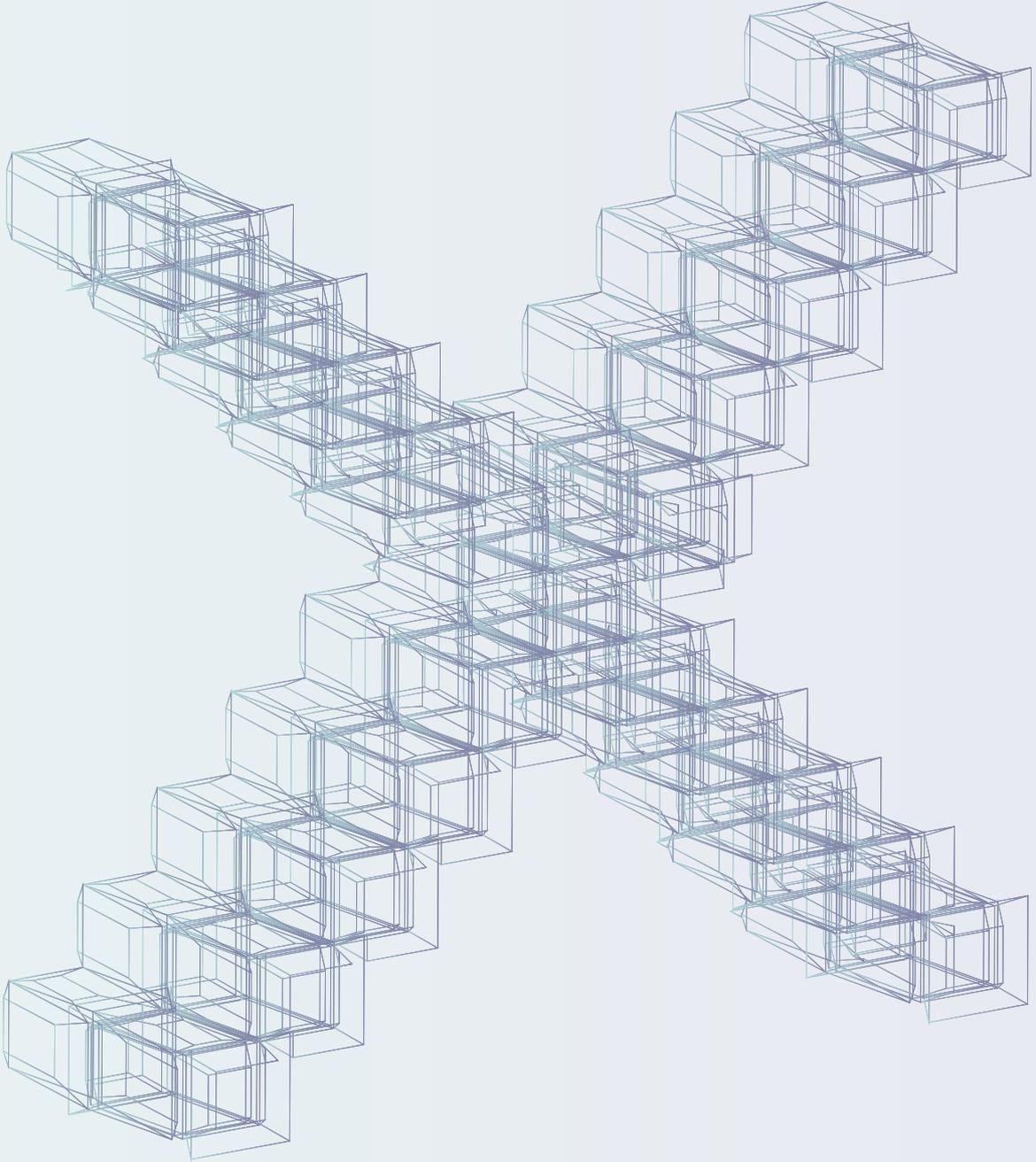


X GRAPHICS

From **X**
To **X**

한국컴퓨터그래픽스학회



KOREA COMPUTER GRAPHICS SOCIETY

KCGS 2018
학술대회 논문집

한국컴퓨터그래픽스학회

2018 학술대회 논문집

2018년 7월 10일 ~ 2018년 7월 13일

여수 디오션리조트

주관기관: 한국컴퓨터그래픽스학회

KCGS 2018 환영사

컴퓨터그래픽스 분야에서 열심히 연구와 교육에 힘쓰시는 교수, 연구원, 학생 여러분. 이제 7월이 되어 각자의 바쁜 업무를 잠시 내려놓고 우리가 한자리에 함께 모이는 학술대회가 열립니다. 올해 우리 학회는 “Graphics from/to X”라는 주제로 진행하고자 합니다. 컴퓨터그래픽스는 어디에서 흘러왔고 앞으로 어떻게 발전할까요? 처음에는 자그마하게 시작하였겠지만 이제 수많은 분야에서 핵심적인 역할을 하고 있는 컴퓨터그래픽스의 역사를 돌이켜보고 앞으로의 발전방향을 함께 고민해보는 좋은 자리가 되길 바랍니다.

KCGS 학술대회는 1993년에 저희 학회가 설립된 이후 끊임없이 개최되면서 국내의 모든 그래픽스 전문가들이 참여하는 대표적인 학술행사로 자리 잡고 있습니다. 매년 컴퓨터그래픽스를 전공한 국내 학계 및 연구계의 전문가들과 관련업계 종사자들, 그리고 열심히 연구에 매진하는 학생들이 한자리에 모여서 최신 연구결과를 공유하고 협력을 모색하는 좋은 기회가 되고 있습니다. 회원 여러분들의 성원에 힘입어 올해에는 구두발표 32편, 포스터발표 27편의 논문이 발표될 예정입니다. 이 중 논문지에 실리는 정규논문은 11편입니다. 그리고 작년과 마찬가지로 학술대회 논문을 인쇄하지 않고 모바일 웹 형태로 제공하여 언제 어디서든 온라인으로 쉽게 접근할 수 있도록 하였습니다.

무엇보다도 올해 새롭게 진행되는 새로운 세션은 교수/박사연구원급 발표 세션입니다. 대부분의 구두발표가 학생들에게 좋은 경험과 기회를 제공해 왔습니다만 학생들만의 발표로는 뭔가 부족한 부분들이 있었던 것이 사실입니다. 교수/박사연구원급 발표 세션에서는 교수님 및 박사연구원들께서 직접 우수한 최신 연구에 대해 발표함으로써 학술대회의 품격을 높이고 전문가들과 학생들에게 보다 유익한 기회를 제공할 것입니다. 이를 위해 수고해 주신 이승용 교수님과 바쁘심에도 흔쾌히 발표를 해 주시는 네 분의 교수님께 진심으로 감사드립니다.

한국컴퓨터그래픽스 대상은 우리 학회가 자랑스러워하고 존경할 만한 연구업적을 이루신 선배 연구자에게 우리가 헌정하는 최고의 상입니다. 올해에는 이 상을 포항공대 이승용 교수님께 드리기로 결정하였습니다. 모두들 진심으로 축하해 주시길 부탁드립니다.

올해에는 엔비디아를 비롯한 여러 기관들의 도움으로 실용적인 내용 중심으로 여름학교 프로그램 구성하였으며, 다양한 분야의 전문가를 모시고 초청강연을 들을 수 있도록 준비하였습니다. 또한 문화체육관광부에서 우리 분야 연구기획을 담당하고 계시는 PD님께서 정부의 R&D 추진전략에 대한 설명하시는 자리도 마련하였습니다.

늘 우리 학술대회에 관심을 가지고 물심양면으로 후원해 주시는 후원사 관계자 분들에게 깊은 감사를 드립니다. 올해도 리더스시스템즈, 소비코, 알체라, 엔비디아, 투썸디지털아이디어, KAI, VIZINF 등 많은 기업들이 후원을 해주셨으며, 세종대학교 최수미 교수님의 모바일 가상현실 연구센터, 이화여자대학교 김영준 교수님의 그래픽스 연구실에서도 후원해 주셨습니다.

우리 학회에서 제가 가장 자랑스러워하는 부분은 회원님들의 헌신입니다. 이번에도 여러 조직위원님들과 프로그램위원님들, 여름학교 준비위원들께서 너무나 많은 수고를 해주셨습니다. 진심으로 너무나 감사드립니다.

이번 학술대회는 우리 학회에서 처음으로 여수에서 개최됩니다. 여수는 EXPO를 치러낸 도시로서 아름다운 관광지임과 동시에 훌륭한 남도음식으로 참가자 여러분들께 새로운 즐거움을 드릴 것으로 확신합니다. 물심양면으로 도와주신 여수시에도 감사드리며, 새로운 장소에서 열린 마음으로 편안하게 공부하고, 교류하며 즐기는 좋은 시간 되시길 기원합니다. 감사합니다.

2018년 7월 10일
KCGS 2018 학술대회 대회장 김동호

학술대회 조직위원회

대회장 김동호 (송실대학교)

조직위원장 이성희 (KAIST), 김준호 (국민대학교)

프로그램위원장 이지은 (한성대학교), 최민규 (광운대학교)

여름학교 준비위원장 김영민 (KIST), 이주행 (ETRI)

수상위원회 위원장 최정주 (아주대학교)

조직위원 계획원 (한성대학교), 김종민 (강원대학교), 김종현 (강남대학교),
박진호 (송실대학교), 이영호 (목포대학교), 최명걸 (가톨릭대학교),
한다성 (한동대학교), 황재인 (KIST)

프로그램위원 고성안 (UNIST), 권지용 (스타십벤처머신㈜), 권태수 (한양대학교),
김구진 (경북대학교), 김민호 (서울시립대학교), 김선정 (한림대학교),
김성기 (계명대학교), 김수균 (배재대학교), 김영준 (이화여자대학교),
김주섭 (서강대학교), 김진모 (부산가톨릭대학교),
김진아 (한국해양과학기술원), 김형석 (건국대학교), 김효용 (한성대학교),
노준용 (KAIST), 문보창 (GIST), 박상일 (세종대학교), 박상훈 (동국대학교),
박진아 (KAIST), 백낙훈 (경북대학교), 손봉수 (중앙대학교),
송오영 (세종대학교), 송창근 (한림대학교), 신병석 (인하대학교),
신현준 (아주대학교), 오경수 (송실대학교), 유민준 (뉴욕대학교),
윤성의 (KAIST), 윤승현 (동국대학교), 윤종철 (강원대학교),
이강훈 (광운대학교), 이성길 (성균관대학교), 이윤상 (한양대학교),
이윤진 (아주대학교), 이인권 (연세대학교), 이제희 (서울대학교),
이택희 (한국산업기술대학교), 이해영 (홍익대학교), 이환용 (아주대학교),
임인성 (서강대학교), 장윤 (세종대학교), 정원기 (UNIST),
조동식 (원광대학교), 조성현 (DGIST), 최유주 (서울미디어대학원대학교),
한정현 (고려대학교), 홍민 (순천향대학교)

목 차

한국컴퓨터그래픽스 대상

A Journey to Graphics and Beyond	2018.07.12.(목) 16:00-16:50
이승용 (포항공과대학교)	1

초청강연

Video Interaction at Scale	2018.07.12(목) 13:00-13:50
김주호 (KAIST)	2
VR 인터랙티브 애니메이션 [미니월드 VR]	2018.07.12.(목) 13:50-14:40
장육상 (중앙대학교)	3
CT R&D 현황과 사업방향	2018.07.12(목) 14:50-15:20
곽재도 PD (문화체육관광부)	4

후원기관 초청 강연

2018.07.12(목) 15:20-15:50

HoloLens를 활용한 MR 콘텐츠 기술 개발	
김종용 (비즈아이엔에프)	5
Realistic Digital Creature Creation with Quadruped Motion Capture	
송재원 (투션디지털아이디어)	6

석사논문상 후보자 발표

2018.07.11.(수) 17:10-18:00

김승욱 (고려대학교)	7
이승환 (서울대학교)	8

논문발표 1: 애니메이션

2018.07.11(수) 13:00-14:40

심층강화학습을 이용한 군중 시뮬레이션	9
	이재동, 이제희 (서울대학교)
물리 시뮬레이션에 기반한 2차원 연체 동물 캐릭터 제어	11
	민세희, 이제희 (서울대학교)
다목적 제어 학습을 이용한 대화형 캐릭터 애니메이션	13
	이경호, 이세영, 이제희 (서울대학교)
사람 애니메이션을 위한 다중 세그먼트 발 모델링	15
	박황필, 유리, 이제희 (서울대학교)
스킨 캐릭터의 공간 관계 유지를 통한 모션리타게팅 방법	17
	진태일, 김미경, 이성희 (KAIST)

논문발표 2: 교수/박사연구원급 발표 2018.07.11.(수) 15:10-16:50

Prediction of Diffusion-Weighted Appearance in Developing Infant Brain using Cycle-Consistent Models 19
김재일 (경북대학교), Geng Chen, Weili Lin, Pew-Thian Yap, Dinggang Shen

High-Quality Hyperspectral Reconstruction Using a Spectral Prior 20
최인창, 전석준, 남길주, Diego Gutierrez, 김민혁 (KAIST)

Retargeting Human-Object Interaction to Virtual Avatars 21
김연준, 박한길, 방승배, 이성희 (KAIST)

고해상도 점군 획득을 위한 카메라 프로젝터 시스템 개발 22
강현민, 박용문, 서용덕 (서강대학교)

논문발표 3: 그래픽스 응용 및 시스템 I 2018.07.12.(목) 09:00-10:20

음함수에 기반한 실시간 GPU 슬라이싱 23
김현준, 로정, 김민호 (서울시립대학교)

3D 공간 모델을 활용한 로봇 기반 정리 자동화 25
강민철, 윤성의 (KAIST)

광선 추적법에 기반한 소리 위치 추정 27
안인규, 윤성의 (KAIST)

딥 러닝을 이용한 변위맵 기반 3차원 기하 디노이징 29
윤여진, 최수미 (세종대학교)

논문발표 4: 가상현실 I 2018.07.12.(목) 09:00-10:20

비대칭 가상현실에서의 현존감을 위한 HMD 사용자와 Non-HMD 사용자간 몰입형 상호작용에 관한 연구 특별호
이지원 (부산가톨릭대학교), 김민규 (고려대학교), 김진모 (부산가톨릭대학교)

FOV Restrictor를 활용한 가상 멀미 저감 요소 기술 연구 특별호
김창섭 (한양대학교), 김소연 (덕성여자대학교), 김광욱 (한양대학교)

가상 홈 트레이닝 - 가상현실 기반 소근육 재활 시스템 특별호
유경호, 김해지, 김한섭 (조선대학교), 이지은 (한성대학교)

가상환경의 비인간 아바타를 이용한 보행에서의 실재감 31
이동용, 이인권 (연세대학교)

논문발표 5: 모델링 및 렌더링 2018.07.12.(목) 10:30-12:10

입자 패치 기반 가상 연필 및 에어브러시 가시화 알고리즘 특별호
이혜린, 오건, 이택희 (한국산업기술대학교)

RGB-D 영상으로 복원한 점 집합을 위한 고화질 텍스처 추출 기법 특별호
서웅, 박상욱, 임인성 (서강대학교)

깊이 기반 실시간 반투명 유리 렌더링	33
이제선, 최재원, 박소영, 최윤지, 김재명, 이성길 (성균관대학교)	
스플라인 인터페이스를 이용한 스키닝 웨이트 계산	35
방승배, 이성희 (KAIST)	
스트로크 와핑을 이용한 회화적 렌더링	37
김범식, 윤경현 (중앙대학교)	

논문발표 6: 가상현실 II 2018.07.12.(목) 10:30-12:10

다중 플레이어들의 팀워크에 기반한 동작-구동 조정 게임	특별호
김혜진, 심재혁, 임승찬, 고영노, 한다성 (한동대학교)	
몰입형 가상현실에서의 협업 기반 다중 사용자 상호작용의 현존감에 관한 연구	특별호
박성준, 박원준, 허하영, 김진모 (부산가톨릭대학교)	
가상현실 기반 범죄 프로파일링 시뮬레이션 교육 및 평가 시스템	특별호
김한섭, 김해지, 이윤식 (조선대학교), 이지은 (한성대학교)	
다트 가상현실 콘텐츠에서의 협업을 위한 기계학습기반 사용자 의사결정 모델	39
봉찬균, 박성준, 박원준, 허하영, 김진모 (부산가톨릭대학교)	

논문발표 7: 이미지/비디오 2018.07.13.(금) 09:00-10:00

이미지 지역 통계를 이용한 모바일 기기로 촬영한 문서에서의 그림자 제거	특별호
나예지, 박상일 (세종대학교)	
GAN을 이용한 자연어로부터 발화 동영상 생성	41
이건희, Jacob Morton, 이승용 (포항공과대학교)	
멀티모달 딥러닝을 이용한 비디오 감정 회귀 분석	43
김하연, 이인권 (연세대학교)	

논문발표 8: 그래픽스 응용 및 시스템 II 2018.07.13.(금) 09:00-10:00

근접한 물체 사이의 공간 관찰을 위한 보조 카메라 위치 최적화	특별호
이지혜, 한윤하, 최명걸 (가톨릭대학교)	
자연재난 데이터 실감 가시화 시스템	특별호
김종용, 정석철, 이계원, 조준영, 김동욱, 박상훈 (동국대학교)	
목소리의 감성적인 필터링을 이용한 의사소통 도우미	45
이수호, 신채영, 신영찬, 지승용, 최지혜, 배대웅, 조윤진, 김지우, 신세영, 김종현 (강남대학교)	

논문발표 9: 시뮬레이션 2018.07.13.(금) 10:20-11:40

의상 시뮬레이션 가속을 위한 패턴 미러링	47
장인서, 이상빈, 고희석 (서울대학교)	
키네틱 프로젝션 매핑을 위한 물리기반 표면변형 시뮬레이션	49
김정민, 김지수 (이화여자대학교), 권오홍 (한국생산기술연구원), 김영준 (이화여자대학교)	

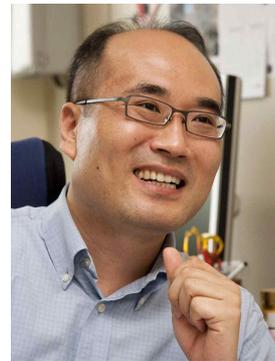
심층 신경망과 계층적 구조를 이용한 옷감 시뮬레이션의 가속화	51
이태민, 오영진, 이인권 (연세대학교)	
자성체 간 상호작용을 위한 자기화량 역학	53
김승욱, 박선영, 한정현 (고려대학교)	

포스터발표 2018.07.12(목) 17:30-19:00

컨볼루션 신경망을 이용한 2차원 연기 시뮬레이션의 보간	55
오영진, 김영준, 손진석, 김성호, 이인권 (연세대학교)	
입 바람 기반 사용자 인터페이스를 이용한 유체 시뮬레이션의 상호작용 기술	57
안준혁, 김종현 (강남대학교)	
컬링 스톤의 빙상 마찰력 분석을 통한 컬링 시뮬레이터 구현	59
이동일, 송오영 (세종대학교)	
제스처 기반의 부동산 거래 VR 보드 게임	61
박고은, 이수아, 김성실, 문주혜, 한다성 (한동대학교)	
홀로렌즈를 이용한 1인칭 슈팅 게임 구현	63
송인석, 전우형, 송창근, 이정, 김선정 (한림대학교)	
Kinect와 Scratch를 이용한 미디어 아트 콘텐츠 제작	65
여주은, 박현주, 정지민, 문주혜, 서충모, 김동주, 이재호, 이원섭 (한동대학교)	
3가지 물질에 대한 반응-확산을 통한 새로운 패턴 생성	67
한도연 (세종대학교), 김병문 (Adobe Research), 송오영 (세종대학교)	
인코더-디코더 네트워크를 이용한 단일 영상의 비균일 블러 제거	69
손형석 (포항공과대학교), 조성현 (대구경북과학기술원), 이승용 (포항공과대학교)	
전경 분리 알고리즘을 활용한 스타일 트랜스퍼에서의 콘텐츠 보존 기법	71
정진서, 김상현, 주은미, 김명진, 김영빈 (중앙대학교)	
모바일 GPU를 위한 새로운 성능 예측 모델	73
윤주원, 이진영 (세종대학교), 김정길 (남서울대학교), 임영규, 나재호 (LG 전자), 김영식 (한국산업기술대학교), 박우찬 (세종대학교)	
대응점 탐색 알고리즘의 GPU기반 연산에 대한 성능 분석	75
김병준, 강현민, 서용덕 (서강대학교)	
제스처 인식 기반 드로잉 시스템의 설계	77
김상준 (서울미디어대학원대학교), 홍민 (순천향대학교), 최유주 (서울미디어대학원대학교)	
물리 시뮬레이션 및 진화 연산을 사용한 진화형 모듈형 로봇 설계	79
박재훈, 이강훈 (광운대학교)	
3D 환경에서의 소리 방향 식별 보정	81
유영제, 이승우, 이정, 송창근, 김선정 (한림대학교)	
VR 인터랙션을 위한 효과적인 자세 인식 및 키 입력 생성 방법	83
권혁주, 권민철, 우다라, 이상훈 (연세대학교)	

가상 교실 내 사회적 상호작용 종류에 따른 정보처리 촉진에 관한 연구: 동공 추적을 중심으로	85
유재환, 허지웅 (한양대학교), 김은주 (연세대학교), 김광욱 (한양대학교)	
360도 비디오 시청에 대한 유저 행동양식 분석	87
유정은, 노준용 (KAIST)	
가상 현실을 이용한 일상생활활동 과제: Virtual Supermarket을 통한 인지기능 훈련 도구 개발	89
김석현, 전주희, 김광욱 (한양대학교)	
Leap Motion 센서를 이용한 드론 애니메이션 제어	91
김하준, 임승찬, 김한설, 황인성, 김시은, 팔롤라 데니스 (Daeyang University), 이원섭 (한동대학교)	
회전 테이블의 기하적 변환을 이용한 점군 정합	93
강현민, 박용문, 김병준, 서용덕 (서강대학교)	
Funny Room: 다중 사용자간 협업을 활용한 캐주얼 가상현실 콘텐츠 제작	95
허하영, 박원준, 박성준, 민성미, 김진모 (부산가톨릭대학교)	
Video see-through HMD 와 optical see-through HMD 간 색상 및 크기 차이를 위한 보정 방법 ..	97
김주환 (원광대학교), 김호원 (한국전자통신연구원), 조동식 (원광대학교)	
가상현실 방향전환보행을 위한 경로예측 시스템	99
조용훈, 이동용, 이인권 (연세대학교)	
가상현실환경에서 사운드를 이용한 시각적 멀미 감소	101
박찬종 (연세대학교), 윤종철 (강원대학교), 이인권 (연세대학교)	
증강현실을 이용한 아동 대상 영어단어 학습 콘텐츠 구현	103
김은경, 유영제, 정지연, 송창근, 이정, 김선정 (한림대학교)	
마커 기반 증강 현실을 이용한 네일 아트 어플리케이션	105
한예진, 서수중, 송창근, 이정, 김선정 (한림대학교)	
시뮬레이터 멀미의 요인 조사 및 분석	107
전한재, 송창근, 이정, 김선정 (한림대학교)	

2018 한국컴퓨터그래픽스 대상



이 승 용

(포항공과대학교 컴퓨터공학과 교수)

한국컴퓨터그래픽스학회는 2018년 한국컴퓨터그래픽스 대상 수상자로 포항공과대학교의 이승용 교수를 선정한다. 이승용 교수는 컴퓨터 그래픽스 분야에서 학문적 영향력과 산업계 파급력을 가진 혁신적인 연구결과를 꾸준히 발표하였으며 세계최고 수준의 연구자들과 다양한 국제공동연구를 통해 상용화 가능한 원천기술 확보와 응용기술 개발에 힘 쏟는 등 한국컴퓨터그래픽스학회가 명실 공히 세계최고 수준의 학술연구 단체로 국내외에서 인정을 받는데 크게 기여하였다.

이승용 교수는 1996년 이후 현재까지 포항공과대학교 컴퓨터공학과에 재직하면서 영상 몰핑 연구를 시작으로 영상 디블러링, 비사실적 렌더링, 메쉬 처리, 삼차원 복원 등의 연구를 다루어왔다. 다양한 주제의 문제를 다루면서도 각 문제에 대한 혁신적인 해결책을 제시해 컴퓨터 그래픽스 분야의 최고 수준 학술지인 ACM TOG (SIGGRAPH, SIGGRAPH Asia), IEEE TVCG, Computer Graphics Forum (Eurographics, Pacific Graphics)에 지속적으로 연구 성과를 게재하였다. 지치지 않는 학문에 대한 열정으로 그는 최근 연구영역을 영상처리와 컴퓨터 비전 분야로 확장하면서 CVPR, ICCV, ECCV를 비롯한 컴퓨터 비전 최고 수준의 국제 학술대회에도 연구 성과를 발표하고 있다.

이승용 교수는 학문적 견고함과 더불어 산업적 파급력이 있는 연구에 매진하였다. 1997년 IEEE TVCG에 발표한 Scattered Data Interpolation 논문은 Google Scholar 기준 현재 1,000회 이상 인용되고 있으며 그래픽스뿐만 아니라 의료영상, 지리학 등 다양한 분야에서 활용되고 있다. 2009년 SIGGRAPH Asia에 발표된 영상 디블러링과 2014년 IEEE TPAMI에 게재된 사진 수직보정 연구는 미국 Adobe Systems에 기술이전 후 최신 제품군에 탑재됨으로써 그의 연구에 대한 산업적 가치를 인정받은 바 있다. 이러한 연구 성과들의 우수성을 인정받아 IT원천 사업 (2008-2011)과 SW스타랩 (2015-) 등 대형국책과제에 선정되기도 하였다.

이승용 교수는 특히 국내외 학술분야에서 활발한 학술활동을 수행하여 왔다. 한국컴퓨터그래픽스학회 2014-2015년 회장으로 활동하면서 학회 발전에 이바지하였으며, Pacific Graphics Steering Committee와 Asia Graphics 초대 부회장을 각각 2008년과 2016년부터 현재까지 역임하며 해외 연구자들과 국내 컴퓨터 그래픽스 연구자들 간의 가교역할을 수행하고 있다. 또한 이승용 교수는 Pacific Graphics 2014 조직위원장, Pacific Graphics 2009 프로그램 위원장, Israel-Korea Bi-National Conference 2003 프로그램 위원장을 역임하면서 활발한 국제학술회의 활동을 하였으며, IEEE TVCG, IEEE CG&A, Computer Graphics Forum, The Visual Computer 등 국제학술지 편집위원을 역임하였다.

이승용 교수의 끊임없는 학문적 열정과 세계 최고 수준의 기술 개발을 위한 실용정신은 컴퓨터 그래픽스 학문분야의 발전에 크게 기여하였다. 이러한 그의 공로에 존경을 담아 이승용 교수에게 한국컴퓨터그래픽스 대상의 자랑스러운 여섯 번째 상을 수여한다.

* 이전 수상자: 2008년 신성용(KAIST), 2009년 김명수(서울대), 2011년 김창현(고려대), 2015년 김명희(이화여대), 2016년 고희석(서울대)

초청강연

Video Interaction at Scale

김주호 교수

KAIST 전산학부

강연 내용

The field of human-computer interaction is witnessing an increasing number of “interaction at scale” systems, in which a massive group of users interact via technology to collaborate, learn, discuss, share information, and make decisions. Three driving forces of interaction at scale are crowds, data, and AI: Crowdsourcing tackles complex problems by involving a large group of people via microtasks and coordination, data-driven interaction techniques leverage large-scale data to create new user experiences, and AI-powered interaction enables more intelligent and adaptive UIs.

In this talk, I'll present a few novel interactive systems designed to support “video learning at scale”. While millions of learners today are watching videos on online platforms, existing video interfaces are not designed to support learning, with limited interactivity and lack of information about learners' engagement and content. Making these improvements requires deep semantic information about video that even state-of-the-art AI techniques cannot fully extract. I take a data-driven approach to addressing this challenge, using large-scale learning interaction data to dynamically improve video content and interfaces. Specifically, my research introduces learnersourcing, a form of crowdsourcing in which learners collectively contribute novel content for future learners while engaging in a meaningful learning experience themselves. I will present learnersourcing applications designed for massive open online course videos and how-to tutorial videos. These systems demonstrate how learnersourcing can enable more interactive, collaborative, and data-driven learning. I will discuss design implications and future directions of relevance to the graphics research community.

강연자 이력

Juho Kim [juhokim.com] is an Assistant Professor in the School of Computing at KAIST, and directs KIXLAB (the KAIST Interaction Lab) [kixlab.org]. His research in human-computer interaction focuses on building interactive systems that support interaction at scale: crowdsourcing and human computation, online education and learning at scale, civic engagement and collective action, and interactive data analytics and mining. He creates interactive and collaborative technology that empowers conventionally passive populations -- students receiving instruction and citizens influenced by social issues -- to be active and self-directed participants who initiate deeper learning and collective action. He earned his Ph.D. from MIT in 2015, M.S. from Stanford University in 2010, and B.S. from Seoul National University in 2008. In 2015-2016, he was a Visiting Assistant Professor and a Brown Fellow at Stanford University. He is a recipient of 10 paper awards from ACM CHI, ACM Learning at Scale, ACM IUI, and AAAI HCOMP, and the Samsung Fellowship.

초청강연

VR 인터랙티브 애니메이션 [미니월드 VR]

장육상 교수

중앙대학교 첨단영상대학원

강연 내용

이번 초청강연에서는 2016년 3월부터 개발을 진행하여 2년 반의 제작과정을 거쳐 중앙대 3D 애니메이션 연구실에서 제작한 VR 인터랙티브 애니메이션 [미니월드]의 제작과정과 그 과정에서 도출된 VR 개발 노하우를 다음과 같은 세부내용들로 공유하고자 한다.

▶ 스팀핑크 세계가 만들어지기까지 ▶ 유저인가? 관객인가? ▶ 스토리 VS 인터랙션 ▶ VR 공간, 이동, 그리고 멀미 ▶ 내손위의 요원, 인터랙션 ▶ 아티스트 & 프로그래머

강연자 이력

▶ 경력

- 중앙대학교 첨단영상대학원 영상학과 교수 (2003-현재)
- 드림웍스 애니메이션 테크니컬 디렉터 (작품 : 신밧드, 7대양의 전설, 2003)
- 20세기 폭스 블루스카이 스튜디오 테크니컬 디렉터 (작품 : 아이스 에이지, 2001)

▶ 학력

- 학사 : 홍익대학교 미술대학 섬유미술 전공 (1991-1995)
- 석사 : 오하이오 주립대학교 Art & Technology/ACCAD M.F.A (1997-1999)

▶ 대표 작품

- In The Forest (2004)
2004 자그레브 애니메이션 페스티벌 'Grand Competition' 진출
2005 안시 애니메이션 페스티벌 "Short film" 본선 진출
- Mom (2010)
2010 안시 애니메이션 페스티벌 "Short film" 공식 진출
유튜브 현재 1500만 뷰
- 북극곰 파오 (2013)
2013 이탈리아 베네통 라이브 윈도우 프로젝트 인터랙티브 애니메이션 <북극곰 PAO> 감독
- VR 인터랙티브 애니메이션 '미니월드' (2018)
2018 부천 판타스틱영화제 공식 초청

초청강연

CT R&D 현황과 사업방향

곽재도 문화기술 PD

문화체육관광부, 한국콘텐츠진흥원

강연 내용

I will introduce the Culture Technology R&D program and strategic policy of the Ministry of Culture, Sports, and Tourism. Also, I will explain the project examples of performing organizations. In addition, we'd like to have a chance to discuss issues and comments for the direction of the program.

강연자 이력

Andrew (Jaedo) Kwak is working on a CT R&D program director at the Ministry of Culture, Sports, and Tourism with various related industrial experiences. He's been had more than 20 years of the working experiences for a content & service business development, UX/UI Design and 3D development. Also, he is actively participating in an expert committee member or BOD of Korea government council and association such as Presidential Council on Intellectual Property, VR Industry Association and so on.

EDUCATION

- Ph.D, Computer Science, SangMyeong University
- MS, Computer Science, Rochester Institute of Technology
- BS, Computer Science, Korea University

PROFESSIONAL EXPERIENCE

- 2017~ Current, CT R&D Director at Ministry of Culture, Sports, and Tourism
- 2006~2017, Evangelist & Director, Business Development, LG Electronics (Seoul, Korea)
- 2004~2006, Co-founder, ANOXIS Inc. (CA, USA)
- 2000~2004, Tech Marketing Director, DTI Inc. (NY, USA)
- 1998~2000, Sr. SW Engineer, ALSTOM Inc. (NY, USA)

OTHERS PROFESSIONAL EXPERINCE

- Expert Committeeman, Presidential Council on Intellectual Property, Korea
- Committee Member of Digital Contents Council, Ministry of ICT & Science, Korea
- Board of Director, Korea VR Industry Association
- Board of Director Korea Picture and Movie Association

후원기관 초청강연

HoloLens를 활용한 MR 콘텐츠 기술 개발

김종용 대리

비즈아이엔에프

khj8499@dongguk.edu

강연 내용

비즈아이엔에프는 창의적 아이디어를 통해 현실과 가상 세계를 융합하여 새로운 미래를 그려나가는 기업입니다. 혼합현실, 증강현실, 가상현실, 컴퓨터그래픽스, 시각화 기술 기반 실감 인터랙티브 콘텐츠 제작 및 창의적 멀티미디어 응용 소프트웨어 개발 서비스를 제공합니다. 이번 강연에서는 비즈아이엔에프에서 개발된 콘텐츠와 기반 기술에 대해 소개합니다.

후원기관 초청강연

Realistic Digital Creature Creation with Quadruped Motion Capture

송재원 CTO

(주)투썸디지털아이디어

jaewon.song,j1@gmail.com

강연 내용

사실적인 디지털 크리처의 제작은 영화 CG VFX 제작에 있어 언제나 도전적인 과제입니다. (주)투썸디지털아이디어에서는 최근 제작중인 영화 <안시성>에 등장하는 사실적인 말 크리처를 위하여 국내 최초로 영화에 활용되는 말의 모션 캡처를 진행하고, 해부학에 기반한 디지털 크리처의 신체 계층구조 컨트롤 기술을 통하여 극사실적인 몸체 표현을 구현하였습니다. 발표를 통하여 세계 수준의 디지털 크리처 기술과 현재 한국의 기술 수준을 비교하고, 작업 공정과 기술 요소들, 또 앞으로 나아가야할 방향에 대하여 토론하는 시간을 갖고자 합니다.

2018 석사논문상 후보

김승욱 (Seung-wook Kim) wook0249@korea.ac.kr

김승욱(영어: Seung-wook Kim, 1994년 2월 18일 ~) 학생은 2016년 봄에 고려대학교 석박사통합과정에 진학하였으며, 현재까지 한정현 교수의 지도를 받으며 5학기를 이수하였다.

이 학생은 컴퓨터 그래픽스 분야 중에서도 특히 물리 시뮬레이션에 많은 관심을 가지고 있다. 이 학생은 SIGGRAPH Asia 2008에 발표된 "Magnets in Motion" 기법의 한계를 발견하였으며, 이를 해결하기 위해 미시자기학 개념을 채택하여 "Magnets in Motion"의 한계를 극복할 새로운 기법을 제안하였다. 이 기법의 결과는 이번 8월 SIGGRAPH에 발표될 예정이다.

또한 이 학생은 서울교대와의 공동 연구를 통해, 실제 과학 실험에서 자기장을 수치해석적으로 분석하여 자석의 자화 상태를 추정하는 기법을 제시하였고, 제3저자로 논문을 작성하여 SCI 학술지인 Journal of the Korean Physical Society에 게재하였다.

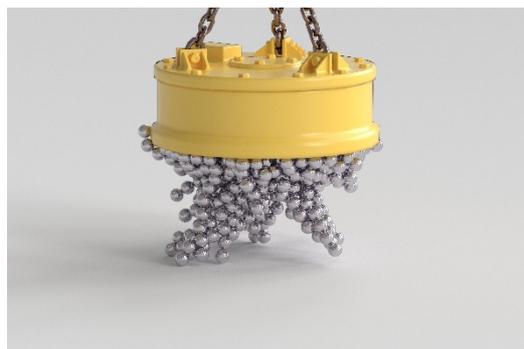
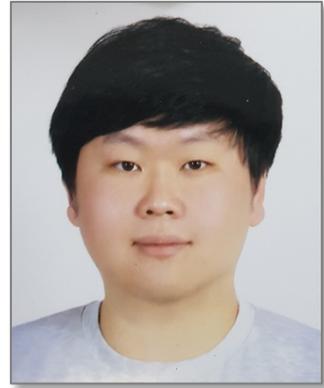


그림 1: 자석 시뮬레이션 연구 결과 (자석 시뮬레이션 장면)

2018 석사논문상 후보

이승환 (Seunghwan Lee) lsw9021@mrl.snu.ac.kr



이승환은 2016년 9월에 서울대학교 컴퓨터공학부에 석박사통합과정으로 입학하였다. 2012년 서울대학교 원자핵공학과에 입학하였고, 핵공학 장학생이 되기 위해 원자로심 시각화(Reactor Visualizer)이라는 주제를 정하였다. 장학생은 떨어졌지만 컴퓨터 그래픽스 분야에 흥미를 느껴 수업을 듣게 되었다. 학기를 마친 후 운동연구실에 방문하였고, 삼고초려 끝에 인턴 자리를 어렵게 차지하며 컴퓨터 공학에 입문하게 되었다.

뒤쳐지지 않기 위해 무수한 노력을 하였고, 약 5개월의 인턴기간 동안에 물리 시뮬레이션에 빠지게 되었다. 그리고 유체, 변형체, 강체 시뮬레이션 등 다양한 가지의 물리를 공부하고 구현을 하던 중에, '부피를 가지는 근육 모델링' 연구를 시작하였다. 부피를 가지는 근육을 시뮬레이션 하는 것은 계산 시간이 오래 걸리고 복잡하다. 그럼에도 불구하고 근골격계가 가지는 고유한 특성들을 재현할 수 있어 실제 사람을 정확하게 재현하기 위해서는 필요한 일이다. 하지만 이 연구의 특성상 다양한 물리 시뮬레이션에 익숙해야 했다. 연구 진행중 변형체 시뮬레이션에 발목이 잡히자, 지도교수님의 추천으로 Wisconsin-Madison 대학에 방문하여 Eftychios Sifakis와 같이 연구를 수행했다. 그 결과, *Volumetric Muscle Controller*(SCA 2017, Poster)를 거쳐 *Dexterous Manipulation and Control with Volumetric Muscles*(SIGGRAPH 2018)를 출판한다.

이 논문을 계기로 사람의 근골격계에 더 관심이 생겨, 병적 보행을 분석하고 재현하는 것을 연구하고 있다. 사람은 신경계, 근골격계가 복잡하게 상호작용하며 걷기 때문에, 이를 컴퓨터에서 모델링하고 재현하는 것은 쉽지 않은 일이다. 특히, 특정 부위가 정상범주에서 벗어나 제대로 걸음걸이를 할 수 없는 경우, 이를 재현하는 것은 더 어려운 일임에 틀림없다. 최근에 떠오르고 있는 강화 학습을 이용하여 효율적으로 근골격계를 제어하는 알고리즘을 개발하고, 다양한 병적 보행을 재현할 수 있는 시스템을 개발할 계획이다.

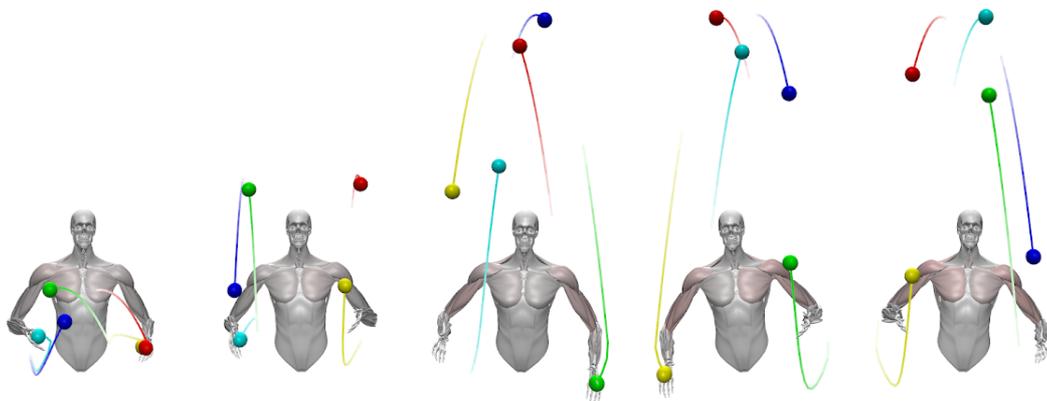


그림 1: Dexterous Manipulation and Control with Volumetric Muscles

심층강화학습을 이용한 군중 시뮬레이션

이재동⁰, 이제희
서울대학교
{jaedong, jehee}@mrl.snu.ac.kr

Learning to Move in Crowd

Jaedong Lee⁰, Jehee Lee
Seoul National University

요약

군중 시뮬레이션의 주된 목표는 수많은 사람들이 주어진 환경에서 각 사람의 움직임을 최대한 사실적으로 묘사하는 것이다. 사람은 주변 환경을 관찰하고 그로부터 자신이 놓여있는 상황을 이해한 다음 어떻게 움직일지를 결정하는 인지-사고-행동의 과정을 통해 움직인다. 우리는 이러한 일련의 과정을 심층강화학습을 통해 구조화하여 군중 속에서의 사람의 움직임을 재현하였다. 심층강화학습의 알고리즘으로는 배우-비평가 모델 중 하나인 심층 결정론적 정책 기울기를 이용하였다. 본 연구에서 제안한 방법론을 통해 장애물 회피, 원형 교차, 복도 교차, 수직 교차의 4가지 시나리오에서 군중들의 움직임을 사실적으로 재현함을 보였다.

1. 서론

군중 시뮬레이션은 컴퓨터 애니메이션 분야에서 오랫동안 활발하게 연구되어온 주제이다. 많은 수의 객체들의 움직임을 시뮬레이션 하는 것은 가상 현실, 영화나 게임 같은 미디어 엔터테인먼트, 재난 대피 시뮬레이션 등 다양한 분야에 활용된다. 이러한 연구들의 공통된 목표는 군중들의 움직임을 자연스럽게 재현해내는 것이다. 사실적인 움직임을 만들어내기 위해서는 객체들 사이의 상호작용을 효과적으로 다루어야 하는데, 가장 중요한 것은 객체들 사이의 충돌을 방지하는 것이다. 또한 인지 과학 분야의 연구에 따르면, 사람은 인지-사고-행동의 과정을 통해 움직이는데 이 과정에서 눈으로 들어오는 시각적 정보에 큰 영향을 받는다. 군중 속에서 각 객체들이 충돌하지 않고 자연스럽게 움직이기 위해서도 시각적 정보를 통해 주변 상황을 인지하고 행동 계획을 수립하는 것이 중요하다.

최근의 급격한 딥러닝의 발전으로 심층강화학습과 관련한 연구가 활발하게 진행되고 있다. 강화학습은 마르코프 결정 과정을 푸는 하나의 방법론이다. 강화학습의

목표는 미래의 상황을 고려하여 현재의 상황에서 최선의 행동을 하도록 하는 정책을 학습하는 것이다. 이러한 정책은 인공 신경망을 통해서 표현될 수 있는데, 인공 신경망은 시각 정보나 시계열 정보 등의 다양한 형태의 입력에 대해서 학습이 가능함이 알려져 있다. 우리는 이러한 특징들을 활용해 강화학습의 구조 안에서 시각정보를 인공 신경망을 통해 처리하며 군중의 움직임을 시뮬레이션하는 시스템을 만들었다. 본 연구에서 이 시스템을 통해 사람의 움직이는 원리를 재현하여 군중들의 움직임을 만들어내는 방법론을 제안한다.

2. 접근 방법

사람의 인지의 과정은 시뮬레이션 상에서 각 객체가 시각 정보를 받아들이는 것으로 재현했다. 우리는 [그림 1]에서와 같이, 각 객체의 이산화된 깊이 정보를 시각 정보로 활용했다. 이산화된 깊이 정보는 객체가 바라보는 방향을 기준으로 190°의 시야에 대해 매 10° 마다 측정된 깊이 정보이다. 사람의 사고-행동 과정은 현재 그 사람이 행동을 통해 이루려는 목표와 관련이 있다. 보통 사람은 자신이 도달하려는 목적지를 향해 움직이며, 움직이는 과정에서 충돌을 피하면서 효율적인 경로를 통해 목적지에 도달하기를 원한다. 우리는 이러한 과정을 강화학습을 통해 구조화 하였다. 강화학습의 수식 내에서 우리는 충돌을 피하며 목적지에 도달할 수 있도록 보상 체계를 설정했으며, 매순간 최선의 행동을 하도록 하는 정책을 학습시켰다. 정책이 학습이 되고 나면, 각 객체는 매 순간 입력으로 들어오는 시각정보를 바탕으로 학습된 정책을 통해 최선의 행동을 하게 된다.

우리는 강화학습의 알고리즘으로 배우-비평가(Actor-Critic)이며, 모델로부터 자유로운 형식인 심층 결정론적 정책 기울기(DDPG)[1]를 이용하였다. DDPG는 행동 공간이 연속적인 경우에 정책 학습이 잘 된다고 알려진 방법론이며 다양한 연구에 활용되었다. Actor-Critic이란 [그림 2]에서 나와있듯이 배우 네트워크(Actor network)와 비평가 네트워크(Critic network)로 구성을 이루며, 각각 행동을 결정하고, 행동을 평가하는 역할을 하는 구조를 말한다.

* 구두발표논문

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW 컴퓨팅산업원천기술개발사업(SW스타랩)의 연구결과로 수행되었음. (IITP-2017-0-00878)

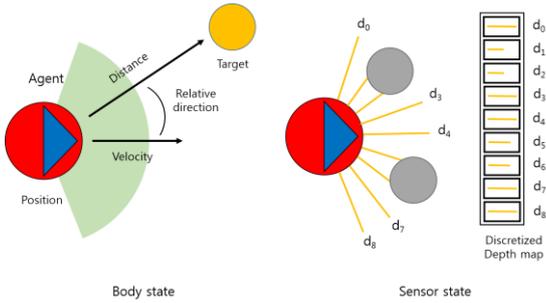


그림 1: 객체의 신체 상태와 센서 상태.

강화학습을 통해 군중들의 움직임을 재현하기 위해 시스템을 정의하였다. 우선 군중 시뮬레이션에 사용될 객체를 정의한다. [그림 1]에서는 같이 객체는 신체 상태와 센서 상태를 자신의 상태로 갖는다. 신체 상태는 각 객체의 위치, 속도, 목적지까지의 방향으로 구성되고 센서 상태는 객체의 시각 정보를 담고 있으며 최근 3개 프레임의 이산화된 깊이 정보로 구성된다. 객체의 상태 정보는 정책 네트워크의 입력으로 주어지게 되며 정책은 이를 기반으로 하여 행동을 출력한다. 행동은 속력의 변화량과 방향의 변화량 값으로 정의하고 이 2차원의 행동 값은 각각 현재 객체의 속력과 방향에 누적된다. 각 객체는 새로 결정된 자신의 속력과 방향에 기반해서 한 단계씩 이동하게 된다.

객체의 정책은 Actor network로 구성이 된다. [그림 2]에서와 같이 Actor network는 객체의 신체 상태를 입력으로 하는 네트워크와 센서 상태를 입력으로 하는 네트워크의 합으로 구성이 된다. Critic network는 Actor network의 구조에 행동 값을 입력으로 하는 네트워크가 추가된 형태이다. Actor network는 2차원의 행동 값을 출력하여 객체의 행동을 결정하고, Critic network는 입력된 행동에 대한 평가 가치를 출력한다.

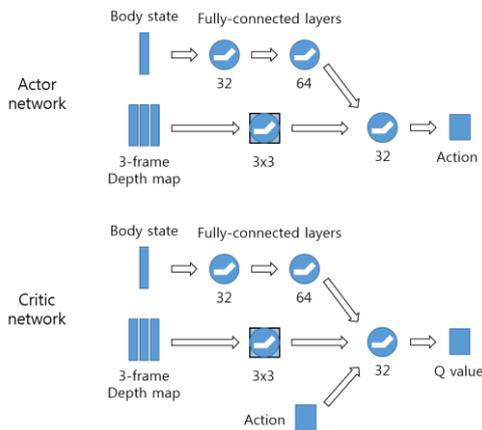


그림 2: 배우-비평가 네트워크.

Critic network가 객체의 행동을 평가하는 기준은 사용자가 설계한 보상 체계에 의해 결정된다. 각 객체는 주어진 환경에서 한 단계씩 행동을 하면서 사용자가 설계한 보상 체계에 의해 보상 값을 받게 되는데,

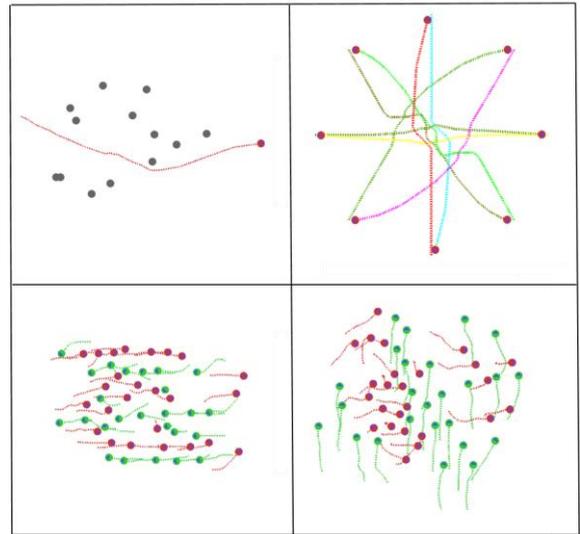


그림 3: 장애물 회피, 원형 교차, 복도 교차, 수직 교차 시나리오에 대한 결과. 빨간색 에이전트는 오른쪽 방향으로, 초록색 에이전트는 왼쪽 혹은 위쪽 방향으로 움직인다.

최종적으로는 이러한 보상의 합이 최대가 되도록 정책 네트워크가 학습이 된다. 이 연구에서 보상체계는 목적지까지 도달하도록 하는 보상과 다른 객체 혹은 장애물과 충돌하지 않도록 하는 보상으로 이루어진다. 목적지까지 도달하기 위한 보상은 목적지까지의 거리의 제곱의 음수값에 비례하도록 설정하여 가까워 질수록 높은 보상을 받도록 하였고 충돌에 대한 보상은 충돌 발생시 고정된 음수값을 보상으로 받도록 하였다.

3. 실험

우리는 잘 학습된 하나의 정책 네트워크를 통해 총 4가지의 시나리오에 대해 실험을 하였다. 시나리오는 순서대로 장애물 회피, 원형 교차, 복도 교차, 수직 교차이다. 학습은 4~5시간이 걸렸고, 네트워크는 가장 학습이 쉬운 시나리오부터 어려운 시나리오 순으로 단계적으로 학습했다. [그림 3]은 실험에 대한 결과이다. 각 시나리오마다 객체의 수는 1, 8, 40, 40이며 다양한 숫자의 객체들에 대해서도 잘 동작함을 확인하였다. 회색의 물체는 장애물을 나타내며, 빨간색의 객체와 파란색의 객체는 서로 다른 방향으로 움직이는 객체들이다. 객체가 움직이는 경로를 출력하였으며 이를 통해 객체들이 충돌 없이 목적지까지 움직이는 것을 알 수 있다. 원형 교차의 경우 서로 다른 색으로 경로를 출력하여 각 객체의 움직임을 표현하였다.

참고문헌

[1] Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nocias Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. 2015. Continuous control with deep reinforcement learning, CoRR abs/1509.02971 (2015). <http://arxiv.org/abs/1509.02971>

물리 시뮬레이션에 기반한 2차원 연체 동물 캐릭터 제어*

민세희⁰, 이제희
 서울대학교 컴퓨터공학부
 {sehee,jehee}@mrl.snu.ac.kr

Realistic Simulation and Control for Mollusks-like Planar Character

Sehee Min⁰, Jehee Lee
 Dept of Computer Science and Engineering, Seoul National University

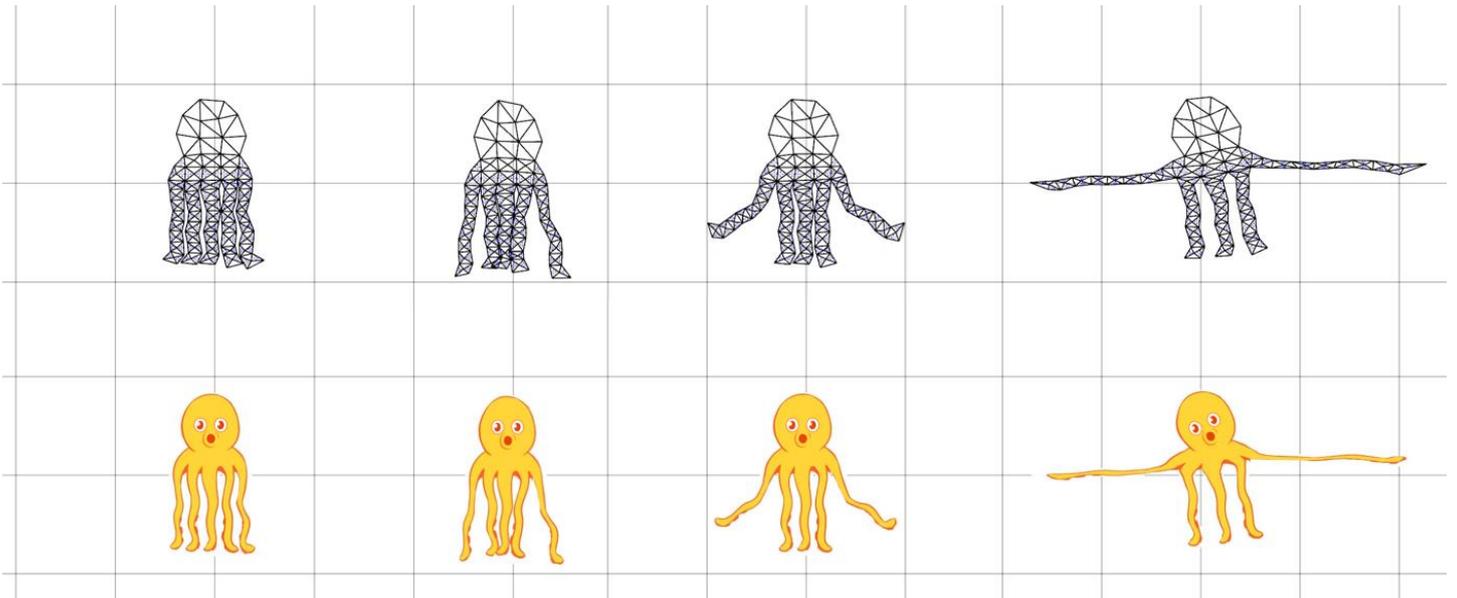


그림 1 2차원 연체 동물 캐릭터의 몸이 잡아당겨지는 모습.

상단: 152개의 꼭지점과 308개 자유도를 가지는 메쉬. 하단: 문어 이미지 텍스처로 렌더링 된 결과.

요약

연체 동물을 시뮬레이션하고 제어하는 것은 높은 차원의 자유도와 복잡하고 완전히 알려지지 않은 변형 가능한 몸체 구성으로 인해 어려운 문제로 여겨진다. 연체 동물 캐릭터의 구성 모델은 근육 수축과 이완을 통해 움직이는 연체 동물의 특성을 나타내면서 동시에 그룹별 부피 보존의 법칙을 반드시 지켜야 한다. 우리는 2차원 상에서 수중 환경에서 연체 동물의 특징을 잘 드러나게 함과 동시에 다양한 움직임을 보여줄 수 있는 구성 모델을 디자인하고 견고하고 빠른 시뮬레이터와 최적 제어를 위한 iLQG 기반 컨트롤러를 제안한다.

