|  |
| --- |
| **1. 주제**  개인 비서 서비스 및 GPT 기반 가상 인간 커뮤니케이터 개발 제안  **분반, 팀, 학번, 이름**  오픈소스기초설계(가), 404 ERROR, 20233083, 김태우 |

|  |  |
| --- | --- |
| **2. 요약**  -목표  이 프로젝트의 주 목표는 비대면 사회와 기술 발전의 흐름 속에서 개인주의와 외로움이 증가하는 현대인들에게 개인화된 가상 인간을 통해 소통의 파트너와 개인 비서를 제공하는 것입니다. 이를 통해 사용자의 일상적인 편의를 증진시키고, 사회적인 외로움을 해소하는 데 기여하고자 합니다.  -핵심 내용  프로젝트는 GPT 기반의 대화형 인공지능을 활용하여 사용자가 외모, 목소리, 성격 등을 맞춤 설정할 수 있는 가상 인간을 제작합니다. 이 가상 인간은 사용자의 일상 활동을 보조하며, 다양한 디지털 플랫폼과 연동되어 어디서든 일관된 서비스를 제공합니다. 예를 들어 NFC 기술을 이용하여 키오스크와의 연동을 통해 개인화된 서비스를 사용자에게 제공하며, 사용자의 디지털 활동과 실생활 사이의 경계를 허물어줍니다.  -중요성  이 프로젝트는 사용자의 외로움 해소와 사회성 강화 도움, 그리고 기술의 접근성을 향상시키는 데 중요한 역할을 합니다. 가상 인간이 제공하는 맞춤형 서비스와 소통은 사용자의 일상에 편의를 제공하며, 동시에 사회적인 소통에 어려움을 격는 사람들에게 따뜻한 말과 피드백을 통해 사회화 지원을 받을 수 있습니다. 또한, 모든 연령대의 사용자가 기술을 쉽게 접근하고 활용할 수 있도록 도와, 디지털 디바이드를 줄이는 데 기여할 것으로 기대됩니다. | **3. 대표 그림**    그림 1. 가정집에서 사용자가 개인 가상 인간과 대화하는 모습 |

|  |
| --- |
| **4. 서론**  -사례 분석  해당 비슷한 대화 상대로 이루다, 스마트폰의 대화형 인공지능, 기업의 인공지능 스피커,미국의 가상 인플루언서 릴 미켈라, 국내 가상 인플루언서 로지를 찾을 수 있었음  이루다  해당 서비스는 자체 개발 플렛폼(Nutty)에서 사용 가능한 서비스로 대한민국에서 인목을 끈 대화형 챗봇이다. 하지만 해당 서비스는 다른 기기를 조작하거나 활용하는 것이 불가능해 정보가 제한적이며 실질적으로 사용자에게 편의를 느끼게 할 수 있는 기능이 존재하지 않음. 이는 단순한 챗봇의 한계를 벗어나지 못한 프로그램으로 생각할 수 있음.  대화형 인공지능 및 스피커  애플의 siri, 갤럭시의 빅스비등의 프로그램은 실제 사람과 같이 말하거나 놀 수 있는 것이 아닌 이미 정해진 siri나 빅스비의 이름으로 사용만이 가능하며 대화하는 주체가 인간이라고 생각할 만한 부분이 적어(얼굴의 부재, 이미 입력된 몇개의 답변들, 자아라고 하는 것을 흉내내지 않으며 성격이라고 할 만한 부분이 없음) 가상 인간과 거리가 멈.  아마존 에코 플러스, 구글 홈, 카카오미니등 다양한 기업들이 제작한 스피커로 사용자와 대화할 수 있지만 앞서 얘기한 것과 비슷한 인간이라고 생각할만한 부분이 적고 기억력이 존재하지 않아 사람과 하는 대화가 사실상 불가능함.  이러한 서비스는 모두 사용자와의 지속적이고 개인적인 관계 형성, 그리고 사용자의 감정과 상황에 따른 섬세한 대응 능력이 아직 미흡한 상태임  가상 인플루언서 3D 그래픽을 활용한 가상 인플루언서는 2021년도부터 국내에도 등장하여 인기를 끌고 있다. 이들이 인기를 끄는 이유는 가상 인간이기 때문에 외적 매력 요소를 분석하고 처음부터 넣을 수 있으며 개인 논란도 없다. 하지만 실시간으로 사용자와 소통할 수 없으며 인격이 존재하지 않는다. 오직 이미지의 형식으로 사용자는 이를 접할 수 있다. 하지만 이러한 선례를 통해 사람들은 사람과 비슷한 가상 인간이 기업에게 이득과 시민들은 환호화 인기를 끄는 것을 확인할 수 있었다.  -문제 정의  현재 시장에서 제공되는 다양한 인공지능 기반 대화 서비스 중 사용자와 사람과 같이 대화하면서 편의적 도움을 주는 가상 인간이 아직 없으며 이는 그 회사의 스마트폰이나 스피커를 사야하는 경제적 부담까지 있음.또한 키오스크와 같은 실생활에 사용하는 기기를 지원하려는 서비스는 아직 존재하지 않으며 일부 사람 형태의 안내 기계를 제작하는 회사기 있으나 고장, 유지비, 사람의 형태로 생겼기 때문에 발생하는 다양한 기술적 문제등 한계가 많고 보급되기에 어려운 기술임. 친구와 같은 현실의 인간 관계는 언제나 바로 연락이 가능한 것이 아니며 시간, 공간의 문제, 성격의 문제, 금전의 여유 차이등 다양한 어려움을 격어 소원해질 수 있음.  -극복 방안  이 프로젝트는 사용자와의 지속적이고 깊이 있는 소통을 가능하게 하는 가상 인간 기술을 제안함. 사용자에 눈높이에 맞게 설명할 수 있으며 카카오톡이나 인스타그램과 같이 개인적으로 소통을 언제나 할 수 있는 친구같은 서비스의 지원이 중요. 현실의 사람과 다르게 언제나 대화하고 연락할 수 있으며 집을 벗어나도 다양한 디지털 플랫폼과 기기에 연동되어, 사용자가 어디서든 일관된 경험을 누릴 수 있도록 함. NFC 기술을 활용하여 키오스크와 같은 실생활 기기와의 연동도 가능하게 하여 사용자의 디지털 활동과 실생활 사이의 경계를 허물어줌. |

