10] Park, Min Jin and Joshua James. “Preliminary Study of a Google Home Mini.” ArXiv abs/2001.04574 (2020): n. pag.
[11] Arya, Shailesh and Samir B. Patel. “Implementation of Google Assistant & Amazon Alexa on Raspberry Pi.” ArXiv abs/2006.08220 (2020): n. pag.
[12] Houde, Stephanie, Steven I. Ross, Michael J. Muller, Mayank Agarwal, Fernando Martinez, John T. Richards, Kartik Talamadupula and Justin D. Weisz. “Opportunities for Generative AI in UX Modernization.” (2022).
[13] Li, Yang, Gang Li, Xin Zhou, Mostafa Dehghani and Alexey A. Gritsenko. “VUT: Versatile UI Transformer for Multi-Modal Multi-Task User Interface Modeling.” ArXiv abs/2112.05692 (2021): n. pag.
[14] Julia, Claudie FAURE-Luc. “An Agent-Based Architecture for a Multimodal Interface.” (2007).
[15] Killoran, Nathan, Leo J. Lee, Andrew Delong, David Kristjanson Duvenaud and Brendan J. Frey. “Generating and designing DNA with deep generative models.” ArXiv abs/1712.06148 (2017): n. pag.
[16] Avdeyev, Pavel, Chenlai Shi, Y.M. Tan, Kseniia Dudnyk and Jian Zhou. “Dirichlet Diffusion Score Model for Biological Sequence Generation.” ArXiv (2023): n. pag.
[17] Chandler, Vicki L. and Virginia Walbot. “DNA modification of a maize transposable element correlates with loss of activity.” Proceedings of the National Academy of Sciences of the United States of America 83 6 (1986): 1767-71 .
[18] Lutova, Ludmila A., M. A. Ramenskaya and S. V. Altukhova. “Late-Blight-Resistant Tomato Plants Obtained by T-DNA Insertion Mutagenesis.” Russian Journal of Plant Physiology 48 (2001): 662-667.
[19] Cullis, Christopher A. “Phenotypic consequences of environmentally induced changes in plant DNA.” Trends in Genetics 2 (1986): 307-309.