1. 서론

물리 시뮬레이션을 기반으로 한 생물체의 현실적인 움직임을 재현하는 것은 애니메이션 연구자들에게 도전

정신을 불러 일으켜 왔다. 인체 뿐만 아니라 생체 모양 및 생존 환경의 차이에 따라 생기는 생물체의 독특한 움직임을 컴퓨터 시뮬레이션에서 재현하는 것이 시도되었다. 대부분의 생물체 재현에서 움직임을 선보이기 위해 골격의 움직임을 통해 토크를 발생한다고 가정 한 후에 시뮬레이션을 수행한다. 이런 가정에서 출발하여 연체 동물 역시 부드러운 몸체 내부에 임의의 뼈대를 가정하고 뼈대의 움직임을 따라 자연스럽게 몸체가 움직이도록 하는 방법이 제시되었다.[1] 하지만 실제 문어와 달팽이와 같은 연체 동물의 경우 이동 메커니즘은 완전히 다르다. 연체 동물은 골격이 전혀 없이 근육의 몸 전체를 지지하고 근섬유의 수축과 이완을 통해 움직이는 근섬유(Muscular Hydrostat) 구조의 중요한 특징을 가지기 때문이다. 연체 동물의 근섬유는 3차원에서 연체 동물의 자유로운 움직임을 유도하는 서로 직각인 횡, 종 및 경사의 세 방향으로 구성된다. 또한 신체의 총 부피가 일정해야 한다는 중요한 조건을 충족시키기 위해 근섬유는 서로 밀접한 상관 관계를 가진다. 이러한 특징을 충족시키는 구성 모델을 만드는 것은 연체 동물의 시뮬레이션에 매우 중요한 역할을 한다.

* 구두 발표논문

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW 컴퓨팅산업원천기술개발사업(SW스타랩)의 연구결과로 수행되었음. (IITP-2017-0-00878)

이 논문에서 우리는 실제 연체 동물과 같은 특성을 재현하기 위해 새로운 재료 모델을 디자인하여 유한 요소 해석(Finite Element Method)로 표현한 후 시뮬레이션을 수행한다. 우리는 2차원에서 디자인된 구성 모델의 무수한 자유도를 안정적이고 빠른 속도로 시뮬레이션하기 위해 준 뉴턴 기법(Quasi-Newton Method)[2]을 구현하고 최적의 제어를 보여줄 iLQG (iterative Linear Quadratic Gaussian)[3] 기반 제어기를 제안한다.

2. 문제 해결 방법

우리는 유한 요소 해석을 사용하여 연체 동물의 몸체를 모델링한다. 캐릭터의 재료 모델은 비선형적인 변형과 부피 보존의 특성을 표현하기 위해서 네오 후키안(Neo-Hookean)[4] 모델을 기반으로 한다. 우리는 문어 팔의 해부학에서 착안한 근섬유를 그룹화하고 그룹당 부피 보존으로 변형하여 설계한다. 2차원 상에서 경사 방향의 근섬유의 역할은 횡, 종 두 방향의 근섬유의 조합으로 설명 가능하기 때문에 그림 2와 같이 근섬유를 디자인한다. 근섬유를 통해 신체가 수축하고 이완할 수 있으며 근섬유의 활성화 단계는 0부터 1까지이며 값이 클수록 수축한다. 각 근섬유가 가져야 하는 활성화 단계의 조합 문제를 해결함으로써 원하는 캐릭터의 모습을 얻어낼 수 있다.

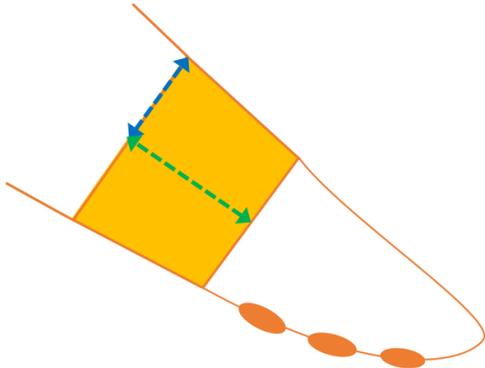


그림 2 그룹당 2종류의 근섬유. 파란색: 횡, 녹색: 종.

메쉬의 자유도가 매우 크기 때문에 캐릭터를 시뮬레이션하는데 많은 시간이 필요하다. 정적(Static) 혹은 준 정적(Quasi-Static)을 사용하여 역기구학(Inverse Kinematics)을 풀어 캐릭터의 자세를 얻어낼 수 있는 방법이 있지만 이 방법은 캐릭터 애니메이션의 풍부한 움직임을 표현하기에는 부족하다. 우리는 준 뉴턴 기법을 사용하여 실시간에 타당한 물리적 정확성을 보이는 동역학(Dynamics) 시뮬레이션을 구현한다. 준 뉴턴 기법은 FEM의 에너지를 표현할 때 제약 조건과의 거리로 측정한다는 제약 조건 투영 (Constraint Projection)의 관점에서 투영 역학 (Projective Dynamics)[3]와 유사하다. 하지만 투영 역학에서는 비선형적인 재료 모델에 대한 제약 조건을 설정하기가 어렵기 때문에 네오 후키안 모델과 같은 기본적인 비선형적인 모델도 사용할 수가 없다는 단점을 가진다. 따라서 준 뉴턴 기법을 사용하여 양 확정 행렬인 특정 선형 연산자를 2계 미분 행렬로 근사하여 실시간에 동작하는 빠른 시뮬레이션과 물리적으로 타당한 시뮬레이션의 장점을 모두 취한다.

연체 동물은 일반적으로 부드러운 몸체와 외투막을 가지고 있으므로 대부분 물 속에서 서식한다. 간단한 유체 역학 공식을 적용함으로써 캐릭터가 유체에 의한 항력과 부력을 가지도록 하여 실제 생물과 비슷한 자연스러운 움직임을 보이도록 설계한다.

대화형 환경에서 캐릭터 모델을 사용자의 입력에 따라 자유자재로 제어하기 위해서 주어진 위치에 대한 최적 제어를 계산해내는 것이 매우 중요하다. 우리는 제적 최적화 중 하나인 iLQG를 사용하여 캐릭터의 움직임을 제어한다. iLQG는 2차 미분값이 아닌 1차 미분만을 사용하기 때문에 보다 빠른 속도로 수렴할 수 있다는 장점을 가진다. 제어기의 성능 향상을 통해 이전 동작을 기반으로 곡선을 통해 기존 동작을 제공하는 방법을 적용한다. 그림 1에서 iLQG 제어기는 사용자로부터 받은 목표 지점으로 1초 동안 잡아당겨지는 자연스러운 움직임을 생성하기 위하여 회당 4.6초의 계산을 총 10회 반복한다.

우리는 추후 연구로 3차원 상에서 연체 동물 캐릭터가 움직이도록 하는 것을 시도하여 실제 생물체의 운동 매커니즘을 재현하는 것을 목표로 한다. 접촉 및 충돌 제약 조건 등을 설정하여 연체 동물이 가지는 특징을 묘사할 것이다. 또한 현재의 iLQG 제어기는 실시간 제어가 불가능했지만 알맞은 최적 값 함수의 근사를 구하는 과정을 설계하여 경로 생성 속도를 높일 수 있을 것이다. 실시간 물리 시뮬레이터를 가지는 사람의 보행과 새의 날갯짓 비행이 최근 심층 강화 학습을 통해 성공적으로 구현되었다. 심층 강화 학습의 경우 실제 생물체가 움직임을 배우는 과정과 매우 닮아있어 적절한 환경과 알맞은 보상 체계를 설계한 후 움직임을 스스로 배우도록 하는데 알맞다. 이에 착안하여 심층 강화 학습을 도입하여 풍부한 움직임 전환과 실시간 제어 속도를 시도할 것이다.

참고문헌

- [1] Hahn, Fabian and Martin, Sebastian and Thomaszewski, Bernhard and Sumner, Robert and Coros, Stelian and Gross, Markus, Rig-space Physics, *ACM Transactions on Graphics(TOG)*, 31(4):1-8, 2012.
- [2] T Liu, Tiantian and Bouaziz, Sofien and Kavan, Ladislav, Quasi-Newton Methods for Real-Time Simulation of Hyperelastic Materials, *ACM Transactions on Graphics(TOG)*, 36(4):1-16, 2017.
- [3] Bonet, Javier and Wood, Richard D., Nonlinear continuum mechanics for finite element analysis, Cambridge University Press, 2008.
- [4] Tassa, Yuval and Erez, Tom and Todorov, Emanuel, Synthesis and Stabilization of Complex Behaviors through Online Trajectory Optimization, IROS, 4906-4913, 2012.
- [5] Bouaziz, Sofien and Martin, Sebastian and Liu, Tiantian and Kavan, Ladislav and Pauly, Mark, Projective Dynamics: Fusing Constraint Projections for Fast Simulation, *ACM Transactions on Graphics(TOG)*, 33(4):1-11, 2014.
- [6] Won, Jungdam and Park, Jongho and Kim, Kwanyu and Lee, Jehee, How to Train Your Dragon: Example-guided Control of Flapping Flight, *ACM Transactions on Graphics(TOG)*, 36(6):1-13, 2017.

다목적 제어 학습을 이용한 대화형 캐릭터 애니메이션

이경호, 이세영⁰, 이제희
 서울대학교
 {khlee, seyounglee, jehee}@mrl.snu.ac.kr

Interactive Character Animation by Learning Multi-Objective Control

Kyungho Lee, Seyoung Lee⁰, Jehee Lee
 Seoul National University

요약

대화형 캐릭터 애니메이션에서 어떻게 가상의 캐릭터가 사람처럼 움직이게 할지, 어떻게 사용자가 원하는 대로 제어할지는 흥미로운 문제이다. 이 논문에서는 사용자와 상호작용 하면서 연속적인 대화형 캐릭터 애니메이션을 생성하는 새로운 학습 모델을 제시한다. 이 모델에서 사용자는 여러가지 제어 목표를 줄 수 있으며, 움직임이 서로 연관되어 있는 여러 명의 캐릭터를 한번에 제어할 수 있다. 연속적 데이터를 처리하기에 유리한 RNN이라는 네트워크를 이용하여 모션 캡처 데이터를 학습하였다. 보행, 테니스, 농구 등 다양한 움직임 생성을 통해서 이 알고리즘이 다양한 제어 목표와 움직임에 적용 가능함을 검증하였다.

1. 서론

가상의 캐릭터가 실제 사람처럼 움직이도록 하는 것과 사용자가 원하는 대로 가상의 캐릭터를 제어하는 것은 대화형 캐릭터 애니메이션의 핵심 문제이다. 가상의 캐릭터와 상호 작용하는 것은 시간 및 공간적 조건 또한 포함한다. 예를 들어서 사용자가 가상의 캐릭터를 제어할 때 어디로 갈지, 언제 도착할지, 어떻게 움직일지, 무엇을 할 지를 정해줄 수 있다. 이 때 캐릭터의 수가 늘어나고 캐릭터들의 움직임이 서로 연관되어 있다면 더욱 어려운 문제가 된다. 최근에 딥러닝 기술이 사람의 움직임을 구현하는 데에도 많이 쓰이고 있는데, 다양한 레퍼토리의 움직임을 다루거나 제어하고 싶은 목표가 여러 개일 때 제어 학습은 잘 이루어지지 않고 있다.

이 논문에서는 사용자와 상호작용 하면서 연속적인 대화형 캐릭터 애니메이션을 생성하는 새로운 학습 모델을 제시한다. 이 모델은 제어 변수가 추가된 형태의 Recurrent Neural Network를 이용하여 시공간적으로 연속된 모션 캡처 데이터를 효율적으로 학습시킨다. 이 모델의 장점은 하나의 네트워크를 통해

다양한 목적에 맞게 행동하도록 제어할 수 있으며, 폭넓은 레퍼토리의 움직임을 학습시키는 것이 가능하다는 것이다.

우리는 보행부터 테니스, 농구와 같은 다양한 레퍼토리의 움직임을 제어하는 것을 통해서 이 알고리즘이 특정한 움직임을 표현하는 데만 국한되지 않고 다양한 목표를 가진 움직임 전반을 생성해낼 수 있음을 검증하였다.

2. 다목적 제어 모델

이 모델에서는 이산적인 시간 단계에서 환경과 상호 작용하는 캐릭터가 있다고 가정한다. 각 시간 단계에서 캐릭터의 모션에 대한 상태와 사용자가 준 제어 목표와 관련된 상태를 가지고 캐릭터의 다음 모션에 대한 상태를 얻어내게 된다. 특정 캐릭터의 모션은 시간에 따른 루트의 위치, 방향과 각 관절의 회전 각도와 발의 접촉 상태로 나타낼 수 있다. 따라서 이 모델에서는 캐릭터의 모션에 대한 상태를 3 차원 상의 전체 좌표계에서의 캐릭터 루트의 방향과 위치, 각 관절의 상대 방향과 위치, 발의 접촉 여부로 정의하고 있다. 관절의 위치와 회전각도는 중복되는 정보이나 동시에 들어갔을 때 학습에 도움이 되기도 한다.

제어 목표와 관련된 상태는 캐릭터의 움직임이 일어나고 있는 환경에 대한 정보를 제공한다. 제어 목표를 달성하는 과정에서 필요한 소도구나 목표 위치나 시간, 액션 타입 등에 대한 정보가 이에 포함되는데, 제어 목표에 따라서 사용자가 다양하게 설정해 줄 수 있다. 예를 들어서 농구에서 공을 던지려고 할 때 캐릭터로부터 목표 지점까지의 상대 위치와 공을 던지기까지 남은 시간 간격이 사용자가 준 제어 목표와 관련된 상태에 해당한다. 우리의 모델은 하나의 네트워크로 다양한 타입의 액션을 생성할 수 있고 각 타입의 액션이 필요로 하는 서로 다른 제어 정보를 유연하게 줄 수 있다.

3. 네트워크 학습

RNN은 인공 신경망의 한 종류로 출력이 현재의 입력과 과거의 입력에 영향을 받는 네트워크이다. 따라서 순차적인 데이터를 학습시킬 때 유용하다. 우리는 인코더-디코더 구조와 lstm 층을 이용하여

* 구두발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 SIGGRAPH Asia. 2018 에 제출 되었음.

* 본 연구는 미래창조과학부 및 정보통신기술진흥센터의 SW컴퓨팅산업원천기술개발사업(SW스타랩)의 연구결과로 수행되었음 (IITP-2017-0536-20170040).

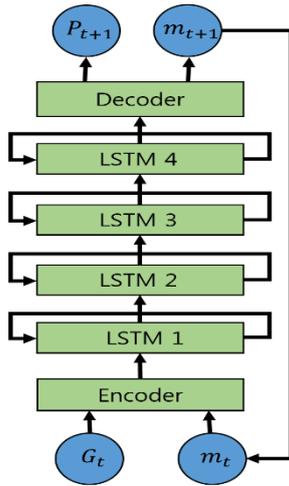


그림 1: 네트워크 구조

입력으로부터 중요한 특성을 뽑아내고 현재로부터 상대적으로 먼 시간 간격에 대한 정보도 이용할 수 있도록 했다. 네트워크의 입력은 2 에서 정의한 캐릭터의 현재 시간 단계 t 에서 모션에 관련된 상태 m_t 와 제어 목표와 관련된 상태 G_t 이며, 네트워크의 출력은 캐릭터의 각 관절의 P_{t+1} 과 다음 시간 단계 $t+1$ 에서의 캐릭터의 모션과 관련된 상태 m_{t+1} 이다. 네트워크의 출력은 다음 시간 단계에서의 캐릭터의 모션을 구할 때의 입력이 된다. 학습은 손실 함수를 최소화시키도록 진행되는데, 손실 함수에는 목표가 되는 모션 정보와의 거리, 발의 미끄러짐, 캐릭터로 복원했을 때의 뼈 길이와 관련된 항이 포함된다.

감독 학습이기 때문에 학습용 데이터를 풍부하게 만들어내는 것이 중요한데, 이 과정은 자동으로 이루어진다. 모션 그래프[1]를 이용해서 연속적이지 않아도 자연스럽게 이어질 수 있는 프레임을 찾아내어 랜덤하게 동작들을 연결시키고, 라플라시안 모션 편집[2]을 통하여 그 데이터를 추가적으로 변형한다. 농구와 같이 규칙이 있는 움직임의 경우는 모션 그래프 대신 동작 문법을 이용하여 규칙을 따르면서 랜덤하게 데이터를 생성할 수 있다[3]. 공과 같은 소도구가 움직임에 포함되는 경우에는 소도구와 관련된 정보가 필요한데 공이 손에서 떨어지는 시간 등을 레이블링 해주는 정도의 간단한 수작업이 요구된다.

4. 결과 및 결론

우리는 위의 모델을 텐서플로우 파이썬 버전을 이용하여 학습시켜서 다양한 데모를 구현했다. 학습에는 NVIDIA GeForce GTX 1080 을 이용했으며 학습 되기까지 평균적으로 12 시간에서 24 시간이 소요되었다.

보행: 이동하고자 하는 위치와 이동 중의 몸의 방향을 제어 목표로 주는 간단한 데모이다. 몸의 방향을 주지 않고 이동하고자 하는 위치만 주는 것처럼 제어 옵션 전체를 다 쓰지 않고 유연하게 끄고 켤 수 있어서 풍부한 움직임을 생성해낼 수 있다. 3 에서 말한 방법을 통해서 짧은 모션 클립으로부터 풍부한 학습용

데이터를 생성하여 학습을 진행한다.

농구: 슛, 패스, 리시브, 캐치, 드리블, 걷기 6 가지의 종류의 움직임을 생성한다. 시간, 위치, 방향 등의 옵션을 조합해서 생성하고자 하는 움직임의 종류에 따라 유연하게 제어 목표를 줄 수 있다. 농구는 규칙이 있는 움직임이기 때문에 동작 문법을 이용하여 학습용 데이터를 생성했다. 사용자가 키보드와 마우스 드래그를 통해서 인터랙티브하게 제어 목표를 주면 네트워크가 바로 제어 목표를 따르면서 규칙에 맞는 움직임을 생성해낼 수 있다.

테니스: 공을 치는 것과 친 후 다시 돌아오는 것 두 가지 종류의 움직임을 생성할 수 있다. 공을 치는 경우 언제 어디로 오는 공을 어느 방향으로 칠지를 제어 목표로 주고, 다시 돌아오는 경우 언제 어느 위치로 캐릭터가 돌아올지를 제어 목표로 주게 된다.

위의 데모를 통해서 우리의 모델이 종류가 다른 다양한 움직임을 생성하는 데에 적용할 수 있으며, 같은 종류의 움직임에서도 여러 가지 제어가 가능하다는 것을 확인할 수 있다.



그림 2: 농구와 테니스 데모 스크린샷

참고문헌

- [1] Jehee Lee, Jinxiang Chai, Paul SA Reitsma, Jessica K Hodgins, and Nancy S Pollard. 2002. Interactive control of avatars animated with human motion data. *ACM Transactions on Graphics* 21, 3 (2002), 491–500.
- [2] Manmyung Kim, Kyunglyul Hyun, Jongmin Kim, and Jehee Lee. 2009. Synchronized multi-character motion editing. *ACM Transactions on Graphics* 28, 3, Article 79 (2009)
- [3] Kyunglyul Hyun, Kyungho Lee, and Jehee Lee. 2016. Motion grammars for character animation. *Computer Graphics Forum* 35, 2 (2016), 103–113.

사람 애니메이션을 위한 다중 세그먼트 발 모델링*

박황필⁰, 유리, 이제희
 서울대학교 컴퓨터공학부
 {hppark, yuri, jehee}@mrl.snu.ac.kr

Multi-Segment Foot Modeling for Human Animation

Hwangpil Park⁰, Ri Yu, Jehee Lee
 Dept. of Computer Science, Seoul National University

요약

우리는 사람의 자세를 자연스럽게 재현하기 위해, 다중 발 모델을 제안한다. 발은 사람이 서 있거나, 걷거나 여러가지 동작을 수행할 때 땅과 상호작용을 하는 주요한 부분이다. 그러나 이전의 연구들은 간단한 발 모델을 사용하였기 때문에 발의 모양이나 움직임의 자세하게 재현할 수 없었다. 우리는 제안된 발 모델을 기존의 균형 제어 시스템에 부착하여 다양한 발 동작을 자연스럽게 보이도록 물리에 기반한 시뮬레이션을 수행하였다.

1. 서론

물리 시뮬레이션 상에서 사실적인 사람의 움직임을 재현하는 것은 컴퓨터 그래픽스와 로보틱스 분야에서 중요한 연구이다. 시뮬레이션에 사용되는 캐릭터는 수십 년 간 간단한 모델에서부터 사람처럼 근육으로 활성화되는 모델까지 발전해 왔다.

발은 지면과 직접적으로 상호작용하는 신체 부위이다. 발은 26개의 뼈, 33개의 관절, 백 여개의 근육으로 이루어져 있는 복잡한 구조를 가지고 있다. 발은 바닥으로부터 오는 충격을 완화하거나, 지지하는 방법을 바꿔가며 중심을 잡기도 하고, 속력을 높이기 위하여 바닥을 밀어 내어 운동량을 얻는 등 여러가지 기능을 수행한다.

발이 보행과 관련된 기능들을 수행하는 중요한 부분임에도, 발 모델링은 상대적으로 간과되어 왔다. 사람의 애니메이션이나 시뮬레이션 연구에서는 대부분 한 개나 두 개의 바디로 이루어진 발 모델이 주로 사용되었다.

이런 발 모델은 자유도가 낮아 사람의 발의 움직임을 사실적으로 재현하기에 어려움이 있다. 우리는 사람 발의 해부학적 구조를 고려하여 여러 개의 세그먼트로 이루어진 발 모델과 발 자세 제어기를 구현하였다.

2. 관련 연구

Meglan과 Berme[1]가 처음으로 땅과 분리된 발 모델을 제안하였다. 이후로 생체역학 분야에서는 6개의 타원체로 이루어진 발 모델[2]이 제안되었으나 모델의 구조가 복잡해 질수록 계산 시간이 증가하였다. Jain과 Karen[3]은 유한요소법(FEM)을 기반으로 한 발 모델을 사용하여 시뮬레이션의 강건성을 높였으나, 발의 모양은 신경 쓰지 않았다.

3. 방법

3.1. 발 모델

우리의 발 모델은 다섯 개의 세그먼트로 이루어져 있다(그림 1 왼쪽). 각 세그먼트에는 발의 뼈와 대응되는 가상의 뼈를 두어 지면과의 충돌을 계산하였다. 각각의 가상의 뼈는 캡슐로 디자인하여 충돌 계산시간을 줄이고자 하였다. 그리고 지면과의 충돌이 없는 노란색 세그먼트를 제외한 나머지 세그먼트를 움직이기 위해 네 개의 관절을 만들었다(그림 1 오른쪽).

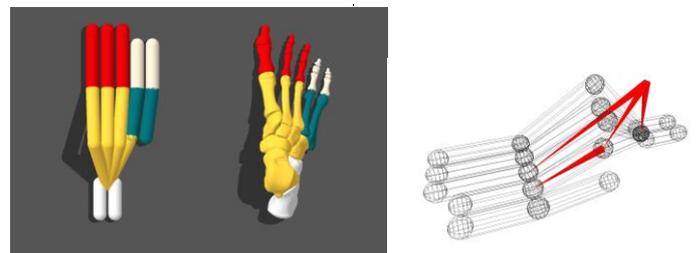


그림 1: 다중 세그먼트 발 모델

3.2. 발 자세 제어기

우리는 발 모션 데이터 없이도 발목관절 정보와 땅과 세그먼트들의 관계가 주어지면 발의 자세를 만들어내는 제어기를 구현하였다.

먼저 기존 모션 데이터의 발목 관절 위치에 맞게 발 모델을 위치시킨다. 다음으로 바닥에 닿을 세그먼트들을 관장하는 관절이 땅에 닿도록 회전축 ω 를 기준으로 ϕ

* 구두 발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 현재 타 학술대회에 제출 중.

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW 컴퓨팅산업원천기술개발사업(SW스타랩)의 연구결과로 수행되었음. (IITP-2017-0-00878)

만큼 발목 관절을 회전시킨다. 이때 회전축 $\hat{\omega}$ 은 지면에 수직한 벡터 \hat{n} 과 발목에서 세그먼트의 조인트로 가는 벡터 $\vec{v}_{sa} = \vec{p}_s - \vec{p}_a$ 와 수직이다. 그리고 회전각도 ϕ 는 회전축 $\hat{\omega}$ 로 발목을 돌렸을 때 세그먼트의 조인트가 땅에 닿도록 결정한다.

마지막으로 땅과 접촉하는 세그먼트들을 땅과 평행하도록 맞추기 위해 세그먼트 관절을 돌려준다. 이때 세그먼트가 향하는 방향이 기존 모션데이터에서 발이 향하는 방향과 최대한 일치하도록 하였다.

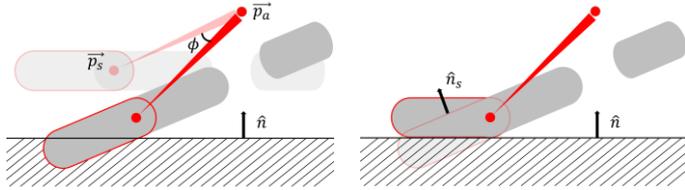


그림 2: 다중 세그먼트 발 모델 자세 제어

우리는 서 있는 동안 사람 캐릭터가 중심을 잡게 하기 위해 Macchietto et al.[4]이 고안한 균형 제어 시스템을 사용하였다. 이 시스템은 캐릭터의 무게 중심 및 압력 중심이 지지면 안에 존재하도록 선운동량 및 각운동량의 변화량을 조절하는 방법을 사용한다.

균형을 제어하기 위해서는 지지면을 이루는 발의 세그먼트가 지면과 평행하게 접촉을 유지하는 것이 중요하다. 우리는 발 자세 제어기의 결과로 나온, 지면에 접촉하는 세그먼트의 자세를 강한 제약조건으로 추종하도록 하여 균형을 잘 유지할 수 있도록 하였다.

4. 결과

다중 세그먼트 발 모델을 사용하여 까치발, 발 기울이기, 발목 돌리기, 한 발 중심 잡기와 같은 자연스럽고 정교한 발 모션을 시뮬레이션 할 수 있다(그림 3).

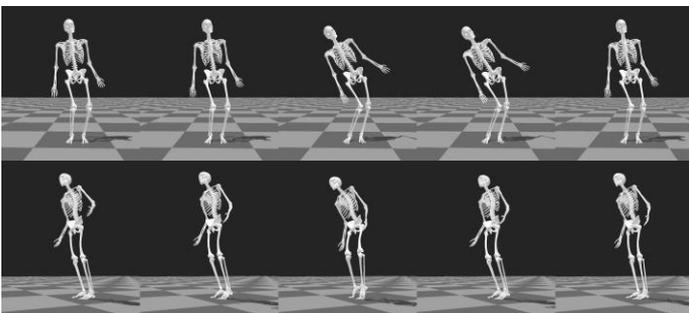


그림 3: 다중 세그먼트 발 모델을 사용한 여러가지 동작 재현

우리는 제안한 발 모델과 발 자세 제어기의 강건함을 증명하기 위해 두 가지 실험을 수행하였다. 첫 번째 실험에서는 서 있는 캐릭터 모델에 여러 방향으로 외부 힘을 적용하였다. 두 번째 실험에서는 울퉁불퉁한 지형을 만들고 그 위에 캐릭터 모델을 서 있게 하였다. 각 실험 결과를 한 개의 바다로 이루어진 발 모델 시뮬레

이션 결과와 비교하였다.

이전 발 모델은 40N의 외부 힘이 작용하였을 때 견디지 못하고 넘어졌다. 반면에 다중 세그먼트 발 모델은 40N의 힘에도 넘어지지 않고 균형을 잡았다(그림 4).

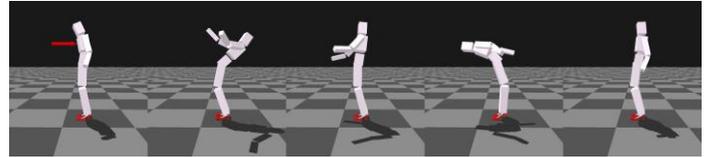


그림 4: 강건함 실험. 40N의 외부힘.



그림 5: 강건함 실험. 울퉁불퉁한 바다

울퉁불퉁한 지면에서는, 이전 발 모델은 자유도가 낮아 지면에 적응하지 못하고 균형을 잃었다. 그러나 우리의 발 모델을 사용한 시뮬레이션에서는 지면에 맞춰 발 자세를 바꿔서 중심을 잡을 수 있었다(그림 5).

5. 결론 및 미래 연구 방향

우리는 우리의 사람의 발을 본 뜬 발 모델과 발 자세 제어기를 통하여 여러가지 모션을 물리적으로 시뮬레이션하고, 실험을 통하여 강건함을 입증하였다. 또한 우리의 발 모델은 발에 대한 별도의 모션 데이터 없이도 발 모션 시뮬레이션이 가능하다. 제안된 다중 세그먼트 발 모델을 사용하여 더 강건하고 풍부하게 사람의 움직임을 시뮬레이션 할 수 있을 것이다.

우리는 사람 발 뼈의 구조를 본 따 발 모델을 제작하였지만, 뼈를 세그먼트라는 단위로 그룹화 하여 제어하기 때문에 신발을 신고 주로 활동하는 사람의 동작과 유사하게 동작할 수 있을 것이라 생각된다.

현재는 서 있는 자세에 한정하여 시뮬레이션이 가능하지만, 앞으로는 걷거나 뛰는 움직임도 재현할 예정이다. 또한 발 모델에 근육을 사용하는 연구도 흥미로운 연구주제가 될 것이다.

참고문헌

[1] Meglan D. and Berme N., A 3d passive mechanical model of the human foot for use in locomotion synthesis, *Journal of Biomechanics*, 26, 3, 331, 1993.
 [2] D. S. Lopes, R. R. Neptune, J. A. Ambrósio and M. T. Silva, A superellipsoid-plane model for simulating foot-ground contact during human gait, *Computer methods in biomechanics and biomedical engineering*, 19, 9, 954-963, 2016.
 [3] Jain S. and Liu C. K.: Controlling physics-based characters using soft contacts. *ACM Trans. Graph.* 30, 6, 163:1-163:10, 2011.
 [4] Macchietto Adriano, Victor Zordan and Christian R. Shelton, Momentum control for balance, *ACM Trans. Graph.*, 28, 3, 2009.

스킨 캐릭터의 공간 관계 유지를 통한 모션리타게팅 방법*

진태일⁰, 김미경, 이성희
 문화기술대학원, 한국과학기술원
 {jin219219, koms1701, sunghee.lee}@kaist.ac.kr

Aura Mesh: Motion Retargeting to Preserve the Spatial Relationships between Skinned Characters

Taeil Jin⁰, Meekyoung Kim, Sung-Hee Lee
 GSCT, Korea Advanced Institute of Science and Technology

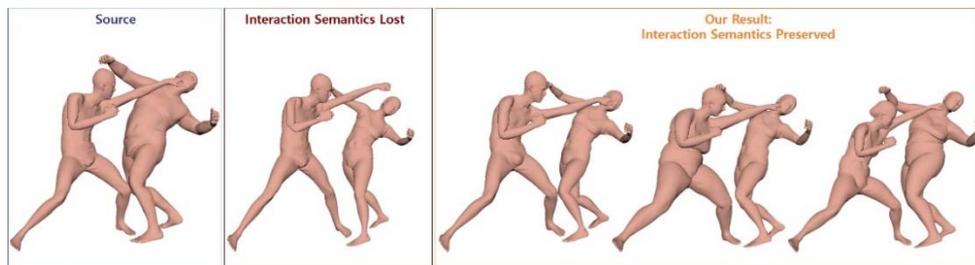


그림 1 본 연구가 제안한 방법을 통해 리타겟된 대상 스킨 캐릭터들은 원본 인터랙션의 의미를 유지한다. 왼쪽부터 원본 인터랙션 동작, 개별적인 리타게팅의 결과, 제안하는 리타게팅 방법을 세가지의 다른 체형 쌍에 적용한 결과이다.

요약

본 논문에서는 캐릭터들의 인터랙션 동작에 대하여, 대상 캐릭터들의 피부 체형까지 고려한 모션 리타게팅 방법을 제안한다. 아우라 메시(aura mesh)라고 명칭 하는 volumetric mesh를 통해, 캐릭터 스킨 주변 인터랙션 공간을 정의한다. 캐릭터간의 공간 관계는 상대방 캐릭터의 아우라 메시와 캐릭터 스킨 메시의 겹쳐짐(overlap)으로 구해지며, 모션 리타게팅은 대상 캐릭터들이 리타게팅의 제약 조건들을 만족하면서 동시에 공간 관계를 가능한 한 유지함으로써 이루어진다.

1. 서론

기존의 모션 리타게팅 방법은 단일 캐릭터의 모션을 다양한 체형과 구조 혹은 다양한 환경에 적응적으로 리타게팅 하였다. 하지만 두 캐릭터의 인터랙션 동작을 대상 캐릭터들에게 적용하는 방법은, 인터랙션의 의미를 보존해야함과 동시에, 캐릭터 스킨들의 비정상적인 침습 방지 등, 리타게팅된 캐릭터의 일반적인 요구조건들을 만족해야 하기 때문에 힘들다. 그 중, 공간 관계를 표현하여 인터랙션의 의미를 효율적으로 정의하고 이를 모션 리타게팅에 적용한 논문들이 존재하였다[1,2].

그러나 해당 방법들은 캐릭터들의 스켈레톤(skeleton)만을 고려했기 때문에, 캐릭터 스킨 체형을 고려하기 위해선 후처리 과정이 다시 요구된다. 본 연구는 스킨 레벨의 공간 관계를 아우라 메시를 이용해 정의하여, 대상 캐릭터들의 다양한 피부 체형까지 고려할 수 있는 모션 리타게팅 방법을 제안하였다.

2. 아우라 메시

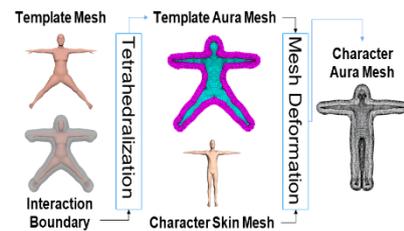


그림 2 아우라 메시 생성 과정

아우라 메시는 캐릭터 스킨 주변을 둘러싼 사면체의 메시들의 집합으로, 캐릭터의 인터랙션 공간을 정의한다. 캐릭터 인터랙션 공간의 외부 경계는 캐릭터 스킨을 기준으로 특정 거리를 떨어진 경계 표면 메시를 생성하여 정의한다. 따라서 경계 메시 밖에서는 인터랙션이 일어나지 않는다고 가정한다.

캐릭터 메시들은 하나의 템플릿 메시지를 통해서 생성되어 같은 메시 구조를 가진다. 이와 같이, 입력 캐릭터의 아우라 메시는 템플릿 아우라 메시에 메시 변형 기법을 적용하여 생성된다. 템플릿 아우라 메시의 안쪽 경계면

* 구두 발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의원본 논문은 Computer Graphics Forum 37(2). 2018에 게재 되었음.

* 본 연구는 한국연구재단 글로벌 프론티어 연구사업 (2015M3A6A3073743) 의 지원으로 수행되었음.

S 는 입력 캐릭터의 스킨 버텍스와 최대한 일치하도록 변형된다.

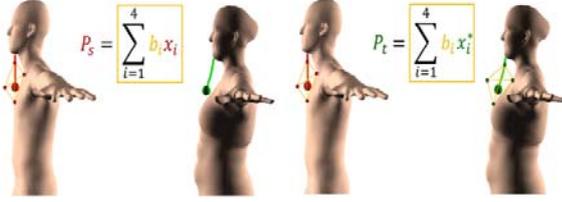


그림 3 두 캐릭터의 인터랙션 공간 속 점의 대응

결과적으로 캐릭터 아우라 메시들은 같은 메시 구조를 가지기 때문에, 두 캐릭터의 아우라 메시 속 대응점은 점이 속한 아우라 메시의 인덱스를 알아내고 동일한 barycentric coordinates를 적용해 알아낼 수 있다.

3. 캐릭터 포즈에 따른 아우라 메시 변형

캐릭터의 포즈가 변해도 아우라 메시는 캐릭터 표면을 완벽히 감싸도록 변형 된다. 이를 만족시키기위해, 표면 메시와 동일하게 linear blend skinning (LBS) 방법으로 아우라 메시지를 변형하였다.

$$x(q) = \sum_{i=1}^m w_i(x) T_i(q) T_i^{-1}(0) x(0)$$

캐릭터의 generalized coordinates 은 $q = (r_0^T, \theta^T)^T$ 이며, $r_0 \in \mathbb{R}^3$ 은 캐릭터의 루트 위치를 의미하고, θ 는 캐릭터의 각도 벡터를 의미한다. $x(q)$ 는 q 에 의해 변형된 아우라 메시 버텍스의 위치이며 T_i 는 캐릭터 각 관절의 transformation matrix 이다.

T_i 가 아우라 메시 버텍스 x 의 위치에 끼치는 영향 정도를 나타내는 $w_i(x)$ 는 bounded biharmonic weight 방법 [3]을 통해 구해졌다. 이를 통해, 아우라 메시 버텍스 중 캐릭터 표면에 해당하는 안쪽 경계면 S는 캐릭터 스킨의 $w_i^*(x), (x \in S)$ 와 같은 skinning weight 을 갖게 되며, 결과적으로 캐릭터의 포즈가 변해도 아우라 메시는 캐릭터 스킨을 완전히 감싼다.

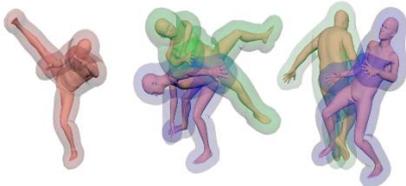


그림 4 포즈에 따라 변형되는 아우라 메시

4. 스킨 레벨의 공간 관계

캐릭터들의 포즈가 변하면, 캐릭터 스킨과 아우라 메시가 동시에 변형된다. 인터랙션 기술자는 아우라 메시와 충돌하는 캐릭터 스킨 버텍스들로 정의된다. 인터랙션 기술자들의 대상 캐릭터 공간으로의 대응은 2 절에서 소개한 대응 관계를 통해 구해진다(그림 5).

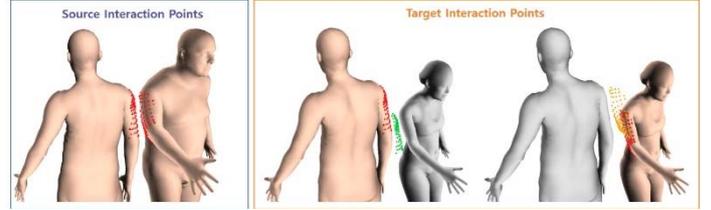


그림 5 인터랙션 기술자들로 표현된 스킨 레벨의 공간 관계

5. 대상 캐릭터 동작 생성

대상 캐릭터들의 포즈는 매 프레임 최적화 기반의 inverse kinematics 방법을 통해 구해진다. $q^* = (q^{t1}, q^{t2})$ 는 다음과 같은 에너지들 합이 최소화 되도록 구해진다.

$$q^* = \underset{q}{arg\ min} \lambda_s E_s + \lambda_d E_d + \lambda_b E_b + \lambda_g E_g + \lambda_\theta E_\theta$$

E_s 는 스킨 레벨 공간 관계 유지에 관한 에너지로, 인터랙션 기술자의 대응점에 대상 캐릭터 스킨 버텍스가 위치하도록 한다. E_d, E_b 는 원본 캐릭터들의 포즈와 루트 위치를 유지하도록, E_g, E_q 는 대상캐릭터의 발과 특정 관절 각도가 정해진 값을 따라가도록 하는 에너지들이다. 동작의 인접 프레임 간 연속성을 위해 이전 프레임의 각도 θ 와 θ^* 간의 spherical linear interpolation된 값을 대상 캐릭터에게 적용하였다.

6. 결과

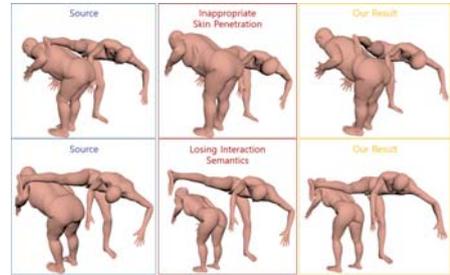


그림 6 모션 리타게팅 결과

제안하는 모션 리타게팅을 두 캐릭터간의 인터랙션 동작에 적용하였다. 캐릭터의 포즈는 동시에 구해지기 때문에 한 명의 캐릭터만 바뀌어도, 모든 캐릭터들의 관절 각도가 조정된다. 생성된 대상 캐릭터는 인터랙션의 의미를 유지하며, 동시에 스킨 간의 비정상적 침습 방지 등과 같은 일반적인 리타게팅 방법의 요구를 만족한다.

참고문헌

[1] Ho, Edmond SL, Taku Komura, and Chiew-Lan Tai. Spatial relationship preserving character motion adaptation. ACM Transactions on Graphics (TOG). (Vol. 29, No. 4, p. 33), 2010.
 [2] KIM, Yeonjoon, et al. Retargeting human-object interaction to virtual avatars. IEEE transactions on visualization and computer graphics, 22.11: 2405-2412, 2016.
 [3] JACOBSON, Alec, et al. Bounded biharmonic weights for real-time deformation. ACM Trans. Graph., 30.4: 78:1-78:8, 2011.

교수/박사연구원급 발표

Prediction of Diffusion-Weighted Appearance in Developing Infant Brain using Cycle-Consist Models

김재일⁰, Geng Chen, Weili Lin, Pew-Thian Yap, Dinggang Shen

경북대학교

본 발표의 원본 논문은 Organization of Human Brain Mapping 2018에서 발표되었음.

교수/박사연구원급 발표

High-Quality Hyperspectral Reconstruction Using a Spectral Prior

최인창, 전석준, 남길주, Diego Gutierrez, 김민혁⁰

KAIST

본 발표의 원본 논문은 ACM Transactions on Graphics (TOG), 36(6), Nov. 27-30, 2017, pp. 218:1-13 (also presented at SIGGRAPH Asia)에 게재되었음.

교수/박사연구원급 발표

Retargeting Human-Object Interaction to Virtual Avatars

김연준, 박한길, 방승배, 이성희⁰

KAIST

본 발표의 원본 논문은 IEEE Transactions on Visualization and Computer Graphics, 22(11), pp. 2405-2412, Nov. 2016 (also presented in the 15th IEEE International Symposium on Mixed and Augmented Reality (ISMAR))에 게재되었음.

교수/박사연구원급 발표

고해상도 점군 획득을 위한 카메라 프로젝터 시스템 개발

강현민, 박용문, 서용덕⁰

서강대학교

본 발표는 "회전 테이블의 기하적 변환을 이용한 점군 정합", 한국컴퓨터그래픽스학회 2018 학술대회, "프로젝터를 이용한 초고해상도 3차원 형상 복원", 한국컴퓨터그래픽스학회 2017 학술대회, "Globally optimal camera-and-rotation-sensor calibration with a branch-and-bound algorithm", Applied Optics 2017 논문들을 기반으로 하고 있으며, 현재 진행 중인 연구임.

음함수에 기반한 실시간 GPU 슬라이싱*

김현준⁰, 로정, 김민호
 서울시립대학교 컴퓨터과학과
 {kamu1324, uosgrad2018038, minhokim}@uos.ac.kr

Real-Time GPU Slicing Based on Implicit Representation

Hyunjun Kim⁰, Lottie Lu, Minho Kim
 Dept. of Computer Science, University of Seoul

요약

본 논문에서는 음함수에 기반한 3D 프린팅용 실시간 슬라이싱 알고리즘을 제안한다. 입력 모델을 SDF(Signed Distance Field) 함수로 변환한 후 지정한 두께를 갖도록 GPU(Graphics Processing Unit)에서 실시간으로 슬라이싱이 이루어진다. 또한 다양한 형태의 내부 지지 구조(internal supporting structure)가 CSG(Constructive Solid Geometry) 연산으로 추가되고 사용되는 재료의 볼륨을 계산한다. 실험결과에 따르면 본 기법은 256³ 해상도의 슬라이싱을 12fps의 속도로 계산하고, 5% 이하의 볼륨 계산 오차를 보인다.

1. 서론

메시(mesh) 기반 모델 표현법은 실시간 3D 렌더링에 최적화된 기법으로, 3D 프린팅에서도 표준 파일포맷으로 널리 사용되고 있으나 실제 부피를 가져야 하는 3D 프린팅에 알맞는 표현법이라고 볼 수는 없다. 이에 반해 음함수 표현법은 날카로운 구조를 표현하기 어렵다는 단점은 있지만 메시 표현법의 여러 단점들을 해결할 수 있어, 최근 이에 기반한 여러 기법들이 등장하고 있다 [1, 2].

본 논문에서는 이러한 음함수에 기반하여 3D 프린팅용 모델을 다루는 프레임워크를 제안하고, 그 응용으로 실시간 GPU 3D 슬라이싱 기법을 제안한다. 입력 모델이 메시 기반 형식일 경우 이를 SDF 함수로 변환하고 방향(orientation)이 바뀔 때마다 실시간으로 GPU에서 셸(shell) 형태의 슬라이싱이 이루어진다. 이 때 모델 내부의 빈 공간은 CSG 연산에 의해 지정한 내부 지지 구조로 채워지고 해당 방향으로 프린팅 할 경우 사용되는 재료의 양 또한 실시간으로 계산된다. 실험결과에 따르면 본 기법은 256³ 해상도의 슬라이싱을 12fps의 속도로 계산하고, 5% 이하의 볼륨 계산 오차를 보인다.

2. 관련 연구

Ezair 등[3]은 z-버퍼를 사용하여 메시 데이터의 지지 구조 볼륨을 GPU에서 실시간으로 계산하는 기법을 제안하였다. Minetto 등[4]은 해시 함수를 사용하여 접근적으로 최적의 시간 복잡도를 가지는 메시 기반 슬라이싱 방법을 제안하였다. Steuben 등[2]은 메시데이터의 각 슬라이스를 음함수로 변환하여 채우기(infill)패턴 등을 계산하는 방법을 제안하였다.

본 논문이 제안하는 음함수 기반의 슬라이싱 방법은 메시 기반의 방법들[3, 4]에 비해 구현이 쉽고 GPU 기반의 연산에 적합한 구조를 가진 장점이 있다. 기존의 음함수 기반의 방법[2]은 레이아웃(layout)이 변경되면 음함수를 다시 생성해야 하지만 본 논문의 방법은 한 번 생성된 음함수를 재사용하기 때문에 계산량이 적은 장점이 있다.

3. Algorithm

본 논문이 제안하는 방법은 음함수를 사용한 GPU 기반의 슬라이싱 프레임워크이다. 이 방법을 통해 사용자는 레이아웃 단계에서 프린팅 시간과 재료의 소비량 등의 정보를 실시간으로 확인할 수 있도록 하였다.

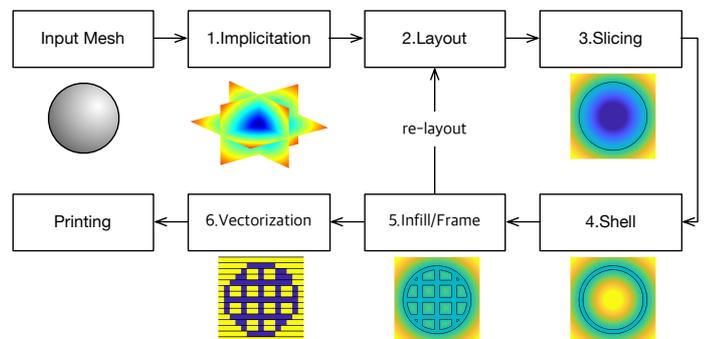


그림 1 제안하는 방법의 순서도. 메시 데이터를 음함수로 변환한 후 3D 프린팅을 위한 슬라이스를 생성한다. 2에서 5단계는 사용자가 모델을 배치하는 과정에서 반복적으로 수행하는데, GPU 연산을 통해 실시간으로 렌더링이 가능하도록 했다.

* 구두 발표논문

* 이 논문은 2016년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임. (No. 201611042004)

그림 1 은 제안하는 방법의 순서도이다. 첫 번째 단계는 메시 데이터를 SDF 함수로 샘플링하여 메시 데이터를 음함수화 하는 단계이다. 이 단계는 계산 비용이 높지만 처음 한 번의 계산만 필요하기 때문에 전체 프로세스에 큰 영향을 미치지 않는다. 두 번째 단계는 레이아웃 단계이다. 사용자가 모델을 출력하기 전에 모델의 크기와 방향성 같은 요소를 수정하는 단계로 모델에 대한 3차원 공간상에서의 변환 행렬을 생성한다. 세 번째 단계는 변환 행렬을 사용하여 SDF에서 슬라이스를 생성한다. 슬라이스의 각 픽셀에는 프린팅 재료의 종류나 셀, 내부 지지 구조 또는 외부 지지 구조 등을 구분하기 위한 유형(type) 값이 저장된다. 네 번째 단계는 셀을 생성하는 단계이다. 주어진 두께의 레벨 값 범위를 선택해서 각 슬라이스에 셀을 생성한다. 이 과정은 두 레벨셋(level set)에 대한 CSG 연산으로 정의할 수 있다. 다섯 번째 단계는 셀 채움 또는 내부 지지 구조를 생성하는 단계이다. 음함수 형태로 정의된 패턴과 SDF의 교집합(intersection) CSG 연산으로 쉽게 생성할 수 있다. 그리고 필요에 따라 외부 지지 구조(external supporting structure)를 동일한 방법으로 생성한다. 이 단계가 끝나면 모든 슬라이스의 유형 값들에 대한 히스토그램(histogram)을 계산함으로써 프린팅에 필요한 재료의 볼륨을 근사할 수 있다. 마지막 단계는 벡터화 단계로 슬라이스들을 하나로 취합하고 프린팅 방식에 따라 알맞은 포맷으로 프린팅 정보를 생성한다. 예를 들어 FDM(Fused Decomposition Modeling) 방식의 프린터를 위해서 툴 경로(tool path)를 생성할 수 있다. 이 단계는 프린팅 방식에 따라 선택적으로 수행될 수 있다.

4. 구현 및 실험

그림 2 는 본 논문의 방법을 구현한 어플리케이션의 실행화면이다. 이를 사용하여 본 논문의 방법을 검증하였다. 실험에 사용한 장비는 Intel® Core™ i5 @ 3.1 GHz CPU와 Intel® Iris™ Plus Graphics 650GPU 사양의 노트북이다.

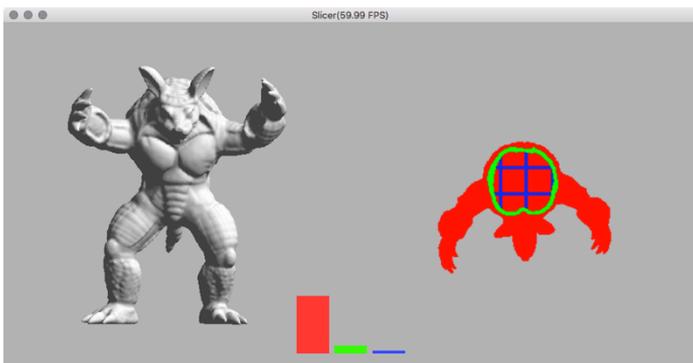


그림 2 어플리케이션으로 구현한 결과. 화면의 왼쪽은 메시 데이터를 렌더링한 것이고, 우측은 슬라이싱한 결과이다. 가운데 그래프는 모든 슬라이스에서 사용된 재료의 비율(적:외부 지지 구조, 녹:셸, 청:내부 지지 구조)을 나타낸다.

그림 3 은 볼륨 계산의 오차를 기존의 어플리케이션과

비교한 결과이다. 비교 대상은 3D Systems사의 ProJet HD 3510 3D 프린터의 모델링 툴에서 예측한 결과이고, 측정 방법은 Armadillo 모델을 회전할 때에 변화하는 외부 지지 구조의 볼륨을 비교했다. 제안하는 방법은 256^2 크기의 슬라이스를 사용했는데, 대체로 5%미만의 볼륨 계산 오차를 보였다.

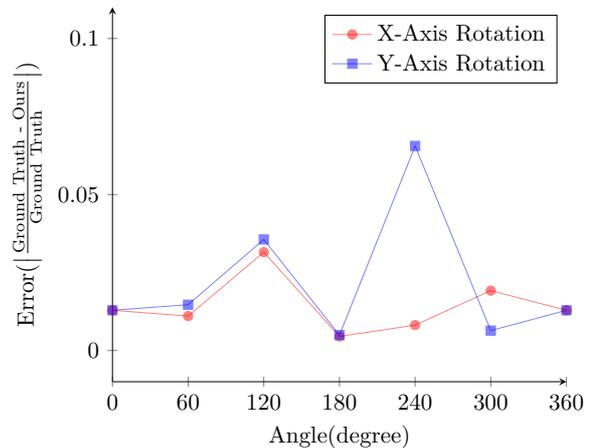


그림 3 모델의 회전에 따른 재료 사용량 예측 결과. Ground Truth는 ProJet HD 3510 3D 프린터가 산출한 소프트 재질의 질량이며, 제안하는 방법이 계산한 볼륨 값을 질량으로 환산해서 비교했다

볼륨의 계산을 포함한 렌더링 속도는 256^3 해상도의 슬라이스를 생성할 때에 12 fps 정도이다. 이를 통해 실시간으로 높은 정확도의 볼륨 계산이 가능함을 알 수 있다.

5. 결론

본 논문에서는 음함수 표현법에 기반한 실시간 슬라이싱 기법을 제안하였다. 실험결과를 통해 실시간으로 정확한 슬라이싱이 가능함을 보였으나, 추후 실제 프린팅한 결과물들을 통한 품질의 검증이 좀 더 필요하다. 후속 연구로 이에 기반한 방향성 최적화 기법을 구현할 계획이다. 또한 물리적인 성질에 기반한 다양한 내부 지지구조 등을 생성하는 기법 또한 연구할 계획이다.

참고문헌

- [1] Livesu, Marco, et al. "From 3D models to 3D prints: an overview of the processing pipeline." *Computer Graphics Forum*. Vol. 36. No. 2. 2017.
- [2] Steuben, John C., Athanasios P. Iliopoulos, and John G. Michopoulos. "Implicit slicing for functionally tailored additive manufacturing." *Computer-Aided Design* 77 (2016): 107-119.
- [3] Ezair, Ben, Fady Massarwi, and Gershon Elber. "Orientation analysis of 3D objects toward minimal support volume in 3D-printing." *Computers & Graphics* 51 (2015): 117-124.
- [4] Minetto, Rodrigo, et al. "An optimal algorithm for 3D triangle mesh slicing." *Computer-Aided Design* 92 (2017): 1-10.