|  |
| --- |
| **5. 본론**  -시스템 개요 그림    -필요한 기술 요소와 구현 방법  gpt 현재 가장 뛰어난 대화형 인공지능으로 사용자에게 가장 만족스러운 결과를 낼 수 있을 것이라 예상함. 해당 gpt에 프롬포트로 사용자가 원하는 인물의 이름, 배경 지식, 성격, 말투등을 설정하여 작동하게 됨.  얼굴 구현  nvidia의 omnverse Audio2Face 기술로 음성에 알맞는 표정을 만들어주며 사용자에게 가상 인간의 얼굴로 사용할 수 있음. 하지만 기기의 성능에 따라 구현을 할 수 없는 사용자도 있을 것이라고 예상됨.  또한 키오스크와 같은 외부 기기에서는 큰 정보를 빠르게 옮기기 힘들어 이를 사진같은 이미지나 데포르메한 2D 이미지를 사용하여 구현할 수 있음. 이러한 경우 단순한 움직이는 얼굴 사진으로 사용자는 사람이라고 인식하는데 도움을 줄 수 있으며 3D라고 체감할 수 있음  2D 이미지의 깊이 생성 및 영상화  NeRF는 신경망 기반 역추적 렌더링으로 기존의 평면 이미지의 좌표에 따라 깊이를 계산하고 이를 통해 작은 움직임을 가진 영상을 생성한다. 이는 컨볼루션 신경망(CNN)이 아닌 이미지 분석에 CNN을 접합한 딥러닝 뉴럴 네트워크(GNN)이다. 해당 기능을 통해 사용자가 원하는 가상 인간을 3D 모델로 제작하지 않아도 사진의 가상 인간이 실제 3차원 공간에 있는 것처럼 움직이는 것을 볼 수 있다. 또한 이러한 방식의 영상 제작으로 인해 3D 모델을 잘 모르는 일반인도 쉽게 접근할 수 있으며 실제 자신의 얼굴을 원할 경우 자신의 얼굴 사진만이 필요하기 때문에 대중화에 적합한 기술이다.  가상 인물의 얼굴 생성  생성적 적대 신경망(GAN)을 통해 사용자는 가상 인물의 얼굴과 목소리를 생성할 수 있다. 이를 통해 사용자가 원하는 특징의 얼굴를 가진 가상 인간이 생성되게 된다.  목소리 구현  diffusion 모델을 통한 사용자가 원하는 인물이나 목소리의 특징을 토대로 학습하거나 다른 목소리를 창조하여 사용할 수 있음. 음성 생성은 tts(text to speach)를 통해 이루어짐.  개발 방향  얼굴은 virtual character(실제 사람의 행동이나 표정을 대신 해주는 캐릭터)같은 프로그램을 통해 가볍게 구현할 수 있음. 모델이 되는 사람이 없어도 2D인만큼 감정이나 말에 따른 입의 움직임을 간단하게 구현할 수 있음. 대화 모델은 gpt를 사용할 것이며 해당 프로젝트를 제안하여 4버전의 api를 얻어 사용할 예정임. 음성 모델은 diff-tts 모델을 사용할 예정이지만 컴퓨팅 파워가 부족할 시 11.ai와 같은 서비스를 이용하여 사용할 예정임. 사용자가 말하는 내용은 google tts로 받아 텍스트화 하여 프롬포트로 들어가게 됨. 사용자가 프로그램을 키면 가상 인간이 나오며 해당 프로그램과 텍스트나 음성으로 대화할 수 있으며 실시간으로 검색하여 정보를 알려주고 몇 프로그램을 조작하여 사용자에게 편의를 줄 수 있는 기능을 실행하도록 할 것이다. |

|  |
| --- |
| **6. 결론**  이 프로젝트는 현대 사회에 증가하는 개인주의와 외로움 문제를 해결하며 디지털 기술에 소외받는 디지털 디바이드 문제 해결에 다가가기 위한 방안이며 여러분의 친구이자 개인 비서, 평생의 동반자로서 다양한 기기에 연동되어 사용자를 도울 것이다. 이를 통해 더 나은 삶을 누릴 수 있으며 이를 구현하기 위해 더 가벼운 얼굴 구현 방식 착안과 우리가 쓰는 메신저 프로그램에서 사용할 수 있도록 메신저와 프로그램의 연결 방식을 공부할 것이다. |

**7. 출처**

[1] 이루다(https://luda.ai/), gpt(https://openai.com/gpt-4), nivdia omniverse(<https://www.nvidia.com/en-us/omniverse/#cid=_so-yout_en-us>), diff-tts(https://github.com/keonlee9420/DiffGAN-TTS), NeRF(https://github.com/bmild/nerf), siri(https://www.apple.com/kr/siri/), googlehome(https://home.google.com/welcome/)

[2] 프로그래밍실전프로젝트(인공지능), 학습자료

[3] 릴 미켈라 인스타그램(https://www.instagram.com/lilmiquela/), 로지 인스타그램(https://www.instagram.com/rozy.gram/)