가상 면접관 생성을 위한 Stable Diffusion 기반의 프롬프트 엔지니어링

김대민[O], 시종욱[**], 손승재[*], 우민식[*], 김성영[*]
[O]금오공과대학교 컴퓨터공학과,
[*]금오공과대학교 컴퓨터공학과,
[**]금오공과대학교 컴퓨터·AI융합공학과
e-mail: daemingim5@gmail.com[O], jwsi425@kumoh.ac.kr[**], aldladdl3@naver.com[*],
alstlr0718@gmail.com[*], sykim@kumoh.ac.kr[*]

Stable Diffusion-based Prompt Engineering for Generating Virtual Interviewers

Daemin Kim[O], Jongwook Si[**], Seungjae Son[*], Minsik Woo[*], Sungyoung Kim[*]
[O]Dept. of Computer Engineering, Kumoh National Institute of Technology,
[*]Dept. of Computer Engineering, Kumoh National Institute of Technology,
[**]Dept. of Computer·AI Convergence Engineering, Kumoh National Institute of Technology

● 요 약 ●

가상 면접은 현대 사회에서 필수적인 기술이지만, 상호작용의 부족으로 인해 한계가 존재한다. 현실적이고 현장감 있는 가상 면접을 구현하기 위해서는 면접관을 자동으로 생성하여 다양한 상황에서의 면접을 진행할 필요가 있다. 본 논문은 Stable Diffusion 기반의 프롬프트 엔지니어링을 통해 가상 면접관 생성에 대한 연구 결과를 제시한다. 프롬프트 엔지니어링은 Stable Diffusion 모델이 생성하는 결과의 품질을 향상시킬 수 있으며 다양한 조건에 따른 실험 결과를 제시한다.

키워드: 가상 면접(Virtual Interview), 생성 모델(Generative Model), 프롬프트 엔지니어링(Prompt Engineering)

I. Introduction

가상 면접에 대하여 면접관과 지원자 모두에게 편리한 방식을 제공하고 제약을 극복할 수 있는 가상 면접 솔루션이 필요하다. 따라서, 다양한 면접관을 통해 면접을 연습할 수 있는 시스템이 필요하다. 실제와 같은 환경에서 가상 면접을 진행하기 위해서 메타버스와 연계할 수 있다.

본 논문에서는 Stable Diffusion[1] 기반의 프롬프트 엔지니어링을 활용하여 가상 면접관을 생성하는 방법을 제안한다. Stable Diffusion은 자연어 처리와 컴퓨터 비전 분야에서 활용되는 모델로 Text to Image의 특징을 가지고 있다. 따라서, 정의된 프롬프트와 함께 텍스트 데이터를 입력으로 하여 가상의 면접관을 다양하게 생성할 수 있다.

프롬프트 엔지니어링은 사전에 정의된 문장 구조를 사용하여 모델 결과의 품질을 개선할 수 있다. 이를 통해 보다 높은 품질의 가상 면접관을 생성할 수 있다. 따라서, 프롬프트 엔지니어링을 통해 Stable Diffusion의 결과를 메타버스 환경에서 활용할 수 있도록 생성하여 면접 시스템을 위한 새로운 연구 방향과 응용 가능성을 제시한다.

II. Generating Virtual Interviewers

메타버스 환경에서 가상 면접관을 생성한 아바타를 활용하기 위하여 필요한 요소들이 존재한다. 가상 면접관의 생성에 있어서 몇 가지 중요한 요구사항이 있다. 첫째로, 가상 면접관은 실제 면접관과 유사한 외모를 가져야 한다. 이를 위해 얼굴과 상반신을 포함한 아바타가 필요하다. 아바타는 깨끗한 배경에서 정장을 입은 모습을 보여주어야 한다. 둘째로, 메타버스 환경에 적용하기 위해서는 정면을 보고 있는 모습이어야 한다. 이는 가상 면접관과 면접자 간의 시선 교환을 자연스럽게 만들어준다. 셋째로, 가상 면접 시스템에서는 한 명의 면접관만 생성되어야 한다.

이를 위해, 이 논문에서는 가상 면접관 아바타의 생성을 위해 8가지 조건들을 정의한다. 이 구성요소는 Stable Diffusion의 모델을 기반으로 프롬프트 엔지니어링을 통해 결과를 생성하고 분석한다. 8가지 조건들은 아래와 같다.

(1) Interviewer in the interview room
(2) High Quality
(3) Realistic
(4) Upper body with all face
(5) Formal Wear
(6) Clean Background
(7) Front View
(8) Alone

실험을 통해 조건 (1)만 사용했을 때의 결과는 그림 1과 같았다. 이는 가상 면접관이 얼굴 및 상반신에만 초점을 둔 것이 아니므로, 배경만 고려된 전체적인 면접장의 모습을 보인다. 가상 면접 시스템에서는 면접자와 면접관 사이의 시각적 상호작용이 중요하므로, 아바타의 생성 성능은 얼굴뿐 아니라 다양한 조건이 필요함을 알 수 있다.

그림 2는 (1)을 제외한 다른 조건들을 하나씩 제외한 결과들과 모든 조건을 사용했을 때의 생성 결과를 나타낸다. 실험 결과, 모든 조건들을 입력으로 하였을 때 가상 면접관의 생성 품질이 가장 높았다. 다양한 조건들을 고려하면서 생성된 결과는 현실적이고 고품질의 가상 면접관을 제공할 수 있음을 보여준다.

Fig. 1. Results of Stable Diffusion with Only Condition (1)

III. Conclusions

본 논문에서 제안하는 Stable Diffusion 기반의 프롬프트 엔지니어링은 다양한 조건들을 고려하여 가상 면접관을 생성함에 있어서 효과적인 방법임을 확인할 수 있다. 메타버스 환경에서 가상 면접관 아바타로 사용하기 위해 2D 이미지를 3D 모델로 생성하는 연구가 진행될 예정이다.

ACKNOWLEDGEMENT

본 연구는 2023년도 한국전파진흥협회 메타버스 랩 지원 사업에 의한 연구임

REFERENCES

[1] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models", Proc. Of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10684-10695, 2023.

Fig. 2. Results of Stable Diffusion with Various Condition Combinations