3D 공간 모델을 활용한 로봇 기반 정리 자동화 *

강민철⁰, 윤성의
한국과학기술원 진산학부
mincheul.kang@kaist.ac.kr, sungeui@kaist.edu

Automated Task Planning Using Object Arrangement Optimization

Mincheul Kang⁰, Sung-eui Yoon
School of computing, KAIST

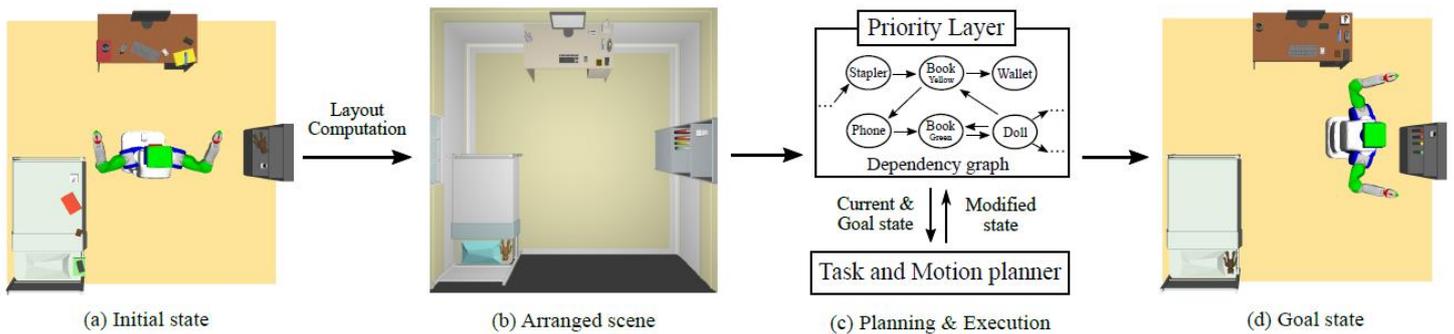


그림 1: 전체적인 시스템 흐름 ([4]의 Fig. 3에서 발췌)

요약

본 논문은 로봇이 자동으로 정돈되지 않은 물체를 정리하는 방법을 제안한다. 물체 정리 자동화를 위해서는 어지럽혀진 물체들을 정돈할 위치를 스스로 결정해야 하며, 각 물체들을 치우기 위해 취할 행동에 대해서도 계획해야 한다. 본 연구팀은 사용자의 선호를 고려하기 위하여, 사용자로부터 입력 받은 정리 예시(positive examples)에서 물체 간의 관계 정보를 추출하고, 이와 인간공학적인 특성을 고려하여 각 물체를 정돈할 위치를 최적화 시킨다. 또한, 물체를 효율적으로 정리하기 위해서 최적화 과정에서 사용된 정보를 이용하여 dependency graph를 구축하고 topological sort를 통해서 물체 정리 문제를 해결한다. 해당 방법으로 다양한 환경에서 로봇이 물체를 정리하는 것을 보이기 위하여 거실, 침실, 부엌에서 정리된 scene을 만드는 것을 보였다. 또한, 물체의 개수를 다르게 하여 기존의 작업 경로 생성 기술에 적용하여 비교한 결과, 제안한 방법은 최적의 행동을 취함에 따라 최대 2.15배 실행 시간을 줄였다.

1. 서론

로봇이 물체를 정리하도록 하기 위해 사용자는 로봇에게 구체적인 목표나 계획을 제공해야 하는데, 이는 사용자에게 매우 불편한 일이다. 게다가, 많은 수의 개체로 구성된 복잡한 경우일수록 물체를 정리하는 것은 어려운 문제이다. 이러한 문제점들을 해결하기 위해, 본 연구팀은 로봇이 단순한 명령에도 자동으로 목표(정돈된 배치)를 생성하고 효율적으로 물체를 정리하도록 하는 방법을 제안한다.

본 논문에서는 로봇이 물체를 정리하는 것을 자동화하기 위하여 목표 배치를 만드는 과정과 효율적으로 연속된 행동을 계획하는 크게 두 가지 과정을 거친다 (그림 1). 목표 배치를 만드는 과정은 사용자의 선호도를 고려하기 위하여, 사용자가 물체를 배치하는 형태를 3D 공간 모델에 옮긴 정리 예시에서 관계 정보를 추출하여 어지럽혀진 물체를 정리할 위치를 최적화 시킨다. 그리고 다수의 물체를 효율적으로 정리하기 위해 dependency graph를 구축하고 topological sort를 통해 정리할 대상 물체를 결정하고, task and motion planning을 통해 로봇의 행동과 움직임을 계획한다.

2. 물체 정리 최적화

우리는 사용자의 선호도를 고려하면서 물체들을 정리하기 위하여, 사용자의 정리 예시에서 추출한 데이터를 기반으로 어지럽혀진 물체들을 정리할 위치를 계산한다.

* 구두 발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 Ubiquitous Robots 2018에 게재 되었음 [4].

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW 컴퓨팅산업원천기술개발사업(SW스타랩)의 연구결과로 수행되었음. (IITP-2015-0-00199)

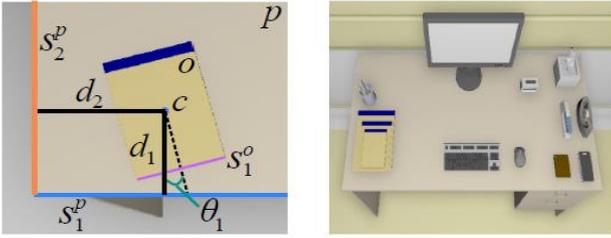


그림 2: 계층적-공간적 관계(좌)와 정리된 상태(우) ([4]의 Fig. 4에서 발췌)



그림 4: 침실(좌), 부엌(중앙), 거실(우)에서의 정리된 상태 ([4]의 Fig. 7에서 발췌)

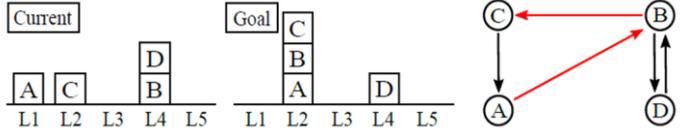


그림 3: 간단한 예제(좌)와 이에 대한 dependency graph(우) ([4]의 Fig. 6에서 발췌)

물체들은 가구 혹은 다른 물체 위에 놓이면서 계층적인 관계를 이룬다. 제안하는 방법은 이러한 물체들의 계층적인 관계와 함께 공간적인 위치 관계도 추출한다. 그림 2를 보면, 공책과 책상 사이의 가장 가까운 모서리까지 거리(d_1)와 각도(θ_1)를 추출하게 된다. 또한, 가장 가까운 모서리에 접하는 모서리 중에서 가까운 다른 모서리까지의 거리(d_2)를 추가적으로 추출한다. 제안하는 기술은 이와 같은 정보들을 이용하여 물체들이 정리되어야 할 위치를 최적화 시킨다.

본 기술은 가상으로 물체를 임의의 위치로 이동 또는 회전시키면서 최적의 정돈된 위치를 계산한다. 물체의 위치에 따라 정돈 예시에서 얻은 정보 사이의 차이를 줄여나가고, 물체가 잡을 수 있는 공간과 모니터와 같이 앞이 보이는 물체에 장애물이 놓이지 않도록 cost 형태로 표현하여 최적화 작업을 수행한다[4]. 본 연구팀은 cost를 빠르게 최소화하기 위하여 very fast simulated annealing(VFSR)[3] 을 이용한다.

3. 로봇의 행동 계획

어지럽혀져 있는 다수의 물체들을 최적화된 위치로 효율적으로 옮기기 위해서 본 연구팀은 dependency graph와 topological sort를 통해 행동의 계획한다. 물체 정리 시 고려할 행동의 제약을 dependency graph로 구축하고 topological sort를 통해서 우선 정리할 물체를 선정한다. Dependency graph는 정리할 물체로 이루어진 node와, 물체를 정리하는 경로에서 방해가 되는 공간적, 계층적 제약으로 구성된 edge로 이루어진다. Edge(장애물->대상)는 물체를 목표 지점에 옮길 때, 방해가 되는 물체가 대상이 되는 물체를 가리킨다. 예를 들어 그림 3에서 물체 A를 목표지점에 놓기 위해서는 C를 먼저 놓아야 하며, 목표지점에서 가장 밑에 있는 물체부터 놓아야한다. 복잡한 경우에는 교착 상태인 circular dependency가 발생할 수 있다. 하지만, 본 연구팀은 목표 배치에서 확인할 수 있는 선행 관계(그림 3의 빨간색 edge)를 고려하여 물체를 다른 곳으로

옮김으로써 문제를 해결한다. 빨간색 edge는 목표 상태의 계층적인 관계를 고려한 것으로 물리적으로 해결할 수 없는 것이기에 물체 C를 다른 곳으로 옮겨 circular dependency를 해결한다. 일반적인 경우에서는 로봇과 가까이 있는 장애물을 옮겨 edge를 제거한다.

4. 결론

본 논문에서는 그래픽 모델을 이용하여 로봇이 물체를 자동으로 정리하는 시스템을 소개하였다. 물체를 정리하기 위해, 두 가지의 과정을 진행을 하였다. 먼저, 3D 모델로 만든 5개의 정리 예시를 통해 물체들을 정리할 위치를 최적화하였다. 그림 4는 3 가지 scene에서 실험한 결과를 보여준다. 이렇게 생성된 scene과 같이 물체를 효율적으로 정리하기 위해, dependency graph와 topological sort를 통해 로봇이 행동해야 할 일의 순서와 움직임에 대해 계획하고 실행하였다. 로봇이 물체 정리하는 예시를 보이기 위해 PR2 로봇을 이용하여 실험하였다. 표 1은 물체의 개수를 변화하며 실험한 결과를 보여준다. 기존의 플래너인 HPN[1]과 CTMP[2]에 본 연구팀의 priority layer(PL)를 적용한 결과, 행동의 수의 감소로 수행 시간이 확연히 감소하였다.

표 1: 실험 결과(수행 시간 및 행동 수) ([4]의 TABLE I에서 발췌)

		HPN [1]	HPN[1] + PL	CTMP [2]	CTMP[2] + PL
10	시간 [s]	391	301	380	315
	행동 수	51	44	50	44
20	시간 [s]	971	693	966	652
	행동 수	116	88	119	88
30	시간 [s]	1952	907	1678	921
	행동 수	214	132	216	132

참고문헌

[1] Kaelbling, Leslie Pack, and Tomás Lozano-Pérez. Hierarchical task and motion planning in the now. Robotics and Automation (ICRA), pp. 1470-1477, IEEE, 2011.
 [2] Srivastava et al. Combined task and motion planning through an extensible planner-independent interface layer, Robotics and Automation (ICRA), pp. 254-258, IEEE, 2014.
 [3] Ingber, Lester. Very fast simulated re-annealing, Mathematical and computer modeling, pp. 967-973, 1989.
 [4] Mincheul Kang, Youngsun Kwon and Sung-Eui Yoon, Automated task planning using object arrangement optimization, Ubiquitous Robots(UR), IEEE, 2018.

광선 추적법에 기반한 소리 위치 추정 *

안인규⁰, 윤성의
한국과학기술원 전산학부
inkyu.an@kaist.ac.kr, sungeui@kaist.edu

Sound Source Localization Based on Ray Tracing

Inkyu An⁰, Sung-eui Yoon
School of computing, KAIST

요약

본 논문에서는 실내 상황에서 지형 정보를 이용한 3차원 소리의 위치 추적 알고리즘을 제시한다. 이를 위해 렌더링 기법 중 하나인 광선 추적법을 이용한 알고리즘을 고안하였다. 광선 추적법은 빛의 효과를 여러 개의 광선을 통해 근사(Approximation) 하는 방법으로 소리가 탐지된 방향에서 본 논문에서 제시하는 역-음향 광선 추적법(Inverse acoustic ray tracing)을 이용해 음원의 위치를 추정하려 하였다. 이를 통해 단일 프레임 안에 발생하는 간헐적인 음원을 탐지할 수 있게 되었다. 역-음향 광선 추적법을 이용해 직접 음향 전달 경로(Direct acoustic path)와 간접 음향 경로(Indirect acoustic path)를 생성한 후, 몬테-카를로 위치 추정 기법을 이용해 3차원 음원의 위치를 찾게 된다. 직접음과 간접음은 로봇에 장착된 정육면체 모양의 마이크 어레이에서 측정되며, 정적인 또는 움직이는 음원이 각각 연속적과 간헐적으로 소리를 발생할 때 모두 결과를 도출하였다. 결과적으로 가로, 세로가 7m이고 천장의 높이가 3m인 실내 공간에서 음원이 움직이는 비가시 선의 상황에서도 거리 오차가 평균 0.8m로 비교적 정확하게 3차원 위치를 추정할 수 있었다.

1. 서론

일상생활의 로봇의 사용이 증가함에 따라 로봇과 인간의 소리를 이용한 상호 작용의 수요는 지속적으로 증가하고 있다. 특히 Google home과 Amazon echo와 같은 스마트 스피커 제품의 인기는 음향 관련 연구에 커다란 도전을 가져왔다. 이러한 음향 관련 제품에서 큰 이슈 중 하나는 실제 환경에서 정확한 음원의 위치를 찾는 문제이다. 이 문제는 소리 위치 추적(Sound source localization)으로 소개되었으며, 많은 연구가 있어왔다. 지난 20년간 소리 위치 추적에서는 소리가 들려오는 방향을 찾기 위한 연구가 진행되었다. 특히, 두 마이크 쌍의 도착 시간 차이(Time Difference of Arrival)를 기반

으로 하는 다양한 기법들이 고안되었다. 이러한 방법은 소리가 들려온 방향을 찾는데 효과적이지만, 정확한 음원의 3D 위치를 찾기에는 어려움이 있다[1]. 최근 연구에서는 소리 측정 장치의 여러 위치와 각도에서 데이터를 측정하여 3D 위치를 추정하고 있지만, 움직이는 음원과 간헐적으로 발생하는 소리 신호의 경우, 또 음원과 로봇 사이에 장애물이 있는 경우 위치를 추정할 수 없다는 한계가 있다[2].

그렇기 때문에 본 논문에서는 실내 상황에서 3D 음원의 위치를 추정하기 위한 방법을 제안한다. 렌더링 분야에서 잘 알려진 광선 추적(Ray tracing) 기법을 통해 문제를 해결하고자 하였다. 먼저 TDOA를 기반으로 하는 기법을 이용해 소리가 들려온 방향을 추정한 후, 역-음향 광선 추적법을 통해 소리의 방향의 역방향으로 직접 & 간접 음향 전달 경로를 생성한다. 이때 3D 실내 공간은 복셀 기반의 Octree로 재구성하여 정반사를 처리하게 하였다. 여기서 생성된 음향 전달 경로들이 하나의 음원에서 전파되었다고 가정하였을 때 역으로 추적된 경로들이 수렴하는 위치가 음원의 위치라고 추정할 수 있으며, 몬테-카를로 위치 추정 기법을 통해 수렴하는 위치를 찾아 결과를 도출하였다.

2. 반사 인지를 통한 위치 추적

마이크 어레이에서 측정된 3D 공간 상의 직접음과 간접음을 이용한 실시간 소리 위치 추적 알고리즘을 소개하고자 한다. 이때 소리의 신호는 회절과 난반사가 적은 고주파 대역의 파형이며, 실내 환경이라고 가정한다. 본 논문은 크게 두 가지 중요한 요소로 나눌 수 있다. 먼저 TDOA 기반 방법에서 측정된 직접음과 간접음의 역방향으로 역-음향 광선 추적법을 수행해 음향 전달 경로를 생성한 후, 경로들의 수렴 지점을 알아냄으로써 음원의 3D 위치를 추정하게 된다. 전체적 알고리즘의 작용 흐름은 Fig.1에 요약되어 있다.

소리 신호들이 마이크 어레이를 통해 측정된 후, TDOA 기반의 방법을 통해 소리의 방향, 크기, 주파수를 알아낸다. 해당 정보를 바탕으로 반사 인지 음향 광선 추적법(Reflection-aware acoustic ray tracing)을 통해 음향 레이(Acoustic ray)를 생성한다. 정확한 음향 레이 계산에 의해 소리의 전파에 따른 에너지 감소, 지도 정보 오차의 보정 등이 고려되며, 모든 연산은 실시간으로

* 구두 발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 일본 논문은 International Conference on Robotics and Automation 2018 에 발표 되었음. [3]

* 이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대정보·컴퓨팅기술개발사업의 지원을 받아 수행된 연구임(No. NRF-2017M3C4A7066317)

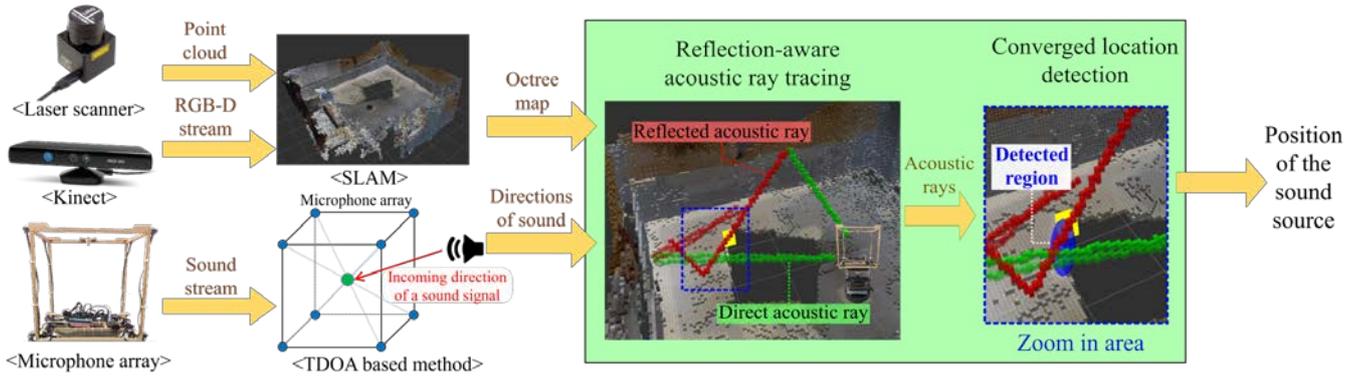


Fig. 1. 반사 인지 소리 위치 탐색 알고리즘의 개요 ([3]의 Fig.2에서 발췌)

동작한다. 또한 반사 인지 음향 광선 추적법을 수행하기 위해서는 3D 공간 정보가 필수적이며, 이를 위해 SLAM 모듈 Octree 기반의 지도를 생성한다.

반사 인지 음향 광선 추적법을 통해 생성된 음향 레이들을 통해 음원의 3D 위치를 추정하게 되며, 음향 레이들은 하나의 음원에서 발생했다고 가정하여, 역으로 추적했을 때 음향 레이들은 음원의 주위로 수렴할 것이라 예상했다. 음향 레이들의 수렴 지점은 몬테-카를로 위치 추정 기법을 적용하여 찾게 되며, 각 파티클의 가중치를 목적에 맞게 정의하였다. 가중치는 파티클과 음향 레이의 거리가 가까울수록 큰 값을 갖게 되며, 이를 위해 여러 기하학적 계산 방법을 사용하였다. 결과적으로 어떠한 파티클 주위에 음향 레이들이 많이 모였을수록 높은 가중치를 갖게 되며, 파티클들은 레이들이 많이 모여있는 지점으로 수렴하게 된다.

3. 결과

실제 실내 환경에서 음원이 움직이고, 소리가 간헐적으로 발생하게 하였을 때 소리의 위치를 추정하게 하여 결과를 측정하였다. 또한 음원과 로봇 사이에 직접음을 막는 장애물을 이동 경로 중간에 설치하여 간접음만 측정이 가능한 상황이 발생하게 제한했다. 이 실험 환경은 매우 도전적인 사례로, 어려운 상황에서도 제시된 방법이 효율적이라는 점을 보이고자 하였다.

실험은 가로와 세로가 7m이며 높이가 3m인 실내에서 진행했으며 Fig.2와 같이 음원이 시작점부터 끝점까지 제시된 이동 경로를 따라서 이동할 때, 자주색의 이동 경로에서만 박수 소리가 간헐적으로 발생한다. 결과적으로 로봇의 왼쪽의 위치에 장애물을 배치하여 직접음을 막았을 때도 파란색 타원으로 비교적 정확히 음원의 위치를 찾는 점을 확인할 수 있다.

실제 음원의 위치와 측정된 타원의 중점의 거리 오차는 Fig.3의 그래프로 볼 수 있다. 1초부터 50초까지의 거리 오차는 음원이 왼쪽 이동 경로에 있을 때이며, 230초부터 280초까지는 오른쪽 이동 경로일 때의 결과이다. 평균 거리 오차는 왼쪽의 경로는 0.7m이고 오른쪽은 0.3m으로, 장애물에 의해 발생한 회절에 의해 왼쪽의 결과가 더 안 좋은 것을 확인할 수 있다. 하지만 제시된 방법을 통해 장애물에 의해 직접음이 차단되는 상황에서도 유의미한 결과를 얻을 수 있었다. 표준편차는 왼쪽은 0.29m, 오른쪽은 0.20m로 비교적 작아 몬테-카를로

위치 추정 기법을 통해 안정적인 수렴한다고 판단할 수 있다. 즉, 음원이 움직이고 장애물의 방해가 있는 상황에서도 소리가 발생했을 경우에 효과적으로 위치를 추정했다고 판단할 수 있다. 제시된 알고리즘을 통해 한 프레임 안에 소리의 위치를 추정할 수 있으며, 움직이는 음원과 장애물이 존재와 같이 도전적인 상황에서도 음원의 위치 추적이 가능하다는 점을 확인할 수 있다.

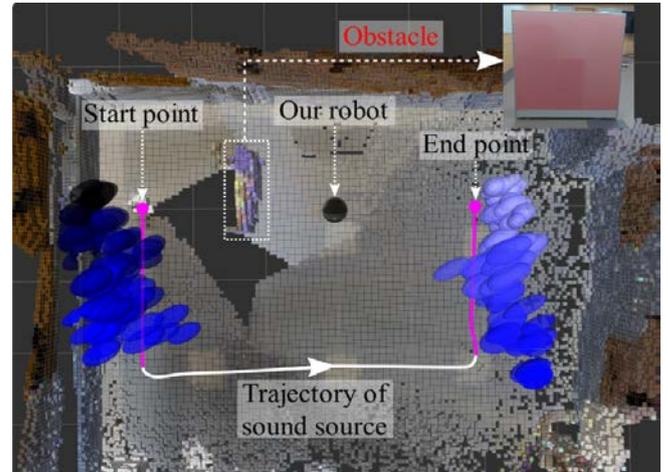


Fig. 2. 음원이 움직일 때 위치를 추정(파란 디스크) ([3]의 Fig.8a에서 발췌)

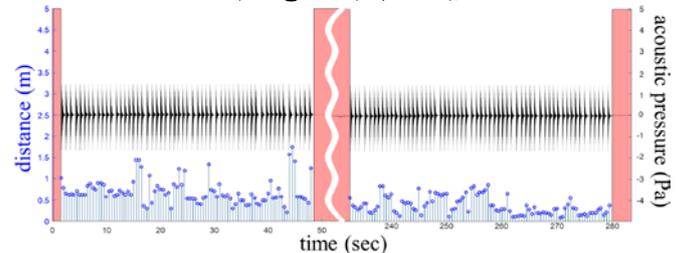


Fig. 3. 추정 위치와 실제 위치의 거리 오차 그래프 ([3]의 Fig.8b에서 발췌)

참고문헌

[1] Knapp, C. & Carter, G. The generalized correlation method for estimation of time delay IEEE Trans. Acoust., Speech, Signal Process. 24, no. 4, 320-327 (1976)
 [2] Su, D., Vidal-Calleja, T., and Miro, J. V. Towards real-time 3d sound sources mapping with linear microphone arrays ICRA (2017)
 [3] Inkyu An, Myungbae Son, Dinesh Manocha and Sung-eui Yoon. Reflection-Aware Sound Source Localization ICRA (2018)

딥 러닝을 이용한 변위맵 기반 3차원 기하 디노이징*

윤여진⁰, 최수미[†]
 세종대학교 컴퓨터공학과
 {yjyoon, smchoi}@sejong.ac.kr

Displacement Map based 3D Geometry Denoising using Deep Learning

Yeo-Jin Yoon⁰, Soo-Mi Choi[†]
 Department of Computer Science and Engineering, Sejong University

요약

다양한 스캐닝 장비 개발에 따라 3차원 모델의 기하 디테일을 보존하면서 노이즈를 제거하는 방법은 기하 처리 분야에서 매우 중요하다. 다차원의 데이터를 포함하는 기하 정보의 효과적인 디노이징을 위해 본 논문에서는 모델의 변위맵을 이용한 기하 학습 방법을 제안한다. 3차원 얼굴 모델을 대상으로 라플라시안 필터 기반의 특징을 추출하고, 패치 단위로 나눈 디테일 레이어 학습을 통해 세 가지 서로 다른 레벨의 노이즈를 가진 얼굴 모델의 노이즈를 효과적으로 제거하고, 그 결과를 검증한다.

1. 서론

최근 Kinect와 같은 3차원 스캐닝 장비의 증가로 사람의 얼굴과 같은 모델 획득이 쉽게 가능해지면서, 센서에서 획득된 노이즈를 제거하는 방법이 중요해지고 있다. 기하 디노이징을 위한 기존의 접근 방법들은 신호 처리 기법을 이용한 데이터 압축 및 복원[1,2], 스캔된 일부 3차원 데이터 정렬과 같은 방법들로 기하 디테일을 보존하면서 노이즈를 제거하기 위한 방법들이 제시되었는데, 최근에는 이미지 디노이징 및 분류를 위한 방법으로 주로 사용되던 딥 러닝을 3차원 기하 모델에 적용해 기존 디노이징 방법들보다 나은 성능을 보여주고 있다. 관련된 연구에서는 2차원 이미지보다 자유도가 높은 3차원 기하의 특징 추출을 위해 열 분포 표식(Heat Kernel Signature), 지오데식 거리 평균(average geodesic distance)과 같은 저레벨 기하 기술자를 이용한다[3]. 이러한 방법들은 트레이닝에 사용될 입력 데이터 생성 과정에서 많은 연산 복잡도와 시간을 소요하므로, 모바일 플랫폼이나 실시간 처리와 같은 빠른 데이터 처리에 적합하지는 않다.

본 논문에서는 3차원 얼굴 모델을 대상으로 변위맵을 추출하고, 이를 바로 학습에 적용하는 디노이징 방법을 제안한다. 먼저 3차원 얼굴모델의 고주파 디테일 획득

을 위해 라플라시안 필터 기반의 변위 맵을 생성하여 노이즈가 포함된 변위 정보를 획득한다. 다음으로 획득한 노이즈 별 변위 맵을 패치 단위로 분할하여 학습을 한 후, 얻어진 디노이징된 변위 정보를 다시 3차원 모델에 재적용 하여 디테일은 보존하면서도 디노이징된 모델을 복원한다.

2. 시멘틱 디노이징 프레임워크

3차원 기하 데이터의 경우, 모델의 부위에 따른 복잡도가 다르고, 자유도가 높기 때문에 로컬 기하에 대한 시멘틱한 정보 분석이 필요하다. 기하 모델의 디테일 정보를 보존하면서 노이즈의 레벨에 관계없이 복원 가능한 데이터 학습을 위해 그림 1과 같이 변위 맵 기반의 학습 네트워크를 구축하였다. 먼저 3차원 모델을 텍스처 공간에 파라미터화 한 후, 라플라시안 필터를 이용한 특징 추출을 수행한다. 다음으로 딥 러닝 네트워크에서 로컬 데이터 학습이 가능하도록 획득된 변위맵을 랜덤하게 일정 사이즈의 패치로 분할하여 학습을 진행한다. 런타임 디노이징 단계에서는 노이즈 레벨에 관계없이 얼굴 모델에 포함된 노이즈를 제거한다.

Runtime semantic denoising

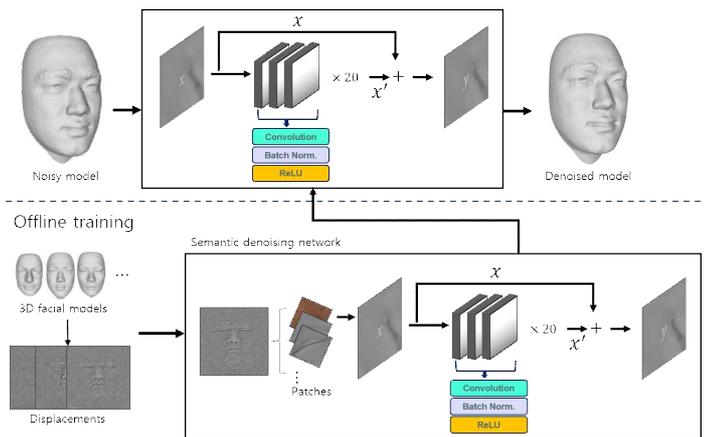


그림 1 시멘틱 디노이징 네트워크 구성

2.1. 디노이징 학습 네트워크

적은 연산량을 사용하면서도 효과적인 기하 디노이징을 위해 변위 데이터를 획득해 학습을 진행한다. 먼저 모델 M 으로부터 고주파 디테일에 해당하는 변위 d 획득을 위해 원본 모델에서 라플라시안 스무딩을 적용한 베이

* 구두 발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 SMA2018 에 발표되었음.

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2018-2016-0-00312).

† 교신저자

스 모델로부터 원본 모델까지의 변위를 계산한다. 라플라시안 스무딩은 원본 모델의 정점 x_i 와 해당 정점을 포함한 i 번째 노드에서 인접한 N 개의 정점들을 이용해 수식 (1)과 같이 추출한다.

$$M_i(x_{i+1}) = \frac{1}{N} \sum_{i=1}^N x_i \quad (1)$$

변위맵은 추출된 베이스 모델 M_i 상의 모든 정점으로부터 원본 모델 혹은 노이즈가 포함된 모델까지 노멀 방향으로 양방향 광선을 이용해 타겟 모델의 폴리곤에 부딪히는 지점까지의 거리를 저장한다. 이때, 저장하는 변위를 측정하는 광선의 갯수는 지정한 변위맵의 해상도와 동일하게 샘플링하고, 모델의 파라미터화 과정에서 할당된 텍스처 좌표에 맞게 이미지 공간에 저장한다.

$$d = M_i - M \quad (2)$$

노이즈가 포함된 모델은 원본 모델에 가우시안 함수를 적용해 각 정점의 노멀 방향으로 랜덤하게 정점의 위치를 변경한다.

학습 네트워크의 구성은 64개의 깊이를 갖는 3x3 컨볼루션 필터와 ReLU를 포함하는 20개의 시퀀셜 블록으로 구성하였다. 기존의 딥러닝 네트워크와 달리, 잔여학습(residual learning)을 이용하는 방식으로, 노이즈가 미포함된 원본 변위맵을 교사 학습의 ground truth로 설정한다. 실험에서는 균일한 사이즈로 분할된 패치를 랜덤하게 선택하고, 0.5의 확률로 회전각을 좌, 우 반전시켜 학습 단계에서 오버피팅을 최소화 하였다. 변위 맵으로부터 로컬 패치가 획득되면 전체 데이터 셋을 기준으로 데이터 정규화를 수행한다. 학습에서는 L_2 손실을 이용해 최적화를 진행하였다.

테스트 단계에서는 학습을 통해 디노이즈된 변위 정보를 3차원 기하 모델의 복원에 사용한다. 각각의 정점은 수식 (3)과 같이 획득된 변위에 따라 정점의 노멀 방향 n_i 으로 앞서 설정한 베이스 모델 M_i 에 적용한다.

$$v_i^{updated} = v_i + n_i \cdot (d \cdot \delta - \alpha) \quad (3)$$

복원된 정점 $v_i^{updated}$ 는 노이즈가 포함된 정점 v_i 로부터 변위 범위 δ , 변위의 최대 음의 값 α 에 의해 갱신된다.

3. 실험 및 결과

본 연구에서는 세 단계로 다르게 적용된 노이즈를 포함하는 얼굴 모델을 이용해 디노이즈 결과를 비교한다. 실험에서는 총 11개의 3차원 얼굴 모델을 사용했으며 각각의 모델은 약 100K개의 정점으로 구성된 모델로, 1024 x 1024의 해상도를 갖는 변위맵 해상도에 맞게 파라미터화 하였다.

노이즈를 포함한 모델 생성은 표준 편차 1인 가우시안 랜덤 함수를 원본 모델에 적용하고, 레벨별로 노이즈를 달리 하기 위해 중간값(mean) 0.05, 0.1, 0.15로 설정하여 세 가지 다른 노이즈 레벨을 포함하는 모델을 추가로 생성해 총 44개 모델에 대한 변위 맵을 획득하였다.

학습환경은 텐서플로우 기반의 Adam 최적화를 이용해

총 50,000번의 트레이닝 스텝 중 매 20,000번째마다 학습률 $5e^{-3}$ 에서 $1e^{-5}$ 로 스텝을 감쇄하는 스케줄로 구성하였다. 총 학습 시간은 4.2GB Intel Core i7-7700K와 16GB 메모리 및 GeForce GTX 1070 그래픽 카드 사양에서 5시간 정도 수행하였으며, 각 노이즈 레벨 별 학습 결과는 그림 2와 같다.

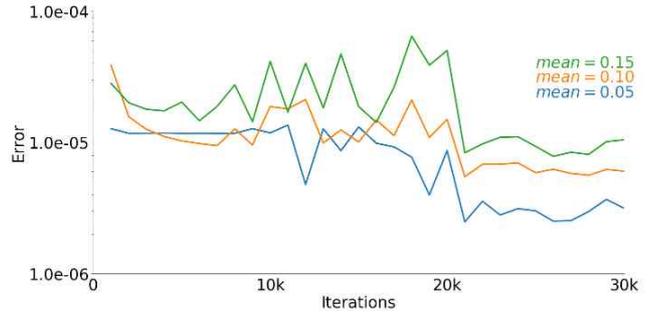


그림 2 노이즈 레벨 별 학습 결과

그림 3의 상단 얼굴 모델 입력 데이터와 복원 결과(하단)에서는 입력 모델에서 확연히 노이즈 레벨이 다르게 나타나지만 디노이즈 결과에서는 시각적으로 노이즈 레벨에 관계없이 얼굴 모델의 디테일은 보존하면서 디노이즈가 유사하게 진행된 것을 알 수 있다.

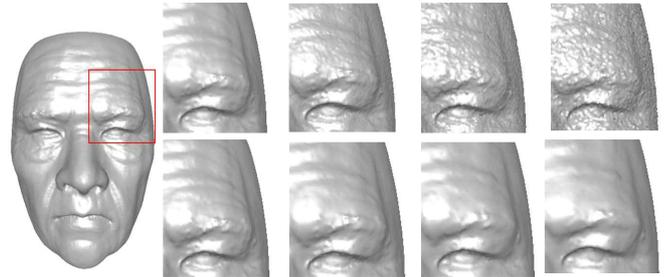


그림 3 얼굴 모델의 디노이즈 입력 데이터(상단)와 디노이즈 결과(하단)

본 논문에서는 노이즈가 포함된 얼굴 모델의 디노이즈를 위해 변위 맵을 추출하고, 이를 딥러닝 기반 학습 모델에 적용해 얼굴 모델의 주름과 같은 고주파 디테일은 보존하면서도 노이즈는 제거하는 방법을 제안하였다. 제안한 방법은 연산량이 적고 빠른 특징 추출이 가능하기 때문에 얼굴 모델을 사용하는 가상현실, 영화, 게임 등 다양한 응용 분야에 활용이 가능하다.

참고문헌

- [1] J. Digne, R. Chaine and S. Valette, Self-similarity for accurate compression on point sampled surfaces, *Computer Graphics Forum*, 33(2):155-164, 2014.
- [2] M. Berger, A. Tagliasacchi, L. M. Seversky, et al., A Survey of Surface Reconstruction from Point Clouds, *Computer Graphics Forum*, 36(1):301-329, 2017.
- [3] K. Guo, D. Zou and X. Chen, 3D Mesh labeling via deep convolutional neural networks, *ACM Trans. Graph.*, 35(12):3:1-3:12, 2015.

가상환경의 비인간 아바타를 이용한 보행에서의 실재감*

이동용⁰, 이인권¹
연세대학교 컴퓨터과학과
{pfldy2850, iklee}@yonsei.ac.kr

Presence in Walking using Non-human Avatar in Virtual Environment

Dong Yong Lee⁰, In-Kwon Lee¹
Dept. of Computer Science, Yonsei University

요약

이 논문에서는 몰입형 가상현실 시스템에서 아바타의 종류와 시점에 따라 사용자가 느끼는 실재감의 차이에 대해서 조사했다. 실재감에 대한 조사를 위하여 인간형 아바타 1종과 비인간형 아바타 2종을 각각 아바타 시점과 사용자 시점으로 실험을 진행하였다. 실험 결과 비인간형 아바타에서는 사용자 시점보다 아바타 시점이 높은 실재감을 제공했고, 아바타 시점을 사용한 비인간형 아바타는 인간형 아바타와 실재감의 차이가 없었다.

1. 서론

몰입형 가상현실 기기의 발전과 가상현실 구현기술이 발전하면서, 가상현실 콘텐츠 시장에는 많은 양의 콘텐츠가 생성되고 있다. 대부분의 콘텐츠가 인간형 캐릭터와 아바타 시점에서 제작되고 있으며, 비인간형 아바타를 사용한 콘텐츠의 제작사례와 관련 연구가 미비하다. 이에 따라 비인간형 아바타를 사용한 콘텐츠에 대한 연구가 필요하며, 본 논문에서는 비인간형 아바타와 인간형 아바타를 아바타 시점과 사용자 시점으로 나누어 실험을 진행하고, 각각의 아바타와 시점이 제공하는 실재감(presence)에 대해서 조사한다.

2. 관련 연구

현실적인 인간형 아바타와 실재감의 상관관계에 대한 분석은 여러 연구에서 진행되었다. 그 중, the illusion of virtual body ownership (IVBO)[1][2]은 사용자가 자신의 몸을 가지고 있다고 느끼는 감각이며, 이는 아바타가 자신의 실제 몸과 비슷하게 보이지는 않더라도 사용자가 아바타를 자신의 몸이라고 느끼는 환상을 말한다. IVBO는 수용도, 조작도, 변화도를 통해 실재감을 나타낸다[2]. 수용도는 해당 아바타가 사용자의 실제 몸

과 같이 느껴졌는지에 대한 지표이다. 사용자가 아바타를 본인의 몸이라고 느낄수록 사용자가 느끼는 수용도가 높아진다. 조작도는 아바타의 움직임이 사용자의 움직임으로 보여지는가에 대한 지표이다. 예를 들어, 사용자가 보행 시 아바타 또한 보행의 행동을 보였을 때, 사용자는 더 높은 조작도를 느끼게 된다. 변화도는 실험도중에 아바타의 행동이나 생김새로 인해, 사용자가 본인의 실제 몸의 모양이나 텍스처의 변화가 있다고 느끼는 정도를 말한다. 예를 들어 사용자가 거구의 아바타를 사용함으로써 자신의 몸이 실제보다 더 무겁고 크다고 느끼게 될 때, 높은 변화도를 가진다.

3. 실험 설계

3.1. 아바타 선택

실험을 위해 세가지 보행군에서 각각 하나의 아바타를 선택했다. 이족보행 동물 중에서 인간형 아바타를 선택했으며, 사족보행 동물 중에서 늑대형 아바타, 그리고 사행운동 동물 중에서 뱀형 아바타를 선택했다 (그림 1).



그림 1: 선택한 세가지 아바타. 인간형 아바타(좌측), 늑대형 아바타(중앙), 뱀형 아바타(우측)

3.2. 시점

사용자가 비인간형 아바타를 사용함에 있어서, 아바타와 사용자의 외형 차이에 의한 시점 차이가 존재한다. 이에 대해서 본 논문은 시점을 사용자 시점과 아바타 시점으로 나누었다 (그림 2). 사용자 시점이란 가상 환경에서의 시점을 실제 사용자의 시점에 맞게 설정하는 것이다. 이는 현실 환경과 가상 환경에서 사용자가 보는 관점을 일치시킨다. 반면에 아바타 시점이란 가상 환경에서의 시점을 아바타의 눈높이에 맞게 설정하는 것이다. 아바

* 구두 발표논문

* 이 논문은 삼성전자 미래기술육성센터의 지원을 받아 수행된 연구임(과제번호 SRFC-IT1601-04).

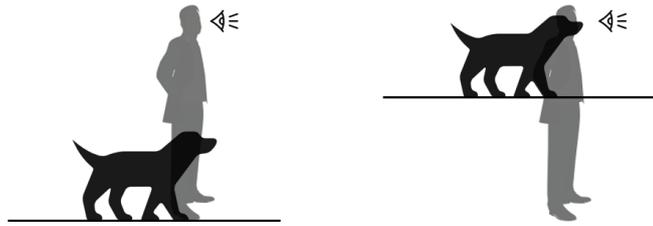


그림 2: 사용자와 비인간형 아바타의 외형 차이로 인한 시점 구분. 사용자 시점(좌측), 아바타 시점(우측)

다 시점은 현실 환경과 가상 환경에서 사용자가 보는 관점의 차이를 발생시킨다.

3.3. 실험 환경

본 논문에서는 실험을 위하여 거울이 있는 가상 공간을 설계했다 (그림 3). 사용자는 주어진 아바타와 시점에 대하여 자유롭게 공간을 돌아다니며 거울을 통해 자신의 모습을 보고 목적지에 도착하도록 지시된다. 인간형 아바타의 경우 사용자 시점과 아바타 시점이 구분되지 않으므로, 아바타 시점으로만 진행했다. 가상환경은 HTC Vive HMD로 구현하였고, 4 × 7m²의 실제 공간에서 걸어서 이동하는 사용자를 트래킹하였다.

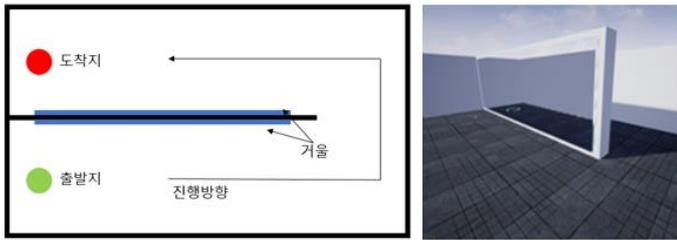


그림 3: 설계된 가상공간의 레이아웃(좌측) 및 구현된 가상공간(우측).

3.4. 측정

실험은 IVBO[2]와 Simulator Sickness Questionnaire (SSQ)에 대한 기준을 바탕으로 7단계 리커트 척도를 사용하여 설문하였다. 실험에 참여한 사람은 총 10명이었고 참여자 중 대부분이 VR 경험이 거의 없었다. 각 참여자는 사전 설문 및 5번의 실험과 설문을 진행하였으며, 전체 실험에서 총 60번의 설문 결과를 수집하였다.

4. 실험 결과

실험 결과를 비교하기 위해 대응 표본 t-검정을 실시하여 분석하였다. 수용도와 조작도, 그리고 변화도를 사용자 시점과 아바타 시점에 대해 비교하였을 때, 각각 통계적으로 유의미하게 다르다($p < 0.05$)는 결과를 얻었다 (표 1). 아바타 시점의 수용도(4.2)는 사용자 시점의 수용도(2.5)보다 높은 점수를 기록했으며, 이는 사용자가 아바타 시점에서 좀 더 아바타와 자신의 몸을 동일시하였다는 것을 알 수 있다. 또한 조작도의 측면에서도 아바타 시점(4.66)이 사용자 시점(3.33)보다 높기 때문에 아바타의 움직임이 사용자의 움직임과 더욱 같다고 느꼈다. 변화도에서도 아바타 시점(3.1)이 사용자 시점

(2.5)보다 높다. 아바타 시점일때 각 아바타의 수용도, 조작도, 변화도에 대해 대응 표본 t-검정을 실시한 결과 세 아바타에서의 통계적으로 유의미한 차이가 없었다 (표 2). SSQ를 이용한 멀미감에 대한 설문에서도 또한 유의미한 결과를 얻을 수 없었다.

수용도	조작도	변화도
p-value = 0.00719	p-value = 0.00397	p-value = 0.01190

표 1: 아바타 시점과 사용자 시점 간의 수용도와 조작도, 변화도의 대응 표본 t-검정 결과

p-value	뱀			인간			
	수용도	조작도	변화도	수용도	조작도	변화도	
누대	수용도	0.1772	-	-	0.3731	-	-
	조작도	-	0.7577	-	-	1.0	-
	변화도	-	-	0.6783	-	-	0.7509
뱀	수용도				0.8722	-	-
	조작도				-	0.8472	-
	변화도				-	-	0.8402

표 2: 각 아바타의 아바타 시점에 대한 IVBO에서의 수용도, 조작도, 변화도에 대한 대응 표본 t-검정 결과

5. 결론

비인간형 아바타와 인간형 아바타에서 사용자가 느끼는 실재감은 통계적으로 유의미한 차이가 없었다. 이는 사용자가 생각하는 자신의 몸의 이미지가 추상적이기 때문에 비인간형 아바타의 경우에도 자신의 몸이라고 느낄 수 있기 때문이라고 생각된다. 이러한 내용을 통해 the illusion of virtual body ownership을 비인간형 아바타까지 확장할 수 있음을 알 수 있다. 그러나 비인간형 아바타에서의 실재감은 시점에 따라 유의미한 차이가 존재했다. 비인간형 아바타의 경우 아바타 시점을 선택했을 때가 그렇지 않을 때보다 더 높은 수준의 실재감을 제공한다. 이에 대해 사용자 시점은 현실 환경과 가상 환경에서 사용자가 보는 관점을 일치시키지만, 아바타의 눈과 시점이 일치되지 않기 때문에, 사용자가 해당 아바타로 실재한다는 느낌을 가지기 힘든 것으로 생각된다. 이러한 결과는 VR 콘텐츠 제작에 있어서 앞으로 비인간형 아바타의 아바타 시점을 선택함으로써 기존의 콘텐츠들과 같은 수준의 실재감을 제공하는 더욱 다양한 콘텐츠 제작이 가능하다는 것을 시사한다.

참고문헌

[1] Wijnand A IJsselstein, Yvonne A W de Kort, and Antal Haans. Is this my hand I see before me? The rubber hand illusion in reality, virtual reality, and mixed reality. *Presence: Teleoperators and Virtual Environments* 15, 455–464, 2006

[2] Daniel Roth, Jean-Luc Lugrin, Marc Erich Latoschik, and Stephan Huber. Alpha IVBO - Construction of a Scale to Measure the Illusion of Virtual Body Ownership. *In Proceedings of the 35th Annual ACM Conference on Human Factors in Computing Systems*, 2017

깊이 기반 실시간 반투명 유리 렌더링

이제선^{1,0}, 최재원¹, 박소영², 최윤지¹, 김재명², 이성길²
¹성균관대학교 전자전기컴퓨터공학과, ²성균관대학교 소프트웨어학과
 {jeseon, jaewonchoi, young505, y.j.choi, jaemyung, sungkil}@skku.edu

Depth-Based Real-Time Translucent-Glass Rendering

Jeseon Lee^{1,0}, Jaewon Choi¹, Soyoung Park², Yoonji Choi¹, Jaemyung Kim², Sungkil Lee²
¹Department of Electrical and Computer Engineering, Sungkyunkwan University
²Department of Software, Sungkyunkwan University

요약

본 연구는 반사와 투과 현상이 발생하는 반투명 유리 렌더링에 대한 효과적 표현 기법을 제안한다. 반투명 평면 물체에 투과되어 보이는 물체는 그 거리에 따라 흐린 정도가 다르게 나타난다. 본 논문은 반투명 평면 유리와 물체간의 거리에 따라 물체의 흐림 정도가 결정되는 것에 착안하여, 이를 물체간의 깊이차를 기반으로 블러의 양을 결정하여 빛의 산란을 근사하는 실시간 반투명 유리 렌더링 기법을 제안한다.

1. 서론

컴퓨터 그래픽스에서 물체의 표현은 물체가 가지는 매질에 의하여 정의된다. 반투명한 물체는 반사와 투과의 성질이 복합적으로 작용하므로 렌더링 시 매질에 의한 빛의 반사, 투과 등을 모델링한 양방향 산란 분포 함수(bidirectional scattering distribution function; BSDF)가 주로 사용된다[1]. 이러한 함수 기반의 광선추적 기법은 사실감이 높은 반면, 계산의 연산량이 커 속도 향상을 위한 기법이 필요하다. 이를 위해 굴절 연산에 필요한 표면 법선 정보를 NDF(normal distribution function)[2]에 저장하여 표면의 거칠기에 따른 굴절을 하는 기법이 제안 되었지만, 거리에 따른 흐림 효과는 고려되지 않았다.

본 논문에서는 빛의 산란 정도를 일정한 상수로 근사하여 연산량을 줄이고, 빛의 산란을 깊이 정보를 기반으로 계층적 블러를 적용하여 근사함으로써 실시간 성능을 달성하면서도 반투명 유리 렌더링 기법을 제안한다.

2. 관련연구

Walter[1]은 실제 굴절 연산과 비교 했을 때 거의 유사하다는 것을 보이기 위해 미세면 이론(microfacet theory)을 도입하여 처음으로 미세면 분산 기법을 제안하였다. Guo[3]는 Walter의 모델[1]을 실시간 렌더링에 최적화 하기 위하여 확장 정규 분포(extended normal distribution)를 도입하여 실시간 렌더링을 하였다.

우리의 연구와 가장 유사한 것은 Charles[2]이다. Charles[2]는 NDF를 기반으로 광선추적을 수행하여 물체의 대한 산란을 연산하는 물리 기반 실시간 렌더링 기법을 할 수 있음을 보이기도 했다.

3. 알고리즘

본 연구에서 투과된 물체의 모양은 블러(blur)의 커널 크기를 물체간의 거리와 반투명 유리의 거칠음으로 근사하여 계층적 흐림 효과를 표현한다. 또한 물체의 경계면이 뚜렷하게 보이는 현상을 커널 크기에 대해 블러 처리를 하여 부드러운 경계면을 생성한다.

3.1 깊이 기반 계층적 흐림 효과

그림 1은 본 논문이 제안하는 깊이 기반 계층적 흐림 효과의 개략도이다. 계층적 흐림 효과는 빛이 공기에서 매질로 입사 시 발생하는 산란과 매질에서 공기로 입사할 때 발생하는 산란을 깊이 정보를 기반으로 정의한다.

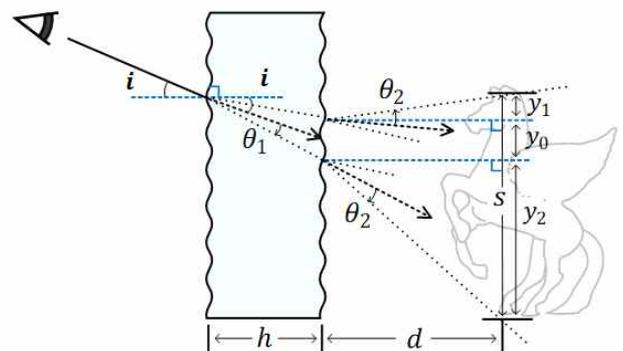


그림 1. 깊이 기반 계층적 흐림 효과 개략도

유리 표면이 거침이 커질수록 투과된 물체 모양의 흐림 정도는 높아진다. 이를 근거리 거침의 정도를 유리를 투과한 빛이 산란되는 각도라고 근사하고 이를 θ_1 과 θ_2 라고 한다. i 는 빛의 입사각을, h 는 유리의 두께를, d 는 유리와 투과되어 보이는 물체 사이 거리를 나타낸다. y_0 은 시점으로부터 나온 빛이 반투명한 물체의 매질과 처음 접하는 면에서 산란 되는 범위이다.

* 구두 발표논문

* 본 연구는 미래창조과학부의 재원으로 한국연구재단의 <실감표현인체감응솔루션> 글로벌프론티어사업, 과학기술인문융합연구사업, 정보통신기술진흥센터 대학 ICT 연구센터육성지원사업으로 수행되었음. (2018M3A6A3058649, 2017M3C1B6070980, ITP-2018-2016-0-00312)



그림 2. 각 모델에 따른 패스 트레이싱(좌)과 본 연구 제안 기법(우)

$$y_0 = h \tan(i + \theta_1) - h \tan(i - \theta_1) \quad (1)$$

y_1 과 y_2 는 초기 산란 범위의 양 끝단에서 최종 산란 범위에 직교하는 선을 그려 측정한 범위를 의미한다. y_0 , y_1 , y_2 은 모두 i 와 θ_1 , θ_2 를 사용하여 계산 한다.

$$y_1 = d |\tan(i - \theta_1 - \theta_2)| \quad (2)$$

$$y_2 = d |\tan(i + \theta_1 + \theta_2)| \quad (3)$$

산란 범위 y_0 , y_1 , y_2 의 합을 카메라와 물체의 거리 z 로 나누고, 스크린 공간의 크기 t 를 곱하여 최종 커널 사이즈를 연산할 수 있다. 유리 표면의 거침이 최대로 늘어나는 블러 커널 사이즈의 최대치를 t 라고 정의 한다. 식 4은 최종 커널 사이즈 식이다.

$$k = t \frac{(y_0 + y_1 + y_2)}{z} \quad (4)$$

3.2 부드러운 경계 적용

블러의 커널 사이즈는 반투명 평면 유리와 투과되어 보이는 깊이 값을 기반으로 결정된다. 따라서 물체의 경계 선에서는 깊이 차이가 커, 커널 사이즈의 차가 크게 발생한다. 우리는 커널 사이즈에 블러를 적용하여 부드러운 경계면을 생성한다[4]. 그림 3는 블러의 커널 사이즈에 흐림 효과를 적용 후 시각화를 한 것이다.



그림 3. 커널 사이즈에 블러 적용 전(좌)과 후(우)

4. 결과 및 토론

본 연구는 Intel Core i7-3700, NVIDIA GeForce GTX 1080 기준으로 진행 되었다. 그림 2 에는 각 모델별로 오프라인 렌더러에서 패스 트레이싱을 수행한 결과 그리고 본 연구에서 제안한 거리 차이에 따른 반투명 유리 결과를 나타내었다. 본 기법 적용 시 거리의

차이에 따른 반투명 평면 유리를 효과적으로 표현할 수 있다.

표 1. 렌더링 성능 측정 결과

모델	패스 트레이싱 (샘플 수: 510)	본 연구 기법		
		블러 연산	렌더링	총합
말 (160,252 tri)	73800 ms	0.157 ms	1.491 ms	1.648 ms
나무 (1,430,206 tri)	253200 ms	0.264 ms	1.795 ms	2.059 ms

표 1에서 보이듯 본 연구가 기존 패스 트레이싱 기반의 반투명 유리 렌더링과 비교했을 때 상당한 성능 향상과 함께 유사한 이미지 품질을 확보할 수 있는 것을 확인 하였다.

그러나 본 연구에서 제안한 방법은 반투명 평면 유리가 여러 개 겹쳐져 있는 경우는 표현하지 못한다는 한계를 지닌다. 추후 연구에서는 이러한 결함을 해결하기 위한 방법을 연구할 예정이다.

참고문헌

[1] Bruce Walter and Stephen R. Marschner, Microfacet Models for Refraction through Rough Surfaces, *In Proceedings of the 18th Eurographics conference on Rendering Techniques*, 195-206, 2007

[2] Charles de Rousiers and Adrien Bousseau, Real-Time Rendering of Rough Refraction, *IEEE Trans. Vis. and Computer Graphics*, 18(10), 1591-1602, 2012

[3] Jie Guo and Jinghui Qian, Rendering Thin Transparent Layers with Extended Normal Distribution Functions, *IEEE Trans. Vis. and Computer Graphics*, 23(9), 2108-2119, 2017

[4] Sungkil Lee and Gerard J. Kim, Real-Time Depth-of-Field Rendering Using Anisotropically Filtered Mipmap Interpolation, *IEEE Trans. Vis. and Computer Graphics*, 15(3), 453-464, 2009.

스플라인 인터페이스를 이용한 스킨링 웨이트 계산*

방승배⁰, 이성희
한국과학기술원
{be2848, sunghee.lee}@kaist.ac.kr

Computation of Skinning Weight using Spline Interface

Seungbae Bang⁰, Sung-Hee LEE
KAIST

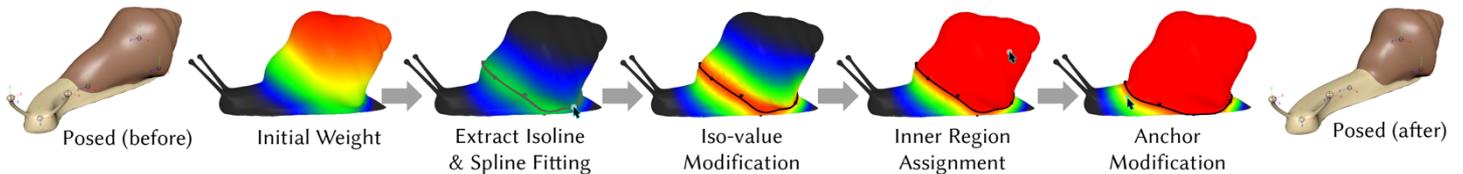


그림1: 왼쪽 두개: 자동 알고리즘으로 계산된 초기 변형 및 스킨링 웨이트. 스플라인 인터페이스는 메쉬 표면에 정의된 소수의 스플라인을 사용하여 스킨링 웨이트를 편리하고 직관적으로 편집할 수 있게 한다.

요약

최근 스킨링 웨이트를 계산하기 위한 자동방법의 발전에도 불구하고 고품질의 캐릭터 변형을 생성하려면 수동 개입이 여전히 필수적이다. 널리 사용되는 페인트 기반 인터페이스는 사용자에게 높은 수준의 자유를 제공하지만 상당한 노력과 시간을 요한다. 본 연구에서 스플라인 기반의 스킨링 웨이트 편집을 위한 새로운 인터페이스를 소개한다.

1. 서론

기하학적 캐릭터 변형 방법론에서 그 품질은 스킨링 웨이트(skinning weight)의 적절한 지정에 크게 의존한다. Maya 혹은 3DS Max 같은 상용 소프트웨어에서 페인트 기반의 인터페이스를 제공하지만, 이는 사용자의 높은 스킬과 노동시간을 요한다. 본 연구에서는 스플라인 기반의 스킨링 웨이트 수정을 위한 새로운 인터페이스를 소개한다. 이 인터페이스는 메쉬에 배정된 스킨링 웨이트의 등고선으로 표현된다. 사용자가 스플라인의 앵커 점을 컨트롤하면 등고선 모양이 업데이트가 되고, 스플라인에 배정된 값을 만족시키는 캐릭터 표면 전체의 스킨링 웨이트를 계산한다.

우리는 인터페이스가 실시간으로 실행되도록 몇 가지 기술을 소개하고 주어진 스플라인 제약의 의한 적절한 스킨링 웨이트 분포를 생성하는 함수의 특정 조합을 제안한다.

2. 스플라인 모델

먼저 기본적인 모델로써, B-Spline 모델을 사용한다. 이는 부드러운 곡선과 쉬운 로컬 제어가 가능하기 때문에 많은 응용분야에서 널리 사용된다. 본 연구 목적에 맞는 스플라인은 표면 도메인에서 정의되어야 하기 때문에, Panozzo et al. [1]의 기술을 사용하여 해결하였다. 해당 연구에서 Euclidean-embedding metric을 이용하여 geodesic metric을 근사 하기 때문에, 빠른 계산이 가능해진다.

3. 스킨링 웨이트 계산

만족되어야 할 주된 조건은 스플라인 제약 조건이다. 즉, 버텍스들에 배정된 스킨링 웨이트는 스플라인에 주어진 값과 일치하도록 결정되어야 한다. 우리는 barycentric coordinate을 이용하여, 각 페이스의 버텍스들의 스킨링 웨이트를 보간한다. 스플라인 점에 배정된 값에 대한 barycentric 식들을 행렬에 쌓아서 다음의 선형식을 구성한다: $Ax = s^i$. 여기서, 행렬 A의 각 행은 barycentric 조건을 의미하고, x는 모든 버텍스들의 스킨링웨이트이며, s^i 는 각 스플라인 점에 배정된 값의 벡터이다. 스킨링 웨이트 w^s 는 스플라인에 의한 선형 등식 제약 조건을 만족하는 에너지 $E(x)$ 를 최소화하여 다음 식을 통해

* 구두 발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 ACM Transactions on Graphics 에 합격 되었음.

* 본 연구는 기가코리아(GK17P0200) 사업부의 지원으로 수행되었음.

계산할 수 있다:

$$w^s = \underset{x}{\operatorname{argmin}} E(x) \quad \text{subject to} \quad Ax = s^i \quad (1)$$

에너지 함수 $E(x)$ 에 의해 스키닝 웨이트의 형태가 결정된다. 우리는 먼저 스플라인 제약을 만족시키는 두 가지 방법, 즉 주어진 값을 보간하는 보간법(interpolation-based method)와, 값을 주변 지역으로 확산시키는 확산법(diffusion-flow method)을 소개한다. 주어진 값이 하나만 있는 경우는 확산법을 사용하여 스키닝 웨이트를 계산한다. 주어진 값이 둘 이상일 경우는 영역을 나누어 보간법과 확산법을 함께 사용한다.

앵커 웨이트(anchor weight)는 행렬 K 로 표현되며, 각 열이 각 앵커 점의 스키닝 웨이트에 끼치는 영향력을 나타낸다. 각 앵커 점에 배정된 값을 c^i 라고 했을 때, 스키닝 웨이트는 다음과 같이 계산된다: $w^s = Kc^i$. 사용자가 앵커 점 위치를 수정할 때, 앵커 웨이트의 non-zero 값을 가진 영역만이 영향을 받는다. 그러한 전제를 바탕으로, 앵커 웨이트를 이용하면 계산 영역을 비약적으로 줄어짐에 기여하게 되고, 이는 빠른 계산 속도로 이어진다.

4. 스플라인 편집

사용자는 GUI를 통해 메뉴에서 앵커 위치를 선택하여 스플라인을 처음부터 자유롭게 만들 수 있다. 또한 스플라인은 주어진 스키닝 웨이트에서부터 추출하여 만들 수도 있다. 사용자는 주어진 스플라인을 자유롭게 위치를 변경할 수도 있고, 삭제, 삽입, 분할 및 병합과 주어진 기능을 통한 수정 작업을 수행할 수 있다. 이러한 수정 과정 중에 시스템은 실시간 상호작용이 가능하게 하기 위한 로컬 계산을 수행한다.

5. 실험

시스템은 독립 실행형 응용 프로그램으로 구현되어 있다. NanoGUI를 이용하여 그래픽 사용자 인터페이스를 구현하였으며, libIGL 라이브러리를 사용하여 기하학 처리 알고리즘을 상당수 이용하였다. 또한 Panozzo et al. [1]에서 제공한 오픈 소스 코드를 사용하여 Euclidean-embedding metric을 계산하였다. 모든 실험은 2.8GHz Intel Core i5 CPU 및 16GB 메모리가 장착된 MacBook Pro에서 수행되었다.

우리 인터페이스의 효과는 6 포인트 컨트롤러로 리깅된 달팽이 예제(그림 1)에서 잘 설명된다. 연체에 해당하는 물체를 변형할 때 강체에 해당되는 껍질 부분은 단단하게 유지되어야만 한다. 이를 위한 이상적인 스키닝 웨이트는 두 부분 사이의 경계를 따라 급격하게 변경되어야 한다. 껍질과 몸체 사이의 경계에서 높은 값의 스플라인 곡선을 생성하여 만족스런 결과를 얻을 수 있었다. 표면의 변형 특성을 고려하지 않는 기존의 자동화 알고리즘 Bounded Biharmonic Weight(BBW) [3]의

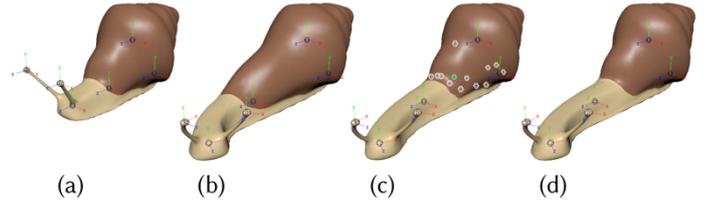


그림 2: 달팽이 모델 초기 형태 (a). BBW를 이용한 변형 결과 (b). BBW에서 13개의 추가 컨트롤러를 삽입 후 변형 결과 (c). 우리의 결과 (d)

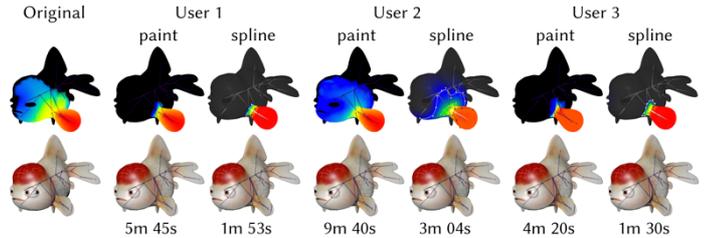


그림 3: 왼쪽: 물고기 모델의 Maya를 이용한 자동 알고리즘으로 인한 스키닝 웨이트(위) 및 그에 따른 변형(아래). 3명의 사용자가 페인트 인터페이스와 스플라인 인터페이스를 사용하여 초기 스키닝 웨이트를 수정함. 아래쪽에는 작업을 완료하는데 소요된 작업시간을 표시.

경우는(그림 2(b))과 같이 부자연스러운 결과를 보여준다. 자동화 알고리즘을 이용하면서 문제를 해결하기 위해서는 훨씬 더 많은 컨트롤러를 추가해야 한다. 껍질 주위에 13개의 추가 컨트롤러를 설치함으로써(그림 2(c)) 우리와 비슷한 결과를 재현할 수 있지만, 스플라인 인터페이스를 사용하여 얻어낸 결과(그림 2(d))만큼 완벽하진 않다.

우리는 또한 자동 알고리즘으로 인해 초래된 문제가 있는 스키닝 웨이트를 기존의 페인트 인터페이스와 우리가 제안한 스플라인 인터페이스를 이용하여 수정하도록 사용자 테스트를 수행하였다. 그림 3(왼쪽)은 Maya에서 제공되는 거리 계산법에 의한 초기 스키닝 웨이트를 보여준다. 지느러미 골격 주변으로 과도하게 스키닝 웨이트가 분포됨으로써, 부자연스러운 변형을 일으키는 것을 확인할 수 있다. 사용자는 페인트 인터페이스와 스플라인 인터페이스를 사용하여 초기 스키닝 웨이트를 각각 수정하도록 했다. 훨씬 짧은 수정 시간에도 불구하고, 스플라인 인터페이스와 페인트 인터페이스를 사용한 결과는 거의 유사함을 확인할 수 있다.

참고문헌

- [1] Daniele Panozzo, Ilya Baran, Olga Diamanti, and Olga Sorkine-Hornung. 2013. Weighted averages on surfaces. *ACM Transactions on Graphics (TOG)* 32, 4 (2013), 60.
- [2] Alec Jacobson, Daniele Panozzo, et al. 2016. libigl: A simple C++ geometry processing library. (2016). <http://libigl.github.io/libigl/>.
- [3] Alec Jacobson, Ilya Baran, Jovan Popovic, and Olga Sorkine. 2011. Bounded biharmonic weights for real-time deformation. *ACM Trans. Graph.* 30, 4 (2011), 78.

스트로크 와핑을 이용한 회화적 렌더링*

김범식⁰, 윤경현
 중앙대학교 컴퓨터공학과
 beomsik@cglab.cau.ac.kr, khyoon@cau.ac.kr

Painterly Rendering with Brush Stroke Warping

Beom-Sik Kim⁰, Kyung-Hyun Yoon
 Dept. of Computer Science Engineering, Chung-Ang University

요약

본 논문은 이미지 와핑(Image Warping)을 사용하여 새로운 스트로크를 생성하는 방법을 제안한다. 기존의 회화적 렌더링은 적은 스트로크 데이터베이스를 보유했으므로 선택의 폭이 좁아 렌더링 결과물에 부자연스러움이 있었다. 하지만 본 논문의 방법으로 기존의 질감이 보존된 새로운 스트로크 데이터베이스를 생성함으로써 기존보다 향상된 렌더링 결과물을 생성할 수 있다.

1. 서론

회화적 렌더링은 비사실적 렌더링의 한 분야로서, 픽셀 기반 혹은 스트로크 기반 등 여러 분야로 연구되어오고 있다. 그 중 본 논문에서는 스트로크 기반 렌더링(stroke based rendering)의 결과를 개선하는 방향을 제시한다. 스트로크 기반 렌더링은 스트로크를 기본 단위로 사용하여 렌더링 결과를 만드는 기법[1]이다. 기존의 스트로크를 생성하는 방법[2]으로는 후면에서 빛을 쏘는 반투명 유리판위에 검은색 유화 안료를 실제로 그리고 캡처하여 얻는다. 그 후 후처리를 통하여 발생한 노이즈를 제거 하여 5개의 스트로크를 생성했다. 그러나 이러한 스트로크 프로파일(Stroke profile)의 방식은 생성하는데 시간이 오래 소요되고 번거로운 단점이 있으므로, 데이터베이스를 확장시키기 위하여 회전, 사이즈 조절 등 이미지 변환을 사용하여 64개로 확장했다. 하지만 가장 근본적인 문제인 다양성이 낮은 부분은 해결하지 못하여, [그림 1]의 실제 회화처럼 여러 모양의 스트로크를 사용하지 못했다. 이런 제한사항을 해결하기 위하여 스트로크를 객체에 알맞게 배치[3]하거나 엣지와 방향성을 기반으로 스트로크를 배치[4]하여 표현했지만 표현방법에 한계가 있었다. 그러므로 본 논문에서는 기존 스트로크의 질감을 보존하며 새로운 모양의 스트로크를 생성할 수 있는 방법으로 이미지 와핑(Image warping)을 사용한 자동 및 수동방식을 제안한다. 그리고 생성된 208개의 스트로크를 포함한 272개의 스트로크를 적합하게 배치하기 위해 새로운 브러시 판단과정을 제시해 개선된 회화적 렌더링 결과를 생성하고자 한다.



[그림 1] 회화 중 스트로크의 다양성
 고호 - 붉은 장미가 있는 꽃병(1), 해바라기(2)

2. 연구내용

2.1 개요

기존 스트로크의 질감을 보존하기 위하여 우리는 이미지 와핑을 사용했다. 그리고 표현하고자 하는 객체와 유사한 브러시 선택을 위하여 이미지와 브러시간의 평균 색상거리와 커브 유사도로 적합한 브러시를 선별했다. [그림 2]는 시스템 개요를 도식화한 것이다.



[그림 2] 시스템 개요

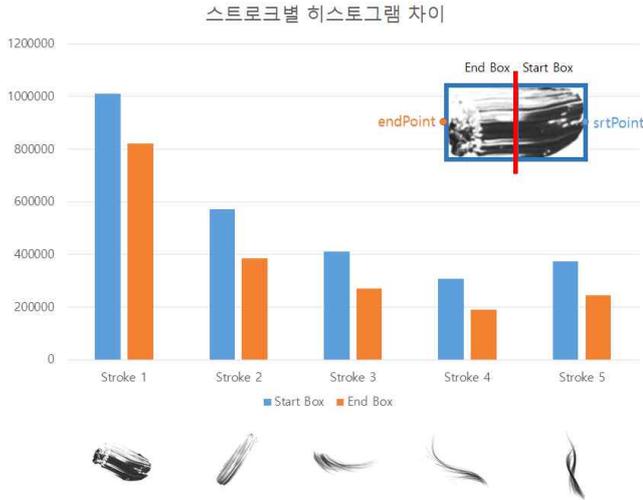
2.2. 커브 방향 및 가중치

스트로크는 중심을 기준으로 위, 아래 두 가지의 커브 방향을 가지며 원하는 방향으로 꺾으려면 스트로크의 시작점(startPoint)과 종료점(endPoint)을 알아야 하므로 스트로크 전체를 포함하는 바운딩 박스(Bounding Box)를 구했다. 그 뒤 가로로 이등분하여 브러시가 시작하는 부분의 박스(Start Box)와 끝나는 부분의 박스(End Box)를 구했다. [그림 3]은 스트로크를 그릴 때 붓을 누르면서 그리고 때면서 완성하므로 시작 부분이 끝나는 부분보다 히스토그램 값이 높다는 것을 보여준다. 그리고 커브 가중치는 사람이 직접 그린 것과 최대한 유사하게 하기 위하여 인체해부학적으로 접근하였다. 일반적으로 팔의 상박을 고정하였을 때 손목의 움직임만으로 원호를 그린다면 최대 내각 130도(배굴 : 60도, 장굴 : 70도)[5]에 해당하는 원호를 그릴 수 있다. 즉, 기존의 스트로크를 손목만의 움직임으로 곡선의 형태로 그린다고 가정한다면 내각이 130도인 원호의 모양으로 그린다는 것이다. 따라서 원호의 내각이 커브 가중치이며 스트

* 구두 발표논문

* 이 논문은 2017년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임.(No. NRF-2017R1A2B4007481)

로크의 중심점들이 와핑 시 출발점들(srcPoint)이다. 목표점들(dstPoint)은 원호의 중점으로부터 출발점의 방향으로 직선을 그렸을 시 원호와 만나는 점들이다. 본 논문에서는 커브 가중치를 0(원본)부터 10(내각 130도)까지 총 11단계로 나누어 커브 가중치를 설정할 수 있다.



[그림 3] 스트로크별 히스토그램 차이

2.3. 브러시 판단과정

본 논문에서는 객체를 표현하기에 적합한 브러시를 고르기 위하여 객체의 이미지와 각 브러시들의 평균 색상 거리와 커브 적합성을 계산한다. 평균 색상거리는 시도 점을 중점으로 브러시 크기만큼의 후보 이미지(I₁)와 브러시 사이(I₂)의 RGB 각 색상의 유클리디안 거리의 평균(D(I₁, I₂))으로 계산된다(수식-1). 그리고 브러시 커브 정도는 2.2에서 목표점들(P_d)과 각 목표점에서 가까운 후보 이미지의 엣지(P_e)의 유클리디안 거리의 평균을 커브 유사도(D(P_d, P_e))로 계산된다(수식-2). 스트로크를 칠하려는 위치에서의 각 브러시별 점수(B_{score})는 각 계산된 점수들의 합과 같다(수식-3).

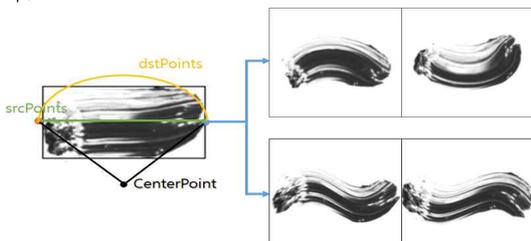
$$D(I_1, I_2) = \frac{\sum_x^w \sum_y^h |I_1(x, y) - I_2(x, y)|}{w * h} \quad (1)$$

$$D(P_d, P_e) = \frac{\sum_s^e |P_d(s) - P_e(e)|}{e - s + 1} \quad (2)$$

$$B_{score} = w_{col} * D(I_1, I_2) + w_{cur} * D(P_d, P_e) \quad (3)$$

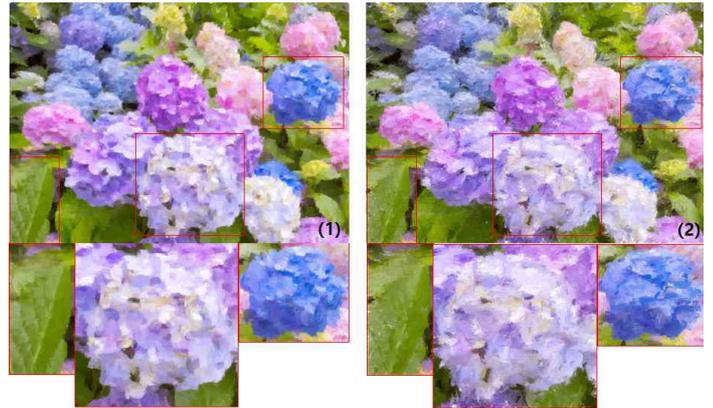
3. 결과

본 논문에서 커브가중치를 주어 스트로크를 생성하는 방법은 기존의 방법[2]보다 생성속도나 편리성이 뛰어난 장점이 있다. [그림 4]에서 상단의 스트로크는 자동으로 생성된 결과이며 하단은 수동으로 생성된 결과이다.



[그림 4] 커브 가중치를 주어 생성된 결과

[그림 5]은 수국의 사진을 확장된 272개의 스트로크 데이터베이스(1)와 기존 64개의 스트로크 데이터베이스(2)로 렌더링을 한 결과이다. 기존의 스트로크 데이터베이스는 반구형의 꽃 모양과 원형모양의 꽃잎을 표현하는데 제한사항이 있어 꽃잎과 꽃 모양이 거칠게 표현된 모습이 보인다. 하지만 확장된 스트로크 데이터베이스는 원호모양 등 기존 스트로크 데이터베이스에는 없는 모양을 포함하고 있으므로 개선된 결과가 보인다.



[그림 5] 렌더링 결과 (1)확장된 DB, (2) 기존의 DB

4. 결론 및 향후연구

본 논문에서 제안하는 새로운 스트로크를 생성하는 방법은 기존의 방법보다 소요시간도 적으며 간단하고 스트로크의 질감을 보존하면서 다양한 브러시를 생성하는 장점이 있다. 하지만 브러시의 크기가 고정되어있으며 자동의 방법으로는 커브 가중치를 1개만 부여하는 단점이 있다. 따라서 이후 스트로크의 크기가 제한이 없으며, 커브 가중치를 2개 이상 부여하여 스트로크를 자동적으로 생성하는 방법이 연구되어야 할 것이다.

참고문헌

[1] A. Hertzmann, "Tutorial: A survey of stroke-based rendering." *IEEE Computer Graphics & Appl.* 23, pp. 70-81, 2003.
 [2] Sang Hyun Seo, "A Study on Stroke Based Rendering Using Painterly Media Profile" *Journal of Korea Multimedia Society* Vol. 12. No. 11. November. pp. 1640-1651, 2009.
 [3] Ho Chang Lee, "A Study on Saliency-based Stroke LOD for Painterly Rendering" *Journal of KIISE : Computer Systems and Theory* Vol. 36. No. 3. June. pp. 199-209, 2009.
 [4] Gyeong-Rok Lee, "Lautrec Portrait Rendering" *Journal of Korea Multimedia Society* Vol. 17. No. 9. September. pp. 1052-1063, 2014.
 [5] "신체 각 관절에 대한 정상인의 표준운동각도 및 운동가능영역" 산업재해보상법(8조2항 관련)

다트 가상현실 콘텐츠에서의 협업을 위한 기계학습기반 사용자 의사결정 모델*

봉찬균⁰, 박성준, 박원준, 허하영, 김진모*
부산가톨릭대학교 소프트웨어학과

rainbon0@naver.com⁰, phs6048@nate.com, zm0nn@naver.com, hayoung6177@hanmail.net, jmkim11@cup.ac.kr*

Machine Learning based User's Decision Making Model for Collaboration in Dart Virtual Reality Contents

Changyun Bong⁰, Seongjun Park, Wonjun Park, Hayoung Heo, Jinmo Kim*
Department of Software, Catholic University of Pusan

요약

본 연구는 여러 사용자가 함께 참여하는 몰입형 가상협업환경에서 컴퓨터가 제어하는 가상의 사용자가 함께 존재함은 물론 사용자간 자연스러운 협업이 가능한 체험환경을 제시하기 위하여 기계학습을 활용한 의사결정 모델을 제안한다. 사용자간 협업에 필요한 구체적인 상황(다트 던지기)을 제시하고, 기계학습을 통해 인간의 사고와 유사하게 행동하며 경험을 공유할 수 있는 심층 신경망 모델을 설계한다. 실험을 위해 제작한 다트 가상현실 콘텐츠가 진행되는 과정 속에서 적절한 의사결정을 수행해 나감을 실험을 통해 확인하였다.

1. 서론

몰입형 가상현실은 사용자의 시각, 청각, 촉각을 포함한 오감을 만족시킴으로써 실제와 유사한 경험을 몰입된 환경에서 제공함을 목적으로 다양한 기술, 상호작용 방법에 대해 연구가 이루어지고 있다[1]. 사용자 한 명의 현존감을 극대화함을 목적으로 연구되던 몰입형 가상현실 기술은 여러 사용자가 동일 공간에서 서로 협업하는 가상협업환경과 함께 발전하며 다중 사용자의 몰입을 높이는 체험환경을 제시하고 있다. 가상현실 기반의 몰입형 가상협업환경에서 사용자의 현존감은 가상 환경 또는 객체와 상호작용하는 기술과 사용자간 의사소통, 협업을 얼마나 실제와 같이 제공하느냐가 중요하다. 특히, 다중 사용자가 생각을 공유하고, 함께 행동함을 인지하는 사회적 상호작용은 몰입형 가상협업환경의 핵심요인 중 하나이다. 이때, 컴퓨터가 제어하는 가상 사용자가 일반 사용자와 유사한 의사결정 모델을 가진다면 가상협업환경에서 제시할 수 있는 경험의 범위가 확대될 수 있으며, 사용자도 다양한 역할을 제공 받을 수 있을 것이다. 따라서 본 연구는 기계학습을 통해 몰입형 가상협업환경에서 사용자 의사결정 모델을 제안한다. 기계학습은 인공지능의 한 분야로 인간이 학습하듯이 컴퓨터가 스스로 학습하게 함으로써 새로운 지식을 배

우도록 하는 것을 의미한다. 인공지능의 개념이 게임과 같은 대화형 콘텐츠에서 NPC의 길 찾기와 같은 기술에 응용되거나 몰입형 가상현실에서 특정 위치로 사용자를 이동시키는 과정에서 학습을 활용하기도 하였다[2]. 기계학습을 장기 게임 개발에 활용[3]하기도 하지만 몰입형 가상협업환경에서 사용자와 함께 협업하는 과정에서 의사를 결정하는데 활용하는 경우는 드물다. 따라서 본 연구는 이러한 한계를 극복하기 위하여 기계학습을 가상협업환경에 활용한 다트 가상현실 콘텐츠를 제작한다.

2. 몰입형 가상협업환경

2.1. 다트 가상현실 콘텐츠

몰입형 가상협업환경이 제시하는 다양한 목적 가운데 본 연구는 사용자간 협업을 통해 공통의 목표를 달성하거나 경험을 공유하는 것에 초점을 맞추었다. 따라서 제안하는 의사결정 모델도 사용자와의 협업에 기반 하여 인간의 사고와 유사하게 행동할 수 있어야 한다. 이를 위해 우선 구체적인 상황을 제시할 수 있는 다트 가상현실 콘텐츠를 제작한다. 그림 1은 제안하는 콘텐츠의 장면으로 의사결정 모델이 적용된 캐릭터는 다트를 던지고, 다른 한 명의 실제 사용자는 다트 판을 들어 정확히 맞히면 된다. 다트 판에 제시된 숫자는 9개의 숫자(1,3,4,6,7,12,13,17,20)로 구성되고, 30~50사이의 임의로 제시되는 목표 값을 빠르고 정확히 맞혀야 한다. 즉, 가상 사용자는 현재 숫자에서 다트 판의 몇 번 숫자를 맞혀야 하는지를 추론해야 한다.



그림 1: 제안하는 다트 가상현실 콘텐츠의 장면

2.2. 기계학습

제안하는 몰입형 가상협업환경에서 다트 사용자의 의사결정은 현재 제시된 숫자에 대하여 다트 판에 나열된

* 구두 발표논문

* 본 연구는 2018년도 정부(교육부)의 재원으로 한국연구재단(NRF) 지원으로 수행되었음. (No. 2017R1D1A1B03030286)

* 학부생 주저자 논문임

숫자 중 최적의 답을 찾는 것이 핵심이다. 이를 위해 우선 학습에 사용할 데이터를 정의한다. 1~50까지의 숫자를 원-핫 인코딩(one-hot encoding) 방식으로 표현(예: 3→[0,0,1,0,⋯,0])하고, 각각의 숫자에 대한 분류 값 데이터([1,3,4,6,7,12,13,17,20])를 구성한다. 분류 값의 경우 입력 데이터(예: 27)에 따라 중복 값 설정(20+7:[0,0,0,0,1,0,0,0,1])이 가능하다. 또한, 하나의 숫자에 대하여 하나의 정답만이 주어지는 것이 아니기 때문에 여러 샘플들을 다양하게 학습시킴으로써 다양한 경우의 수가 가능하도록 한다. 다음으로 예측 값과 실제 값 사이의 확률 분포 차이를 계산하여 손실(cost)을 계산하고 학습을 통해 이를 최소화시키며, 배열로 제시되는 결과 값을 전체 합이 1이 되도록 만들어 확률로 해석한다. 이때 확률 값의 차이가 15%이내인 결과 값만 분류하여 임의로 선택하게 함으로써 인간과 같이 다양한 사고를 하도록 한다. 그림 2는 제안하는 심층신경망 모델이다. 50개의 특징에 9가지 분류 수로 구성되며 은닉층의 뉴런 수는 300으로 설정하여 가중치(w_1, w_2)와 편향 값(b_1, b_2)을 설정한다. 활성화 함수는 실험을 통해 확인한 최소 손실을 가지는 시그모이드(sigmoid)를 사용한다.

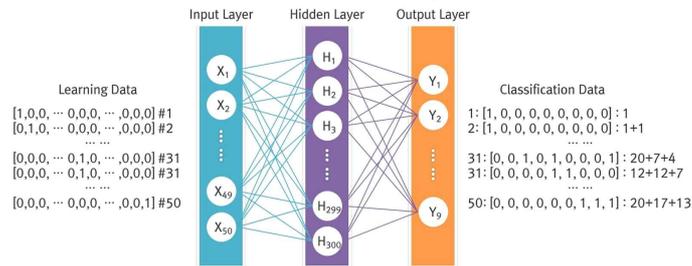


그림 2: 의사결정 모델에 적용된 심층신경망 설계

3. 실험 및 분석

실험을 위해 제작한 콘텐츠는 Unity 3D 2017.3.1f1 엔진에 SteamVR 플러그인을 통합하여 구현하였다. 의사결정 모델 설계를 위한 기계학습은 Python 3.6.5에 TensorFlow 1.8을 활용하였다. 실험에 사용된 PC환경은 Intel Core i7-6700, 16GB RAM, Geforce GTX 1080 GPU를 탑재하고 있다.

그림 3은 제작된 콘텐츠의 가상협업환경에서 사용자와 함께 컴퓨터가 의사 결정을 수행하는 과정을 나타낸 것이다. 구체적인 상황을 예로 설명하면 목표 숫자가 30~50 사이의 숫자 가운데 31이 제시되었고, 컴퓨터는 학습된 정보를 바탕으로 다트로 맞춰야 할 숫자로 7을 추론하였다. 일반적으로 가장 큰 숫자인 20이 아니고 7을 추론한 이유는 31이라는 숫자의 조합을 만들기 위하여 20,7,4 3개의 숫자를 맞추는 경우도 있지만, 7,12,12라는 숫자의 조합 역시 가능하기 때문에 20과 7 또는 7과 12사이의 확률 값이 15% 이내로 발생하였고, 7이 임의로 선택된 경우라 볼 수 있다. 이와 같이 제시되는 숫자에 대해 항상 같은 답을 제시하는 것이 아닌 다양한 경우의 수를 고려하게 함으로써 협업하는 사용자의 참여와 몰입을 더욱 높이는 가상협업환경이 제시됨을 확인할 수 있었다. 학습된 데이터를 토대로 컴퓨터가 답을 제시한다 하더라도 함께 협업하는 과정에서 다트 판 사용자가 실수 하는 경우에도 다시 현재 상황에

맞는 답을 새롭게 제시함으로써 만족하는 경험을 제공할 수 있었음을 알 수 있었다.

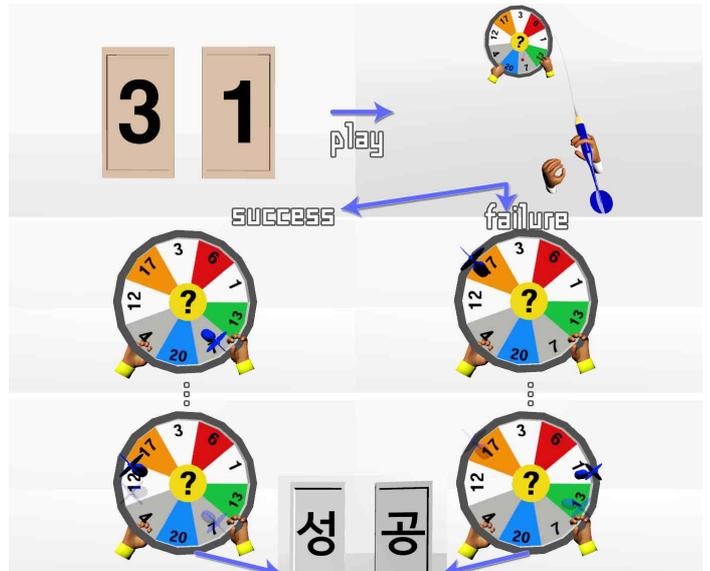


그림 3: 제안한 사용자 의사결정 모델을 통해 사용자와 함께 협업하여 목표를 달성하는 결과의 예

4. 결론

본 연구는 다중 사용자가 참여하여 함께 소통하고 협업하는 몰입형 가상협업환경에 컴퓨터가 제어하는 사용자까지 참여의 폭을 넓힐 수 있도록 기계학습을 활용한 사용자 의사결정 모델을 제안하였다. 제안한 모델은 가상현실 환경에서 주어진 목표를 일반 사용자와 함께 협업을 통해 달성해나가는 과정에서 행동을 결정하는 방법으로 구체적인 상황을 다트 가상현실 콘텐츠를 통해 제시하고, 적절한 결과를 다양한 경우로 나누어 학습을 시킨 다음 행동 패턴을 분석하였다. 실험 결과 제안한 의사결정 모델이 적용된 가상 사용자가 주어진 상황에서 적합한 답을 추론함과 동시에 다양한 경우의 수를 제시하는 것을 확인할 수 있었다.

향후 보다 다양한 가상협업환경의 상황을 제시하고, 참가자들을 대상으로 가상 사용자와 함께 협업하는 실험 환경에서 설문을 진행하며 이를 토대로 현존감, 경험에 대한 변화, 차이를 분석할 계획이다.

참고문헌

[1] M. Kim, C. Jeon and J. Kim, A Study on Immersion and Presence of a Portable Hand Haptic System for Immersive Virtual Reality, *Sensors*, 17(5), 1141, 2017.
 [2] A. Rovira and M. Slater, Reinforcement Learning as a tool to make people move to a specific location in Immersive Virtual Reality. *Int. J. Hum.-Comput. Stud.* 98(C), pp.89-94, 2017.
 [3] M. Jang et al., Development of Artificial Intelligence Janggi Game based on Machine Learning Algorithm, *Journal of Information Technology Services*, 16(4), pp.137-148, 2017.

GAN을 이용한 자연어로부터 발화 동영상 생성*

이건희⁰, Jacob Morton, 이승용
포항공과대학교
{victorlee, jacob, leesy}@postech.ac.kr

TellGAN: GAN based Speech Video Generation From Natural Text

Gunhee Lee⁰, Jacob Morton, Seungyong Lee
POSTECH

요약

본 논문은 인물 영상과 자연어를 통해 다음 발화 영상을 예측하여 동영상을 생성하는 방법을 제안한다. 자연어를 이용한 시계열에 따른 움직임 예측은 LSTM 구조를 활용하고, 이를 토대로 적대적 신경 생성망(GAN) 구조를 활용하여 예측된 움직임이 반영된 동영상을 제작한다. 최종적으로 자연어에 따라 동영상의 움직임이 다르게 변화하고, 움직임 속도를 조절하여 실제와 같은 동영상을 생성한다.

1. 서론

물체의 움직임을 예측하는 연구는 오랫동안 시도되어 왔다. 기존의 optical flow를 이용하여 다음 프레임의 움직임을 예측하는 연구부터 시작하여, 최근 딥러닝 연구에서는 시계열 학습에 사용되는 LSTM 구조를 이용하여 움직임을 예측하기도 한다[1]. 하지만 이러한 연구는 다음 움직임을 예측할 수는 있어도 예측을 바탕으로 동영상 제작하기까지는 어려움이 따랐다.

한편 딥러닝 분야에서 적대적 신경 생성망(GAN)을 활용하는 연구 또한 왕성하게 진행되고 있다. 적대적 신경 생성망은 서로 다른 두 종류의 네트워크가 최소극대화(Minimax) 경쟁을 통하여 데이터의 분포를 배우게 한다. 이러한 방법을 통하여 데이터 분포와 유사한 영상을 생성할 수 있고, 최근에는 특정 guide가 주어졌을 때 그에 맞는 영상 또한 생성할 수 있다[2].

본 논문에서는 이러한 LSTM 구조와 GAN 구조를 활용하여 한 장의 영상과 특정한 조건을 주었을 때 다음 움직임을 예측하고, 이 예측을 토대로 제한적인 동영상을 생성한다. 동일한 구조로 다양한 영상과 이에 맞는 조건이 가능하나, 이번 연구에서는 얼굴 영상과 자연어로 조건을 제한한다.

2. 네트워크 구성

일반적인 GAN 구조와 같이 독립적인 생성자(Generator)와 판별자(Discriminator)를 사용하여 네트워크를 학습시킨다. 생성자는 움직임 예측자(Motion Predictor)와 영상 생성자(Image Generator)로 구성되고, 판별자는 시계열 판별자(Sequence Discriminator)와 쌍 판별자(Pair Discriminator)로 이루어진다(그림 1 참조).

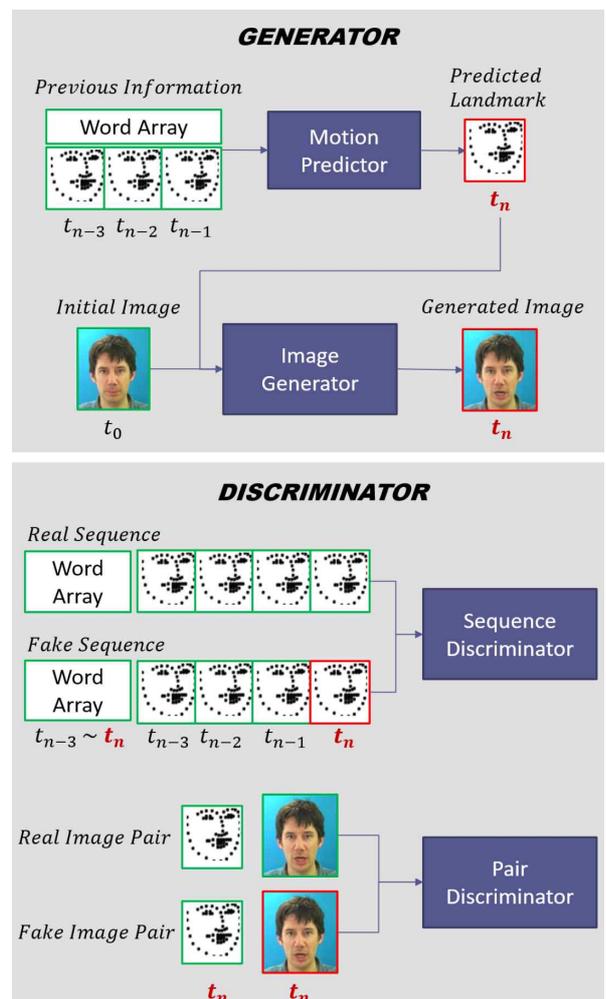


그림 1: 전체적인 네트워크 구성.

* 구두발표논문

* 본 연구는 과학기술정보통신부의 재원으로 차세대정보·컴퓨팅 기술개발사업(NRF-2017M3C4A7066317)과 정보통신기술진흥센터(SW 스타랩, IITP-2015-0-00174)의 지원을 받아 수행되었음

2.1. 생성자(Generator) 구조

생성자는 움직임 예측자와 영상 생성자 두 모듈로 구성된다. 움직임 예측자는 주어진 텍스트와 과거 랜드마크 정보를 통하여 다음 랜드마크 정보를 예측한다. 영상 생성자는 예측된 랜드마크 정보와 주어진 인물 영상을 통하여 이를 사실적인 영상으로 변환시켜준다.

움직임 예측자는 기본적으로 LSTM 구조로 이루어져 있다. 이때, LSTM은 해당 단어마다 다른 랜드마크를 예측하게 되는데 항상 일정한 수의 예측을 하게 된다. 즉, 단어에 따른 랜드마크의 변화를 일정한 수의 프레임으로 정규화시킨다. 이렇게 일정한 간격으로 정규화된 예측값은 실제 데이터가 있을 때 오류 역전과 과정을 통하여 학습을 진행하게 된다.

영상 생성자는 U-Net 구조를 기본으로 한다. 이를 위하여 얼굴 정보와 랜드마크 정보를 concatenate 한 후 네트워크로 통과시킨다. U-Net 구조는 skip connection을 가진 구조로 영상 생성 시에 랜드마크 변화만 반영한 정확한 영상을 만들어 준다.

2.2. 판별자(Discriminator) 구조

판별자는 시계열 판별자와 쌍 판별자로 구성된다. 시계열 판별자는 만들어진 예측값이 이전 값과 연속성이 있는지를 판별하고, 쌍 판별자는 랜드마크 값과 만들어진 인물 영상이 서로 쌍을 이루는지를 판별한다.

시계열 판별자는 텍스트 정보와 이전 모든 랜드마크 정보, 그리고 현재 시각의 실제 랜드마크 정보에 대해서는 Real 데이터로 판별해야 하고, 현재 시각의 예측된 랜드마크 정보에 대하여는 Fake 데이터로 판별해야 한다. 이전 랜드마크 정보는 매 순간 달라질 수 있으므로 이를 반영할 수 있는 LSTM 구조를 사용한다.

쌍 판별자는 실제 랜드마크 정보와 실제 인물 영상이 주어질 때 Real 데이터로, 실제 랜드마크 정보와 예측된 인물 영상이 주어질 때 Fake 데이터로 판별해야 한다. 쌍 판별자는 GAN의 가장 기본적인 판별자와 마찬가지로 convolution 연산으로 구성된다.

3. 실험

본 논문에서 제시한 네트워크 구조를 통하여 실험을 진행한 결과는 그림 2와 같다. 움직임 예측자는 단어에 따라 유사하지만 미세한 차이로 다르게 랜드마크를 예측한다. 영상 생성자는 이렇게 예측된 랜드마크 정보에 상응하는 실제 같은 영상을 생성한다. 또한 예측시간 정규화를 통하여 움직임의 속도를 조절할 수 있다.

4. 결론 및 향후계획

본 논문에서는 한 장의 인물 영상과 특정 자연어 조건을 이용하여 다음 순간의 얼굴 랜드마크를 순차적으로 예측하고, 이에 상응하는 얼굴 영상들을 생성하여 실제

와 같은 동영상을 만드는 방법을 제안하였다.

실험 결과를 개선하기 위하여 자연어 조건에 따라 다양한 모양을 더욱 정확히 예측할 수 있도록 움직임 예측자의 수정이 필요하다.

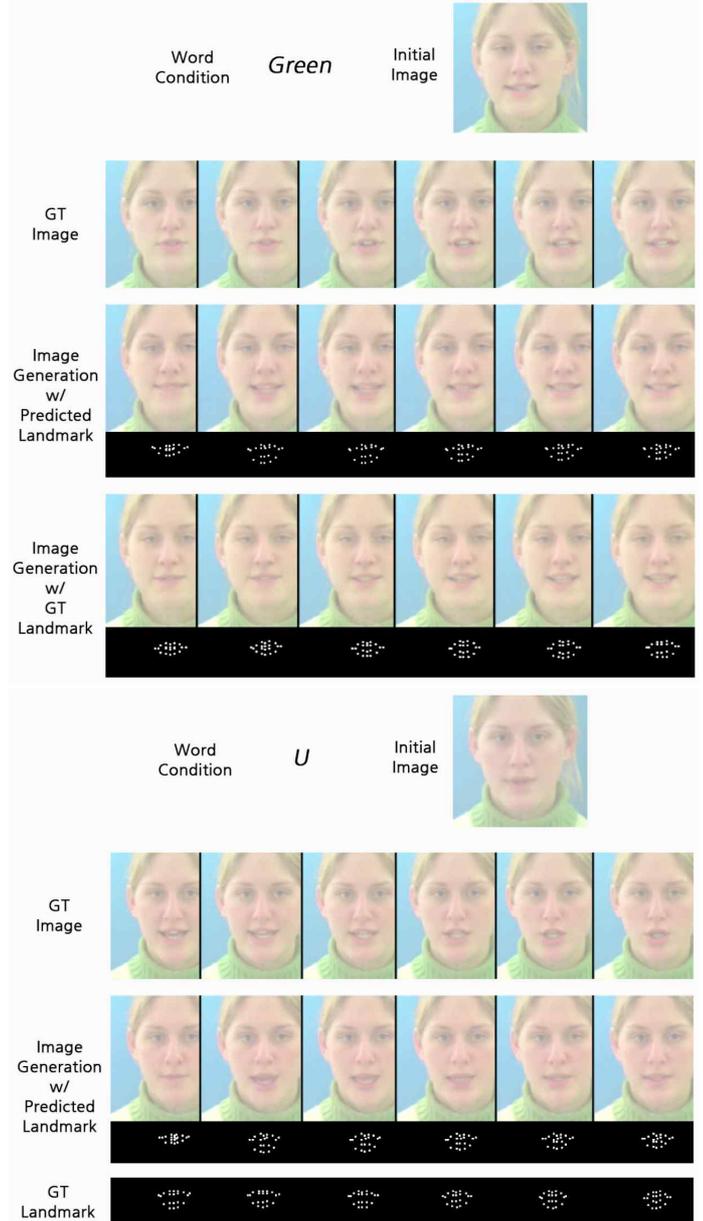


그림 2: 한 장의 초기 영상과 자연어를 통한 동영상 생성 결과. 단어 'Green'과 'U'가 주어졌을 때 예측된 랜드마크 모양의 변화가 유사하지만 미세한 차이가 있다. 또한 초기 영상만을 이용하여 예측된 랜드마크 모양에 상응하는 실제 같은 얼굴이 생성된다.

참고문헌

[1] J. Martinez, M. J. Black, and J. Romero, On human motion prediction using recurrent neural networks, *CVPR*, 2017.
 [2] L. Ma, X. Jia, Q. Sun, B. Schiele, T. Tuytelaars, and L. V. Gool, Pose guided person image generation, *NIPS*, 2017.

멀티모달 딥러닝을 이용한 비디오 감정 회귀 분석*

김하연⁰, 이인권
연세대학교 컴퓨터과학과
quocrab@gmail.com, iklee@yonsei.ac.kr

Video Emotion Regression Analysis using Multi-modal Deep Learning

Hayeon Kim⁰, In-Kwon Lee
Dept. of Computer Science, Yonsei University

요약

비디오 콘텐츠 분류 및 검색에서 감정은 중요한 구성 요소 중 하나이기 때문에 비디오 감정 인식은 꾸준히 연구되고 있는 분야이다. 본 논문은 최근 다양한 인식 분야에서 좋은 성과를 거둔 딥러닝(deep learning)을 이용한 멀티 모달 학습 구조의 새로운 비디오 감정 회귀 분석 모델을 제안한다. 그리고 본 논문에서 제안한 모델이 이전 다른 형태의 멀티 모달 구조의 회귀 분석 모델들보다 성능의 향상이 있었음을 보인다.

1. 서론

비디오는 대표적인 복합 데이터로, 영상과 소리로 이루어져 있다. 특히 감정과 같은 복합적인 정보 기반의 요소를 기계학습 방식으로 분석하고자 할 때는 두 개 이상의 데이터 모달리티(modality)를 학습하는 멀티모달 학습(multimodal learning) 방식으로 주로 접근한다.

현재 비디오 감정 분석을 위해 멀티모달 러닝 구조로 접근하는 모델들은 주로 피쳐 (feature) 수준에서의 융합 방법과 (feature-level fusion) 결정 수준에서의 융합 방법 (decision-level fusion)으로 나누어진다. 전자는 모든 피쳐들을 하나의 벡터로 연결하여 분류기(Classifier) 혹은 Regressor에 넣어 인식하는 방법이고, 후자는 각 요소의 분류 혹은 회귀 결과를 합치는 방법이다.

본 논문에서는 최근 이미지 분류 및 인식에서 좋은 성능을 보였던 DenseNet [1]과 사운드 데이터를 학습시켜 레이블이 지정되지 않은 비디오를 분류하는 모델 SoundNet [2]을 fine tuning하고 피쳐 수준에서의 융합 방법을 이용한 회귀 분석 모델을 제안한다. 그리고 감정 비디오 데이터셋인 LIRIS-ACCEDE 데이터베이스를 이용한 실험 결과로 제안된 방법의 우수성을 입증한다.

2. 제안 방법

본 연구에서 제안하는 모델은 그림 1과 같이 총 3개의 스트림으로 나누어진다. 첫 번째 옵티컬 플로우 스트림

에서는 정밀한 움직임 특성을 추출할 수 있는 dense optical flow [9]를 전체 프레임 영역에서 추출한 후 ImageNet 데이터셋에 사전 훈련된 5 layer의 ConvNet 모델에 통과시킨다. 두 번째 플로우 스트림에는 10프레임 간격으로 샘플링된 연속 RGB 프레임들 ImageNet 데이터셋으로 사전 훈련된 DenseNet 모델에서 fine tuning한다. 두 스트림에서 얻은 피쳐를 하나의 특징 벡터로 연결한 후 1 layer ConvNet과 글로벌 풀링에 통과시킨다.

마지막 스트림에는 각 비디오 파일에서 얻은 sound waveform을 추출한 후 라벨이 없는 200만개의 동영상으로 미리 훈련된 SoundNet 모델에서 피쳐를 얻는다. 그리고 이전 글로벌 풀링에 통과시킨 특징 벡터와 결합한다. 이 최종 특징 벡터를 회귀 계층(FC Layer)의 입력으로 사용하였다. 그리고 loss 함수는 MSE (mean squared error)를 사용하였고 optimizer는 Adam (adaptive moment)을 사용하였다.

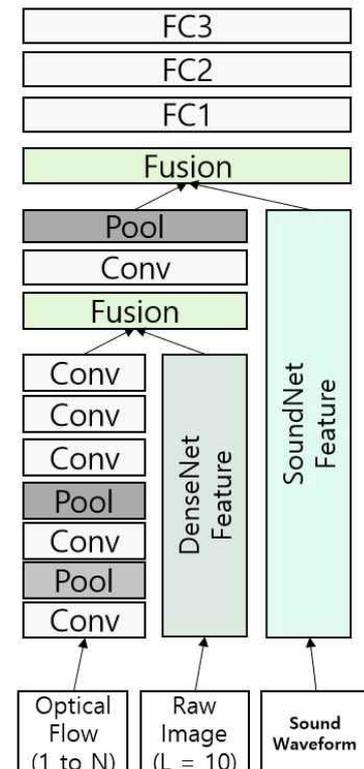


그림 1: 제안 모델

* 구두 발표논문
* 본 논문은 연구초기결과의 요약논문 (Extended Abstract)임
* 이 논문은 삼성전자 미래기술육성센터의 지원을 받아 수행된 연구임(과제번호 SRFC-IT1601-04).

3. 실험 및 결과

3.1. 감정 디스크립터

본 연구에서는 비디오에 사용자의 주관적인 감정 평가를 매기기 위해 차원적(dimensional) 접근의 디스크립터를 사용한다. 특히 Valence, Arousal로 구성된 2차원 레이블 집합을 이용해 감정을 측정한다. Valence란 즐거움의 정도 혹은 긍정/부정의 감정을 나타내는 척도이고, Arousal는 감정의 활성화 즉 흥분의 척도를 나타낸다.

3.2. 데이터 세트

MediaEval 2016[3]에서 제공하는 LIRIS-ACCEDE Discreate 데이터셋을 실험에 사용하였다. 160개의 장편 영화 및 단편 영화에서 추출한 9800개의 클립으로 이루어진 트레이닝 데이터셋과 테스트를 위한 별도의 1200개의 클립이 주어진다. 트레이닝 셋의 경우 원본은 0위부터 9799위까지 Arousal 값과 Valence 값 순위가 매겨져 있다. 해당 데이터셋을 사용하였던 MediaEval 2016에서는 랭킹 값을 이용해서 Gaussian process regression 모델로 회귀 분석을 위한 감정 스코어 값을 추출했다. 본 실험에서는 이 MediaEval 2016 ground truth 값을 기준으로 회귀 분석을 진행하였다.

3.3. 실험 조건

9800개의 영상 중 2450개의 영상을 Validation Set으로, 7350개를 Train Set으로 사용하였다. 또한, 성능 증가를 위해서 Data Augmentation을 적용하였다. 224 × 224 사이즈로 Random cropping을 하였고, 무작위 좌우 반전(left-right flipping)을 적용하였다.

3.4. 실험 결과 및 분석

우리의 실험 결과를 같은 데이터셋을 사용한 MediaEval 2016에서 제안한 비디오 회귀 분석 작업과 비교하였다. 또한 회귀 분석 결과에 대한 척도로 평균 제곱 오차 (MSE, Mean Squared Error)를 채택하였다. 그 비교 결과는 표1에서 확인할 수 있다.

	Valence	Arousal
RUC(run1) [4]	0.218	1.479
THU-HCSI [5]	0.214	1.531
BUL [6]	0.231	1.413
Liu's [7]	0.240	1.185
Gan [8]	0.33	0.77
Ours	0.269	0.628

표 1. MediaEval 2016 Movie Task 참가 모델 및 최근 모델과의 결과 예러 비교 (MSE)

우리의 모델이 상대적으로 좋은 성능을 보이고 있음을 볼 수가 있는데 특히 다른 모델들과 비교했을 때 흥분의 척도인 Arousal 값에서 더 좋은 성능을 보이고 있음을 확인할 수 있었다.

화면의 물체나 사람의 움직임 강도가 증가하면 청중의 Arousal 이 증가한다는 연구가 있다[10]. 그러므로 프레임 장면에서 이미지 객체의 정밀한 움직임을 측정할 Optical Flow의 사용으로 인해 성능 향상이 있었던 것

으로 보인다. 또한, SoundNet은 오로지 사운드를 통해서 비레이블 비디오 내의 행동에 대해 좋은 인식률을 보이고 있는 모델이다. 200만개의 비레이블 비디오를 미리 학습한 이 모델을 이용하여 추출한 피쳐를 사용하였기에 흥분의 척도인 Arousal이 더 잘 인식된 것으로 보인다. Valence의 경우 상대적으로 낮은 인식률을 보이고 있음을 볼 수 있는데, 이것은 기존의 멀티모달 모델들이 Valence 인식을 위해 추가하는 색상 피쳐, 얼굴 표정 피쳐의 부재 때문으로 판단된다.

4. 결론

본 연구에서는 비디오 감정 인식을 위한 새로운 형태의 멀티모달 딥러닝 구조를 제안하였다. 피쳐 수준의 융합 모델로, SoundNet과 DenseNet을 적용한 모델이다. 이 모델은 이전 최신 비디오 감정 인식 모델들과 비교하였을 때 비교적 적은 레이어 개수나 피쳐 개수를 사용한 간단한 구조임에도 불구하고 특히 Arousal 값에 대한 인식률이 개선되었고 전반적으로 성능이 뛰어난 것을 확인할 수 있었다. 이는 연구 초기 결과로서 그 의의가 있다고 볼 수 있다.

이후 본 연구를 확장하여 LIRIS-ACCEDE Continuous 데이터셋에도 적용하여 모델 성능을 검증해보고, Valence 인식률을 개선하기 위해 사람 얼굴 표정 피쳐를 추가하는 방식을 검토해 볼 것이다.

참고문헌

- [1] G. Huang, Z. Liu. "Densely connected convolutional networks." In CVPR, 2017.
- [2] Y. Aytar, C. Vondrick, A. Torralba. "SoundNet: Learning Sound Representations from Unlabeled Video." In NIPS, 2016
- [3] Y. Baveye, E. Dellandrea, C. Chamaret, and L. Chen. "LIRIS-ACCEDE: A Video Database for Affective Content Analysis." in IEEE Transactions on Affective Computing, 2015.
- [4] S. Chen and Q. Jin. Ruc at mediaeval 2016 emotional impact of movies task: Fusion of multimodal features
- [5] Y. Ma, Z. Ye, and M. Xu. Thu-hcsi at mediaeval 2016: Emotional impact of movies task
- [6] A. Jan, Y. F. A. Gaus, F. Zhang, and H. Meng. Bul in mediaeval 2016 emotional impact of movies task.
- [7] Y. Liu, Z. Gu, Y. Zhang, and Y. Liu. Mining emotional features of movies
- [8] Q. Gan, S. Wang, L. Hao, and Q. Ji. "A Multimodal Deep Regression Bayesian Network for Affective Video Content Analyses" In ICCV, 2017.
- [9] T. Brox, A. Bruhn, N. Papenber, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In Proc. ECCV, 2004
- [10] B. H. Detenber, R. F. Simons, and G. G. Bennett Jr. "Roll em!: The effects of picture motion on emotional responses." J. Broadcasting Electron. Media, vol. 42, no. 1, pp. 113-127, 1998.

목소리의 감성적인 필터링을 이용한 의사소통 도우미*

이수호^{0,1}, 신채영¹, 신영찬¹, 지승용¹, 최지혜^{0,2}, 배대용¹, 조윤진², 김지우¹, 신세영¹, 김종현^{1*}
 강남대학교 소프트웨어응용학부¹, 강남대학교 중등특수교육과²
 jonghyunkim@kangnam.ac.kr*

Communications Helper with Emotional Filtering of Voice

SooHo Lee^{0,1}, ChaeYoung Shin¹, YoungChan Shin¹, SeungYong Ji¹, JiHye Choi^{0,2}, DaeWoong Bae¹, YunJin Jo²,
 JiWoo Kim¹, SeYoung Shin¹, Jong-Hyun Kim^{1*}
 Dept. of Software Application¹, Dept. of Special Education in Middle School², Kangnam University

요약

본 논문은 일상적인 대화를 사운드와 텍스트로 분리하고, 각각에 대해서 감성을 추출 및 분석함으로써 상황에 맞는 대화의 내용과 분위기를 전달해줄 수 있는 의사소통 도우미를 제안한다. 텍스트는 대화의 내용을 직관적으로 전달하되 상황에 따라 다르게 표현 또는 해석된다. 예를 들어 말하는 화자의 목소리 톤에 따라서 일상적인 대화도 짜증이나 슬픔의 내용으로 변경될 수 있으며, 그 변화는 주파수 형태로도 나타난다 (그림 1 참조). 본 연구는 목소리의 감성 분석을 통해 정확한 의사소통을 도와줄 수 있는 기반 프레임워크를 제안한다. 결과적으로는 청각장애인 및 미취학 아동들이 화자의 대화를 정확하게 인지하여 의사소통 환경을 개선하고자 한다.

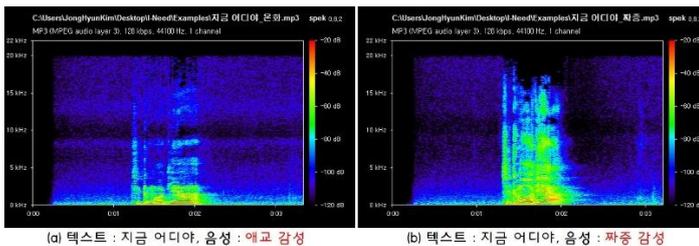


그림 1. 감정에 따라 달라지는 음성 주파수.

1. 서론

청각장애아동은 청각적 신호를 전달받지 못하기 때문에 화자의 말 전체를 이해하기 보다 자신이 들은 단어에 초점을 맞춘 대화를 시도하여 대화 유지가 어렵다. 이러한 청력 손실을 보상해주는 지원에는 보청기, 수화나 구화, 필담 등이 있다. 그러나 청각적 보상을 받았음에도 불구하고 주위의 소음과 말소리에 해당되는 주파수 일부를 왜곡해서 듣거나 듣지 못하게 되어 의사소통

이나 학습 등에서 방해를 받는다[1]. 그러므로 일반아동에 비해 자기표현이 어려운 청각장애아동은 언어적으로 제한되어 있기 때문에 일상생활에서의 기본욕구의 좌절 등에서 초래되는 부정적 심리를 감소시켜주는 것이 중요하다. 본 연구는 이러한 청각장애인의 일상생활에서 필요한 요구들을 바탕으로 이에 적합한 새로운 기능과 형태의 보조공학기기 기획을 통하여 장애인들의 의사소통에 도움을 줄 수 있는 새로운 프레임워크를 제안한다.

2. 제안하는 프레임워크

제안하는 시스템의 알고리즘 개요는 일반적인 대화 음성 파일을 입력으로 아래와 같은 순서로 수행된다. 기존 접근법과의 가장 큰 차이는 목소리를 사운드와 텍스트로 분할하여 각 감성을 추출한 뒤 최종적으로 사용자에게 상대방의 감성을 동시에 제공해주어 텍스트만 가지고 있는 정보의 한계를 개선시켰다는 점이다 (그림 2 참조).

1. 대화로부터 텍스트와 사운드를 추출하여 각각에 대해서 감성을 분석
2. 텍스트와 사운드의 감성을 비교/분석하여 말했을 때의 화자의 감성을 예측
3. 대화의 내용과 이를 분석했을 때의 화자의 감성을 동시에 전달

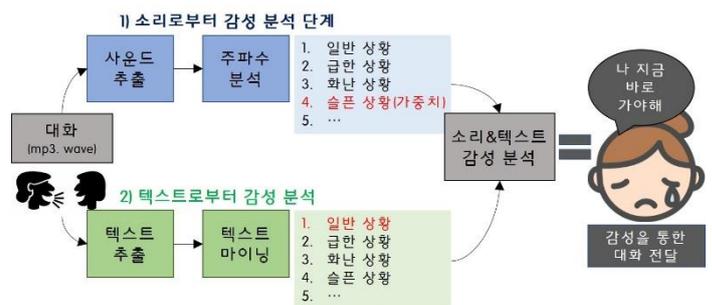


그림 2. 의사소통 도우미 프레임워크의 알고리즘 개요.

* 구두 발표논문
 * 학부생 주저자 논문임

2.1. 사운드로부터 감성 분석

우리는 5개의 감성 카테고리(평범, 행복, 슬픔, 화남, 무서움)를 이용하여 사운드로부터 감성을 추출한다. 사운드로부터 감성을 추출하기 위해 5가지 감정의 평균과 표준편차를 이용하여 감성을 분류하였으며, 본 연구에서는 이를 수행하기 위해 *OpenVokaturi* 라이브러리[2]를 이용하였다 (그림 3 참조).

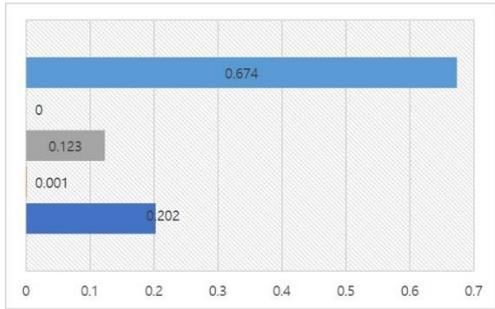


그림 3. 입력된 사운드로부터 분석된 감성 그래프 (X축: 감성 카테고리, Y축: 각 감성의 수치 값).

계산된 감정은 긍정(평범, 행복)인 s^p 과 부정(슬픔, 화남, 무서움)인 s^n 으로 카테고리를 분류였고, 각 카테고리에 속해 있는 감정은 평균으로 최종 사운드의 감성 수치 값을 계산한다 (그림 4 참조).

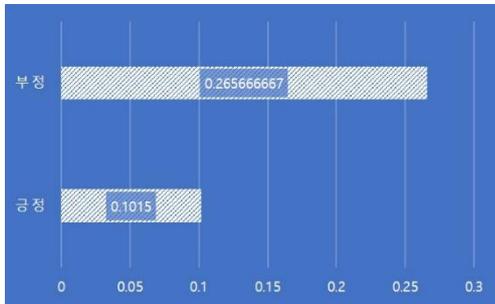


그림 4. 분류된 감성 카테고리에 따라서 계산된 사운드의 감성 수치 값.

2.2. 텍스트로부터 감성 분석

본 연구에서는 대화를 텍스트로 변환하기 위해 IBM의 왓슨 음성 인식[3]을 활용하였으며, 조건부 확률 모델인 나이브 베이즈(Naïve Bayes) 분류 기법을 이용하여 텍스트 내 감정을 분석하였다. 그 중에서 감성 단어 사전과 단어를 비교 대조하여 감성 점수를 산출하는 기법을 활용한다 (수식 1 참조).

$$p(c|x) = \frac{P(x|c)P(c)}{P(x)} \cong \operatorname{argmax} \prod P(x|c)P(c) \quad (1)$$

수식 1은 베이즈 정리를 수학적으로 표현한 것이며 보통 $p(A|B)$ 라고 할 수 있는데 B 가 나왔을 때 A 일 확률을 계산하는 것이다. 여기서 x 는 분류될 문장이고, c 는 분류할 클래스이다. 우리는 이 기법을 기반으로 긍정과 부정 감성 단어 사전과 대조하여 긍정적인 단어와

부정적인 단어의 빈도를 추출한 후 긍정 단어가 나타나면 증가, 부정 단어가 나타나면 감소의 점수를 부여해 텍스트의 감성을 수치화 한다. $p(c|x)$ 가 0으로 나오는 문제를 피하기 위해 라플라시안 스무딩을 적용하여 최종 텍스트의 감성 수치 값을 계산한다 (수식 2 참조).

$$p(c_i|x_k) = \frac{\operatorname{count}(c_i, x_k)+1}{|x_k|+|u|} \quad (2)$$

여기서 $\operatorname{count}(c_i, x_k)$ 는 k 번째 클래스에서 c_i 가 나타난 횟수이다. $|x_k|$ 와 $|u|$ 는 각각 k 번째 클래스의 모든 요소의 수와 모든 클래스에서 유일한 요소의 수를 나타낸다.

2.3. 텍스트와 사운드의 감성 합성

우리는 사운드와 텍스트의 긍정/부정에 대한 수치 값을 이용하여 최종 감성을 사용자에게 제안한다. 텍스트는 감성 단어 횟수에 의존하는 결과는 나타내지만, 사운드의 감성은 소리에 따라 변화가 크게 나타난다. 그렇기 때문에 본 연구에서는 각 감성을 합성할 때 사운드의 감성에 좀 더 높은 가중치로 설정하여 실험하였으며, τ_p 와 τ_n 를 비교하여 최종 감성을 결정한다.

$$\tau_p = \log\left(\frac{s^p\alpha+t^p\beta}{2}\right), \tau_n = \log\left(\frac{s^n\alpha+t^n\beta}{2}\right) \quad (3)$$

여기서 s^p 와 s^n 는 사운드의 긍정과 부정에 대한 감성 수치이며, 텍스트도 t^p 와 t^n 도 마찬가지이다. α 와 β 는 사운드와 텍스트에 대한 가중치이며, 본 연구에서는 각각 0.6, 0.4으로 설정하였으며, 사용자에게는 최종 감성과 개별적인 사운드와 텍스트의 감성도 같이 제공한다.

3. 결론

본 논문에서는 텍스트와 사운드의 감성을 이용하여 의사소통에 도움을 줄 수 있는 프레임워크를 제안하였으며, 일상적인 대화를 긍정/부정에 따라 실험했으며(표 1 참조), 대화의 감성 인식 측면에서 유의미한 결과를 얻었다.

지금 몇 시야, 지금 뭐해, 빨리 와, 빨리 왔네, 이게 뭐야, 잘했네, 등

표 1. 본 연구에서 실험한 일상적인 대화 템플릿.

참고문헌

[1] Kyoungghan Park, A study on the characteristics of connective ending use by connective ending forms for the students with hearing impairment, *The Korean Society of Education for Hearing-Language Impairments*, 5(1):1-22, 2014.
 [2] OpenVokaturi, <https://developers.vokaturi.com/getting-started/overview>, 2018.
 [3] IBM, Watson Speech to Text, <https://www.ibm.com/watson/kr-ko/developercloud/speech-to-text.html>, 1996.

의상 시뮬레이션 가속을 위한 패턴 미러링*

장인서⁰, 이상빈, 고희석
 서울대학교 전기정보공학부 그래픽스 및 미디어 연구실
 {isjang, sblee, ko}@graphics.snu.ac.kr

Pattern Mirroring for Fast Cloth Simulation

Inseo Jang⁰, Sangbin Lee, Hyeong-Seok Ko
 Graphics and Media Lab, Seoul National University

Abstract

This paper describes the method to reduce simulation time for cloth-body simulation. This method is only applicable for symmetric panel and symmetric body meshes which are centered on the Y-Z plane. The proposed method reduces simulation time up to 37%, by halving the matrix size of the linear equation.

1. Introduction

Stable and fast cloth simulation has been a problem. Prior researches have been made to overcome this limitation [1], proposed implicit method to stably simulate cloth in large time steps. Collision handling methods ([2], [3]) and acceleration techniques have been introduced. Algorithm to process contact and friction between cloth and obstacle is presented [4].

The pattern mirroring algorithm is based on the insight of the cloth garment's structure. Since human body model and basic cloth mesh are commonly symmetric, we use this characteristic to speed up the simulation. Cloth garments are symmetric to the center plane, which can be observed on read-to-wear garments. For garments that have additional layers such as pocket, which harm the symmetry, a simple post-process step is needed; stitching the additional layer on simulated basic mesh.

To set the positions of the vertices on the cloth mesh, velocity change amount(Δv) of each vertices is obtained from following linear equation[1].

$$A \cdot \Delta v = b$$

A and b are expressed by following equation where h is time step, f is force, x is position, v is velocity on vertices.

$$A = M - h \frac{\delta f}{\delta v} - h^2 \frac{\delta f}{\delta x}$$

$$b = h(f_o + h \frac{\delta f}{\delta x} v_o)$$

Assuming total number of the particle in a cloth mesh is n , size of matrix A is each $(3n*3n)$ and column vector Δv and b is $3n$.

One of the time-consuming part is solving the large linear equation. The larger n is (which means cloth has a larger number of vertices), the longer it takes to solve this linear equation because the time complexity to get the inverse of matrix is $O(n^3)$. Instead of using A^{-1} , methods for solving linear equation is used such as Newton's linear method and conjugate gradient method. Time complexity of conjugate gradient method is $O(n^{1.5})$ [5]. With the proposed mirroring method, the expected speed up for solving the linear equation grows exponentially as the number of vertices increases.

The other time-consuming part is collision detection and handling. Collisions can occur between cloth-body, between different cloth meshes, and between same cloth mesh. This method also reduces the number of searching tree, leading to reduction of simulation time.

2. Symmetric panel method

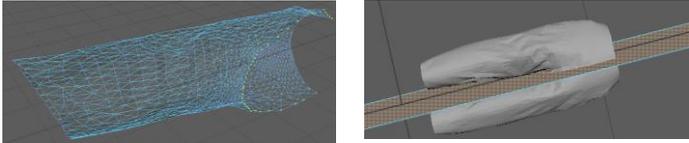
2.1 Pattern Mirroring Method

The main idea of mirroring method is to solve the linear equation using only the symmetric left half of the mesh (M^l) and duplicating the result to the right (M^r).

The first step of mirroring is solving linear equation [1] only for M^l . Vertices on the middle plane should be constrained to have 2 degree of freedom, which means they can move only on y-z direction, not x direction. We call this plane the constraint plane.

* Verbal presentation paper

* This work was supported in part by the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (MEST) (NO.2015R1A2A1A10055178), the Brain Korea 21 Plus Project in 2018, and ASRI(Automation and Systems Research Institute at Seoul National University).



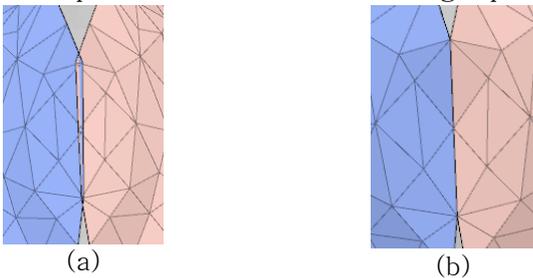
(Fig. 1) Constrained vertices(yellow) on M^l of pants (left) and the constraint plane(right)

The second step is reflecting step. Copy the position value of M^l to that of M^r about the constraint plane. The last is optional post-processing step for additional layers such as pocket, button. Attach additional layer on the simulated mesh, and simulate again with seam.

Unfortunately, comparing with the whole mesh (M) Fig.5(a), our method Fig.5(b) has a slightly different draping at middle part. This is caused by unconstrained vertices passing the constraint plane.

2.2. Following Artifact and Solution

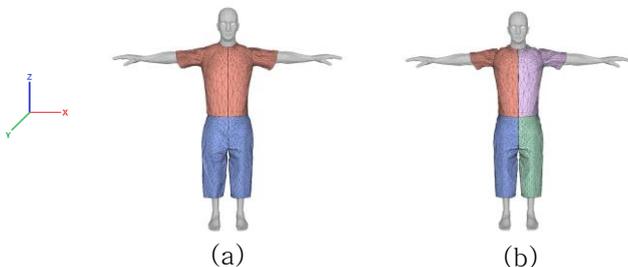
When the unconstrained vertices passed constraint plane, they can penetrate the other side mesh. Some edge lines belonging to blue panel penetrate red panel near the chest(Fig.2(a)). In this case, apply attraction force between these vertices and the constrained plane to move them to the right position.



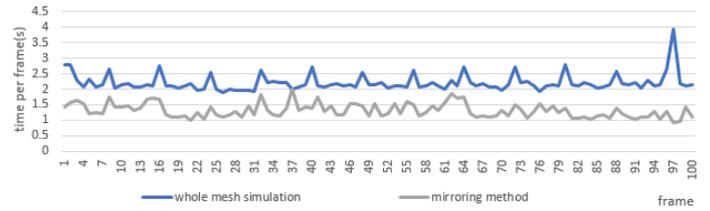
(Fig. 2) Penetrated edges (a) and wrong vertices pulled to the constrained plane (b)

3. Result

We performed 2 experiments on ARCSim which is open source simulator from University of California-Berkeley to compare result. The first one is pants and t-shirts simulation, which has 1,194 vertices, 2,290 vertices respectively and body has 5,400 vertices. Fig4 shows simulation time for each frame. Using mirroring method is x1.4 faster than simulating whole mesh on average.

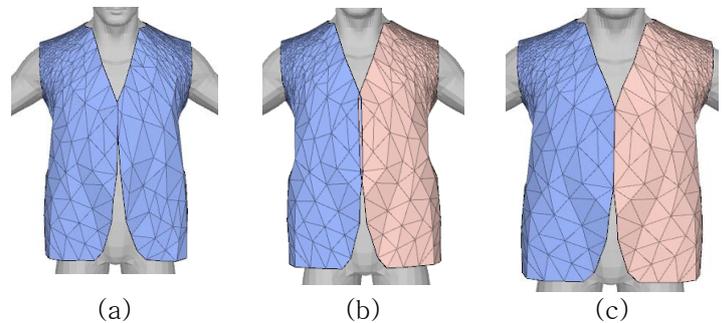


(Fig. 3) Pants and t-shirts scene by simulating whole meshes (a) and mirroring method after simulation half meshes (b).



(Fig.4) Comparing time for simulating whole mesh and mirroring method

The second experiment is a body with jacket which has the constrained plane only on the back. The number of vertices for M^r is 532 and 7482 vertices for body. Average time spent per frame on original mehs, half-mesh, symmetric-mesh are 0.27s, 0.18s, 0.19s respectively. In Fig5, panels colored in blue is the actually simulated panel, and reds are the mirrored panel.



(Fig. 5) Jacket scene by simulating whole meshes (a) and mirroring method with artifacts (b), and artifacts solved by attraction force (c)

4. Conclusion

We propose a fast simulation method based on implicit method under a condition that cloth and body is both symmetric. We developed our method for the symmetric body and garment. Considering basic layer for complex garment is usually symmetric, this method can be used widely.

Reference

- [1] David Baraff, Andrew Witkin, Large Steps in cloth simulation, *SIGGRAPH*, 1998.
- [2] Xavier PROVOT, Collision and self-collision handling in cloth model dedicated to design garments, pp.177-189, *Computer Animation and Simulation*, 1997.
- [3] David Baraff, Andrew Witkin Michael Kass, Untangling Cloth, *ACM SIGGRAPH 2003*.
- [4] Robert Bridson, Ronald Fedkiw, John Anderson, Robust Treatment of Collisions, Contact and Friction for Cloth Animation, *ACM TOG SIGGRAPH* pp.594-603, 2002.
- [5] J. Shewchuk. An Introduction to the Conjugate Gradient Method Without the Agonizing Plan. Technical Report CMU-CS-TR-94-125, Carnegie Mellon University, 1994.

키네틱 프로젝션 매핑을 위한 물리 기반 표면변형 시뮬레이션*

김정민^{0,1}, 김지수¹, 권오홍², 김영준¹
 이화여자대학교 컴퓨터공학과¹, 한국생산기술연구원²
 {zxij|cywh5526}@ewhain.net, ohung@kitech.re.kr, kimy@ewha.ac.kr

Physics-based Surface Deformation for Kinetic Projection Mapping

Jungmin Kim^{0,1}, Jisu Kim¹, Ohung Kwon², Young J. Kim¹
 Dept. of Computer Science and Engineering, Ewha Womans University¹
 Korea Institute of Industrial Technology²



그림 1: (왼쪽부터) 실제 카멜레온 서피스 구성 및 동적 프로젝션 매핑 결과와 가상 환경에서의 카멜레온 서피스 구성

요약

본 연구에서 제안된 키네틱 프로젝션 매핑 시스템인 카멜레온 서피스 시스템에서는 프로젝션 매핑 콘텐츠의 3차원 정보에 따라 프로젝션 스킨의 표면이 변형된다. 이러한 변형은 카멜레온 서피스의 액츄에이터에 부착된 제한된 수의 말단작용체에 의해 제어되며, 본 논문에서는 프로젝션 스킨의 물성치를 반영한 실시간 가상 직물 시뮬레이터를 구현하여 원하는 가상의 매핑 콘텐츠 형상과 카멜레온 서피스와의 차이를 최소화하도록 말단작용체의 위치를 최적화하는 방법을 제안하여 카멜레온 서피스를 효과적으로 제어할 수 있도록 하였다.

1. 서론

키네틱 아트에서 영상을 입체적으로 가시화하기 위해, 영상이 투사되는 표면(스킨)을 물리적으로 변형하는 기술들이 개발되고 있다. 본 연구에서 제안된 키네틱 아트 시스템인 카멜레온 서피스에서는 선운동을 하는 로봇릭 액츄에이터의 전면부에 신축성과 복원력이 좋은 직물 스킨을 부착한 형태로 동적 프로젝션 스킨을 구성한다(그림 1). 이때, 액츄에이터의 선운동 이동거리는 스킨에 투사되는 타겟 콘텐츠의 3D 깊이 정보에 의해 결정되고 이는 부착된 스킨을 변형하여 그 표면에 프로젝션

매핑을 통해 영상을 가시화한다. 카멜레온 서피스는 어떤 의미에서는 실시간 텍스처 매핑 기술인 변위매핑(displacement mapping)의 기계적인 구현이라고 볼 수도 있다.

한편 카멜레온 서피스에서 액츄에이터는 일정한 간격으로 배치되어 있고 그 개수가 제한되어 있기 때문에 투사하려는 타겟 메시의 변위량과의 차이가 항상 존재한다. 본 연구에서는 카멜레온 서피스를 물리 기반 가상 직물 시뮬레이션으로 구현하여, 직물 스킨 표면이 타겟 메시와 최대한 유사하게 변화하도록 액츄에이터의 위치를 최적화하는 것을 목표로 한다. 구현된 시뮬레이터는 물리 기반 시스템으로 실제 직물 스킨의 물성치가 적용되어 있으며 실시간으로 동작한다.

2. 물리 기반 표면변형 시뮬레이션 및 최적화

본 절에서는 카멜레온 서피스의 가상 시뮬레이터의 구현에 사용된 물리 시뮬레이션 시스템과 액츄에이터의 위치를 최적화하는 방법에 대해 알아본다.

2.1. Position-based Dynamics

물리 기반 실시간 가상 직물 시뮬레이션 중 대표적 방법인 Position-based dynamics(PBD)는 기하학적 제약조건에 따라 가상 직물 노드의 위치 정보를 직접적으로 제어한다 [1][2]. 본 논문에서 사용된 시뮬레이터는 스킨 서피스의 변형을 액츄에이터의 말단작용체와의 접촉에 의한 위치제약조건으로 모델링한다.

* 구두발표논문, 요약논문 (Extended Abstract).

* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2018년도 문화기술연구개발지원사업의 지원으로 수행되었음.

2.2. 가상 직물 물성치 조정

본 연구에 사용된 PBD의 경우에 고성능의 시물레이션이 가능하다는 장점이 있지만, 엄밀한 의미에서의 물리 기반 시물레이션이라고 보기는 힘들다. 하지만 기하학적 제약조건을 직접적으로 제어할 수 있으므로, 임의의 물성치조건을 만족시키게 할 수 있는 장점이 있다.

직물은 탄성체로서 이의 물성치는 직물의 변형을 결정하는 요소이다. 물성치는 직물이 늘어나고 줄어드는 변형이 선형적인지 비선형적인지, 혹은 등방성인지 비등방성 성질을 갖는지에 따라 다르게 표현될 수 있다 [3].

일반적으로 유한요소법(FEM)에서 직물의 물성치는 외력에 의한 직물의 변형율(strain)과 그로 인해 발생하는 응력(stress) 사이의 관계인 강성(stiffness)에 따라 결정된다. PBD에서는 거리 제약 조건을 통하여 직물 노드가 변형될 수 있는 길이를 제한하므로 [1] 거리제약 조건 변수의 강성치를 조절한다.

카멜레온 서피스 프로젝트에서 사용된 스킨은 비선형적 변형을 겪으며 일반적인 직물에 비해 신축성이 높고 등방성을 띤다 (그림 2). 이 물성치를 반영하기 위해 거리 제약 조건 변수(k)를 직물의 변형율(x)에 따라 다음과 같이 모델링 한다. 이때 상수 c 는 모델링 된 그래프가 실제 그래프의 기울기를 반영하도록 선택한다.

$$k(x) = e^{cx^2} - 1, \quad c \in R^+$$

2.3. 가상 직물 변형 최적화

액츄에이터에 부착된 직물 스킨이 타겟 메시와 유사한 형태를 이루게 하기 위해 액츄에이터의 위치를 최적화 해야 한다.

우선, 최적화 전의 가상 직물 시물레이션의 결과와 타겟 메시의 깊이 차이를 가상 직물의 모든 노드에 대해 구한다. 이후에 각 노드에 대해 가장 가까운 위치의 액츄에이터에 타겟 메시와의 차이를 더하여 모든 액츄에이터에 대해서 주변 노드들의 타겟 메시와의 차이에 대한 평균 값을 얻는다. 이 평균 값을 액츄에이터가 이동해야 할 값으로 할당을 하며 이때 액츄에이터의 움직임은 PBD의 위치제약조건에 의해 제어된다. 이 과정을 타겟 메시와의 차이가 일정 값 이하가 될 때까지 반복한다.

3. 결과

3.1. 가상 직물과 실제 직물 물성치 비교

2.2절에 제시된 방법을 통해 가상 직물에 실제 직물의 물성치를 반영한 후 이를 검증하기 위해 실제 직물의 인장실험과 같은 환경의 시물레이션 구성을 통해 가상 직물에도 동일한 인장 실험을 진행하여 가상 직물의 변

형율-응력 그래프를 얻었다. PBD 상에서 응력이 직물의 변형 가속도가 아닌 변형율에 의거하기 때문에 그림 2와 같은 결과를 확인할 수 있다.

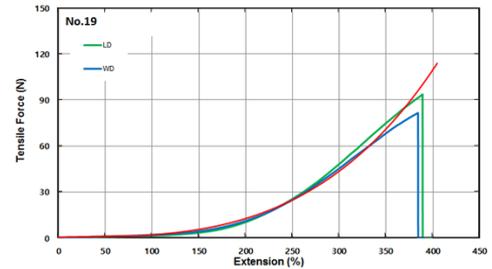


그림 2: 실제 직물의 가로, 세로 방향의 인장 실험 결과 그래프(파랑, 초록)와 등방성 가상 직물의 인장 실험 결과(빨강)

3.2. 가상 직물 형상 최적화 분석

카멜레온 서피스에서는 액츄에이터 말단작용체로 반구를 이용한다. 2.2절에 제안된 최적화 알고리즘을 테스트 하기 위해, 반구의 반지름이 1인 시물레이션 결과를 타겟 메시로 삼고 동일한 깊이 정보를 가지며 반지름이 0.5인 시물레이션을 이에 대해 최적화하였다. 그림 3은 타겟 메시와의 최대 차이를 기준으로 차이의 정도를 시각화 한 결과이며, 이 값이 0.002%이하이면 차이가 없는 것으로 간주하고 파란색으로 표시하였다.

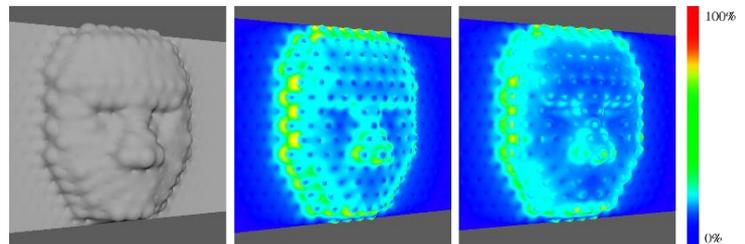


그림 3: 타겟 메시(좌)와 최적화 전(중간)과 후(우)

4. 결론

키네틱 아트 시스템인 카멜레온 서피스에서 액츄에이터에 부착된 직물 스킨은 동적 프로젝션 매핑을 효과적으로 하기 위해서 실제 타겟 메시와 유사하게 변형되어야 한다. 본 논문에서는 물리 기반 가상 직물 시물레이션에 실제 직물의 물성치를 반영하는 수학적 모델과, 직물 스킨 표면이 실제 타겟 메시와 유사하게 변형되도록 액츄에이터의 위치를 최적화하는 방법을 제시하였다.

참고문헌

- [1] M. Müller, B. Heidelberger, M. Hennix, and J. Ratcliff, Position based dynamics, *Proc. Virtual Reality Interactions and Physical Simulations (VRIPhys)*, Eurographics, 71-90, 2006.
- [2] M. Müller, J. Stam, D. James, and N. Thürey, Real Time Physics, *In ACM SIGGRAPH Courses*, 2008.
- [3] E. Sifakis, and J. Barbič, Fem simulation of 3d deformable solids: A practitioner's guide to theory, discretization and model reduction, *In ACM SIGGRAPH Courses*, 2012.

심층 신경망과 계층적 구조를 이용한 옷감 시뮬레이션의 가속화*

이태민⁰, 오영진, 이인권
연세대학교 컴퓨터과학과
{dnflxoals, skrcjstk, iklee}@yonsei.ac.kr

Acceleration of Cloth Simulation using Deep Neural Networks and Hierarchical Structure

Tae Min Lee⁰, Young Jin Oh, In-Kwon Lee
Dept. of Computer Science, Yonsei University

요약

빠르고 안정적인 물리 기반 시뮬레이션 기술은 컴퓨터 그래픽스 콘텐츠의 유연한 시각 효과를 보이는데 필수적인 기술이다. 이 논문에서 제안하는 심층 신경망을 이용한 계층적 옷감 시뮬레이션 방법은 기존의 물리 기반 시뮬레이션 계산방법과 심층 신경망(Deep Neural Networks)을 동시에 사용하는 방법으로 빠르고 안정적인 결과를 생성한다. 계층적 옷감 모델의 가장 하위 레벨의 시뮬레이션은 기존의 물리적 기반 시뮬레이션을 사용하여 계산하며, 상위의 더 세밀한 레벨의 다음 위치는 심층 신경망을 이용한다. 우리는 다양한 조건에서 실험을 진행하여 제안한 방법이 안정적이면서 빠른 옷감 시뮬레이션 결과를 생성할 수 있음을 증명하였다.

1. 서론

빠르고 안정적인 물리 기반 시뮬레이션 기술은 영화, 게임 등 컴퓨터 그래픽스 콘텐츠의 풍부하고 유연한 시각적 효과를 제공하는데 필수적인 기술이다. 특히 옷감을 위한 물리 기반 시뮬레이션 기술은 보다 빠르고 안정적인 시뮬레이션 방법을 위해 오랫동안 연구되어온 분야로 향상된 implicit Euler integration[1], 계층적 옷감 모델을 이용한 시뮬레이션[2] 등 다양한 사전 연구들이 존재한다.

최근 들어 물리 기반 시뮬레이션 기술연구에 기계학습 방법 중 하나인 심층 신경망(Deep Neural Networks) 기술을 이용한 연구들이 발표되고 있다. 유체와 연기 시뮬레이션의 일부 복잡한 계산과정을 심층 신경망 기술을 이용하여 대체하거나[3], 심층 신경망만을 이용하여 물체의 물리적인 움직임을 예측하는 연구[4]가 발표되어 심층 신경망을 이용한 물리 기반 시뮬레이션 기술이 안정적인 결과를 생성할 수 있음을 보였다.

사전연구는 주로 유체와 연기 시뮬레이션 분야에서만 수행되었으며 옷감 시뮬레이션 분야 역시 심층 신경망 기술을 이용한 연구가 요구되고 있다. 또한 심층 신경망

을 이용한 물리 기반 시뮬레이션 연구는 심층 신경망 추론에서 발생하는 오류가 만들어내는 불안정한 시뮬레이션을 보완할 수 있어야 하며, 기존의 물리 기반 시뮬레이션 방법보다 빠른 예측이 가능해야 한다.

본 논문에서 제안하는 심층 신경망을 이용한 계층적 옷감 시뮬레이션 방법은 물리 기반 시뮬레이션 계산방법과 심층 신경망을 동시에 사용하여 빠르고 안정적인 결과를 생성한다. 가장 하위 계층의 다음위치는 기존 물리 기반 시뮬레이션 기술로 계산하고, 나머지 상위 계층은 학습된 심층 신경망을 이용하여 추론한다. 가장 하위 레벨을 물리 기반 시뮬레이션 방법으로 계산하므로 거듭된 시뮬레이션에서도 안정적이며, 단순한 구조로 구성된 심층 신경망 모델은 기존 물리시뮬레이션 방법들 보다 시뮬레이션 결과를 빠르게 예측할 수 있게 한다. 우리는 실험을 통해 제안한 방법이 안정적이고 빠른 옷감 시뮬레이션 결과를 생성함을 확인하였다.

2. 심층 신경망과 계층적 구조를 이용한 옷감 시뮬레이션의 가속화

2.1. 계층적 옷감 모델 생성

그림 1은 계층적 옷감 모델의 구성 및 초기위치를 결정하는 과정이다. 옷감 내 계층 i 에 속한 한 삼각형 세 점(masses)의 초기위치를 $\hat{x}_p^i, \hat{x}_q^i, \hat{x}_r^i$ 라 할 때, 상위 계층 $i+1$ 은 세 점의 에지(Edge) 중점에 초기위치 $\hat{x}_p^{i+1}, \hat{x}_q^{i+1}, \hat{x}_r^{i+1}$ 의 세 점을 추가하여 구성한다. 이 과정은 계층적 옷감 모델 생성에 모든 계층 및 삼각형에 동일하게 반복적으로 수행되어 계층적 옷감 모델을 생성한다.

2.2. 계층적 옷감 모델의 시뮬레이션 계산

가장 하위 계층 l_0 의 시뮬레이션은 기존의 물리 기반 시뮬레이션 방법인 ADMM 방법을 적용한 the projective dynamics 방법[1]을 이용하여 계산한다. l_0 의 계산이 끝나면 l_0 의 정보를 이용하여 차 상위 계층들의 시뮬레이션은 심층 신경망으로 추론한다. 본 연구에서 사용한 심층 신경망 구조는 3개의 완전연결(Fully-connected) 레이어와 ReLu(Rectified Linear Unit) 활성화 함수로 구성된다(그림 2 참조). 심층 신경망의 입력 및 출력 특징 벡터는 9차원 크기의 벡터로 상위 계층 $i+1$ 에 위치한 세 점의 다음위치 $x_p^{i+1}, x_q^{i+1}, x_r^{i+1}$ 를 예측하기

* 구두발표논문

* 본 논문은 요약논문(Extended Abstract)으로서, 본 논문의 원본 논문은 Proceedings of Computer Graphics International 2018 에 게재 되었음.

* 본 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임.
(No. NRF-2017R1A2B4005469)

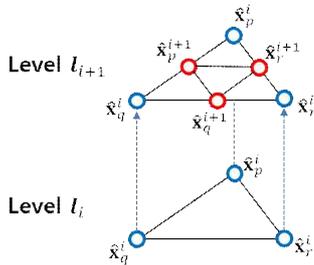


그림 1: 옷감 모델의 계층 생성

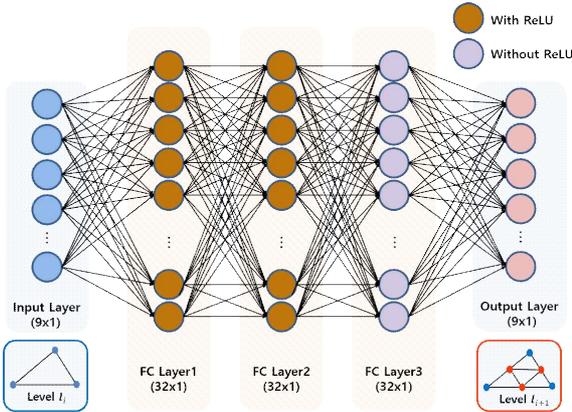


그림 2: 옷감 시뮬레이션 예측을 위한 심층 신경망 구조

위한 입력 벡터 δ_{pqr}^{i+1} 는

$$\delta_{pqr}^{i+1} = [(x_p^i - \hat{x}_p^i), (x_q^i - \hat{x}_q^i), (x_r^i - \hat{x}_r^i)], \quad (1)$$

대응되는 하위 계층 i 내 세 점의 초기 위치 $\hat{x}_p^i, \hat{x}_q^i, \hat{x}_r^i$ 와 현재 위치 x_p^i, x_q^i, x_r^i 의 차이로 정의된다(그림 3 참조).

심층 신경망의 출력 벡터 o_{pqr}^{i+1} 는

$$o_{pqr}^{i+1} = [(x_p^{i+1} - \hat{x}_p^{i+1}), (x_q^{i+1} - \hat{x}_q^{i+1}), (x_r^{i+1} - \hat{x}_r^{i+1})] \quad (2)$$

로 정의되며, 최종 상위 계층 $i+1$ 의 다음 위치는

$$[x_p^{i+1}, x_q^{i+1}, x_r^{i+1}] = [\hat{x}_p^{i+1}, \hat{x}_q^{i+1}, \hat{x}_r^{i+1}] + o_{pqr}^{i+1} \quad (3)$$

로 초기 위치와 출력 벡터의 값을 더한 값이다.

N 개의 상위 계층을 가진 옷감 모델은 N 개의 분리된 심층 신경망을 사용하여 추론하며, 추론의 참값(Ground truth)를 g 이라고 할 때 심층 신경망 학습과정에서 사용하는 Loss는

$$Loss = \sqrt{\frac{1}{n} \sum_{i=1}^N \sum_j (g_j^i - o_j^i)^2} \quad (4)$$

로 모든 상위 계층 내 모든 삼각형 j 에 대한 참값과 추론결과와의 차이의 총 합으로 계산한다.

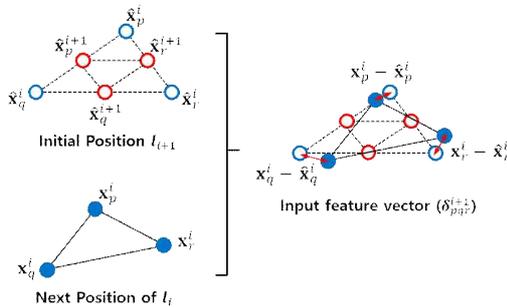


그림 3: 심층 신경망의 입력 특징 벡터

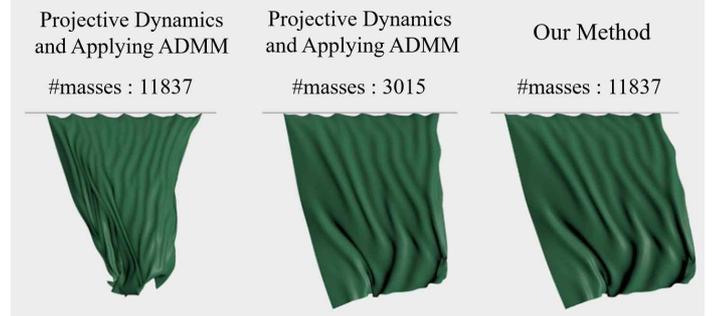


그림 4: 심층 신경망을 이용한 계층적 옷감 시뮬레이션(커튼)

모델	#masses	CG	ADMM	Ours
커튼	11837	416.2	367.1	132.7

표 1: 기존 물리 시뮬레이션 방법과 성능 비교(ms)

3. 결과

그림 4는 제안한 방법으로 생성한 시뮬레이션 결과를 기존의 물리 시뮬레이션 방법으로 생성한 결과와 비교한 결과를 보여준다. 왼쪽의 시뮬레이션 결과와 우리의 결과는 결과가 다르나 가운데의 계층적 옷감 모델의 l_0 를 기존의 방법으로 시뮬레이션 한 결과와는 유사하게 안정적으로 생성되었다. 표 1은 그림 4의 모델을 Conjugate Gradient 방법을 포함한 기존의 2가지 물리 기반 시뮬레이션 방법과 성능을 비교한 결과로 제안한 방법이 빠른 시뮬레이션이 가능한 것을 보여준다.

4. 결론

본 논문에서는 빠르고 신뢰할 수 있는 심층 신경망을 이용한 계층적 옷감 시뮬레이션 방법을 제안하였다. 기존의 물리 시뮬레이션 방법으로 계산한 결과와는 다른 시뮬레이션 결과를 생성하나 거듭되는 시뮬레이션에도 안정적인 결과를 생성하며 기존의 방법보다 빠르게 동작한다.

그러나 학습데이터에 포함되지 않은 상황에 대해서는 불안정한 결과를 보이며, 고주파 영역의 세부적인 움직임이 사라지는 단점을 가지고 있어 다른 신경망 구조 등을 이용한 추가적인 연구를 수행 중이다.

참고문헌

- [1] R. Narain, M. Overby, and G. E. Brown. ADMM \ni Projective Dynamics: Fast Simulation of General Constitutive Models. In Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation(SCA '16), pp. 21–28, Aire-la-Ville, Switzerland, 2016.
- [2] D. Zhang and M. M.F. Yuen. Cloth simulation using multilevel meshes. Computers & Graphics 25, 3 (2001), 383 – 389.
- [3] J. Tompson, K. Schlachter, P. Sprechmann, and K. Perlin. Accelerating Eulerian Fluid Simulation With Convolutional Networks. ArXiv 2016.
- [4] M. B Chang, T. Ullman, A. Torralba, and J. B Tenenbaum. 2016. A Compositional Object-Based Approach to Learning Physical Dynamics. ArXiv 2016.

자성체 간 상호작용을 위한 자기화량 역학*

김승욱⁰, 박선영, 한정현
고려대학교 컴퓨터학과
{wook0249, iamsun0720, jhan}@korea.ac.kr

Magnetization Dynamics for Magnetic Object Interactions

Seung-wook Kim⁰, Sun Young Park, JungHyun Han
Dept. of Computer Science and Engineering

요약

이 논문에선 물리적으로 올바른 방법으로 자성체 간의 상호작용을 시뮬레이션하는 기법을 다룬다. 논문에서 제안하는 자기 시뮬레이션은 자기 쌍극자 모멘트의 시간에 따른 변화를 나타내는 자기화량 역학을 기반으로 한다. 자기화량 역학을 다루기 위해 미시자기학에서 널리 사용되는 Landau-Lifshitz-Gilbert 방정식을 사용한다. 자기 쌍극자 모멘트를 사용하는 자석 모델은 거시적 범위로 확장되어 실시간 강체 시뮬레이션과 결합된다. 자기 시뮬레이션은 안정적이며 상호 자기 유도 및 잔류자를 표현할 수 있다.

1. 서론

컴퓨터 그래픽스 분야에서 자기 효과는 [1-3]에서 주로 다루어졌다. 그 중 [1]은 자성체의 움직임을 시뮬레이션하는 방법을 제안하였다. 하지만 이 연구는 자성체 간 상호 자기 유도와 잔류자를 고려하지 않았다는 한계를 가지고 있다. 또한 자석이 매우 센 자기장에 노출되었을 때, 자석의 자기화량이 매우 커져 자기력이 비물리적으로 과도하게 커지는 한계를 가지고 있다. 올바른 자석의 움직임을 예측하기 위해서는 앞서 언급한 문제점들을 해결해야 한다. 이 논문에서는 물리적으로 올바른 방법으로 자성체 간의 상호작용을 효과적으로 구현하기 위해 Landau-Lifshitz-Gilbert 방정식을 사용하여 자기화량을 시뮬레이션한다.

2. 자석 모델

2.1. 유효 자기장

한 점에서의 유효 자기장(effective magnetic field)인 \mathbf{H}_{eff} 는 외부장인 \mathbf{H}_E , 자기 소거장인 \mathbf{H}_D , 자기 결정 이방성 자기장인 \mathbf{H}_K 의 합이다. \mathbf{H}_E 는 지구 자기장과

* 구두발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 ACM Transaction on Graphics 37(4). 2018 에 게재 확정 되었음.

* 이 성과는 2016년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2016R1A2B3014319)

같이 외부로부터 작용하는 자기장이다. \mathbf{H}_D 는 모든 자성체 내 자기 쌍극자 모멘트로부터 발생하는 자기장이다. \mathbf{H}_K 는 자성물질의 자기 결정구조에 의해 정의되는 특정한 축인 용이축(easy axis)들에 의해 결정되며, 이는 자기화량이 용이축으로 정렬되도록 돕는다.

2.2. Landau-Lifshitz-Gilbert 방정식

한 점에서의 자기화량인 \mathbf{M} 의 시간에 따른 변화량은 아래와 같이 Landau-Lifshitz-Gilbert 방정식[4]에 의해 결정된다.

$$\frac{d\mathbf{M}}{dt} = -\gamma \left(\mathbf{M} \times \mathbf{H}_{\text{eff}} - \eta \left(\mathbf{M} \times \frac{d\mathbf{M}}{dt} \right) \right)$$

위 식의 γ 는 자기 회전 비율이며 η 는 자성물질의 감쇠 상수이다.

위 식의 첫 번째 항은 자기화량의 세차 운동을, 두 번째 항은 자기화량의 감쇠 운동을 나타낸다. 자기화량은 유효 자기장의 주변을 맴돌며 최종적으로 유효 자기장에 정렬되는 나선 모양의 움직임을 갖는다. 이 때, 자기화량의 크기는 고정된 채로 방향만 변한다.

2.3. 자석의 셀 분할

자석은 균일한 부피 V 를 갖는 구형 셀로 나뉘며, 유효 자기장이 가해지는 각 셀의 중심에는 두 개 이상의 정규화된 자기 쌍극자 모멘트인 \mathbf{m}_i 가 배치된다. 셀에 n 개의 정규화된 자기 쌍극자 모멘트가 배치되었다고 가정할 때, 이 셀의 자기화량은 아래와 같이 정의된다.

$$\mathbf{M} = \frac{1}{n} \sum_{i=1}^n \mathbf{m}_i$$

셀 내부의 자기화량이 균일하다는 가정 하에, 셀의 알짜 자기 쌍극자 모멘트 $\boldsymbol{\mu}$ 는 $\mathbf{M}V$ 로 정의된다.

셀의 알짜 자기화량에 의한 \mathbf{H}_D 를 고려하기 위해, \mathbf{H}_D 를 셀 내부 자기 쌍극자 모멘트에 의한 자기 소거장인 $\mathbf{H}_{\text{endo}_D}$ 와 셀 외부 자기 쌍극자 모멘트에 의한 자기 소거장인 $\mathbf{H}_{\text{exo}_D}$ 로 나뉜다. 구형 셀의 경계값 조건을 고려한 $\mathbf{H}_{\text{endo}_D}$ 와 $\mathbf{H}_{\text{exo}_D}$ 는 다음과 같다.

$$\mathbf{H}_{\text{exo}_D} = -\frac{1}{3} \mathbf{M}$$

$$\mathbf{H}_{\text{endo}_D} = \frac{1}{4\pi} \sum_i \left(\frac{3\mathbf{r}_i(\mathbf{r}_i \cdot \boldsymbol{\mu}_i)}{\|\mathbf{r}_i\|^5} - \frac{\boldsymbol{\mu}_i}{\|\mathbf{r}_i\|^3} \right)$$

여기서 $\boldsymbol{\mu}_i$ 는 i 번째 셀의 자기 쌍극자 모멘트이고, \mathbf{r}_i 는

i 번째 셀의 중심으로부터 해당 셀까지 연결하는 벡터다. 셀의 중심에 작용하는 다른 자기장과 달리, \mathbf{H}_K 는 셀 내부의 정규화된 자기 쌍극자 모멘트마다 다르게 작용한다. z 축이 용이축인 단축 비등방성 강자성체에 대해, \mathbf{m}_i 에 작용하는 자기 결정 이방성 자기장은 다음과 같이 정의된다.

$$\mathbf{H}_{K,i} = -2K_u \begin{pmatrix} m_{ix} \\ m_{iy} \\ 0 \end{pmatrix}$$

여기서 K_u 는 비등방상수이며 m_{ix} 와 m_{iy} 는 각각 \mathbf{m}_i 의 x 축 성분과 y 축 성분이다.

2.4. 시뮬레이션

정규화된 자기 쌍극자 모멘트들은 나노초 단위의 타임 스텝 사이즈를 사용하는 자기 시뮬레이션에 의해 갱신된다. 자기 시뮬레이션의 성공적인 종료를 보장하기 위해 자기 쌍극자 모멘트와 유효 자기장이 이루는 각도 차이를 오차로 사용하며, 이 오차가 일정 크기보다 작을 경우 자기 시뮬레이션을 종료한다.

자기 시뮬레이션에서 \mathbf{H}_E 와 $\mathbf{H}_{\text{ex0}_D}$ 의 합인 셀 외부로부터 가해지는 자기장 \mathbf{H}_A 에 따른 자기화량의 변화는 아래와 같은 자기화량 곡선으로 표현된다.

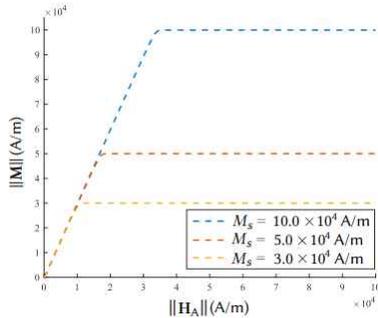


그림 1: 자기화량 곡선

\mathbf{H}_A 가 약할 경우 자기화량의 변화는 알려진 선형 자기화량 구성식과 일치하며, \mathbf{H}_A 가 셀 경우 자기화량이 포화 자기화량에 도달한다. 이는 실제 물리 실험으로부터 얻어진 [5]의 결과와 일치한다.

최종적으로 결정된 자기 쌍극자 모멘트로부터 자기력과 자기토크를 [1]에서 제안된 방법을 통해 계산한다.

3. 결과

아래 그림은 실제 자석 실험 장면과 논문에서 제안하는 기법의 결과를 보여준다. 쇠공의 자기화량은 막대자석뿐만 아니라 모든 자석에 의해 결정되기 때문에, 상호 자기 유도에 의해 주변 쇠공의 영향을 받아 충분한 자기화량을 갖는다. 따라서 그림과 같은 구슬 줄이 형성된다.



그림 2: 자화된 쇠구슬의 실제 실험과 시뮬레이션

아래 그림은 상호 자기 유도로 인한 자기장 차폐 현상을 보여준다. 그림의 왼쪽엔 막대자석과 쇠공 사이에 나무판을, 그림의 오른쪽엔 철판을 배치하였다. 나무판은 자기 효과를 발생시키지 않기 때문에 쇠공이 막대자석에 의해 이끌려 올라오지만, 철판은 자화되어 쇠공에 가해지는 자기장을 감쇠하기 때문에 쇠공이 끌려오지 못한다.

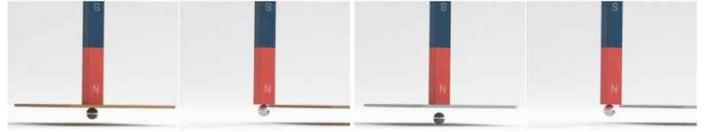


그림 3: 차폐 판 재질별 자기장 차폐 시뮬레이션

아래 그림은 잔류자기 시뮬레이션을 보여준다. 자화되지 않은 쇠못은 영구 자석인 막대자석에 의해 자화된다. 쇠못의 자기화량은 자기 결정 이방성에 의해 유지되며, 이로 인해 작은 쇠구슬들이 쇠못에 의한 자기 유도에 의해 자화되어 쇠못에 달라붙는다.

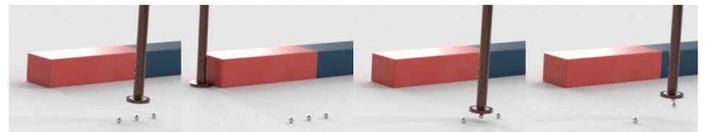


그림 4: 잔류자기 시뮬레이션

4. 결론

이 논문에선 자기화량 역학을 구현하는 새로운 방법을 제시한다. 자석을 균일한 크기의 셀로 분해하고, 셀 내부에 다중 자기 쌍극자 모멘트를 배치한다. 배치된 다중 자기 쌍극자 모멘트는 Landau-Lifshitz-Gilbert 방정식에 의해 시뮬레이션된다. 이러한 방법을 통해 유효 자기장에 노출된 각 셀의 자기화량을 조절할 뿐만 아니라 자기력과 자기토크의 과도한 증가를 방지하여 강체 시뮬레이션을 안정화시킨다. 또한 상호 자기 유도와 잔류 자기의 구현을 가능하도록 한다.

참고문헌

- [1] Thomaszewski, Bernhard, et al. "Magnets in motion." ACM Transactions on Graphics (TOG) 27.5 (2008): 162.
- [2] Park, Jiyoung, KyungOk Lee, and JungHyun Han. "Interactive visualization of magnetic field for virtual science experiments." Journal of Visualization 19.1 (2016): 129-139.
- [3] Yoon, Wonbae, et al. "Computer-generated iron filing art." The Visual Computer 30.6-8 (2014): 889-895.
- [4] Gilbert, T. L. "A Lagrangian formulation of the gyromagnetic equation of the magnetization field." Phys. Rev. 100 (1955): 1243.
- [5] Abbott, Jake J., et al. "Modeling magnetic torque and force for controlled manipulation of soft-magnetic bodies." IEEE Transactions on Robotics 23.6 (2007): 1247-1252.

컨볼루션 신경망을 이용한 2차원 연기 시뮬레이션의 보간*

오영진⁰, 김영준¹, 손진석², 김성호³, 이인권⁴
연세대학교 컴퓨터과학과

(⁰skrcjstk, ⁴iklee)@yonsei.ac.kr, ¹rct1236@naver.com, (²sjs2666, ³wastenotears)@gmail.com

Interpolation of 2D Smoke Simulation using Convolutional Neural Networks

Young Jin Oh⁰, Young Jun Kim¹, Jin Seok Son², Sung Ho Kim³, In-Kwon Lee⁴
Dept. of Computer Science, Yonsei University

요약

빠르고 정확한 계산이 요구되는 물리 기반 시뮬레이션 분야 중 특히 유체 시뮬레이션 분야는 대량의 계산비용을 낮추기 위해 다양한 연구들이 수행되었다. 본 논문에서는 큰 시간 간격으로 연기 시뮬레이션을 계산하고 그 사이의 보간 영상을 컨볼루션 신경망 모델로 예측하는 방법을 제안한다. 작은 시간 간격을 이용하여 매 프레임 시뮬레이션 방법으로 계산하는 것과 비교했을 때 유사한 시뮬레이션 결과를 보다 빠르게 생성할 수 있다.

1. 서론

물리 기반 시뮬레이션 기술을 보다 폭넓은 컴퓨터 그래픽스 콘텐츠 분야에 사용하기 위해서는 빠르고 정확한 방법들이 요구된다. 특히 대량의 계산비용이 필요한 유체 시뮬레이션은 이러한 요구사항을 만족하기 위해 다양한 연구들이 수행되었다[1, 2]. 최근 들어 기계학습 방법 중 하나인 신경망(Neural Networks) 기술을 이용한 연구들이 발표되고 있으며, 유체 시뮬레이션의 가장 계산비용이 높은 과정을 신경망 예측으로 대체하거나[3], 저해상도 시뮬레이션의 결과를 고해상도 시뮬레이션 결과로 생성하는 방법이 제안되어[4] 신경망 기술이 빠르고 안정적인 시뮬레이션 결과를 생성하는데 사용될 수 있음을 보였다.

본 논문에서 제안하는 컨볼루션 신경망을 이용한 연기 시뮬레이션의 보간은 빠른 연기 시뮬레이션 생성을 위해 물리 시뮬레이션 방법과 신경망(Neural Network) 방법을 동시에 사용하는 방법이다. 큰 시간 간격의 시뮬레이션을 기존의 방법을 이용하여 계산하고(키 프레임 시뮬레이션), 두 개의 키 프레임 시뮬레이션의 결과 사이에 작은 시간 간격으로 연속되는 보간 영상을 컨볼루션 신경망으로 예측하는 방법이다. 키 프레임 시뮬레이션 계산 및 신경망 모델을 이용하여 보간 영상을 예측하는 방법은 작은 시간 간격으로 매 프레임 시뮬레이션 계산하는 경우와 비교했을 때 유사한 결과를 보다 빠르게 생성하는 이점을 가지고 있다.

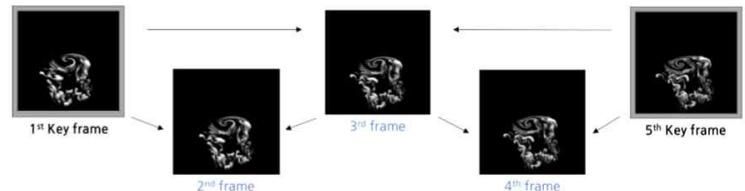


그림 1: 보간을 이용한 빠른 연기 시뮬레이션 생성

2. 보간 영상을 이용한 빠른 연기 시뮬레이션 생성

제안한 방법은 두 개의 키 프레임 시뮬레이션 사이에 3개의 보간 영상을 생성하여 연속된 5 프레임 시뮬레이션 영상을 생성한다. 기존의 연기 시뮬레이션 방법을 이용하여 큰 시간 간격을 가진 시뮬레이션을 계산하고(그림 1의 1, 5번 이미지), 인접한 두 키 프레임 시뮬레이션의 결과를 학습된 컨볼루션 신경망 모델에 입력하여 3번 보간 영상을 생성한다. 그 다음 두 개의 키 프레임 시뮬레이션 영상과 1번 보간 영상을 이용하여 추가로 2, 4번 보간 영상을 생성한다(그림 1 참조).

3. 보간 영상 예측을 위한 컨볼루션 신경망 모델

보간 영상 예측을 위해 사용된 컨볼루션 신경망 모델은 다수의 컨볼루션 신경망 레이어로 구성되어 있다. 두 영상 사이의 보간 이미지를 생성하는 연구인 U-Net[5]의 모델을 참고하였다. 입력 데이터는 보간 영상의 전·후 프레임 영상의 밀도 정보 이미지를 붙인 결과이다. 입력 데이터는 4단계의 컨볼루션, ReLU 활성화 함수, max-pooling의 과정을 거쳐 각각의 결과 영상을 만들고, 각 단계의 결과 데이터는 그 다음단계의 결과를 만드는데 함께 사용한다(그림 2 참조). 최종 예측 결과를 y , 참 값을 \hat{y} 라고 할 때, 신경망 모델 학습에 사용한 Loss 함수는

$$Loss = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

로 정의하며, n 개의 학습 예제에 대해 MSE(Mean Square Error)를 의미한다.

신경망 학습을 위한 학습 데이터는 연기 시뮬레이션 오픈소스 라이브러리인 MantaFlow[6]를 사용하여 생성하였다. 우리는 이 라이브러리를 이용하여 256×256 크기 2차원 격자구조를 이용한 연기 시뮬레이션을 대상으로 50개의 초기 조건을 다르게 만들어 결과를 계산하였다. 작은 시간 간격(0.5초)으로 각각 100 frame 생성하

* 포스터 발표논문

* 본 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임.
(No. NRF-2017R1A2B4005469)

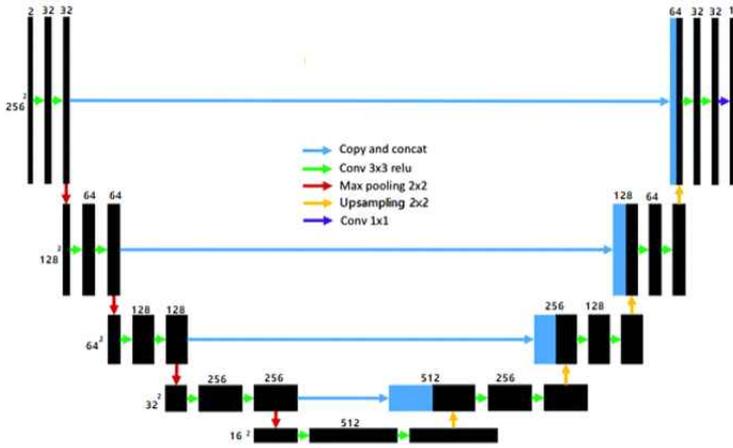


그림 2: 보간을 위한 컨볼루션 신경망 모델

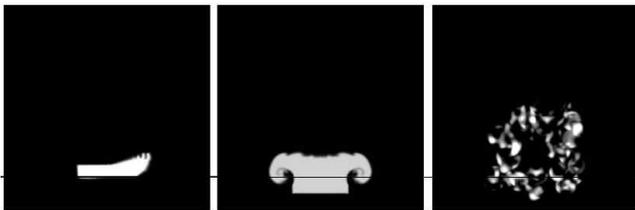


그림 3: 초기조건이 다른 학습 데이터의 예제

여 수집하였다. 그림 3은 학습 데이터 중 3가지의 연기 시뮬레이션의 초기 조건을 보여준다.

4. 결과

큰 시간 간격(2초)으로 계산되는 키 프레임 연기 시뮬레이션은 앞서 학습 데이터 생성에 사용한 MantaFlow를 이용하여 계산하며, 키 프레임 시뮬레이션 사이의 보간 영상은 제안한 모델로 예측하여 생성하였다. 그림 4의 왼쪽 열은 제안한 신경망 모델로 시뮬레이션을 예측한 결과이며, 오른쪽 열에 제시된 실제 시뮬레이션 계산결과와 비교했을 때 매우 유사한 형태의 연기 시뮬레이션을 보인다. 그러나 예측된 결과는 시뮬레이션 계산 결과에 비해 디테일이 사라져 다소 번져있는 결과를 보이기 때문에 추후 개선이 필요하다.

표 1은 기존 시뮬레이션 방법과 제안한 방법으로 동일한 길이의 시뮬레이션을 생성하는데 소요된 시간을 비교한 결과다. 키 프레임 계산 및 키 프레임간의 보간 영상을 예측하기 위해 소요된 시간은 기존의 방법을 이용하여 계산한 경우보다 빠른 것을 확인할 수 있다.

5. 결론

본 논문에서 우리는 컨볼루션 신경망을 이용한 연기 시뮬레이션의 보간 모델을 제안하고, 기존의 시뮬레이션의 키 프레임 계산 및 보간 영상을 예측하는 방법을 제안하였다. 우리의 신경망 모델은 연기의 전체 모양과 흐름은 시뮬레이션 계산한 결과와 비교했을 때 매우 유사하게 생성하여 빠르고 안정적인 시뮬레이션 결과를 만들어 냈다. 그러나 정확한 시뮬레이션 방법에 비해 다소 디테일이 감소된 결과를 생성하기 때문에 사라진 디테일을 복원하기 위한 추가적인 신경망 구조 연구가 필요하다. 또한 현재 2차원 시뮬레이션의 신경망 구조를

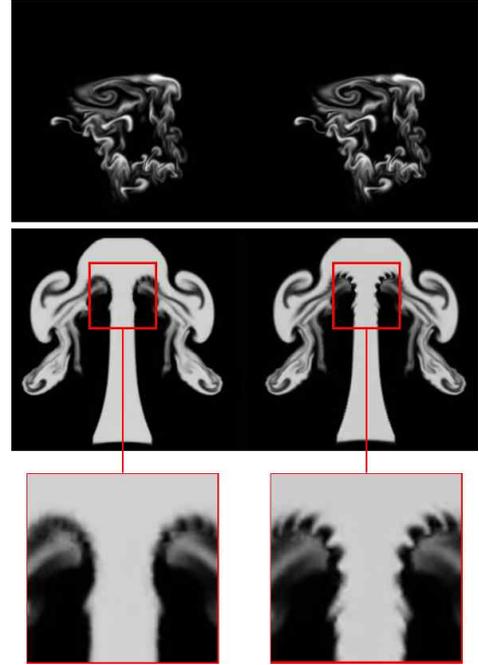


그림 4: 시뮬레이션 보간 결과(왼), 계산 결과(오)

#frames	기존 시뮬레이션	키 프레임 생성 (A)	보간 (B)	제안방법 (A+B)
101	33.432s	9.880s	8.937s	18.817s
201	73.462s	19.080s	9.432s	28.512s

표 1: 시뮬레이션 생성 시간 비교

3차원 연기 및 액체 시뮬레이션 분야로 확장하는 연구와 실험에서 임의로 설정한 3프레임 보다 다수의 프레임 보간하는 강인한 신경망 구조에 대한 연구를 계획하고 있다.

참고문헌

[1] M. Ihmsen, J. Cornelis, B. Solenthaler, C. Horvath, and M. Teschner. 2014. Implicit Incompressible SPH. IEEE Transactions on Visualization and Computer Graphics 20, 3 (March 2014), 426-435.

[2] N. Chentanez and M. Müller. 2011. Real-time Eulerian water simulation using a restricted tall cell grid. In ACM SIGGRAPH 2011 papers (SIGGRAPH '11).

[3] Tompson, J., Schlachter, K., Sprechmann, P., & Perlin, K. (2016). Accelerating eulerian fluid simulation with convolutional networks. arXiv preprint arXiv:1607.0359.

[4] Xie, Y., Franz, E., Chu, M., & Thurey, N. (2018). tempoGAN: A Temporally Coherent, Volumetric GAN for Super-resolution Fluid Flow. arXiv.

[5] R. Olaf, P. Fischer, and T. Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.

[6] N. Thurey and T. Pfaff, MantaFlow, 2017, <http://mantaflow.com>.

입 바람 기반 사용자 인터페이스를 이용한 유체 시뮬레이션의 상호작용 기술*

안준혁⁰, 김종현*
 강남대학교 소프트웨어응용학부
 ksjun4463@gmail.com, jonghyunkim@kangnam.ac.kr*

Interaction Technique in Fluid Simulations using Mouth-Wind Interface

Junhyeok An⁰, Jong-Hyun Kim*
 Dept. of Software Application, Kangnam University

요약

본 논문은 사용자의 입 바람을 이용하여 모바일 디바이스에서 동작하는 새로운 인터페이스 기술을 제안한다. 우리는 이 기술을 물리 기반 시뮬레이션에 통합시켜 상호작용 기술을 요구하는 산업에 다양한 형태의 솔루션으로 확장할 수 있는 기반기술을 개발한다[1]. 큰 동작을 요구하는 손, 발 등의 직접적인 상호작용 기술과 달리 상세하게 콘텐츠를 제어할 수 있는 상호작용 기술을 입 바람을 통해 개발한다. 실시간으로 상호작용되는 결과를 보여주기 위해 우리는 제안하는 모바일 상호작용 기술을 유체 시뮬레이션에 적용하여 입 바람에 의해 실시간으로 유체를 제어할 수 있는 결과를 보여준다.

1. 서론

최근에 안드로이드 모바일 앱과 가상현실 분야가 이슈가 되면서 사용자의 상호작용 기술은 날이 발전되고 있고, 다양한 상호작용 기술들이 요구되고 있다[2,3]. 다양한 기술과 장비들이 제안되었지만 대부분 손과 발의 움직임에 초점이 맞춰져 있다. 가상현실뿐만 아니라 상호작용의 중요성은 모바일 시장에서도 그 영향을 크게 미치고 있다. 대표적으로 사용자의 동작을 요구하는 콘텐츠는 스케치 툴이나 게임이 있으며, 이것들은 대부분 손가락의 움직임을 기반으로 실행되는 구조이다. 손을 이용한 인터페이스는 직관적이라는 장점이 있지만, 상세하게 콘텐츠를 컨트롤하기 어렵다는 단점이 있다. 본 연구에서는 사용자가 직접 바람을 불어 실시간으로 상호작용하는 새로운 인터페이스 프레임워크를 제안한다. 우리는 제안하는 기법을 모바일 디바이스에서 실험하지만 향후 가상현실 환경에서도 쉽게 확장이 가능한 방법이다.

* 포스터 발표논문
 * 학부생 주저자 논문임

2. 제안하는 프레임워크

제안하는 시스템의 알고리즘 개요는 사용자가 바라보고 있는 모바일 디바이스와 각도와 위치를 기반으로 아래와 같은 순서로 수행된다.

1. 모바일 디바이스를 바라보는 사용자 시점을 기반으로 입 바람의 방향을 결정
2. 모바일 디바이스의 사운드 센서를 이용하여 입 바람의 크기를 계산
3. 바람의 방향과 크기를 유체 방정식의 외력에 통합시켜 바람 시뮬레이션을 실시간으로 시각화

2.1. 입 바람을 통한 바람의 방향 결정

바람의 방향을 결정하기 위해 우선 모바일 디바이스와 사용자의 시점 간의 각도를 계산한다. 여기서 시점은 광선추적법과 같이 사용자가 모바일 디바이스를 바라보는 정확한 시점이 아니라, 각도를 기반으로 근사한 값을 이용한다.

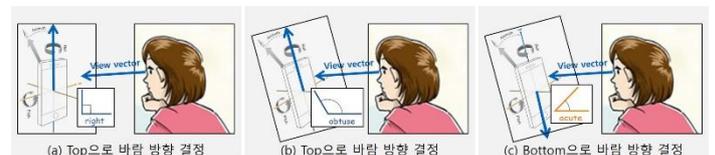


그림 1. 시점에 따른 바람의 방향 결정.

우리는 안드로이드의 가속도 센서를 이용하였으며, 이 센서 값은 3차원 공간상에서의 방향과 움직임에 대한 값을 알려준다. 가장 기본적인 상태는 사용자의 시점과 가속도 센서의 Y축 사이각이 직각인 상태이며, 이 상태에서 바람의 방향은 위쪽을 향한다(그림 1a 참조). 일반적으로 사용자가 모바일 디바이스를 잡을 때 직각 또는 둔각 타입으로 잡기 때문에 우리는 이 두 가지 유형에 대해서는 바람의 방향을 위쪽으로 설정하였다(그림 1b 참조). 모바일 디바이스를 예각으로 잡는 경우는 평상시에 흔치 않은 경우이며 이 경우는 바람의 방향 백

터를 아래쪽 방향으로 설정한다 (그림 1c 참조). 각도는 가속도 센서 값을 이용하여 계산하였으며 (그림 2 참조), 직각에서 둔각으로 각도가 커질 때마다 바람 방향에 가중치를 두어 더욱 강하게 바람이 불도록 방향 벡터의 크기를 증가시켰다. 모바일 디바이스가 움직이지 않는 정지 상태라면 각 축에 측정되는 가속도의 값은 $9.8m/s^2$ 이다. 따라서 디바이스가 정지상태라고 가정하고, 각 축에서 측정되는 가속도의 벡터 값을 이용하여 스마트폰의 기울기를 계산한다 (수식 1 참조).

$$\tan(\theta) = \frac{x}{z} = \arctan\left(\frac{x}{z}\right) \quad (1)$$



그림 2. 모바일 디바이스에서 회전 각도.

그림은 Y축을 기준으로 설명했지만, X축도 Y축과 마찬가지로 계산한다. 결과적으로 우리는 사용자와 모바일 디바이스 간의 상태를 직각, 둔각, 예각으로 정의하고, 이로부터 크게 벗어날 경우 방향을 반대로 전환해주는 규칙을 사용한다.

2.2. 사운드 센서로부터 입 바람의 크기 추출

우리는 바람의 크기를 계산하기 위해 모바일 디바이스에 내장되어 있는 사운드 센서를 이용한다. 디바이스에서는 저가용 사운드 센서를 사용하기 때문에 노이즈 제거를 위해 우리는 시간적 필터인 평균 값 필터를 이용하여 입력 사운드의 크기를 정제하여 사용한다. 본 연구에서는 시간 간격을 이전 5프레임으로 설정하여 평균 값 필터를 적용하였다 (그림 3 참조).

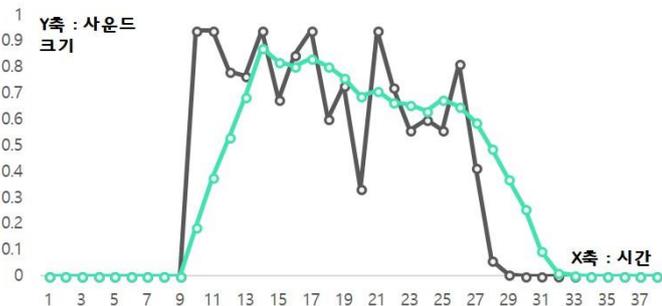


그림 3. 평균 값 필터를 이용한 정제된 사운드 크기(회색: 원본 사운드, 청록색: 정제된 사운드).

3. 결론

본 연구에서는 삼성 갤럭시 S6와 갤럭시 노트8에서 실험했으며 입 바람의 방향과 크기를 통해 모바일 디바

이스와 실시간으로 상호작용 되는 결과를 얻었다. 그림 4는 입 바람을 통해 벡터 장이 확산되는 결과이다. 디바이스를 직각으로 두고 불었기 때문에 부력과 같이 위쪽 방향으로 벡터 장이 흘러가는 모습을 보였다.



그림 4. 입 바람을 이용한 벡터 장 제어.

그림 5는 연기 시뮬레이션으로 확장한 결과이며 입 바람의 방향과 크기에 맞게 실시간으로 상호작용되는 결과를 잘 보여준다.

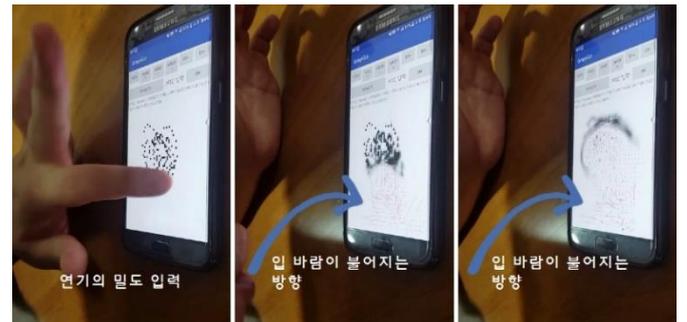


그림 5. 입 바람을 통한 연기 시뮬레이션 제어.

본 연구의 결과는 가상환경에서 큰 동작이 필요한 손과 발의 움직임과는 달리 상세한 상호작용이 필요한 곳에(게임, 예술과 관련된 애플리케이션 등) 활용될 수 있다. 뿐만 아니라 손과 발의 움직임이 불편한 사람들인 특정 계층(아동, 장애인, 노인 등)을 위한 상호작용 기술이 필요한 헬스케어, 미래 복지·기술 융합에도 활용될 수 있을 거라 생각한다.

참고문헌

[1] Jos Stam, Stable Fluids, ACM SIGGRAPH, *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, pp.121-128, 1999.
 [2] Z. Wenping, C. Jinxiang, X. Ying-Qing, Combining Marker-based Mocap and RGB-D Camera for Acquiring High-fidelity Hand Motion Data, *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp.33-42, 2012.
 [3] Cakmak, Tuncay, Hager, Holger, Cyberith Virtualizer: A Locomotion Device for Virtual Reality, *ACM SIGGRAPH*, pp.6-12, 2014.

컬링 스톤의 빙상 마찰력 분석을 통한 컬링 시뮬레이터 구현*

이동일⁰, 송오영¹
세종대학교

ehddlf47@naver.com, oysong@sejong.ac.kr

Curling Simulator based on Ice Friction Analysis of Curling Stone

Dong-Il Lee⁰, Oh-young Song¹
Sejong University

요약

동계 스포츠 중 컬링은 최근 국내에서 인기가 많은 종목 중 하나이다. 단순히 당구처럼 컬링 스톤을 던져 다른 스톤을 맞추는 것처럼 보이지만 컬링 스톤의 움직임을 모사하기 위해서는 마찰력을 포함한 복잡한 많은 물리 법칙이 적용되어야 한다. 본 연구에서는 컬링 스톤의 빙상 마찰력 분석을 통한 VPython 기반의 컬링 시뮬레이터를 소개한다.

1. 서론

컬링(Curling)은 직사각형의 긴 빙판 위에 컬링 스톤(Curling Stone)이라 불리는 원기둥 모양의 돌을 미끄러뜨려 하우스라 불리는 4개의 동심원 안쪽으로 스톤을 위치시키는 스포츠이다. 컬링은 두 팀이 각 8개씩의 스톤을 투구하고 마지막에 위치한 스톤들로 점수를 계산한다. '컬링 시트(Curling Sheet)'라 불리는 경기장은 '페블(Pebble)'이라 불리는 작은 얼음알갱이를 포함하고 있는 얼음 바닥이기 때문에 스톤의 움직임을 모사하는 것은 쉽지 않다. 본 연구에서는 컬링 스톤이 움직이는 원리를 분석하여 컬링을 위한 시뮬레이터를 구현하는 것을 보여준다.

2. 컬링 시뮬레이터 구현

물리 관련 시뮬레이션 결과를 컴퓨터 그래픽으로 표현해주는 도구는 많이 존재한다. 본 연구에서는 컬링 시뮬레이터의 결과를 표현하기 위해 VPython[1]을 이용하였다. VPython은 Python 기반의 물리 시뮬레이션을 지원하는 모듈이며, 3D 애니메이션과 그래프, 각종 출력값을 확인하는 것도 쉽다.

VPython의 box(), cylinder()를 이용하여 실제 컬링 시

트 규격과 같은 컬링 시트를 만들었고, 구현한 컬링시트는 그림 1과 같다. 컬링 시트는 9000개의 구역으로 나누었으며, 각 구역마다 조금씩 마찰 계수를 다르게 적용하였다.

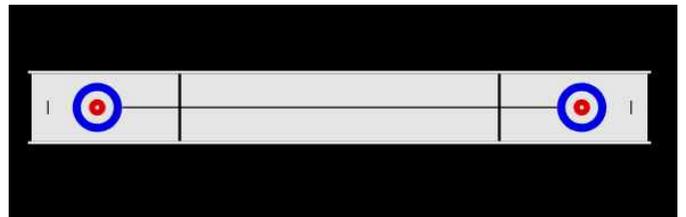


그림 1. VPython으로 구현한 컬링 시트

컬링 스톤은 컬링 시트와 마찬가지로 cylinder()와 box()를 이용하여 만들었으며, 실제 컬링 스톤의 규격인 질량 19.96kg, 원둘레 36인치(914.4mm), 높이 4.5인치(114.3mm)로 실제 컬링 스톤과 동일한 규격을 가진 스톤을 만들어 사용하였다.

3. 컬링 스톤의 움직임

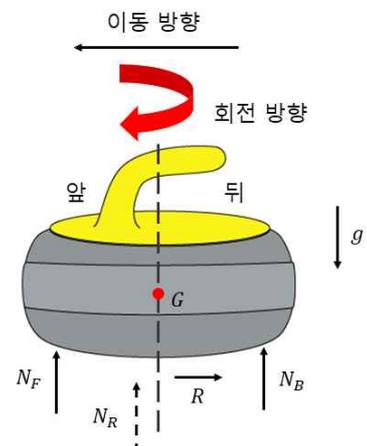


그림 2. 컬링의 회전 원리(정면)

컬링 선수들은 스톤을 투구할 때 스톤을 약간 돌리기 때문에 스톤이 원하는 방향으로 회전하여 이동하게 할

* 포스터 발표논문

* ¹교신저자

* 본 연구는 미래창조과학부 및 정보통신기술진흥센터의 SW중심대학의 지원사업의 연구결과로 수행되었음(2015-0-00938).

* 본 연구는 신진연구지원사업의 연구결과로 수행되었음(NRF-2015R1C1A1A02036501).

수 있다. 그림 2에서 N_R 은 켈링 스톤 바닥 전체에 가해지는 수직항력으로 켈링 스톤의 진행 방향에서 앞쪽에 가해지는 수직항력 N_F 와 켈링 스톤의 진행 방향에서 뒤쪽에 가해지는 수직항력 N_B 의 합으로 나타낼 수 있다($N_R = N_F + N_B$). G는 스톤의 중심, R은 스톤 전체에 작용하는 마찰력의 결과 성분으로, 스톤의 진행 방향과 반대로 작용한다. 이 마찰력으로 인해 스톤은 속도가 느려지게 된다. 켈링 스톤은 페이지의 평면에서 회전 평형 상태에 있기 때문에 이 평면에서 G에 대한 모멘트의 합은 0이 되어야 한다. 이는 N_R 이 R에 의해 야기된 힘과 반대 방향으로 힘을 가해야 한다는 것을 의미한다. 이는 N_R 이 작용하는 위치가 G보다 앞쪽에 있다는 것을 의미한다($N_F > N_B$). 물체가 얼음과 접촉할 때, 접촉 압력이 얼음 표면을 녹이면서 수막을 형성하고 이 수막이 물체와 얼음사이의 마찰력을 낮춘다. 이는 마찰력이 얼음과의 접촉 압력의 양에 반비례 하는 결과이다. 회전하는 스톤의 경우 $N_F > N_B$ 이므로, 스톤 이동방향의 앞쪽에 가해지는 마찰력이 뒤쪽에 가해지는 마찰력보다 작게 된다.

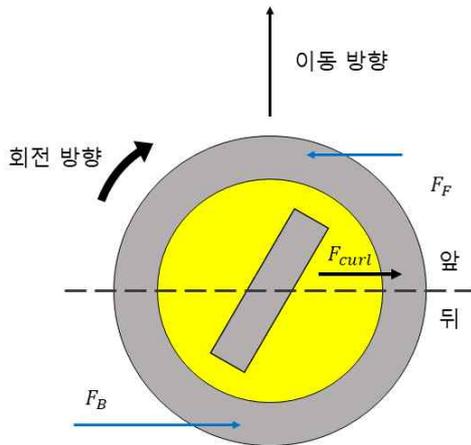


그림 3. 켈링의 회전 원리-2(단면)

그림 3에서 F_F 는 켈링 스톤의 앞쪽 절반에 작용하는 마찰력의 결과이고, 회전 방향과 반대 방향이다. F_B 는 켈링 스톤의 뒤쪽 절반에 작용하는 마찰력의 결과이고, 회전 방향과 반대 방향이다. 마찰력과 회전의 결과로 생기는 힘 \vec{F}_{curl} 은 아래의 식과 같다.

$$\vec{F}_{curl} = k(\mu_F N_F - \mu_B N_B)(\hat{v} \times \vec{w}) \quad (1)$$

$$F_F = \mu_F N_F, \quad F_B = \mu_B N_B$$

k 는 비례상수이며, μ_F 와 μ_B 는 켈링 스톤의 앞쪽과 뒤쪽의 마찰계수이다. \hat{v} 는 스톤의 이동방향이고, \vec{w} 는 각 속도이다. 그림 3에서는 $F_F < F_B$ 이고, 식 (1)의 결과로 켈링 스톤은 F_B 의 방향으로 휘어지게 된다.

4. 켈링 경기 시뮬레이션

본 연구에서는 켈링 스톤의 움직임의 원리를 분석하여 시뮬레이션 하였다. 켈링 시트를 그리드(Grid) 형태로 9000개로 나누어 0.05 ~ 0.06 사이의 값으로 마찰 계수를 설정하였으며, 켈링 스톤은 이동하면서 계속 다른 마찰 계수를 가지는 구역을 지나치게 된다. 그 결과로 켈링 스톤의 속도는 마찰력에 따라 다르게 변화되는 것을 확인 할 수 있었다.

켈링 경기를 시뮬레이션 하기 위해 실제와 같은 규칙을 적용하였다. 규칙은 각 8개의 스톤을 두 팀이 번갈아가면서 투구하는 방식이고, 점수 계산도 실제와 같은 점수 집계 방식을 적용하였다. 1팀은 분홍색 켈링 스톤을 사용하였고, 2팀은 하늘색 켈링 스톤을 사용하였다.

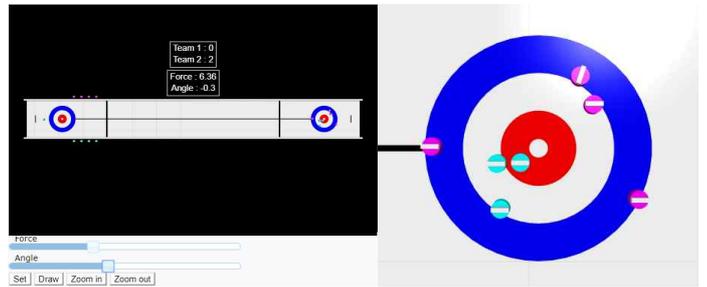


그림 4. 켈링 시뮬레이터(왼: 전체, 오: 확대)

5. 결론

본 연구에서는 켈링 스톤의 빙상 마찰력 분석을 통한 켈링 시뮬레이터를 구현하였다. 켈링 시트와 켈링 스톤은 경기를 진행할수록 스크래치가 생긴다. 이로 인해 예상치 못한 방향으로 스톤이 움직이는 경우도 있다. 추후 연구에서는 경기 도중에 생길 수 있는 스크래치와 그에 따른 회전과 이동 방향의 변화를 모사할 것이다[3]. 그리고 페블의 분포와 ‘켈링 브룸(Curling Broom)’을 이용한 ‘스융(Sweep, 페블을 닦아냄으로써 온도가 올라가고 그 결과 시트의 마찰력을 낮춰줌)’을 고려한 정교한 모델을 적용하고, 실제 켈링 경기에서의 스톤의 궤적과 비교하여 정확성을 확인할 예정이다.

참고문헌

[1] R.CHABAY, B.SHERWOOD, Computational physics in the introductory calculus-based course. *American Journal of Physics*, vol.76, Nos. 4 & 5, 2008.
 [2] F. Normani, REAL WORLD PHYSICS PROBLEMS. The Physics of Curling, Retrieved June 8, 2018, from <https://www.real-world-physics-problems.com/physics-of-curling.html>
 [3] H.Nyberg, S.Alfredson, The asymmetrical friction mechanism that puts the curl in the curling stone, *Wear of Materials* 2013, Volume 301, Issues1-2, pp.583-589, April-May 2013,

제스처 기반의 부동산 거래 VR 보드 게임 *

박고은⁰, 이수아, 김성실, 문주혜, 한다성
 handong대학교

21500250@handong.edu, 21700525@handong.edu, sskim5272@gmail.com,
 answngp1@naver.com, dshan@handong.edu

A Gesture-based VR Board Game for Real Estate Transaction

Goeun Park⁰, Sua Lee, Seongsil Kim, Juhye Mun, Daseong Han
 Handong Global University

요약

최근 게임 패드와 같은 별도의 컨트롤러 없이 모션센서 기술을 통해 직접 몸을 움직여 조작하는 체감형 게임에 대한 시장의 수요가 증가하고 있다[1]. 이와 관련하여 본 논문에서는 다양한 연령층에게 다가갈 수 있는 부동산 거래 보드게임에 모션센서와 가상현실기술을 접목시켜 몰입도를 높인 체감형 부동산 거래 보드게임을 제안한다.

1. 서론

부동산 거래 보드게임은 플레이어 각자가 자본을 가지고 시작하고, 땅을 사고 팔면서 자본을 늘려가는 게임이다. 대표적인 게임은 Parker Brothers사의 'Monopoly'로 출시 이후 오랜 시간 꾸준히 대중의 관심을 받아왔다. 그러나 인터넷 기반 게임시장의 확장으로 인해 점차 수요가 줄어들게 되자, 디지털화된 보드게임이 등장하면서 그 범위를 확장하기 시작했다[2]. 한국에서는 지난 2013년 모바일 버전 부동산 거래 보드게임인 '모두의 마블'이 출시되어 흥행에 성공했다. 2014년 Berserk Games에서 출시한 'Tabletop Simulator'는 온라인상에서 여러 명이 보드게임을 즐길 수 있는 시뮬레이션 비디오 게임으로, VR버전의 Monopoly 게임도 지원하고 있다. 이 게임은 가상 공간에서 간편하게 보드게임을 즐길 수 있다는 장점이 있지만, 컨트롤러로 주사위, 말, 카드 등을 일일이 포인팅 해야 해서 조작이 쉽지 않다. 이러한 까다로운 정밀성을 극복하기 위해 본 논문에서는 제스처를 이용해 손쉽게 조작하고, 사용자가 직접 말이 되어 보드 위를 걸어가는 듯한 시점을 보여주는 인터랙티브한 부동산 거래 보드게임을 제안한다.

2. 시스템 개요

본 논문에서는 먼저 사용자의 손목과 어깨에 마커를 부착하고, 모션 캡처 시스템을 통해 그 마커의 움직임을

추적한다. 추적된 마커 데이터는 현재 게임 상태에 따라 분석되어 그 상태에 맞는 제스처로서 인식된다. 이 과정에서 인식된 제스처 정보는 보드 게임 시뮬레이션 컴포넌트로 이동되어 게임 상태를 업데이트 하는데 사용된다. 게임의 시점은 주사위를 던질 차례인 플레이어의 일인칭 시점으로, 사용자가 말판 위에 서서 직접 이동하는 것처럼 보이도록 설정한다. 설정된 화면은 게임에 대한 플레이어의 몰입도를 높이기 위해 대형 곡면 스크린으로 출력된다. [그림 1]은 시스템 개요의 도식화이다.

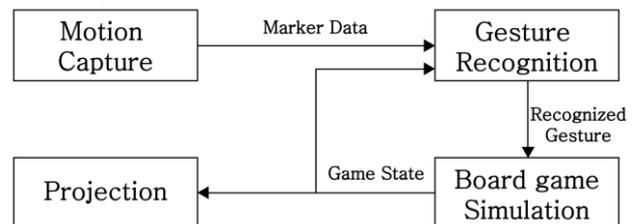


그림 1: 시스템 개요.

3. 모션 캡처

본 논문에서는 간편성과 빠른 응답속도를 갖는 체험형 게임을 제작하기 위해, 많은 마커를 부착하여 전신 동작을 추적하기 보다는 신체 일부의 동작을 추적하여 다른 부분의 동작을 추정하는 방식을 취한다 [3]. 이를 위해, 제스처 인식의 범위를 사용자의 한쪽 팔로 한정하고, [그림 2]와 같이 여러 개의 마커를 손목과 어깨 부위에 부착하여 손목과 어깨 부위가 각각 별도의 강체(rigid body)로 인식되도록 한다.



그림 2: (좌)마커를 손목과 어깨 부위에 부착한 모습, (우)마커

* 포스터 발표논문

* 학부생 주저자 논문임

4. 제스처 인식

4.1. 주사위 던지기 제스처

본 논문에서는 손목과 어깨 부위에 대한 강제 정보가 주어지면, 이 정보를 분석하여 손을 뒤에서 앞으로 흔드는 팔 제스처를 주사위를 던지는 제스처로 인식한다. 이를 위해 어깨 부위의 방위를 이용하여, 사용자의 전면 쪽을 향하면서도 지면과 수평인 기준 벡터(X)를 계산하고 그 벡터와 어깨에서 손목까지의 변위벡터 사이의 각도 θ 를 계산한다 [그림 3]. 그 각도 값을 검사하여 사용자가 충분히 팔을 뒤로 젖힌 경우 ($\theta > \theta^1$) 주사위 던지기 제스처를 시작한 것으로 인식하고 팔을 다시 충분히 내밀 때는($\theta < \theta^2$) 제스처가 끝난 것으로 간주한다.

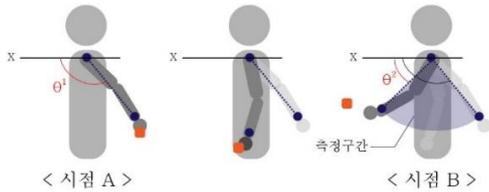


그림 3: 주사위 던지기 제스처 인식: 각도 θ^1 와 θ^2 일 때를 각각 제스처의 시작과 끝으로 간주함.

4.2. 선택하기 제스처

주사위를 던져 이동한 후 사용자는 도착한 칸과 건물의 매매에 관해 선택하는 과정을 거쳐야 한다. 본 논문에서는 사용자가 상하좌우 중 어느 방향으로 손을 짓는지를 인식해 사용자의 의사를 결정한다. 이를 위해 센서의 속도가 특정 값을 넘었을 때의 손목 강제 위치(b)와 어깨 강체를 기준으로 설정한 기준점(a) 위치를 비교해 오차 값이 가장 큰 방향을 선택한다. 예를 들어 [그림 4]와 같이 사용자가 왼쪽 방향으로 손을 저었을 때 수평방향의 오차 값(c)보다 수직 방향의 오차 값(d)가 작으므로 수평방향을 선택하고, - 방향이기 때문에 왼쪽으로 간주하여 사용자의 선택은 '이전 건물'이 된다.

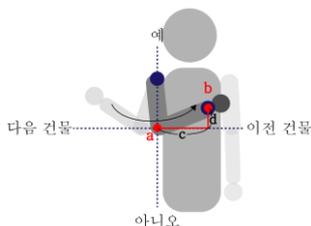


그림 4: 사용자 의사 결정 제스처 인식.

5. 보드 게임 시뮬레이션

주사위 던지기가 진행되는 동안에는 사용자가 실제로 주사위를 던지는 느낌을 주기 위해 주사위가 손목의 위치와 방위 정보를 반영하여 움직이도록 하고 제스처가 끝날 때는 그 순간의 손목의 이동 속도를 주사위에 반

영하여 주사위가 보드 위로 던져 지도록 한다. 물리 기반 시뮬레이션을 통해 던져진 주사위 움직임이 생성된 후, 시점을 사용자의 말로 옮겨서 [그림 5] 사용자가 실제 이동하는 듯한 느낌을 준다.

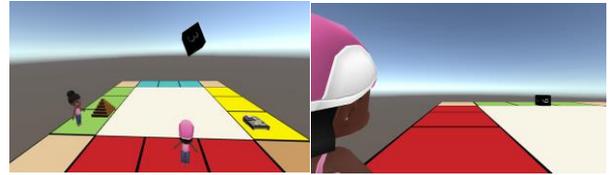


그림 5: 게임 화면.

토지와 건물의 매매 과정에서는 사용자의 의사를 물어보는 메시지 화면을 보여주고, 사용자의 선택에 따라 적절한 피드백을 제공한다.

6. 실험 결과

본 논문에서 제안한 VR 보드 게임은 Unity3D 게임 엔진을 기반으로 구현되었고 몰입감을 높이기 위해 게임 상태를 대형 곡면 스크린에 프로젝션하여 VR 공간을 구축하였다. 보드판, 건물, 말, 주사위 등 보드 게임에 필요한 소도구들은 직접 모델링하거나 Unity Asset Store의 모델들을 활용하였다. 어깨 부위와 손목에 부착한 마커의 개수는 각각 4개씩이며, 마커 인식을 위해 총 12개의 광학식 카메라를 사용했다.

7. 결론

본 논문에서는 모션센서 기술을 이용해 사용자의 모션을 인식하고 인식한 모션 데이터를 주사위를 던지거나 옵션을 선택하는 두 가지 상황에 따라 분석해 적절한 피드백을 제공하였다. 그리고 이를 보드게임 프로그램에 적용해 제스처로 조작되는 새로운 형식의 보드게임을 제작하였다. 향후 말판의 국가별로 건물과 배경에 특색을 추가하거나 다채로운 애니메이션을 적용하는 식으로 발전시켜 나간다면 기존 보드게임 시장에서는 볼 수 없었던 체험형 부동산 거래 보드게임을 개발할 수 있을 것으로 보인다. 또한 기존 VR 보드게임과의 비교를 통해 제안된 시스템의 효용성을 분석하는 것도 흥미로운 향후 연구가 될 수 있을 것이다.

참고문헌

- [1] 이대현, 모션센서 기반 체험형 액서게임의 동향 및 개발 사례, *한국콘텐츠학회지*, 5(2), 95-101, 2007.
- [2] 한은, 장준영, 손주영, 이동운, 조혜련, 조준동, 마라카스 : 신체적 능력 증진을 위한 활동 기반 보드게임 개발, *한국HCI학회 학술대회*, 965-968, 2018.
- [3] 구교호, 이승룡, 체험형 게임을 위한 복합센서 및 인체 움직임구조 모델 기반 실시간 모션캡처, *한국통신학회 학술대회논문집*, 1085-1086, 2010

홀로렌즈를 이용한 1인칭 슈팅 게임 구현*

송인석⁰, 전우형¹, 송창근¹, 이정¹, 김선정¹

한림대학교 컴퓨터공학과⁰, 융합소프트웨어학과¹

{20145134, M18058, cgsong, airjung, sunkim}@hallym.ac.kr

Implementation of a FPS Game Using HoloLens

InSeok Song⁰, WooHyung June¹, Jung Lee¹, Chang-Geun Song¹, Sun-Jeong Kim¹

Dept. of Computer Engineering⁰ and Dept. of Convergence Software¹, Hallym University

요약

마이크로소프트사에서 출시한 홀로렌즈(HoloLens)는 현실 세계에 3D 객체를 증강시켜 실제로 해당 객체가 존재하는 것과 같은 시각적 효과를 주는 혼합현실(Mixed Reality) HMD(Head-Mounted Display)이다. 본 논문은 홀로렌즈를 이용하여 현실 공간에 몬스터를 증강시켜 사방에서 몰려오는 몬스터를 무찌르는 1인칭 슈팅 게임을 개발한 내용을 포함한다.

1. 서론

혼합현실 기술이란, 현실의 정보 위에 가상의 오브젝트를 증강. 현실 공간과 가상의 공간을 혼합하는 기술이다. 본 논문은 대표적인 혼합현실 장비, 홀로렌즈를 사용하여 현실 공간에서 몬스터를 증강시켜 실제로 몬스터가 나온 것 같은 시각적 체험을 주어 몰입감 있는 1인칭 슈팅 게임을 개발한다 (그림 1).



그림 1. 게임속 몬스터가 증강된 화면

2. 본론

2.1. 게임 제작 및 내용

본 게임은 Windows 10 환경에서 Unity3D와 Windows Mixed Reality SDK를 이용하여 제작되었다. 게임 시작 시 플레이어가 인지할 수 있는 범위 내에서 충분한 거리를 가지고 몬스터가 무작위로 생성된다 (그림 2). 몬스터는 생성된 직후 플레이어를 향해 일정 속도로 다가온다.

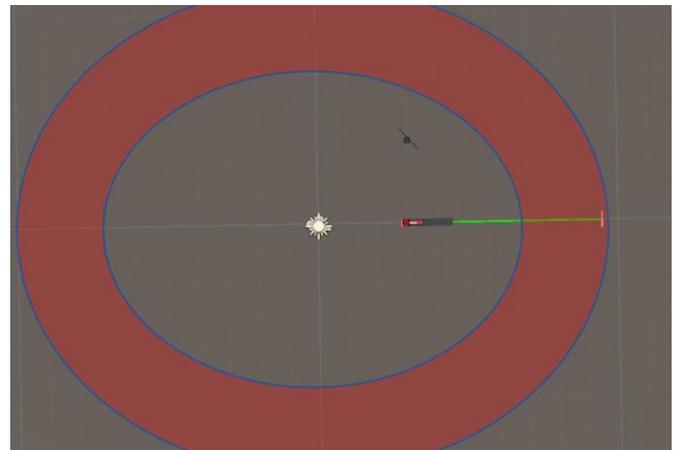


그림 2. 플레이어를 기준으로 몬스터가 생성되는 영역

플레이어는 몬스터를 머리의 움직임으로 총을 조준하고 'Tap gesture'를 취할 시, 몬스터를 맞춰 파괴함으로써 몬스터가 가까이 오지 못하도록 해야 한다. 몬스터 파괴 시 점수가 누적되고 몬스터와 접촉 시 게임이 종료된다.

2.2 콘텐츠 제어

게임의 입력 방법은 HMD를 이용한 머리 움직임 기반 조준과 Tap 제스처를 이용한 발사로 나뉜다.

2.2.1 HMD

HMD는 사용자 머리의 움직임을 추적하여 가상의 오브젝트를 이동시키고 화면에 증강시킬 수 있게 한다.

* 학부생 주저자 논문임

* 포스터 발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 제12회 한국정보과학회/한국빅데이터학회 공동 학술 심포지움에 발표되었음.

* 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2017R1D1A1B 03035576).

(그림 3) 이를 통해 가상의 오브젝트가 마치 현실의 공간에 있는 느낌을 주어 몰입도를 증가시킨다.

본 논문에서는 이러한 HMD 기술을 이용하여 슈팅을 위한 총을 사용자의 머리의 정면에 고정함으로써 다가오는 몬스터를 사용자의 머리 움직임을 이용하여 조준할 수 있도록 구성하였다. 머리를 움직이면 고정된 총과 총의 조준점이 같이 움직여 언제나 시선의 정면에 자리 잡고 있어 사방에서 다가오는 몬스터를 조준할 수 있도록 구현하였다.

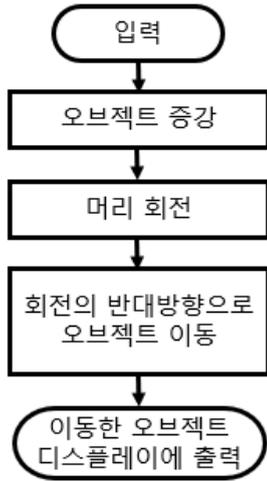


그림 3. HMD 동작 순서도

2. 2. 2 제스처

홀로렌즈는 마우스나 키보드와 같은 입력기기를 사용하지 않고 자제적인 컨트롤 방법을 가지고 있다. 본 논문에서 사용하는 방법은 손가락의 움직임을 추적하는 제스처 인식이다. 사용자의 머리를 움직여 조준점을 몬스터에게 맞추고 난 뒤 사용자의 손을 홀로렌즈의 카메라에서 보일 수 있도록 들어올린 뒤 ‘Tap’ 제스처를 취함으로써 조준하고 있는 총을 발사한다. 또한 ‘Bloom’ 제스처를 통하여 윈도우 메뉴를 불러 응용프로그램을 강제종료 할 수 있다.

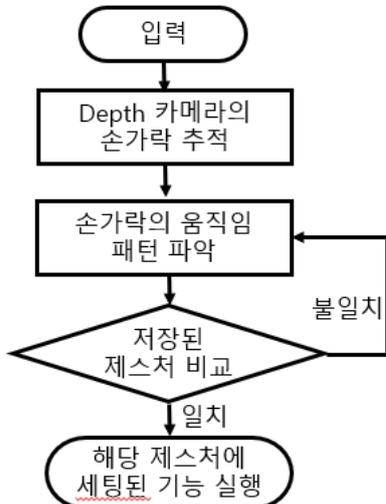
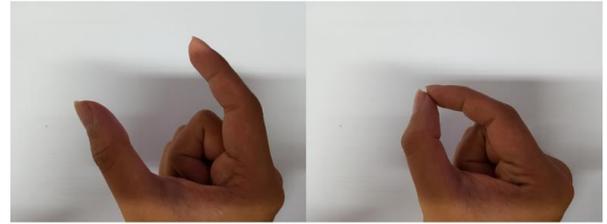


그림 4. 제스처 동작 순서도



(a) Tap



(b) Bloom

그림 5. 홀로렌즈에서 인식 가능한 제스처의 예

3. 결론

본 논문에서는 혼합현실 기기인 홀로렌즈를 이용하여 기존의 모니터를 사용하는 1인칭 슈팅 게임과는 달리 현실 공간에서 사용자가 가상의 증강물체와의 상호작용을 체험할 수 있는 1인칭 슈팅 게임을 제작하였다. 사용자는 몬스터에게 실제로 습격 당하는 듯한 시각적인 경험을 하게 되고 기존의 화면으로만 볼 수 있던 1인칭 시점을 실제 자신에게 대응시켜 더욱 현장감 있는 체험을 할 수 있도록 하였다.

향후 연구로는 게임에 주변 지형을 맵핑하여 플레이어에게 다가오는 몬스터에게 좀 더 현실감을 부여하고 실제로 총을 조작하는 듯한 경험을 줄 수 있는 방법을 고안할 예정이다.

참고문헌

[1] Beomjun Son, Eunjung Lee, Hongkyu Jung, Hyuna Lee, Changhoon Park. A Development of Augmented Reality Simulation Game Using HoloLens, Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology Vol.7, No.11, November (2017), pp. 897-904

[2] Jea-Ryun Chung, Jong-Hoon Song, Jeong-Ho Lee, Seung-Hyun Yoon. Human Anatomy Education App for HoloLens, 한국컴퓨터그래픽스학회 2017년 KCGS 학술대회 논문집, 2017.7, 139-140

[3] Xiao, R., Schwarz, J., Throm, N., Wilson, A. and Benko, H. 2018. MRTouch: Adding Touch Input to Head-Mounted Mixed Reality. In IEEE Transactions on Visualization and Computer Graphics (TVCG) Special Issue, Volume 24, Number 4, 1653-1660.

Kinect와 Scratch를 이용한 미디어 아트 콘텐츠 제작*

여주은, 박현주, 정지민, 문주혜, 서충모, 김동주^o, 이재호, 이원섭
 한동대학교 ICT창업학부

(21500854, 21400344, 21600656, 21500243, 21500852, 21600056, 21200578, w.lee)@handong.edu

Development of Media Art Contents using Kinect and Scratch

Jueun Yeo, Hyunjoo Park, Jimin Jeong, Juhye Mun, Chungmo Seo, Dongju Kim, Jaeho Lee, Wonsup Lee
 School of Global Entrepreneurship and Information Communication Technology, Handong Global University

요약

인체의 자세와 동작을 측정하는 기기인 Kinect 센서는 컴퓨터 그래픽 연구, 제품개발, 그리고 미디어 아트와 같은 예술 등 다양한 분야에 활용되고 있다. 본 연구는 Scratch를 활용하여 Kinect 센서로 측정되는 인체의 동작을 이용하는 3가지의 서로 다른 미디어아트 콘텐츠를 제작하는 연구를 수행하였다.

1. 서론

Kinect는 사람의 형상과 동작을 인식하여 관절 부위를 추정하고, 관절을 기준으로 하는 인체의 움직임을 측정한다. Kinect를 이용하는 다양한 측정 및 응용 기술들이 개발되었고 다양한 연구 사례들이 제시되었다. 또한 Kinect를 쉽게 사용할 수 있도록 하는 open source 응용 프로그램 및 library들도 사용자들에 의해 많이 제공되고 있어 장비를 사용하기 위한 진입 장벽이 낮다. 한편, Scratch는 MIT에서 개발한 visual programming 언어로, C나 Python 등에 비하면 프로그래밍이 쉬우면서도, 다채로운 컴퓨터 그래픽스 및 시각적 효과를 간단하게 창출할 수 있다. 본 연구는 Kinect를 Scratch와 연동시켜주기 위해 개발된 ScratchX라는 응용 프로그램을 이용하여 미디어 아트 콘텐츠들을 기획하고 제작하였다.

2. 미디어 아트 콘텐츠 제작

본 연구는 세 가지 서로 다른 주제 및 형식의 미디어아트를 제작하였다. 세 가지 형식은 각각 (1) 스토리텔링식 visual novel 게임, (2) 영화홍보 게임, (3) 공익광고 게임 형식이다.

2.1. 이상한 나라의 엘리스 스토리텔링식 미디어 아트

첫 번째 콘텐츠로서, 이상한 나라의 엘리스 동화를 스토리텔링 형식의 visual novel game 형식으로 제작하였다. 게임은 5가지 정도의 장면으로 구성되는데, 첫 번째 장면에서는 Kinect 앞에서 손을 움직여 열쇠로 문을 연다.

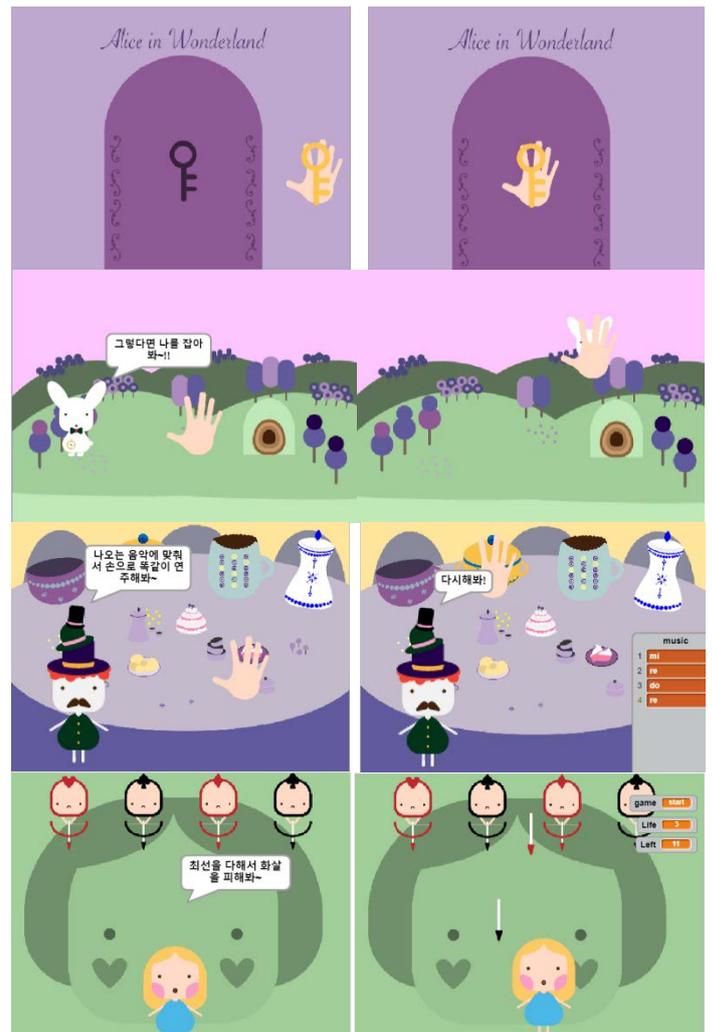


그림 1: 이상한 나라의 엘리스 스토리텔링식 미디어 아트 콘텐츠 예

* 포스터 발표논문

* 학부생 주저자 논문임

* 본 논문은 요약논문(extended abstract)으로서, 본 논문의 원본 논문은 현재 타 학술대회에 제출 중

두 번째 장면에서는 시계를 가지고 뛰어다니는 토끼를 손으로 잡는다. 세 번째 장면에서는 계이를 다르게 내는 찻잔을 가상으로 터치하여 “뺏다 뺏다 비행기” 음을 만든다. 네 번째 장면에서는 온 몸을 움직여 카드 병정들이 쏘는 화살을 피한다. 화살을 다 피하였으면 게임이 끝나고 주인공 엘리스는 잠에서 깨어난다.

2.2. 영화홍보 게임식 미디어 아트

두 번째 콘텐츠는 영화 홍보를 위한 게임 형식의 미디어 아트이다. Kinect가 인체의 관절 부위를 파악하는 것을 이용하여 인체의 관절과 관절 사이에 아이언맨의 머리, 몸통, 팔, 다리 세그먼트들을 배치하였다. 그리고 몸을 움직여 아이언맨을 조작하여 악당을 무찌르는 간단한 시나리오를 미디어 아트 형식의 게임으로 제작하였다. 이와 같은 기술을 이용하면 간단하게 영화 홍보를 위한 재미있는 콘텐츠를 제작할 수 있을 것이다.



그림 2: 아이언맨 영화 홍보에 사용될 수 있는 미디어 아트 게임 콘텐츠 예

2.3. 사회적 메시지를 담은 공익광고식 미디어 아트

마지막으로, 사회적 메시지를 담은 공익광고식 미디어 아트로서, 북한이탈주민의 아픈 이야기를 전달하고 북한이탈주민을 향한 희망의 메시지를 전해주는 콘텐츠를 개발하였다. 북한이탈주민에게 무심코 던지는 우리들의 말이 그들에게 큰 상처가 된다는 것을 표현하고자 하였다. 캐릭터 이름은 북한에서 많이 사용되는 이름인 ‘인남이’이며, 인남이는 위에서 떨어지는 북한이탈주민에 대한 상처의 표현들을 열심히 피해야 하지만 절대로 모

든 말을 피할 수 없게 되어 있다. 인남이가 상처의 말들을 세 번 맞게 되면 다음 화면으로 넘어가며 북한이탈주민들에게 상처의 말을 하지 말자는 공익광고 메시지를 띄운다.



그림 3: 공익광고 형식의 미디어 아트 콘텐츠 예

3. 토의

사람의 동작을 측정하여 이를 interactive 시각화에 적용하는 것에 관한 기술은 이미 잘 알려져 있기 때문에 새로운 기술적 제안은 쉽지 않아 보인다. 카메라나 Kinect를 이용하여 측정된 사람의 동작에 디자인적 및 예술적 심미성을 더하여 사용자에게 감성적 만족감을 전달하는 미디어 아트 분야에 적용하기 위하여 이 기술은 점차 사용이 간단해졌고, Kinect와 Scratch를 간편하게 연결할 수 있는 소프트웨어 toolkit이 개발되기 이르렀다[1]. 따라서 앞으로의 사람 동작 기반의 미디어 아트는 기반기술보다는 기술을 응용한 콘텐츠의 창의성 측면이 주로 회자될 것으로 생각된다.

Kinect와 Scratch의 조합은 이미 알려진 기술로, 상대적으로 손쉽게 사용자 참여식 미디어 아트 콘텐츠를 제작할 수 있다는 장점이 있다. Kinect와 Scratch를 이용하면 창의적이고 흥미로우면서도 효과적으로 비주얼 효과를 낼 수 있는 콘텐츠 및 인터랙션을 비교적 쉽고 빠르게 제작할 수 있다는 장점이 있다. 비록 화면의 크기, 해상도, 세부적인 구현 방식 등에 있어 기술적인 제약이 있기 때문에 전문적인 미디어 아트 제작에는 한계가 있지만, 간단한 미디어 아트 콘텐츠를 제작하거나 전문적인 인터랙티브 미디어 아트의 프로토타입을 목적으로 이용하기에 충분한 것으로 판단된다.

동영상의 시연 영상은 아래와 같다.

- 엘리스: <https://youtu.be/tAQJ2TnOeU8>
- 아이언맨 홍보: <https://youtu.be/-siTn7P17wc>
- 공익광고: <https://youtu.be/XQY1y7qoQhw>

4. 참고문헌

[1] 김현태, 티동국, 손재현, 김성중 (2016) Scratch 미디어 아트, 퍼플.

3가지 물질에 대한 반응-확산을 통한 새로운 패턴 생성*

한도연^{0,1}, 김병문², 송오영^{*.1}
 세종대학교¹, 어도비 연구소²

hando715@gmail.com, bmkim@adobe.com, oysong@sejong.ac.kr

New Pattern Generation from 3-State Reaction Diffusion*

Do-yeon Han^{0,1}, Byung-moon Kim², Oh-young^{*.1}
 Sejong University¹, Adobe Research²

요약

반응-확산 모델은 화학 반응 과정을 시뮬레이션 하는 것으로부터 시작되었으며, 그래픽스 분야에서는 반응-확산 모델을 통해 생성된 패턴을 텍스처 및 스타일 맵으로 활용하거나 유체 시뮬레이션 기법과 결합하여 다양한 효과를 생성하는 방식으로 이용해오고 있다. 지금까지 널리 활용되는 반응-확산 모델들은 2가지 물질에 대한 반응-확산 모델에 집중 되어 있으며 3가지 이상의 물질에 대한 패턴 생성 연구는 드물고, 반응-확산을 위한 지배방정식도 한정되어 있다. 이에 본 논문에서는 광범위한 탐색을 통해 3가지 물질에 대한 새로운 반응-확산식을 찾고 이에 따른 패턴 결과를 제시한다.

1. 서론

반응-확산은 화학 반응 과정 모델로 알려져 있으며, 넓은 관점에서는 치타 또는 호랑이 가죽 무늬 같은 생물학적 패턴을 생성하는 것도 포함한다[1]. 그레이-스콧[2], 피츠휴-나구모[3], 앨런-튜링[1], 바클리[4] 모델은 2가지 물질에 대한 반응-확산식이다. 벨루주브-자보틴스키[5,6] 반응식은 3가지 물질에 대한 반응-확산식이나 그 내용이 공식적으로 출판되지 않았다. 하지만 이러한 반응-확산식에 대한 연구는 다양한 패턴 생성을 위한 것이라기보다는 화학 반응 과정을 충실히 반영하여 모델링한 결과이다.

한편, 컴퓨터 그래픽스 분야에서는 반응-확산 모델의 시간가변성 패턴을 텍스처 및 스타일 맵으로 활용하거나 유체 시뮬레이션 기법과 결합하여 다양한 효과를 생성하는 방식으로 이용해오고 있다. 특히, 반응-확산 모델로부터 시간에 따라 변하는 패턴을 생성해 내고 곡면에서의 텍스처 합성에 대한 응용과 관련하여 연구가 이루어져 왔다[7,8].

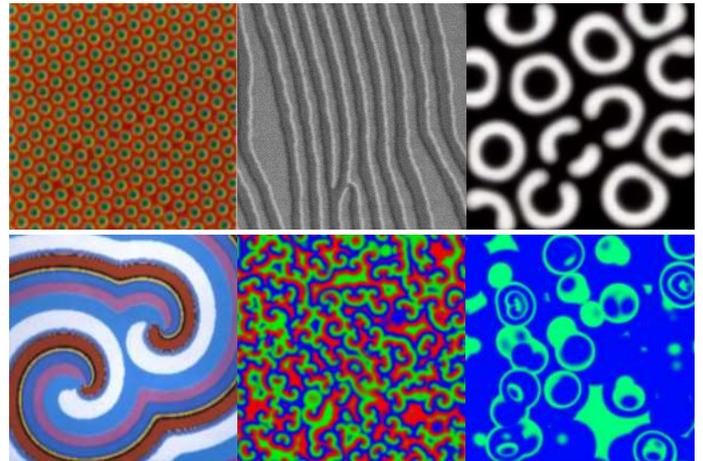


그림 1: (위) 그레이-스콧 모델의 패턴, 피츠휴-나구모 모델의 패턴, 앨런-튜링 모델의 패턴, (아래) 바클리 모델의 패턴, 벨루주브-자보틴스키 모델의 패턴 2가지

2. 반응-확산 모델

반응-확산 모델은 반응과정과 확산과정으로 나눌 수 있다. 확산과정에서 물질은 주변 셀로 이동하고, 반응과정에서 서로 다른 물질이 반응계수에 따라 반응을 일으키며 다른 물질로 변환되어 물질의 양이 증가하거나 감소한다.

반응-확산 모델에서의 확산과정은 확산 커널을 적용하는 것으로도 충분하고 반응과정은 물질의 농도에 따른 다항식으로 표현 가능하므로, 반응-확산식은 유한차분 기반으로 구현하기 간단하다. 이러한 장점에도 불구하고 반응-확산식의 근본적인 한계점은 지금까지 연구된 반응-확산식의 종류가 적다는 점이다. 이 논문에서는 광범위한 탐색을 통해 새로운 반응-확산식을 찾고 이에 따른 패턴 결과를 제시한다.

3. 3가지 물질에 대한 반응식

기존 연구 조사에 따르면 지금까지 알려진 3가지 이상의 물질에 대한 반응-확산 모델은 드물다. 특히 3가지 물질에 대한 벨루주브-자보틴스키 반응 모델은 단순하여 탐색 범위에 한계가 있으므로 모든 물질들 간의 반응이 일어날 수 있도록 항들을 추가하여 탐색할 필요가 있다. 본 연구의 실험에서는 각 물질의 농도를 기본 색

⁰ 발표자

* 교신 저자

* 포스터 발표 논문

* 본 연구는 신진연구지원사업의 연구결과로 수행되었음 (NRF-2015R1C1A1A02036501).

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW 중심대학지원사업의 연구결과로 수행되었음(2015-0-00938).

상 3가지로 표현하여 각 물질이 어떻게 반응하고 움직이는지 시각적으로 쉽게 확인할 수 있도록 설계하였다. 식 (1)에서 r, g, b 는 각 물질에 대한 농도이며, 각각 빨강, 초록, 파랑으로 나타난다.

$$\begin{aligned} \frac{\partial r}{\partial t} &= \mu_r \nabla^2 r + C_r(r, g, b) \\ \frac{\partial g}{\partial t} &= \mu_g \nabla^2 g + C_g(r, g, b) \\ \frac{\partial b}{\partial t} &= \mu_b \nabla^2 b + C_b(r, g, b) \end{aligned} \quad (1)$$

μ_r, μ_g, μ_b 는 각 물질에 대한 확산계수이고 $C(r, g, b)$ 는 반응식이며, 식 (2)와 같다.

$$\begin{aligned} C_r(r, g, b) &= \sum_{i,j,k=0}^3 \alpha_{ijk} r^i g^j b^k \\ C_g(r, g, b) &= \sum_{i,j,k=0}^3 \beta_{ijk} r^i g^j b^k \\ C_b(r, g, b) &= \sum_{i,j,k=0}^3 \gamma_{ijk} r^i g^j b^k \end{aligned} \quad (2)$$

α, β, γ 는 반응계수이며 각 물질마다 64개 즉, 총 192개의 반응계수가 사용된다. i, j, k 값의 범위는 0에서 3이며, 각각의 반응식은 i, j, k 의 모든 조합을 적용시켜 광범위한 탐색을 할 수 있도록 설계하였다.

4. 결과 및 향후 연구

반응계수의 수가 192개로 탐색 영역이 방대하므로 모든 경우의 수를 실험하는 것은 상당히 어렵다. 따라서, 세 물질의 확산계수를 포함한 계수 195개의 값을 매 시뮬레이션마다 무작위로 설정하여 실험하였다(확산계수의 범위는 0~1.0, 반응계수의 범위는 -0.1~0.1). 확산과정으로는 3x3 커널을 사용하였으며, 시뮬레이션 시간 간격은 1/60, 반응-확산 시뮬레이션 영역은 256으로 설정했다. 초기조건으로는 빨강, 파랑, 초록색이 무작위로 분포 되어 있는 텍스처를 사용하였으며, 모든 시뮬레이션에 동일한 텍스처가 적용되었다. 총 58,000개의 시뮬레이션 중 303개의 흥미로운 패턴이 발견되었으며, 기존에 연구된 모델의 패턴과 유사성을 보이는 것도 있지만 전반적으로 다양성이 더 크고 더 다양한 색상을 사용하며 무질서한 패턴들이 많이 발견되었다. 이 후 연구로 반응-확산 모델과 초기조건의 상관관계에 대한 실험을 진행 중이며, 일부 반응계수의 값을 고정시키는 실험 등을 통해 패턴 생성에 큰 영향을 끼치는 반응계수를 찾아내 보다 적은 반응계수를 사용하면서 패턴 분포 영역이 넓은 반응식에 대한 연구를 진행할 예정이다.

참고문헌

[1] A. M. Turing : The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 237(641):37-72, 1952.

[2] J. E. Pearson, Complex patterns in a simple system. *SCIENCE-NEW YORK THEN WASHINGTON-* 261:189-189, 1993.
 [3] R. Fitzhugh, Mathematical models of excitation and propagation in nerve. *Publisher Unknown*, 1966.
 [4] D. Barkley, A model for fast computer simulation of waves in excitable media. *Physica D: Nonlinear Phenomena* 49(1-2):61-70, 1991.
 [5] B. P. Belousov, *Radiates. Med.* 145, 1959.
 [6] A. M. Zhabotinsky, Periodic oxidizing reactions in the liquid phase, *Dokl. Akad. Nauk SSSR*, 157(2):392-395, 1964.
 [7] G. Turk, Generating textures on arbitrary surfaces using reaction-diffusion. In *ACM SIGGRAPH Computer Graphics*, 25:289-298, 1991.
 [8] A. Witkin, and KASSM. Kass, Reaction-diffusion textures. *ACM Siggraph Computer Graphics* 25(4):299-308, 1991.

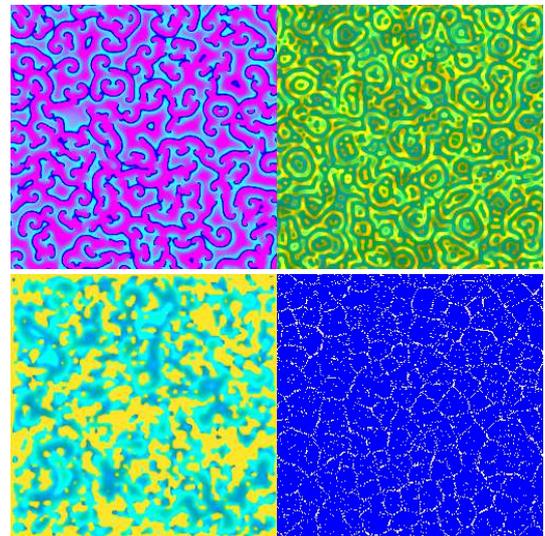


그림 2: 뚜렷한 무늬를 보이는 패턴. (왼쪽 위) 나선형 모양의 패턴, (오른쪽 위) 지문 모양의 패턴, (왼쪽 아래) 두 가지 패턴이 겹쳐진 형태의 패턴, (오른쪽 아래) 거품 모양의 얇은 선 구조물의 패턴

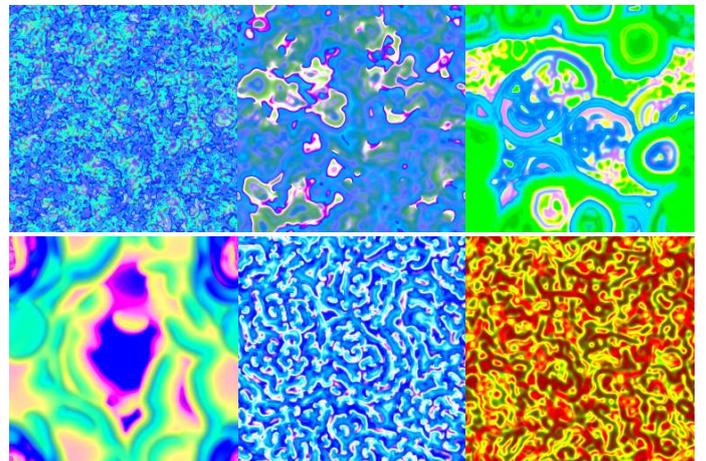


그림 3: 기타 컬러 노이즈 패턴

인코더-디코더 네트워크를 이용한 단일 영상의 비균일 블러 제거*

손형석^{1,0}, 조성현², 이승용¹

¹포항공과대학교 ²대구경북과학기술원

(sonhs, leesy)@postech.ac.kr scho@dgist.ac.kr

Non-Uniform Blur Removal for Single Image using Encoder-Decoder Network

Hyeongseok Son^{1,0}, Sunghyun Cho², Seungyong Lee¹

¹POSTECH ²DGIST

요약

본 논문에서는 딥 러닝을 이용하여 비균일하게 블러된 영상으로부터 자연스럽게 선명한 영상을 복원하는 방법을 제안한다. 영상에서 나타나는 다양한 크기의 비균일 블러를 제거하기 위해서는 다양한 크기의 구조 정보를 고려해야 한다. 본 논문에서는 여러 크기의 구조 정보를 효율적으로 다룰 수 있는 인코더-디코더 구조의 네트워크를 사용하여 효과적으로 비균일 블러를 제거한다. 또한 대립적 손실 함수를 이용하여 정량적 성능을 유지하면서 시각적 성능을 높일 수 있음을 보인다.

1. 서론

단일 영상 디블러링은 컴퓨터 그래픽스 및 비전 분야에서 활발하게 연구되고 있는 주제이다. 그 중에서 영상 전체가 한 가지 블러 커널에 의해서 블러된 영상을 복원하는 균일 블러 제거 연구의 경우 많은 발전이 있었다. 하지만 영상이 영역마다 다른 블러 커널에 의해서 블러된 비균일 블러 제거 연구의 경우 높은 난이도로 인해서 일반적인 상황을 처리할 수 있는 방법에 대한 연구가 많지 않다.

최근 초해상도 영상 생성과 같은 다양한 영상 처리 관련 연구들이 딥 러닝을 사용하여 큰 성능 향상을 보여 주고 있다. 비균일 블러 제거 연구도 마찬가지로 딥 러닝이 활용되고 있다. Nah et al. [1]은 비균일 블러 제거를 위한 데이터 셋을 생성하고 이를 이용해 네트워크를 학습하여 기존 방법들보다 훨씬 높은 성능을 보여주었다. 이 방법의 경우, 효과적으로 다양한 크기의 구조 정보를 고려하기 위해서 여러 개의 네트워크를 사용하는 cascade 구조를 사용한다. 그러나 이러한 방법은 컨볼루션 레이어의 수가 많아져서 수행 속도가 느리다는 문제점이 있다.

본 논문에서는 한 개의 네트워크로 다양한 크기의 구조 정보를 고려할 수 있는 인코더-디코더 구조의 네트워크를 사용하여 비균일 블러 제거 문제를 학습한다. 이를 통해 기존 방법과 비슷한 정확도를 보이면서 처리 속도를 향상시킨다. 또한 대립적(adversarial) 손실 함수를 통해서 시각적 성능을 더 높일 수 있음을 보인다.

2. 비균일 블러 제거 방법

2.1. 비균일 블러 제거 네트워크

본 논문에서는 다양한 크기의 모션 블러를 효과적으로 처리하기 위해서 인코더-디코더 구조의 네트워크를 사용한다. 전체적인 구조는 U-Net[2]과 비슷하다. 인코더 네트워크, 디코더 네트워크가 존재하고 두 네트워크의 각 블록마다 skip-connection이 존재하여 인코더 네트워크의 중간 특징 맵을 디코더 네트워크로 바로 전달한다. 이러한 skip-connection을 통해 큰 구조 정보를 다루는 디코더가 작은 구조 정보를 함께 사용할 수 있도록 한다. 이 때 [2]와 다르게 인코더 네트워크와 디코더 네트워크의 특징 맵을 결합할 때 concatenation 연산이 아닌 요소별 합(element-wise summation) 연산을 사용한다. 이러한 인코더-디코더 구조는 적은 수의 레이어를 사용하기 때문에 실제 수행 과정에서 메모리 사용량 및 수행 시간 측면에서 이득이 있다.

2.2. 데이터 셋

본 논문에서는 Nah et al. [1]이 제작한 짝 데이터 셋을 사용하여 네트워크를 학습한다. 해당 데이터 셋은 GOPRO를 사용해 240fps로 동영상을 촬영한 뒤, 인접한 15 프레임들을 섞어 자연스러운 비균일 블러 영상을 만든다. 15프레임마다 샘플링한 영상을 선명한 영상(블러 영상을 만드는 데 사용한 프레임들 중 중간 프레임)으로 사용해 짝 데이터 셋을 구성한다. 이러한 방법은 자연스러운 비균일 블러 영상과 그에 해당하는 선명한 GT 영상을 제작할 수 있다는 장점이 있다.

* 포스터 발표논문

* 본 연구는 과학기술정보통신부의 재원으로 정보통신기술진흥센터(SW 스타랩, IITP-2015-0-00174)와 차세대정보·컴퓨팅기술개발사업(NRF-2017M3C4A7066317)의 지원을 받아 수행되었음



그림 1: 비균일 블러 제거 방법들의 시각적 결과 비교

2.3. 손실 함수

본 논문에서 사용한 손실 함수는 평균 제곱 오차 (MSE, mean squared error) 손실 함수, 인지적(perceptual) 손실 함수, 대립적(adversarial) 손실 함수 세 가지이다. MSE 손실 함수는 가장 기본적인 데이터 손실 함수로 네트워크의 결과와 GT의 픽셀 별 차이의 제곱이 줄어들도록 함으로써 네트워크는 입력 영상을 GT와 비슷하게 만들도록 학습된다. 데이터 내에 노이즈가 존재하기 때문에 MSE 손실 함수를 0으로 만드는 것은 불가능하다. 이 때, MSE 손실 함수는 각 픽셀 마다 손실 함수에 기여하는 정도가 비슷해지도록 만들기 때문에 결과 영상이 흐려지는 경향이 있다. 따라서 PSNR과 같은 정량적인 측정 방법으로는 높은 성능을 보이지만 실제 시각적 결과는 이에 못 미치는 경우가 많다.

본 논문에서는 이러한 문제를 완화하기 위해서 GAN 방법을 사용한다. 최근 Ledig et al. [3]이 GAN 방법을 이용하여 초해상도 생성 기술의 시각적 성능을 크게 개선시켰다. 본 논문에서도 [3]과 같이, GAN 방법으로 네트워크를 학습하기 위해서 평균 제곱 오차 손실 함수 대신 인지적 손실 함수와 대립적 손실 함수를 사용한다. 대립적 손실 함수는 네트워크의 결과 영상들의 통계적 분포가 GT 영상들의 통계적 분포와 비슷해지도록 만든다. 이러한 방법은 픽셀 별로 강제로 비슷하게 만드는 평균 제곱 오차와 다르게 시각적으로 비슷한 특징을 가지도록 만든다. 데이터 손실 함수로는 인지적 손실 함수를 사용한다. 인지적 손실 함수는 의미론적인 구조에 집중을 하기 때문에 평균 제곱 오차 손실 함수보다 대립적 손실 함수와의 충돌이 적고 적절히 조화된 결과를 보여준다.

3. 결과

테스트 셋으로 Nah et al. [1]에서 제안한 데이터 셋의 1111장의 테스트 이미지들을 사용했다. 기존 최신 기술인 [1]과 비교한 결과는 표1과 그림1에서 확인할 수 있다. [1]의 경우 평균 제곱 오차 손실 함수를 이용해 학습된 모델의 결과이다. 실험에서 사용된 PC는 Intel i7 CPU 및 NVIDIA Titan X GPU를 사용했다.

표 1: 정확도 및 속도 비교. 1280×720 해상도 기준.

	[1]	본 방법 (MSE)	본 방법 (GAN)
PSNR	28.62dB	29.12dB	28.68dB
SSIM	0.909	0.918	0.906
속도	3.09s	0.40s	0.40s

표1을 보면 본 방법이 [1]과 비슷한 성능을 보이면서 훨씬 더 빠르다는 것을 확인할 수 있다. 이는 본 방법의 네트워크 구조가 더 효율적으로 블러를 제거할 수 있다는 것을 보여준다. 그림1을 보면 정량적 성능이 비슷함에도 불구하고 본 방법이 시각적으로 GT와 더 비슷한 결과를 보여준다는 것을 확인할 수 있다. 또한 대립적 손실 함수를 사용하여 학습했을 때 시각적 성능이 향상된다는 것을 확인할 수 있다.

4. 결론

본 논문에서는 딥 러닝을 이용해서 비균일 블러를 효율적으로 제거할 수 있는 방법을 제안하였다. 큰 구조 정보와 작은 구조 정보를 효과적으로 결합할 수 있는 인코더-디코더 네트워크를 사용하여 기존 최신 방법들과 비슷한 성능을 유지하면서 처리 속도를 향상 시켰다. 또한 GAN 방법을 사용해서 시각적인 성능을 향상시킬 수 있다는 것을 확인했다. 하지만 GAN 방법의 효과가 초해상도 영상 생성 연구 등에서 보여준 효과에 비해 떨어진다는 한계가 있다. 이는 고프레임률의 동영상을 이용해 데이터 셋을 생성하기 때문에 GT 영상의 화질이 충분히 좋지 않기 때문이다. 이러한 문제점을 해결하는 방법을 찾는 것은 향후 좋은 연구주제가 될 것이다.

참고문헌

[1] S. Nah, T. Kim and K. Lee, Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring, *CVPR*, 2017.
 [2] O. Ronneberger, P. Fischer and T. Brox, U-net: Convolutional Networks for Biomedical Image Segmentation, *MICCAI*, 2015.
 [3] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang and W. Shi, Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, *CVPR*, 2017.

전경 분리 알고리즘을 활용한 스타일 트랜스퍼에서의 콘텐츠 보존 기법*

정진서¹, 김상현¹, 주은미¹, 김명진¹, 김영빈^{2*}

¹중앙대학교 디지털이미징공학과, ²중앙대학교 첨단영상대학원
ecrime@gmail.com, kraills@naver.com, wer9957@naver.com, rlaudwls96@naver.com, ybkim85@cau.ac.kr*

Preserving Content in Style Transfer via Foreground Separation Algorithm

Jin-Seo Jeong¹, Sang-Hyun Kim¹, Eun-Mi Joo¹, Myung-Jin Kim¹, YoungBin Kim^{2*}

¹Dept. of Integrative Engineering, Chung-Ang University, ²Graduate School of Advanced Imaging Science, Multimedia & Film, Chung-Ang University

요약

본 논문에서는 전경 분리 알고리즘을 활용하여 스타일 트랜스퍼에서의 콘텐츠 보존 방법을 제안한다. 제안된 방법은 실제 이미지의 콘텐츠 중심과 원하는 스타일 중심의 결과 영상 2가지를 이용한다. 두 결과 영상을 GrabCut 알고리즘을 이용하여 콘텐츠 중심의 영상이 전경, 스타일 중심의 영상이 배경이 되도록 합성하여 입력 영상의 콘텐츠 정보가 보존된 새로운 스타일 트랜스퍼 영상을 생성한다. 제안한 방법은 기존의 방법에 비해 영상의 내용적 손실을 줄여 콘텐츠 인식률을 높일 수 있다.

1. 서론

입력 영상에 참조 스타일을 전이해 새로운 영상을 생성하는 스타일 트랜스퍼 기술[1, 3]은 영상처리, 컴퓨터 비전 분야에서 활발히 연구되고 있다. 하지만 기존 스타일 트랜스퍼 기술은 참조 스타일 전이 과정에서 입력 영상의 내용 정보가 유실되는 문제가 발생한다. 이러한 문제점을 해결하기 위해 본 논문에서는 Jonson et al[1]의 기법 기반하여 스타일 트랜스퍼를 수행하고, GrabCut 알고리즘[2]을 이용하여 스타일 트랜스퍼의 결과 영상에서 스타일 정보와 이미지 콘텐츠 정보를 적절히 보존하여 사용자가 영상의 스타일과 이미지 콘텐츠 내용을 정확히 인식하도록 체계를 구축하는 알고리즘을 제안한다.

2. 본문

* 포스터 발표논문

* 학부생 주저자 논문임

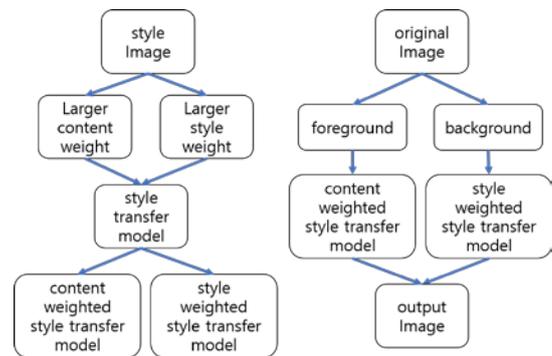


그림 1: 제안하는 방법의 순서도
(왼쪽부터 학습 방법, 이미지 획득 방법)

본 알고리즘은 총 3단계로 구성된다[그림 1]. 먼저, 참조 영상의 스타일 정보와 입력 영상의 내용 정보에 가중치를 준 학습 정보를 각각 추출한다. 이를 입력 영상에 각각 적용하여 스타일 중심적인 영상과 내용 중심적인 영상을 생성한다. 마지막으로 GrabCut 알고리즘을 이용한 이미지 합성을 통해 내용 중심적인 영상이 전경, 스타일 중심적인 영상이 배경으로 설정된 새로운 영상을 생성함으로써 연구의 목적인 입력 영상의 내용 정보가 보존된 스타일 트랜스퍼 결과를 창출해낸다.

2.1. 전경 배경 분리

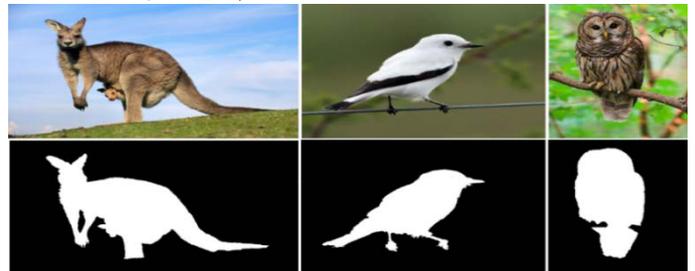


그림 2: 콘텐츠 분리를 위한 마스크링 이미지

GrabCut 알고리즘을 이용하여 콘텐츠와 배경을 분리하는 마스크를 생성한 뒤, 입력 영상과의 행렬을 요소별 곱연산 함으로써 콘텐츠만 존재하는 영상을 얻는다[그림 2]. 배경만 존재하는 영상을 얻기 위해 반전된 이진화 마스크 영상을 사용한다. 콘텐츠 영상과 배경 영상은 다음 수식에 의해 얻을 수 있다.

$$(1) I_c = I_o \odot B$$

$$(2) I_b = I_o \odot (1 - B)$$

여기서 I_c 는 콘텐츠만 존재하는 영상을 나타내며, I_o 는 입력 영상, I_b 는 배경만 존재하는 영상, B 는 콘텐츠를 나타내는 이진화된 마스킹 영상을 의미한다.

2.2. 모델 학습



그림 3: 콘텐츠 중심과 스타일 중심의 학습 모델을 통한 결과 (스타일 입력 영상, 콘텐츠 중심 영상, 스타일 중심 영상)

기존 방법의 경우, 전이된 영상에서의 내용적 손실이 생기는 문제가 있다. 이를 해결하기 위해 2가지 모델을 학습한다. 콘텐츠에 큰 가중치를 두는 모델과 스타일에 큰 가중치를 두는 모델을 학습함으로써 전경 영상에는 전자의 모델을 사용하여 내용적 손실을 줄임과 동시에, 배경 영상에는 후자의 모델을 사용하여 원하는 스타일로의 변화를 수행한다[그림 3].

3. 실험결과

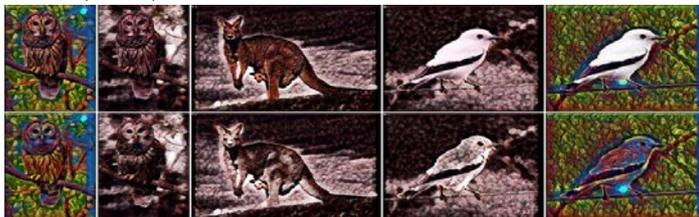


그림 4: 제안하는 방법(위)과 기존 방법(아래)의 결과 영상



그림 5: Inception Resnet V2를 이용한 콘텐츠 보존 확인. 제안한 기법의 콘텐츠 분류 결과(좌측 3번째, 5번째)가 원본 콘텐츠(좌측 1번째)를 유지하면서 스타일 트랜스퍼가 되는 것을 확인할 수 있음. 기존 스타일 트랜스퍼 기법이 great grey owl에 대해 comic book, magpie에 대해 weevil, goldfinch 등으로 콘텐츠를 잘못 파악하는 것(좌측 2번째, 4번째)에 비해 제안된 기법은 원본 콘텐츠의 정보를 제대로 파악하는 것을 확인할 수 있음.

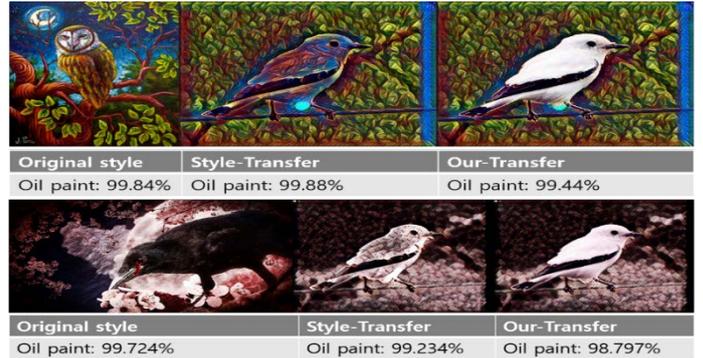


그림 6: vgg19를 이용한 입력 스타일 보존 확인. 제안한 기법이 콘텐츠를 유지함과 동시에 원본 스타일(좌측 1번째, Oil paint)도 유지하고 스타일 트랜스퍼가 이뤄지는 것(좌측 3번째)을 확인할 수 있음.

[그림 4]는 제안된 방법으로 스타일 변화를 주었을 때와 기존 방법을 이용한 결과를 보여준다. 육안으로도 확연하게 기존 방법보다 제안한 방법이 영상의 내용적 손실이 적은 것을 확인할 수 있다. 이를 객관적으로 확인하기 위해 [그림 5]와 같이 개발되어 있는 분류기인 inception Resnet V2를 이용하여 이미지 콘텐츠를 분류하였다. 결과는 육안으로 확인한 것과 마찬가지로 제안한 방법이 기존 방법에 비해 영상 내용을 그대로 보존한 것을 알 수 있다. 또한 [그림 6]의 vgg19를 이용한 스타일 분류 결과를 보면 콘텐츠 뿐만 아니라 기존의 스타일 또한 잘 보존되어 변경된 것을 확인할 수 있다.

4. 결론

본 연구에서는 입력 영상의 콘텐츠를 보존하기 위해 GrabCut 알고리즘을 활용하는 방법을 제안하였다. 스타일 전이를 위한 학습 단계에서 콘텐츠 손실과 스타일 손실 각각에 가중치를 둔 모델을 학습하였다. GrabCut 알고리즘을 이용해 입력 영상을 전경과 배경으로 분리하였으며, 전경은 콘텐츠 중심 모델에, 배경은 스타일 중심 모델에 입력하였다. 콘텐츠 중심으로 전이된 전경 영상과 스타일 중심으로 전이된 배경 영상을 합성하였으며, 이 영상이 콘텐츠를 보존함과 동시에 원하는 스타일로 변경되었는지 확인을 위해 Inception Resnet V2와 vgg19를 사용하여 합성된 영상을 분류하였다. 분류한 결과, 기존 방법에 비해 콘텐츠 인식이 올라간 것을 확인하였으며, 원하는 스타일 방향으로도 인식되었음을 보였다.

참고문헌

[1] J. Johnson, A. Alahi, L. Fei-Fei. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. arXiv:1603.08155, 2016
 [2] C. Rother, V. Kolmogorov, A. Blake, "GrabCut": interactive foreground extraction using iterated graph cuts, ACM Transactions on Graphics (TOG), v.23 n.3. doi:10.1145/1015706.1015720, 2004.
 [3] H. Zhao, PL. Rosin, YK. Lai. Automatic Semantic Style Transfer using Deep Convolutional Neural Networks and Soft Masks. arXiv:1708.09641, 2017

모바일 GPU를 위한 새로운 성능 예측 모델*

윤주원^{0,1}, 이진영¹, 김정길², 임영규³, 나재호³, 김영식⁴, 박우찬¹
 세종대학교 컴퓨터공학과¹, 남서울 대학교 컴퓨터과학과², LG 전자³, 한국산업기술대학교 게임공학과⁴
 (jwyun,jylee)@rayman.sejong.ac.kr, cgkim@nsu.ac.kr, postrain@yonsei.ac.kr,
 na2@naver.com, kys@kpu.ac.kr, pwchan@sejong.ac.kr

A Novel Performance Prediction Model for Mobile GPUs

Juwon Yun⁰, Jinyoung Lee, Cheong Ghil Kim, Yeong-Kyu Lim, Jae-Ho Nah, Youngsik Kim, Woo-Chan Park
 Department of Computer Engineering, Sejong University,
 Department of Science, Namseoul University,
 LG Electronics,
 Department of Game & Multimedia Engineering, Korea Polytechnic University

Abstract

Recent mobile application processors enable the incorporation of high-end 3-D graphics into mobile devices. The problem is that such features consume large amounts of power. To address this deficiency, we proposed a novel performance prediction model for mobile GPUs on Adreno. The average error rate for the proposed performance model was 3.32% for the three applications running on the four mobile devices.

1. Introduction

High-quality graphics require substantial power consumption [1], [2]. but mobile platforms have less physical storage than PC platforms. Therefore, the analytics model to predict electronic power consumption during 3D graphics processing on mobile GPUs can lead to significant competitiveness in product design and manufacturing. The power consumption of mobile GPUs is directly related to processing performance. Accurately predicting power consumption therefore necessitates the identification of factors related to the performance of mobile GPUs. A mobile GPU generally consists of a US, a texture mapping unit (TMU), and a render output processor (ROP). Fig. 1 shows the architecture of an Adreno GPU. In a standard GPU, much of the graphics pipeline operations are performed by shader cores, which accordingly

occupy the largest portion of the GPU [3]. Therefore, this research focused on predicting the performance of Adreno GPUs and deriving a performance prediction model for a unified shader (US) that may be the main factor for decisions on mobile GPU performance.

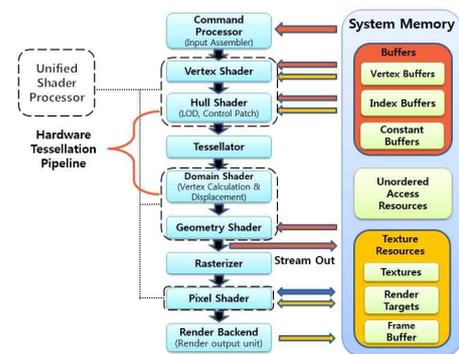


Figure 1: Adreno 420 architecture and graphics pipeline

2. The Proposed Performance Model

2.1. Unified Shader Instruction

As previously stated, the US accounts for the greatest portion of a GPU, thus significantly affecting performance. A micro-benchmark test was carried out to define the relationship between the performance of the US and Adreno GPUs. Through this, The number of vertex and fragment instructions per second in the US exerted the strongest influence on the performance of the Adreno GPUs. In this study, USI (i.e., the number of instructions processed in the US per second) was defined as an index of US performance. However, because of the variable performance caused by Dynamic voltage and frequency scaling (DVFS) and throttling, determining

* 포스터 발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 IEEE Access, Vol. 6, Issue. 1, pp.16235-16245, Mar 2018 에 온라인 게재 되었음.

* 본 연구는 과학기술정보통신부의 재원으로 정보통신기술진흥센터의 지원을 받아 수행되었음.

performance-related factors using the measured USI throughput is difficult. According to the results of the utilization experiment, utilization could be influenced and changed by DVFS and throttling, and shader performance changed in accordance with utilization as shown fig 2.

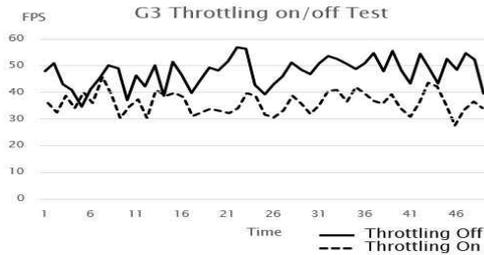


Figure 2 FPS on an application when the throttling is on/off

All utilization factors should be revised into their maximum value to calculate the USI throughput under constant conditions and accordingly solve the utilization issues as shown fig 3.

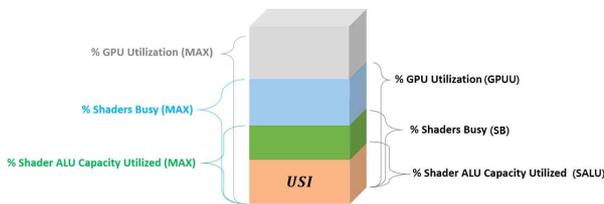


Figure 3: Normalization for GPU related utilization factor

2.2. Proposed Performance Prediction Model with GFLOPS

Through experiments result as shown table 1, the research on calculated maximum USI throughputs and H/W specifications indicated that GFLOPS and US processing performance were related. Therefore, GFLOPS value was defined as a principal factor in the performance prediction model proposed in this work.

Table 1 : Maximum USI throughput results

The maximum USI throughput per second			
Vertices	G4	G3 CAT6	G FLEX2
100k	82.98G	82.97G	164.33G
200k	84.54G	84.55G	169.14G
300k	85.25G	85.30G	170.16G
400k	85.65G	85.74G	171.60G
500k	85.93G	85.96G	171.99G
600k	86.11G	86.06G	172.46G
700k	86.26G	86.27G	172.79G
800k	86.37G	86.38G	172.86G
900k	86.46G	86.47G	168.79G
Average USI throughput	85.510G	85.527G	170.461G
Shader performance 32 bit	163.2	172.8	326.4
Average USI throughput	523.96M	494.95M	522.24M

calculating the difference value between an index device and a prediction device for which the performance prediction model of the Adreno GPUs was used. We also adopted the number of instructions that can be maximally processed per 1 GFLOPS in the US to calculate the maximum USI throughput of the prediction device. Through this, we proposed performance prediction model with GFLOPS.

3. Experimental Results

Table 2 shows the verification results on the application of the Adreno GPU performance prediction model to the scenes, which represent high-level benchmarks from real-world gaming conditions. For the MANHATTAN scene, the error rate of each device was 0.95% to 4.32%. For the TRENCH scene, the error rate was 0.10% to 6.22%. For the BOOTCAMP scene, the error rate was 0.02% to 5.91%, and the average error rate of all the devices was 3.32%. Additionally, the latest Adreno GPU model, Adreno 530, showed an average error rate of 2.96%. These findings indicated that the proposed performance prediction model is applicable not only to experimental devices but also to new devices.

Table 2: Verification results of the proposed performance model

Category	G4	G3 CAT6	G FLEX2	G5
Predicted USI _{max}	82.3G	87.4G	170.4G	177.5G
MANHATTAN Actual USI _{max}	84.16G (Error rate 2.21%)	83.78G (Error rate 4.32%)	165.75G (Error rate 2.81%)	175.83G (Error rate 0.95%)
TRENCH Actual USI _{max}	82.22G (Error rate 0.10%)	82.92G (Error rate 5.40%)	160.42G (Error rate 6.22%)	171.19G (Error rate 3.69%)
BOOTCAMP Actual USI _{max}	82.28G (Error rate 0.02%)	82.52G (Error rate 5.91%)	163.82G (Error rate 4.02%)	170.28G (Error rate 4.24%)

References

[1] R. Wang, B. Yu, J. Marco, T. Hu, D. Gutierrez, H. Bao, "Real-time rendering on a power budget", ACM Trans. Graph., vol. 35, 2016.

[2] X. Ma, Z. Deng, M. Dong, L. Zhong, "Characterizing the performance and power consumption of 3D mobile games", IEEE Computer, vol. 46, no. 4, pp. 76-82, Apr. 2013.

[3] L. Samuli, K. Tero, "High-performance software rasterization on GPUs", Proc. ACM SIGGRAPH HPG, pp. 79-88, 2011.

We calculated the valid GFLOPS value by

대응점 탐색 알고리즘의 GPU 기반 연산에 대한 성능 분석*

김병준⁰, 강현민, 서용덕
서강대학교 영상대학원

arrbem@sogang.ac.kr, khm8206@sogang.ac.kr, yndk@sogang.ac.kr

Performance Analysis about GPU based Correspondence Matching

Byungjoon Kim⁰, Hyunmin Gang, Yongduek Seo
Graduate School of Media, Sogang University

요약

본 연구는 다시점 영상의 3D 복원 및 실시간 4D 복원을 비롯해 다양하게 개발 되고 있는 증강현실 및 가상현실 환경에서의 대응점 정보를 얻기 위해 진행된다. 이를 위해 전통적인 대응점 탐색 알고리즘의 CPU기반의 결과와 GPU기반 결과를 비교하고 성능 및 효율을 분석한다.

1. 서론

영상의 물체의 움직임을 분석하거나 영상의 구조에 대한 정보를 획득하기 위한 대응점 탐색에 대한 많은 연구가 진행되었다. 이런 연구는 광학흐름(Optical Flow) [1,2] 및 스테레오 매칭의 응용 연구 분야에서 핵심 기반 기술로 볼 수 있다. 본 논문에서는 전통적인 기법의 알고리즘들이 GPU기반 병렬 처리를 통해 향상 되는 성능과 학습 기반으로 Decision Forest를 생성하여 대응점 탐색 및 흐름 추정하는 경우의 성능 및 효율에 대해 연구하여 실시간 시스템에 대응점 탐색 기술의 사용 가능성을 찾는다.

2. 본론

대응점 탐색은 Horn-Schunck[1]와 Lucas-Kanade [2]의 연구를 기반으로 응용되며 단점을 보완하며 진행되어 왔다. 이를 학습기반의 연구 이전의 전통적인 알고리즘이라 하고 본 논문에서는 피라미드 Lucas-Kanade [3]에 대한 실험을 진행한다. 대응점 탐색은 크게 픽셀의 특징점을 이용하는 방법과 시공간 경사를 이용하는 방법이 있는데 전자의 경우에는 특징점을 구하는 다양한 방법이 이용된다. 영상 간의 특징점에 대해 기술자를 연산하고 대응되는 픽셀을 탐색하여 대응점에 대한 정보를 얻게 된다. 후자의 경우에는 영상 간 밝기 분포가 일정하다는 가정 하에 프레임간의 수평, 수직, 시간에 대한 변화를 연산하여 최적화(3) 시켜 추정이 가능하다

$$f(x,y,t) - f(x-u,y-v,t-1) = 0 \quad (1)$$

$$f_x u + f_y v + f_t = 0 \quad (2)$$

$$\iint d(u,v)^2 + \alpha^2 (|\nabla u|^2 + |\nabla v|^2) dx dy \quad (3)$$

두 가지 방법 모두 각 픽셀에서의 주변 정보를 획득하여 특징점이 되는 픽셀의 정보를 획득하게 된다. 이는 정보획득 과정에서 병렬처리가 가능하다는 것을 예상할 수 있다. 하지만 연산에서 적분영상에 대한 정보를 사용할 경우에는 순차적 연산이 필요한 경우가 있다.

학습기반의 대응점 탐색 알고리즘으로는 The Global Patch Collider(이하 GPC)[4]를 사용하는 연구에 대해 실험을 한다. GPC는 학습을 통해 Decision Forest를 생성하고 영상 내의 모든 픽셀에 대해 특징 기술자(descriptor)를 연산한다.

$$C_{T(X)}(I, I') = (x, x') | T(x) = T(x') \quad (4)$$

$$|S_{T(X)}| = |S'_{T(X)}| = 1 \quad (5)$$

$$C_F(I, I') = (x, x') | F(x) = F(x') \quad (6)$$

$$|S_{F(X)}| = |S'_{F(X)}| = 1 \quad (7)$$

모든 픽셀을 Decision Forest에 학습에 대한 data로 사용하며 위의 (4)~(7) 식과 같이 리프노드(Leaf node)에서 하나의 같은 인덱스를 갖는 픽셀을 대응점으로 판단한다.

3. 실험 결과 및 평가

3.1. 실험환경

표1. 실험환경

CPU	i7-8700, 3.2GHz
GPU	Nvidia GTX1080
운영체제	Window 10
CUDA toolkit	version 8.0
RAM	16GB
입력영상 크기	1080 × 1920

3.2. GPU기반 대응점 탐색 및 구현

본 논문에서 실험은 특징점 기반으로 대응점을 찾는 대표적인 알고리즘인 피라미드 Lucas/Kanade[5]를 분석한다. 입력영상은 30fps의 영상에서 인접한 두 프레임으로 대응점을 탐색하고 시간을 측정했다.

* 포스터 발표논문

* 이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 '범부처 Giga KOREA 사업'의 지원을 받아 수행된 연구임 (201839022.01, (4D-2세부)초실감 서비스를 위한 동적 객체의 실시간 4D 복원 기술)



그림1. (좌)입력영상에 대한 CPU를 이용한 순차적 연산
(우)동일 입력 영상에 대한 GPU 병렬처리 연산결과

그림1.은 PyrLK 알고리즘을 이용하여 대응점 탐색을 한 결과이다. 전체 1080×1920영상에서 가장 특징점이 잘 나타나는 부분을 비교한 것이다. 단순 픽셀 연산이 주로 이뤄지기 때문에 전체 픽셀에 대한 정보를 GPU에 할당한 후 CPU연산과 동일한 연산과정을 가진다. 연산 속도는 평균 350ms와 58.8ms로 측정되었다. CPU연산의 16.8%의 속도로 약 5.95배 빠른 연산 속도를 보이고 있다.

3.3. GPC 분석

일반적으로 대응점 탐색 및 광학흐름 추정에서 픽셀에 대해 일정 거리 내에서의 대응점을 탐색한다. 하지만 이는 움직임이 큰 물체나 비정형적 움직임에 대한 대응점 탐색 연산이 어렵다. 이를 극복하기 위해 GPC는 전체 픽셀에 대하여 기술자를 비교한다. 그림 2.는 GPC 알고리즘의 전체적인 도식으로 GPC 대응점 탐색연산을 단계적으로 분석한 것이다.



그림 2. GPC 알고리즘 순차적 연산 순서

3.4. GPU 기반 GPC 성능

표2. 1920×1080 이미지에 대한 연산속도 비교

	CPU기반	GPU기반
이미지 채널 분리	61.37ms	0.66ms
기술자 계산	180.15ms	23.51ms
Forest 연산	1029.09ms	61.33ms
데이터 정렬	302.69ms	121.55ms
대응점 획득	42.39ms	22.18ms
전체 연산 시간	1615.68ms	229.23ms

분석한 GPC 알고리즘을 CPU와 GPU에 동일한 과정으로 실험을 진행하여 연산 속도를 비교한다. 픽셀 기술자 연산에 있어서 적분영상이 이용되는데 이는 부분적

으로 순차적 연산이 필요하다. 그리고 Index에 따라 픽셀 정보를 정렬하는 경우에도 병렬처리가 쉽지 않음을 예상할 수 있다.

그림3.에서 보듯이 정확도 측면에서 같은 결과를 보이며 연산 속도는 약 7배 더 빠른 결과를 얻었다. 단순한 픽셀 정보를 이용한 연산의 경우에 병렬처리로 연산 속도를 극대화 되었으며 픽셀 간 정보를 순차적으로 처리하는 경우에는 병렬처리 연산으로 인한 연산 속도 이득의 한계를 확인 할 수 있었다.



그림3. (좌)CPU 기반 대응점 탐색 (우)GPU기반

4. 결론 및 토의사항

3D 영상의 물체 및 움직임을 복원하는 실시간 4D 시스템을 구현하기 위해서는 GPU기반의 연구가 필수이기에 GPU 병렬처리 연산에 맞춰 알고리즘 구현을 부분적으로 응용하여 사용 할 수 있어야 할 것이다. 위 실험을 예를 들면 데이터 정렬 및 적분 영상에 대한 연산이 비교적 효율이 떨어지는 것을 확인할 수 있다. 병렬처리 연산의 효과를 극대화하기 위한 추가적인 연구가 필요할 것이다.

이는 SLAM, 실시간 4D 복원을 비롯한 많은 응용 분야의 시스템을 설계하는데 바탕이 되는 연구로 사용하려 한다.

참고문헌

[1] B.K.P. Horn and B.G. Schunk, "Determining optical flow", *Artificial Intelligence*, Vol. 17, pp. 185-201, 1981.
 [2] B. Lucas and T. Kanade, "An iterative image registration technique with applications in stereo vision", *In Proc. of the DARPA Image Understanding Workshop*, pp. 121-130, 1981.
 [3] Jean-Yves Bouguet, "Pyramidal Implementation of the Affine Lucas Kanade Feature Tracker Description of the algorithm", *Intel Corporation*, 2001.
 [4] Shenlong Wang, Sean Ryan Fanello, Christoph Rhemann, Shahram Izadi, Pushmeet Kohli, "The Global Patch Collider", *CVPR*, 2016.

제스처 인식 기반 드로잉 시스템의 설계*

김상준⁰, 홍민¹, 최유주^{0,2}

⁰서울미디어대학원대학교 뉴미디어학부 미디어공학전공

¹순천향대학교 컴퓨터소프트웨어공학과

²서울미디어대학원대학교 실감미디어연구소, 교신저자
gogo5911@naver.com, mhong@sch.ac.kr, yjchoi@smit.ac.kr

Design of the Drawing System based on Gesture Recognition

Sang-Joon Kim⁰, Min Hong, Yoo-Joo Choi

Dept. of New Media, Seoul Media Institute of Technology

Dept. of Computer Software Engineering, Sun Chun Hyang University

Immersive Meida Lab. Seoul Media Institute of Technology

요약

본 논문에서는 키넥트(Kinect)의 인체 추적 기능을 이용하여 드로잉을 위한 사용자의 제스처를 정의하고, 제스처를 통하여 자유롭게 그림을 그릴 수 있는 드로잉 시스템을 설계 구현하였다. 제안 시스템은 그림을 그리는 붓의 굵기 선택, 지우개를 이용한 그림 지우기, 사물을 이용한 물감색 선택등의 세부 기능들을 포함하고 있다.

1. 서론

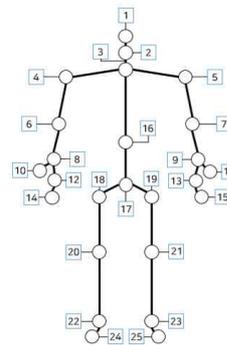
최근 콘텐츠 시장은 사용자와 상호작용이 가능한 단말기의 등장과 발전으로 체험과 참여를 유발하는 상호작용 콘텐츠가 주를 이루기 시작하였다. 상호작용 콘텐츠는 다른 콘텐츠보다 사용자에게 몰입감을 주기 때문에 게임 교육 분야를 벗어나 광고, 영화, 전시, 공연 등에서 까지 쓰이기 시작하였다. 그 중 가장 간단하게 사람의 몸을 이용하여 상호작용하는 제스처 기반 콘텐츠는 키넥트, 립모션(Leap Motion)등과 같은 단말기를 이용하여 사람의 몸과 상호작용한다. 키넥트는 컨트롤러 없이 이용자의 신체를 이용하여 게임과 엔터테인먼트를 경험할 수 있는 엑스박스 360 주변기기이다[1]. 립모션은 두 손과 열 손가락의 여러 움직임을 감지하여 그 움직임을 통해 컴퓨터를 제어할 수 있게 도와주는 장치이다. 이와 같은 단말기를 이용한 콘텐츠는 저스트랜스[2], 키넥트어드벤처[3], 모션비트 '팝업'[4], 후르츠닌자[5] 등이 있다.

본 논문에서는 키넥트의 제스처 인식 기능을 이용하여 사용자의 좌우 양손의 움직임 인식하여 드로잉에 필요한 기능을 선택하고, 사용자의 오른손의 움직임을 통하여 그림을 그릴 수 있는 드로잉 시스템을 설계 구현하였다.

2. 키넥트 V2 제스처 인식기능

키넥트는 0.5~4.5m의 거리의 사용자를 최대 6명까지 인식할 수 있으며 인식된 사용자 각각의 관절 25개

를 획득할 수 있다. 그림 1은 키넥트를 통하여 인지할 수 있는 인체 관절을 보여 주고 있다. 획득한 관절의 정보를 이용하여 관절의 각도, 관절의 위치와 관절의 동작 유무를 감지할 수 있고 그에 따른 상호작용이 가능하다. 또한 손의 개폐상태와 표정을 인식할 수 있는데 손의 개폐상태는 가위, 바위, 보 3가지를 인식할 수 있고 표정은 웃고 있는지 무표정인지 입을 벌리고 있는지 눈을 뜨고 있는지 등을 인식할 수 있다.



번호	구분	번호	구분
1	Head	14	HandTipLeft
2	Neck	15	HandTipRight
3	SpineShoulder	18	SpinMid
4	SholderLeft	17	SpineBase
5	SholderRight	18	HipLeft
8	ElbowLeft	19	HipRight
7	ElbowRight	20	KneeLeft
8	WristLeft	21	KneeRight
9	WristRight	22	AnkleLeft
10	ThumbLeft	23	AnkleRight
11	ThumbRight	24	FootLeft
12	HandLeft	25	FootRight
13	HandRight		

그림 1. 키넥트에서 인식할 수 있는 25개 관절

3. 제안 시스템

본 논문의 제안시스템은 다음과 같이 동작한다. 키넥트에서 사용자의 골격이 인식되면 사용자의 손짓과 손의 개폐상태를 인식한다. 메뉴 조작기에서는 상단에 있는 메뉴와 오른손, 왼손의 개폐상태, 손의 위치에 따라 어떤 기능을 사용자에게 전달할지를 검사한다. 상단의 메뉴는 색상인식 영역, 붓과 지우개, 붓과 지우개의 굵기 지정, 초기화 기능이 있다. 메뉴는 손의 위치가 메뉴 속에 들어갔을 때 동작하며 손의 개폐상태에 따라 기능이 달라진다. 색상인식 영역은 오른손의 개폐상태가 가위일 때 영상의 지정된 영역에서 매 프레임마다 픽셀 정보를 가져와 RGB의 평균을 구하고 평균의 RGB값을 붓의 색상으로 변경하여 실제 사물을 이용하여 여러 가지 색상을 사용자가 자유롭게 정의 할 수 있도록 한다. 지우개와 붓의 선택의 경우 왼손의 위치가 메뉴 속에 들어가 있고 왼손의 개폐상태가 주먹일 때 지우개가 선택되고, 보자기일 때 붓으로 변경된다. 붓의 경우 일반적인 방법과 같이 영상에 RGB 컬러의 색을 뿌려 그림을 그린다. 제스처를 이용하여 그려지는 모든 그림은

* 포스터 발표논문

* 본 논문은 한국연구재단 이공학개인지초연구지원사업 (NRF-2017R1D1A1B03035718)의 지원에 의하여 수행됨

OpenGL FBO(Frame Buffer Object)에 RGBA 컬러 포맷으로 저장된다. 그리고, 카메라 영상 앞에 FBO로 정의된 텍스처를 렌더링 함으로써, 카메라 영상과 렌더링 그림이 함께 보이도록 한다.

지우개의 경우 일반적인 바탕화면색의 색을 가져와 덧칠하는 방식으로 구현이 불가능하다. 그렇기 때문에 다음과 같은 방법을 제시했다. RGB 컬러 중 임의의 특정 색을 지정한다. 예를 들어 RGB(255,0,0)으로 임의의 색을 정하고 오른손의 위치에 붓의 크기에 맞는 원을 그리고, 미리 정한 색 RGB(255,0,0)으로 색칠한다. 원의 그림은 FBO에 RGBA 포맷으로 저장된다. 이때 FBO의 모든 픽셀에서 RGB(255,0,0) 컬러의 픽셀을 찾아 알파 값을 모두 0으로 주어 투명 픽셀로 정의한다. 그리고, 이를 카메라 영상과 겹쳐 렌더링 함으로써 지우개에 의해 그려진 픽셀에는 카메라 영상 픽셀이 보이도록 함으로써 그려진 그림을 지우는 효과를 가지게 한다. 만약 붓의 색으로 RGB(255,0,0)의 색이 들어왔을 경우에는 RGB(254,0,0)으로 변경한다.

붓의 굵기의 경우 오른손의 위치가 메뉴 속에 들어가 있고 왼손의 개폐상태가 주먹일 경우 굵기가 작아지고 개폐상태가 보자기일 경우 굵기가 커진다. 한 번에 붓의 굵기가 굵어지고 작아지는 것을 막기 위해 굵기의 크기가 변하면 왼손의 개폐상태를 가위로 만들어 카운터 값을 초기화 시켜줘야 한다. 초기화 메뉴는 오른손의 위치가 메뉴 속에 들어가 있을 때 렌더링 되는 FBO를 Clear시켜 그려던 그림을 모두 사라지게 한다. 그림 2의 플로우차트는 제안 시스템의 처리 절차를 보여주고 있다.

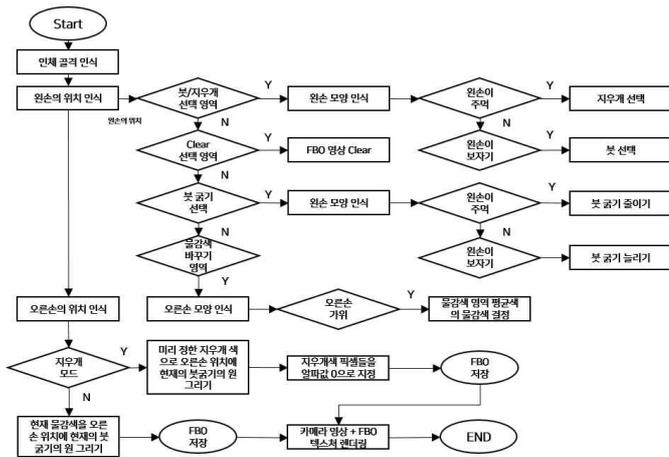


그림 2. 제안 시스템의 처리절차

4. 구현 결과 및 분석

제안된 드로잉 시스템의 인터페이스는 그림3과 같이 좌측 상단에 색상확인창과 중심에 메뉴조작기, 좌측 하단에 인체골격 확인 창으로 구성된다. 그림 4는 제안 시스템에서 사용하는 사용자 동작의 종류를 보여주고 있다. 사용자는 중심의 메뉴조작기를 이용하여 그림5와 같이 그림을 그릴 수 있다. 그림5의 좌측 상단의 모습은 초기화 버튼을 사용하여 그림을 지우는 모습이다. 우측 상단의 모습은 색상을 사용자가 원하는 색상을 입력하는 모습이다. 좌측 하단의 모습은 붓의 굵기를 조정하는 모습이다. 우측 하단의 모습은 붓에서 지우개로 변경

하는 모습이다. 다음과 같이 사용자는 자신의 손짓과 개폐상태만으로 그림을 그릴 수 있었다.



그림 3. 드로잉 시스템 인터페이스

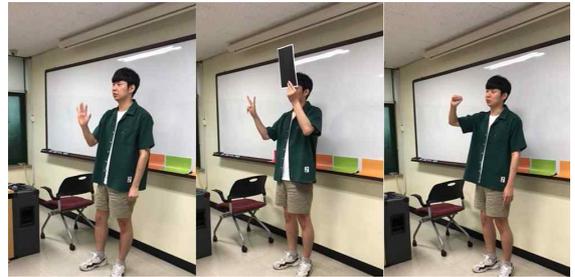


그림 4. 드로잉 시스템 사용자 동작



그림 5. 드로잉 시스템을 이용한 그림 그리기

5. 결론 및 향후연구

본 논문에서는 사용자와 상호작용이 가능한 단말기 키넥트를 이용하여 자유롭게 물감색을 선택하고 그림을 그릴 수 있는 드로잉 시스템을 설계 구현하였다. 사용자의 제스처로 그림을 그리도록 하여 몰입감과 사용자의 흥미를 높였지만 색을 입력하는 과정과 개폐상태의 숙지 어려움으로 불편함을 겪었다. 향후 연구로 인터페이스 개선과 여러 사람들이 그림을 그릴 수 있도록 하며 그림을 따라 그릴 수 있는 도안을 추가할 예정이다.

참고문헌

[1]wikipedia kinect, <https://ko.wikipedia.org/wiki/%ED%82%A4%EB%84%A5%ED%8A%B8>
 [2]Justdancenow, <https://justdancenow.com/>
 [3]Kinect Adventures, https://www.youtube.com/watch?v=nec86i_vInE
 [4]Motion Bit 'Pop Up', <https://lm-s-apps-amnesia.leapmotion.com/apps/pop-up/windows>
 [5]Fruit Ninja, <https://fruitninja.com/>

물리 시뮬레이션 및 진화 연산을 사용한 진화형 모듈형 로봇 설계*

박재훈⁰, 이강훈
 광운대학교
 (techiky, kang)@kw.ac.kr

Evolutionary Modular Robot Design Using Physics Simulation and Evolutionary Computation

Jaihoon Park⁰, Kanghoon Lee
 Kwangwoon University



Figure 1 각 세대별 높은 적합도를 보인 모듈형 로봇들의 형태. 왼쪽부터 오른쪽으로 각각 1, 5, 20, 40, 100세대에서 발견된 모듈형 로봇이다. 점차 세대가 내려감에 따라 부적절한 모듈이 제거되고 이동에 적합한 구조가 형성되는 것을 확인할 수 있다. 100 세대의 로봇 구성에서 바퀴가 달린 모터가 원동력을 제공하고, 피벗이 장애물을 유동적으로 대응할 수 있도록 도와주는 것을 확인할 수 있다.

요약

본 연구에서는 다양한 환경 및 직무에 최적화된 로봇을 찾기 위해, 가상환경에서 모듈형 로봇을 구성하여 시뮬레이션 한 후, 얻은 적합도를 바탕으로 유전 알고리즘 연산을 적용하여 최적화된 로봇을 찾는 자동 설계 시스템을 제안한다. 기존 연구들이 단일한 모듈로 모듈형 로봇을 구성하여 진화를 수행하였다면, 본 연구에서는 각기 다른 형태와 움직임을 가지는 다양한 모듈들로 구성된 모듈형 로봇을 구성하였다.

1. 서론

모듈형 로봇은 각기 독립적인 형태와 상태로 구성된 모듈들로 이루어져 있어 진화학습을 위한 유전형 및 표현형간의 전환이 용이하고, 이러한 특성은 교차, 변이 등의 유전 알고리즘 연산에도 적합하다. 또한, 모듈형 로봇은 실세계에서 물리적으로 조립 및 제어가 용이하여 진화 로봇 공학에 용이하다 [1].

본 연구에서는 모듈형 로봇을 구성하고 시뮬레이션 할 수 있는 가상환경에서 유전 알고리즘 연산을 사용하여 목표함수에 최적화된 모듈형 로봇을 자동 설계하는 시스템을 제안한다. 제약 조건 내에서 구성된 초기 모듈형 로봇 해들을 시작으로,

물리 시뮬레이션을 수행하여 적합도를 계산한 뒤, 반복적으로 유전 알고리즘 연산을 수행하여 새로운 모듈형 로봇 개체를 만들어낸다. 시뮬레이션 및 진화 연산 수행 횟수가 많아질수록 최적화된 모듈형 로봇을 얻는다.

2. 모듈형 로봇



Figure 2 사용된 모듈의 형태: 기본(Base), 모터(Motor), 피벗(Pivot), 바퀴(Wheel), 그래버(Grabber) 모듈.

	기본	모터	피벗	바퀴	그래버
연결부 개수	4	4	2	1	1
작동기	X	O	O	X	O

Table 1 각 모듈의 정보. 연결부 개수는 다른 모듈이 연결될 수 있는 위치의 개수, 작동기는 모듈이 자기 자신 혹은 연결된 다른 모듈을 움직이고 제어할 수 있는지 여부를 나타낸다.

본 연구에서 제시하는 모듈형 로봇의 주요 모듈에 대한 정보는 Figure 2 및 Table 1에서 확인할 수 있다. 각 모듈은 다른 모듈과 연결될 수 있으며 연결될 수 있는 지점은 모듈마다 상이하다. 예를 들어, 모든 모듈은 정육면체 형태의 기본, 모터(Motor)의 어느 면에든 연결될 수 있지만, 피벗(Pivot)의 경우에는 위 혹은 아래 부분

* 포스터 발표논문

* 이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 이공학 개인기초연구 성과임(No. 2015R1D1A1A01059596).

에만 연결될 수 있다. 피벗과 모터는 작동장치로서 연결된 다른 모듈에 힘을 가하여 모듈형 로봇의 움직임을 유도한다.

3. 유전 알고리즘 적용

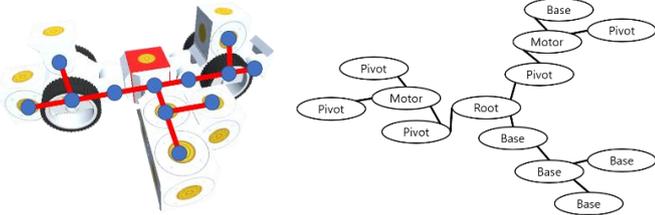


Figure 3 모듈형 로봇의 예시와 트리 구조 표현

유전 알고리즘은 자연계의 진화 메커니즘을 모방한 최적화 방법론이다. 유전 알고리즘에서는 문제의 해를 진화 연산이 가능한 자료구조로 표현하여 초기 해집합을 구성한 후, 매 세대마다 각 해의 적합도를 계산하여 우수한 해를 선택해 나감으로써 점차 최적의 해에 가까워지도록 한다.

3.1 유전자 표현

본 연구에서는 모듈형 로봇의 유전형질을 트리 형식으로 표현한다. 트리의 노드는 로봇을 구성하는 각 모듈을 표현한다. 노드에 모듈의 상태 값과 모듈과 연결된 부모 및 자식 모듈에 대한 정보가 저장된다. Figure 3에서 모듈형 로봇의 트리 표현을 시각적으로 볼 수 있다.

3.2 진화 연산

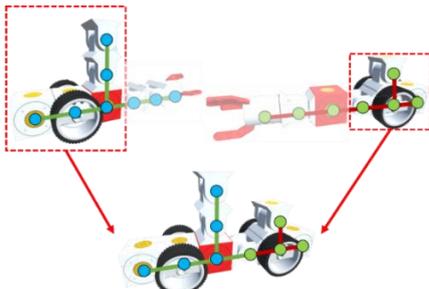


Figure 4 두 모듈형 로봇의 교차연산 결과.

본 연구에서는 일점 교차 연산을 적용하였다. 먼저 아래의 계산식을 통해 해집단의 각 해들의 적합도를 구한다.

$$O_i = w_D D_i + w_{N_1} \exp\left(-\frac{1}{w_{N_2}} * N_i\right) \quad (1)$$

$$f_i = (O_i - O_w) + (O_b - O_w)/(k - 1), k > 1 \quad (2)$$

식 (1)은 해집단의 해 i 의 품질값인 O_i 를 계산하는 수식으로, D_i 는 로봇 i 가 이동한 거리, N_i 는 로봇을 구성하는 모듈의 개수이며 C 값은 상수로 본 연구에서는 25를 사

용하였다. 로봇의 이동거리가 길수록, 구성 모듈의 개수가 적을수록 좋은 품질값을 받는다.

f_i 는 해 i 에 대한 적합도로, O_b 는 해집단에서 가장 좋은 해의 품질값, O_w 는 가장 나쁜 해의 품질값으로 본 연구에서 k 값은 2.5를 사용하였다.

계산 후 품질 비례 룰렛휠 방법을 사용하여 연산을 수행할 두 개의 해를 선택한다. 하나의 해에서 임의의 에지를 선택하여 잘라낸 후, 다른 해에서 임의로 에지를 선택 후 앞에서 선택한 것으로 대체하여 새로운 해를 만든다. 교차연산 결과는 Figure 4에서 확인할 수 있다.

4. 실험 결과

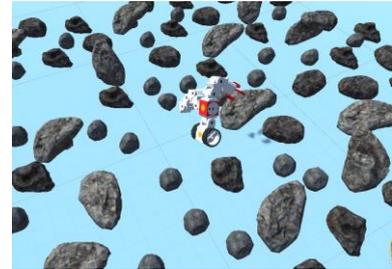


Figure 5 여러 종류의 장애물들이 배치된 물리 기반 시뮬레이션 환경에서 동작하고 있는 모듈형 로봇.

본 연구에서는 무작위로 생성된 25개의 모듈형 로봇을 해로 가지는 모집단에서 유전 알고리즘을 시작했다. 서로 다른 부피, 질량을 가진 장애물이 다수 배치된 환경에서 로봇을 30초의 시간동안 구동시켜 물리 기반 시뮬레이션을 수행한다. 시뮬레이션이 끝난 후, 로봇의 이동거리 및 구성 모듈 개수를 측정하여 적합도를 계산한다. 진화 연산을 통해 20개의 해를 생성하고, 무작위로 5개의 새로운 해를 생성하여 기존의 모집단을 교체하여 시뮬레이션을 반복하였다. 각 세대별 높은 적합도를 보인 결과는 Figure 1에서 확인할 수 있다.

5. 결론

현재 연구 단계에서는 모듈의 움직임은 규칙, 반복적으로 로봇의 형태에 대해서만 진화를 수행하였다. 로봇의 상태 및 주변 환경을 감지하여 동적으로 움직임을 수행하도록 모듈의 제어부도 진화 연산을 수행하면 좀 더 최적화된 모듈형 로봇을 만들 수 있을 것으로 기대한다 [2].

참고 문헌

- [1] Veenstra, F., Faina, A., Risi, S., & Stoy, K. (2017, April). Evolution and morphogenesis of simulated modular robots: a comparison between a direct and generative encoding. In *European Conference on the Applications of Evolutionary Computation* (pp. 870-885). Springer, Cham.
- [2] Sims, K. (1994, July). Evolving virtual creatures. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques* (pp. 15-22). ACM.

3D 사운드에 반응하여 소리가 발생했다고 생각되는 방향을 보고 대응 사격을 하도록 지시하였다. 6개의 3D사운드에 대한 대응 사격을 하는 동안 발생한 오차 값을 이용하여 보정 값을 생성한다. 보정 값을 적용하여 다시 6개의 3D 사운드에 대한 대응 사격을 진행한다.

초기 실험을 통해 발생한 오차와 보정 값이 적용된 실험에서 발생한 오차 값을 비교하여 보정 값이 존재할 때 성능이 향상되었는지 확인한다.



그림 4 사용자가 보는 실험 화면

2.3. 실험 결과

실험을 통해 얻은 값들의 평균 제곱 오차 값으로 보정 값을 정하였다. 오차가 나는 절댓값의 크기가 5미만인 경우 0.1의 값으로, 5이상 15미만인 경우 1.0의 값으로 그리고 15이상인 구간은 1.5의 값으로 가중치를 정하였다.

대응 사격 시 소리가 난 지점과 다른 방향을 가리킬 때 정 위치에서 벗어난 각도의 방향에 반대 방향으로 평균 제곱오차 값과 가중치를 곱한 값을 적용하여 오차를 보정 하였다.

표1의 결과를 보면 좌측 초기 실험에 비하여 실험 데이터를 이용해 보정 값을 적용한 실험의 결과에서 평균 오차가 줄어드는 것을 확인할 수 있다.

표 1 3D 사운드 실험 결과(평균 제곱 오차 값, 단위: 도)

	N	좌측 오차	우측 오차
초기 실험	10	10.864	11.615
보정 값 적용 실험	10	8.017	2.803

3. 결론

우리는 연구를 통해서 3D 환경에서 높이의 차이가 없이 좌, 우측 방향에서 들리는 3D 사운드를 듣고 사용자

가 가리키는 지점과 실제 소리가 발생한 위치와의 차이가 있음을 발견하였고, 평균 제곱 오차를 이용해 보정해주었다.

그러나 본 연구에 참가한 실험자의 집단이 작았기 때문에 실험의 결과가 정확하다고 단정 지을 수 없다. 때문에 향후 본격적인 연구에서는 실험자 집단의 수를 늘려 객관성을 확보 해야겠다. 또한 몰입형 가상현실 환경으로 환경을 확장하여 좌, 우측 방향뿐만 아니라 높이의 차이까지 고려된 사운드 보정 시스템을 개발하여 실험을 진행해야 할 것이다. 이 시스템을 이용하여 몰입형 가상현실 환경에서 진행되는 게임 플레이어가 조금 더 원활한 게임을 진행할 수 있을 것으로 기대한다.

참고문헌

[1] Shyam Prathish Sargunam; Kasra Rahimi Moghadam; Mohamed Suhail; Eric D. Ragan, Guided head rotation and amplified head rotation: Evaluating semi-natural travel and viewing techniques in virtual reality, *IEEE Virtual Reality (VR)*, pp.19-28, 2017

[2] Zhi-Chao Dong; Xiao-Ming Fu, Chi Zhang, Kang Wu, Ligang Liu, Smooth Assembled Mappings for Large-Scale Real Walking, *ACM Transactions on Graphics(SIGGRAPH Asia)*, 36(6), 2017

[3] 정대영; 서재일; 이용주; 유재현; 박태진; 이태진, UHDTV를 위한 실감 오디오 재현 기술, *방송 공학회 방송공학회논문지*, 20(1), pp.68-81, 2015

[4] 강호준; 박선빈; 김지수; 박병화, FPS 게임의 성공 요인 분석 : 배틀그라운드, 카운터스트라이크, h1z1를 통한, *한국지능정보시스템학회*, 98-99, 2017

[5] 이재현, *절대강좌! 유니티5 Unity 전문 개발자가 알려주는 효과적인 게임 제작 기법*, pp.211-218, 2016

VR 인터랙션을 위한 효과적인 자세 인식 및 키 입력 생성 방법*

권혁주⁰, 권민철, 우다라, 이상훈
연세대학교 전기전자공학과

(hyuckjookwon, yhsf99, udara27, slee)@yonsei.ac.kr

Effective Pose Recognition and Key Input Generation Methods for VR Interaction

Hyuck-Joo Kwon⁰, Mincheol Kwon, Udara Darshana Panamulle Arachchige, Sanghoon Lee
School of Electrical and Electronic Engineering, Yonsei University

요약

본 논문에서는 자유로운 VR 인터랙션을 위해 다수의 Kinect를 기반으로 사용자의 자세를 인식하고 키 입력을 생성하는 효과적인 방법을 제안한다. 제안된 자세 인식 알고리즘은 사용자와 카메라 간의 특징을 효과적으로 반영하여 자세 인식 성능을 향상시켰다. 또한 키 입력 생성은 94.7%의 모션 인지 정확도와 274ms의 낮은 인지 지연시간으로 수행되었다. 이런 실험결과를 통해, 우리는 제안된 기법이 다양한 VR 어플리케이션에 적용될 수 있음을 확인하였다.

1. 서론

현재 대부분의 VR 인터랙션은 모션 캡처 장비를 몸에 장착하고 주변의 센서들이 이를 감지하여 사용자가 어떠한 행동을 하고 있는지 인식하는 방법으로 수행되어 왔다. 이러한 장비 착용 기반의 인터랙션은 사용자가 자유로운 자세를 취하는데 있어서 방해요소가 될 수 있다. 최근 발표된 연구에서는 이런 문제를 해결하기 위하여 multi-view 기반의 VR 인터랙션 환경을 제안하였다 [1][2]. Multi-view 기반의 시스템은 다수의 카메라로부터 얻은 정보를 기반으로 fusion을 수행하여 자세를 인식한다. 따라서 높은 정확도를 갖는 카메라 정보를 사용하는 것이 매우 중요하다. 본 논문에서는 이런 특성을 이용하여 자세 인식의 정확도를 향상시키고 VR 인터랙션을 효과적으로 제공할 수 있는 기법을 제안한다.

2. 제안하는 자세 인식 기법

기존 자세인식 기법에서는 다수의 카메라 fusion 문제를 촬영 중인 여러 대의 Kinect에서 출력된 skeleton과 fusion skeleton의 차이를 구하고 모든 오차의 합이 최소가 되도록 최적화 기법으로 해결하였다. 하지만 각각의 Kinect는 정확하지 않은 skeleton 정보를 제공할 수 있으므로 모든 Kinect의 좌표 정보를 이용하는 기존 방식은 자세인식 정확도를 저하시킬 수 있다. 본 논문에서는 이런 문제를 해결하기 위해 자세인식의 정확도를 향상시킬 수 있는 Kinect를 선택하고 fusion 시 해당 Kinect

정보만을 사용함으로써 전체적인 성능을 향상시켰다.

제안하는 자세인식 기법은 [1][2]에서와 같이 Kinect의 특성을 반영하여 사용자의 전방벡터를 찾고 후면에서 촬영된 데이터 및 관절 상태 신뢰 값이 낮은 정보에 대한 페널티를 부여한다. 그 다음 신체부위별 반대편에 있는 Kinect에 페널티를 부여하는 과정을 거친다.

그림 1과 같이 신체 부위 반대편에 있는 Kinect의 경우 다른 신체부위에 의해 촬영하고자 하는 관절부위가 가려짐으로써 정확도가 낮을 수 있다. 범위는 전방벡터 기준 왼쪽 신체부위의 경우 $-90^{\circ} \sim 30^{\circ}$, 오른쪽 관절일 경우 $-30^{\circ} \sim 90^{\circ}$ 사이에 들어올 때 해당 방향의 Kinect로 판단한다. 또한, 사용자의 모션에 따라 상체와 하체의 바라보는 방향이 다를 수 있으므로 상체와 하체의 전방벡터를 따로 산출하여 상, 하체 별로 반대편 Kinect를 각각 선택한다. 이런 방식을 통해 해당 범위 안에 들어오는 Kinect 정보들에 대해 평균값 가중치를 취해줌으로써 fusion 과정을 진행하고 최종 skeleton을 생성한다.

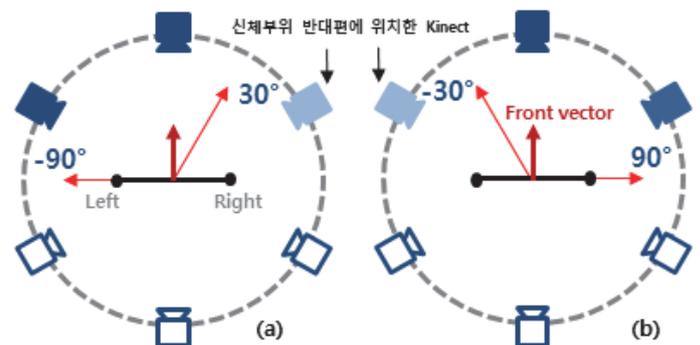


그림 1: 신체부위의 반대편에 위치한 Kinect의 예.
(a) 왼쪽 신체부위 기준. (b) 오른쪽 신체부위 기준

3. 제안하는 키 입력 생성 기법

3.1. Feature extraction

사람의 모션 인지를 수행하는데 있어서, 이전 다수의 연구에서는 3차원 관절위치와 그 관절의 이동속도 관련 정보가 사용되었다. 본 논문에서는 이런 모션 인지에 유용한 관절위치, 관절의 이동속도, 그리고 frame에 따른 관절의 위치와 이동속도들의 평균 정보를 기반으로 300 차원의 feature를 생성하였다.

* 포스터 발표논문

* 본 연구는 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2016-0-00204, 극사실적인 실시간 가상현실을 위한 모바일 GPU 하드웨어 개발).

표 1: 제안하는 기법을 통해 측정된 자세 오차 및 키 입력 성능 표

		Change Pistol	Change Rifle	Change Grenade	Change Sword	Throw High	Throw Low	Reload Pistol	Shoot Sword	Lean	Open	Average
Pose	[1](cm)	90.8	90.9	62.6	62.7	67.9	109.1	77.5	113.8	80.4	71.4	75.28
	ours(cm)	47.0	59.5	51.5	59.9	56.9	108.1	62.7	72.2	80.2	57.8	65.58
	improvement(%)	48.2	34.5	17.7	4.5	16.2	0.9	19.1	36.6	0.2	19.0	12.9
Key Input	accuracy(%)	87.5	100	100	100	88.1	90.7	83.7	100	100	97.6	94.7
	delay(frame)	12.4	7.1	5.9	8	7.2	8.9	10.4	8.1	7.8	7.6	8.3

3.2. Long Short-Term Memory(LSTM) 모션 인지 모델

그림 2는 제안된 LSTM 기반의 모션 인지 모델이다. 우리는 모션 인지 처리 시간에 따른 인터랙션의 delay 발생 시간을 최소화하기 위하여 LSTM 모델을 간결하게 구성하였다. 제안된 모델은 연속되는 모션의 복잡한 관계를 학습하기 위하여 2개의 LSTM cell이 사용되었으며 평균 30ms로 수행된다. 우리는 모션 인식을 위하여, 순차적인 자세 정보(30 frame)와 이에 대한 모션 label 정보를 기반으로 학습을 수행하였다. 여기서 인터랙션이 발생되지 않는 모션은 “모션 없음”으로 학습된다.

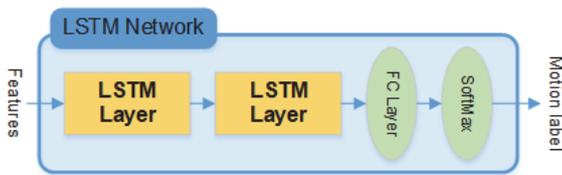


그림 2: 제안하는 LSTM 기반의 모션 인지 모델

3.3. 키 입력 생성 알고리즘

사람의 움직임(던지기, 열기 등)에 대한 모션 인지 결과는 연속되는 시퀀스 형태로 출력된다. 우리는 이런 시퀀스의 결과를 기반으로 VR 인터랙션을 수행하기 위하여, 그림 3과 같은 방법으로 특정 모션 구간을 대표할 수 있는 키 입력을 생성하였다.

먼저, 제안하는 키 입력 생성 알고리즘은 이전 30 frame의 자세 정보를 취합하여, 현재 frame에서 모션 정보를 인지한다. 만일, 최근 인지된 10 frame 중 동일한 모션 label이 70% 이상 발생된다면, 그 label이 유효한 값을 갖는지 확인한다. 여기서 유효한 값이란 VR 인터랙션을 위한 키 입력을 발생시키는 모션 label을 의미한다. 유효한 값을 갖는 경우 해당 모션은 키 입력으로 생성된다.

우리는 동일한 모션 구간에서 다수의 키가 입력되는 것을 방지하기 위하여 enable check 단계를 수행하였다. Enable check는 키 입력이 수행되면 “모션 없음” label이 70% 이상 발생되기 전까지 키 입력을 어플리케이션으로 전송하는 것을 제한하는 방식으로 동작된다.

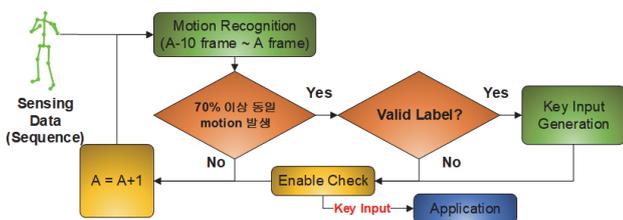


그림 3: 키 입력 생성 기법의 처리 흐름도

4. 실험 결과

우리는 성능 분석을 위하여 기존에 연구된 360° 가상 현실 트레이닝 시스템의 환경에 제안하는 기법을 적용하고 pose와 모션 인지 정확도에 대한 실험을 수행했다. 실험 DB는 특정 10개 모션에 대한 자세의 시퀀스로 이뤄진 3D Kinect skeleton 데이터를 사용하였다.

제안하는 자세 인식 기법의 정확도는 Opti-track으로 촬영한 Ground Truth와 Kinect로부터 도출한 fusion skeleton의 오차값으로 성능을 측정하였다. 자세 인식에 대한 오차값은 프레임별로 오차를 계산하고 모든 프레임별 오차의 합을 평균 내어 산출하였다. 제안된 기법은 [1]기법 대비 모든 자세에서 정확도가 향상된 되었고 평균 9.7cm개선 되어 12.9%의 성능향상을 보였다.

제안하는 모션 인지 기법은 평균 94.7%의 키 입력 정확도와 8.3 frame delay의 성능을 보였다. 여기서 delay는 유효 모션의 시작부터 키입력이 생성된 시점까지의 frame 개수로 판단한다. 본 논문에서 사용한 키넥트 센서의 프레임 취득 속도가 30 FPS인 것을 반영하면, 제안하는 기법은 사용자에게 274ms의 낮은 지연시간으로 몰입감 있는 VR 인터랙션을 제공할 수 있다.

5. 결론

본 논문에서 제안한 기법은 게임, 가상훈련, 테마파크 등 다양한 산업에 적용 가능 하지만, kinect를 기반으로 사용자의 자세 정보를 취득하기 때문에(30 FPS), 고속의 인터랙션을 요구하는 스포츠 분야(골프, 야구 등)에 적용하는데 한계가 존재한다. 추후 연구에서는 고속촬영 가능한 카메라 시스템을 기반으로 제안하는 기법의 성능을 향상 시킬 것이다.

참고문헌

[1] Beom Kwon, Junghwan Kim, Kyoungoh Lee, Yang Koo Lee, Sangjoon Park, and Sanghoon Lee, "Implementation of Virtual Training Simulator Based on 360° Multi-view Human Action Recognition," *IEEE Access*, vol. 5, pp. 12496-12511, 2017.

[2] Junghwan Kim, Inwoong Lee, Jongyoo Kim, and Sanghoon Lee, "Implementation of an Omnidirectional Human Motion Capture System using Multiple Kinect sensors," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 98, no. 9, pp. 2004-2008, Nov. 2015.

가상 교실 내 사회적 상호작용 종류에 따른 정보처리 촉진에 관한 연구: 동공 추적을 중심으로*

유재환⁰, 허지웅, 김은주, 김광욱
한양대학교 컴퓨터·소프트웨어학과, 연세대학교 정신건강의학과
(zealot0724, heo0614, kenny)@hanyang.ac.kr, EJKIM96@yuhs.ac.kr

Research on Information Processing of Virtual Classroom According to Social Interaction Types with Eye-tracking Technology

Jaehwan You⁰, Jiwoong Heo, Eunjoo Kim, Kwanguk(Kenny) Kim
Dept. of Computer Science and Engineering, Hanyang Univ,
Dept. of Psychiatry and Institute of Behavioral Science in Medicine, Yonsei Univ.

요약

가상교실은 주의력 결핍 과잉행동 장애, 자폐 스펙트럼 장애 등의 평가와 훈련에 활용되는 중요한 가상현실 응용프로그램이다. 그러나 기존 연구에 따르면 공동주의를 포함한 가상 교실내 사회적 상호작용에 대한 연구는 매우 제한적이다. 본 연구에서는 가상교실내에서 공동 주의가 교사 아바타와의 상호작용에 어떠한 영향을 주는지 확인하기 위해, 상호작용 종류에 따른 공동주의 효과를 연구하고, d-prime, 전체 머리 움직임, 전체 눈 움직임을 측정하였다. 본 연구 결과에 따르면, 상호작용의 종류에 따라 피험자의 학습 효과, 눈 및 머리의 움직임양이 통계적으로 유의미한 차이가 있음을 확인하였다. 이러한 연구 결과는 향후 환자군 연구를 위해 중요한 시사점을 가질 것으로 보인다.

1. 서론

주의력 결핍 과잉행동 장애, 자폐 스펙트럼 장애를 가진 학령기 아이들은 학교에서 시간을 많이 보낸다는 점에서 기존 연구진들은 가상교실을 그들의 행동 패턴 그리고 인지 과정 평가의 도구로 사용을 했다[1]. 이러한 연구들은 가상교실이 실제 교실과 유사하다는 점에서 생태학적 타당성을 가지고 있지만, 실제 학교생활과 유사한 사회적 상호작용이 없다는 점에서 한계를 가지고, 이전 연구에서 교실 내에 상호작용 방식 중 하나인 공동주의를 추가한 연구가 있다[2]. 공동주의란 둘 이상의 사람이 서로의 주의를 공유한 상태에서 사물 또는 사람에 공유된 주의를 형성한 것을 의미한다. 눈, 머리, 손 등의 행동을 통해서 다양하게 아바타와 상호작용할 수 있지만, 기존 연구는 이러한 점을 고려하지 않았고, 측정 변인 측면에서 사회적

상호작용의 중요한 요소 중 하나인 눈 움직임[3]을 고려하지 않았다. 따라서, 본 연구는 가상교실 내에서 교사 아바타의 사회적 상호작용 방법에 따라서 피험자가 효과적으로 상호작용을 할 수 있는지를 동공 추적을 포함하여 연구하고자 한다. 환자군 적용에 앞서 정상 성인군 대상으로 User Study를 수행하였으며, 연구 가설은 다음과 같다. (1) 교사 아바타의 상호작용 종류에 따라 피험자의 정보처리에 주는 영향이 다를 것이다. (2) 기존 연구를 근거하여 전체 머리 움직임 양과 전체 눈 움직임 양이 교사 아바타의 상호작용 종류에 따라 차이가 있을 것이다.

2. 실험 내용

2.1. 피험자

본 실험은 피험자를 모집하기 위해서 교내 광고를 통해 총 20명의 피험자 (Age: M=24, SD = 2.67; Female 50%)를 모집을 하였다. 실험 전 SCL-90-R (Symptom Checklist-90-Revised)[4] 통해서 정신 질환이 없음을 확인하였다.

2.2. 장비 및 개발 환경

본 실험은 3D 개발 플랫폼인 Unity 2017.1.0f3 (Unity Technologies, USA)을 사용하여 개발을 하였다. 피험자들은 Oculus rift CV1(Oculus, USA) HMD(Head Mounted Display; HMD), 그리고 Tobii Pro Glasses2 (Tobii Technology, USA) Eye tracker를 착용하였다.

2.3. 가상교실 과제 및 실험 절차

본 실험에 사용된 과제는 가상교실 환경에서 진행이 되며 Learning phase와 Testing phase 순으로 진행을 한다. Learning phase는 가상교실 내 교사 아바타가 TV를 가리키며 총 12번의 과제를 진행을 한다. 피험자들은 그림 자극을 외워야 하며, Testing phase에서 학습된 자극을 기억하는지 확인한다. 자극은 12개의 Target 그림, 12개의 Non-target 그림, 그리고 12개의 Novel 그림을 포함하여 총 36개의 그림이 나온다. 모든 피험자들은 상호작용 방법에 따라 Eye

* 포스터 발표논문.

* 이 논문은 2017년도 정부(미래창조과학부)의 재원으로 한국연구재단-과학기술인문사회융합연구사업의 지원을 받아 수행된 연구임(2017M3C1B6071069), Correspondence to K. Kim (Kenny@hanyang.ac.kr)

조건, Head 조건, 그리고 Gesture 조건 총 3가지 조건을 counter balance하여 진행을 한다. 실험 전 연습시행을 수행하였고, 각 조건진행전 정확한 동공추적을 위해서 calibration을 진행을 하였다. 매 조건 후 2분의 쉬는 시간을 가졌으며 총 실험은 50분이 소요되었다.



그림 1: 가상교실 환경

3. 결과

3.1. d-prime 값

d-prime (계산식: $Z(\text{Hit Rate}) - Z(\text{False Alarm})$; Z-score: Gaussian 분포의 역)은 피험자들이 얼마나 그림을 잘 기억을 했는지에 대한 척도로 사용하였다. ANOVA 분석한 결과 d-prime은 조건간 주 효과가 있었다 ($F(2, 38)=8.90, p<0.01, \eta^2 = 0.319$). 사후 검증 결과 Eye 조건과 Head 조건 ($p<0.001$) 그리고 Eye 조건과 Gesture 조건 ($p<0.01$)에서 유의미한 차이가 있었다.

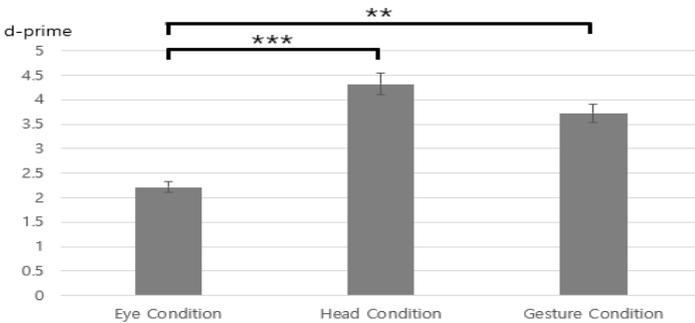


그림 2: 각 조건 별 d-prime

참조: ** $p < 0.01$; *** $p < 0.001$.

3.2. 전체 머리 움직임

전체 머리 움직임은 HMD에 내장된 센서를 통해 Yaw, Pitch, Roll를 1/60초 마다 기록을 하여 프레임별 변화량을 합하여 구했다. 데이터 저장의 문제로 2명은 제외했다. ANOVA 분석한 결과 조건간 주 효과가 있었다 ($F(2, 34)=3.540, p<0.05, \eta^2 = 0.172$). 사후 검증 결과 Eye 조건과 Hand 조건($p<0.05$)에서 유의미한 차이가 있었다.

3.3. 전체 눈 움직임

전체 눈 움직임은 eye tracker를 통해 1/60초 마다 피험자가 바라보는 X좌표, Y좌표를 기록하여 프레임별 변화량을 모두 합하여 구했다. ANOVA 분석한 결과 조건간 주효과가 있었다 ($F(2, 38)=8.530, p<0.01, \eta^2 =$

0.310). 사후 검증 결과 Eye 조건과 Head 조건 ($p<0.01$) 그리고 Eye 조건과 Gesture 조건 ($p<0.01$)에서 유의미한 차이가 있었다.

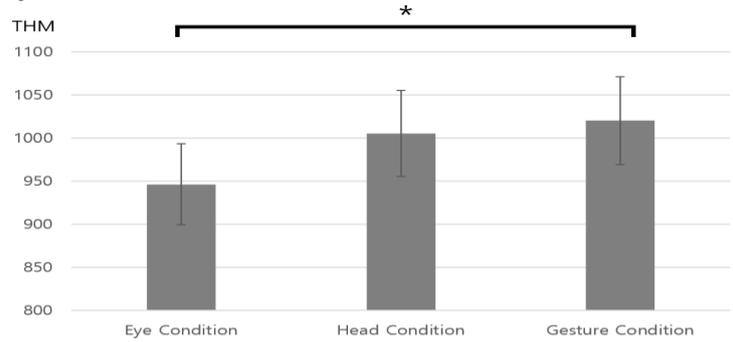


그림 3: 각 조건 별 THM(Total Head Movement)

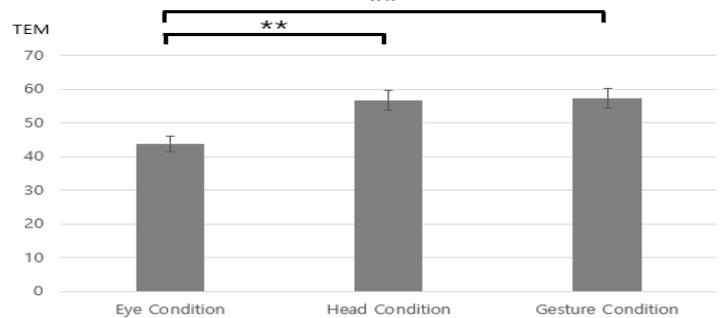


그림 4: 각 조건 별 TEM(Total Eye Movement)

참조: * $p < 0.05$; ** $p < 0.01$.

4. 논의

본 연구의 결과에 따르면 교사 아바타의 상호작용 방법에 따라 정보처리 축진의 정도가 다르다는 것을 확인하였다. 또한 기존 연구에서는 공동주의를 위해서 머리의 움직임이 정보처리에 긍정적인 영향을 준다는 결과를 확인하였다[2]. 본 연구 결과는 여기에 사회적 상호작용 요소 중 하나인 피험자의 눈 움직임을 확인한 결과 전체 눈 움직임이 아바타의 상호작용 방법에 따라 다름을 확인하였고, 특히 d-prime 점수가 높을수록 눈의 전체 움직임이 더 많아 진다는 점을 확인하였다. 따라서 피험자의 전체 눈 움직임은 가상교실내 공동주의 연구에 있어 활용되어야 할 핵심 지표이며, 향후 피험자 연구를 위한 평가 지표로써 가치를 가짐을 확인하였다.

참고문헌

[1] Rizzo, A. A., Bowerly, T., Buckwalter, J. G., Klimchuk, D., Mitura, R., & Parsons, T. D., A virtual reality scenario for all seasons: the virtual classroom, *CNS Spectrums*, 35-44, 2006.
 [2] 허지웅, 김은주, 김광욱, 공동주의 가상교실을 이용한 ADHD 아동군과 정상 대조군 연구, 한국HCI학회 학술대회, 599-602, 2017.
 [3] Bayliss, A. P., Paul, M. A., Cannon, P. R., Tipper, S. P., "Gaze cuing and affective judgments of objects: I like what you look at," *Psychonomic Bulletin & Review* 13(6), 1061-1066, 2006.
 [4] 김광일, 원호택, 이정호, 김광윤, 간이정신 진단검사 (SCL-90-R)의 한국판 표준화연구, *신경정신 의학* 17 (4) : 449-458, 1978.

360도 비디오 시청에 대한 유저 행동양식 분석*

유정은⁰, 노준용
 카이스트 비주얼미디어연구소
 (jey920, junyongnoh)@kaist.ac.kr

Analyzing User's 360° Video Watching Behavior

Jungeun Yoo⁰, Junyong Noh
 KAIST, Visual Media Lab.

요약

본 연구는 360도 영상 시청 시 유저가 보일 행동패턴을 정보탐색행동 연구에 기반하여 정의하고 그 행동이 360도 영상 시청경험에 주는 영향을 분석하였다. 360도 영상의 보이지 않는 영역의 정보를 찾아 나서는 행동을 브라우징(browsing), 찾아낸 흥미로운 정보에 집중하는 행동을 모니터링(monitring)이라고 정의하였다. 실험을 통해 사용자들이 360도 영상 시청 시 해당 행동을 보이는 것을 확인하였고 이 행동들이 360도 영상 시청 경험의 만족감 및 몰입감으로 이어지는 것을 확인하였다.

1. 서론

360도 영상은 전 방향으로 시각적인 정보를 포함하고 있어 2D 영상보다 몰입감 있고 실존감 있는 영상 시청을 가능하게 한다. 하지만 인간의 시야각은 제한적이기 때문에 360도 영상을 시청할 때 전체의 일부만을 받아들일 수 있다. 이를 보완하기 위해 360도 영상 시청자들은 눈을 이용한 시선의 움직임뿐만 아니라 마우스 또는 머리의 움직임을 통해 시야를 움직여 보이지 않은 영역을 탐색한다. 하지만 여전히 한 번에 받아들일 수 있는 시각적 정보량은 시야각에 제한되어있고, 360도 영상 관람시 불가피하게 정보를 놓치는 부분이 생기게 된다.

360도 정보의 경험적 손실을 줄이기 위한 방안 중에는 가장 중요한 정보만을 보여주는 것이 있다 [1]. 이 경우 중요한 정보의 손실은 최소화되지만 그 외의 정보는 모두 무의미해지는 단점이 존재한다. 한편, 유저가 영상을 시청하는 행동패턴에 따라 화면을 구성하여 간접적으로 정보 손실을 줄이고 360도 영상만이 가지는 경험을 극대화하는 방법도 있다. 이를 위해선 유저가 360도 영상 시청시 어떤 행동을 하는지, 그 동기가 무엇인지, 그리고 어떤 부분에서 만족감을 느끼는지에 대한 연구가 필요하다.

본 논문에서는 정보탐색행동 연구에 기반하여 유저들의 360도 시청패턴이 크게 정보 탐색과 집중으로 구성되며 이 행동들의 결과가 유의미한 정보습득으로 이어지면 시청경험에 긍정적인 영향을 줄 것이라는 가설을 세웠다. 제시한 가설들은 실험과 인터뷰를 통해 확인하였다.

2. 실험 구성

2.1. 정보탐색행동 연구 기반 가설 설정

정보탐색행동 관련 연구에 의하면 사람은 정보가 부족한 상황에서 호기심과 불확실함을 경험한다. 이를 해소하고자 사람은 명확한 목표 없이 정보 탐색에 나선다. 그런 과정에서 집중하고자 하는 정보를 찾으면 비로소 불확실함이 사라지게 된다 [2, 3].

이와 같은 행동양식이 방대한 시각정보로 이루어진 360도 영상 시청 과정에서도 나타날 것이라 예측하였다. 360도 영상 속 보이지 않는 부분에 무엇이 있는지 알지 못하더라도 유저들은 호기심에 의해 그것을 찾아 나설 것이며 (브라우징), 그로 인해 집중할 (모니터링) 정보를 찾으므로써 만족감을 느낄 것이다.

2.2. 실험 환경

가설 검증에 필요한 데이터 수집을 위해 다음과 같은 실험이 이루어졌다. 22살에서 32살 사이의 360도 영상 시청 경험이 있는 21명의 실험 참가자들이 각각 10개의 소리 없는 영상을 무작위로 시청하였고, 그들이 시청하면서 보인 머리의 움직임은 HMD를 통해 기록되었다. 영상들은 실사 또는 가상 공간에서 촬영된 홍보, 다큐멘터리, 영화 속 장면, 그리고 액션 스토리 영상으로 이루어졌으며 시청 후에는 인터뷰가 진행되었다.

3. 결과 분석

3.1 데이터 분석 방법 및 결과

실험 데이터들을 표 1과 같이 영상 구간별로 나누어 비교하였다. 각 상황에 해당하는 영상 구간들의 브라우징

* 포스터 발표논문

* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2018년도 문화기술연구개발 지원사업으로 수행되었음.

또는 모니터링 정도를 측정하고 그 차이가 유의미한지 비교하였다. 브라우징의 경우 시야 방향의 변화에 따른 중심각의 크기를 구해 유저별로 평균을 낸 뒤 두 그룹 간의 유의미한 차이를 윌콕슨 순위합 검정을 통해 확인했다. 모니터링의 경우 유저들의 시야 방향의 밀집도를 분산을 통해 측정하였다. 두 그룹의 측정된 값들이 유의미한 차이를 갖는지 독립표본 T검정을 통해 확인했다.

표 1 각 행동방식이 나타나는 영상 구간 및 검정 결과

상황	브라우징(탐색)		모니터링(집중)	
	새로운 정보 등장	맥락적으로 새로운 정보 암시	흥미로운 정보 발견	흥미로운 움직임 발견
해당 구간	컷 시작부터 3초	야외 배경	주요 피사체 존재	카메라 움직임 존재
대조 구간	컷 시작 3초 후	실내 배경	주요 피사체 없음	카메라 움직임 없음
차이 (p)	0.1058 rad/frame (0.0089)	0.1133 rad/frame (3.13e-08)	-31.29 (2.63e-06)	-130 (0)

분석 결과 유저들은 영상에서 컷이 발생해 새로운 정보가 주어졌을 때의 초반 3초 동안 더 활발하게 시야를 움직였다는 것을 알 수 있었다. 또한 유저들은 시야 밖의 정보에 대한 맥락적 암시가 주어지는 야외 배경 장면에서 더 활발하게 시야를 움직였다는 것을 알 수 있었다. 주요 피사체 또는 카메라 움직임이 존재하는 장면에서는 유저들의 시야 방향이 더욱 밀집되어 있다는 것을 알 수 있었다. 카메라 움직임의 유무는 주요 피사체의 유무보다 밀집도의 차이가 큰데, 이는 카메라 움직임 방향으로 새롭고 흥미로운 정보가 유입되어 유저의 집중을 지속시켜주기 때문이라고 여겨진다.

3.2 인터뷰 결과 및 분석

정량적 분석을 통해 특정한 상황 속에서 시야의 움직임 또는 밀집도의 크고 작음을 알 수 있었다. 그러한 행동이 나타난 이유에 대해 알아보고자 인터뷰를 통해 무엇이 유저로 하여금 시야를 한곳에 집중하게 했는지, 다른 곳으로 시야를 돌리게 했는지 질문하였다. 또한 영상 안에서 360도 영상으로써 제일 좋았던 장면과 기대에 못 미치는 장면에 대해 질문하였다.

첫 번째 질문에 대하여 많은 유저들이 호기심을 이유로 내세웠고, 두 번째 질문의 경우 ‘흥미’와 ‘기대감’이 시야를 한 곳에 집중하게 하는 주된 이유로 나타났다.

만족스러운 장면의 경우 유저들마다 선택한 장면이 달랐지만 선택의 이유에서 ‘다양하게 둘러 볼거리가 있다’라는 유사점을 찾을 수 있었다. 반면 불만족스러운 장면으로 자주 언급된 영상은 *pokemon in 360 degree video*였고 ‘둘러볼 것이 없었다’가 주된 이유였다.

인터뷰 결과 ‘둘러본다’, 즉 브라우징 행동이 유저의 360도 영상 시청경험에 영향을 준다는 것을 알 수 있었다. 불만족스러웠다고 언급된 영상 장면들은 그림 1(b)와 같이 주요 정보가 찾기 쉬운 중앙에 위치해있지만 그 외의 부분에는 흥미로운 요소가 존재하지 않아 유저로 하여금 둘러볼 필요가 없게끔 구성되어있었다. 보기

엔 편하지만 그것이 영상을 시청하는 유저들 다수에게 긍정적인 영향을 주지 않았다는 것이다.

이는 브라우징 행동을 통해 얻은 정보들이 실존감 증가에 기여하기 때문이라고 생각할 수 있다. 실존감은 유저가 가상공간 안에서 자신이 실제로 ‘존재’한다고 여기는 감각을 뜻한다. 첫 번째 질문에서 한 유저는 “내가 어디 있는지, 뭘 하고 있는지 알고 싶었다”라고 답하였다. 즉 유저는 자신을 영상 속 시점의 주인공과 동일 시 하여 브라우징을 통해 영상 속 ‘내’가 마주친 상황에 대한 정보를 알아가고 싶어 하는 것이다. 브라우징으로 정보를 얻으므로써 영상 속 자신을 확립해나가고 이는 실존감, 더 나아가 몰입감의 증가로 이어진다. 그림 1(a) 장면들의 경우 유저는 브라우징으로 얻은 정보를 통해 자신을 잠수부, 돼지위의 야만인, 또는 열기구 속 여행자로 정체화한다. 반면 그림 1(b)의 경우 브라우징 할 유의미한 정보가 부족하고 유저는 스크린 뒤 제3자에 머문다. 이는 실존감과 몰입감의 하락으로 이어지고 유저로 하여금 360도 영상이 주는 매력을 느끼기 힘들게 한다.

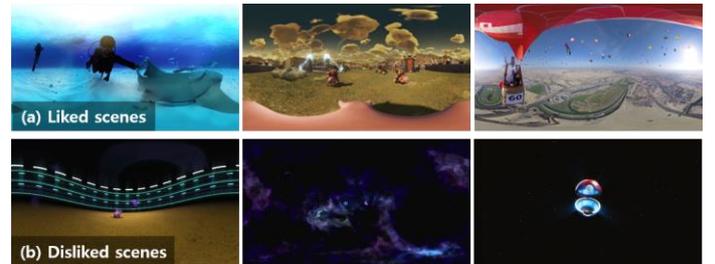


그림 1 만족스러운 장면들과 불만족스러웠던 장면들

4. 결론

본 연구는 360도 영상 시청시 유저들이 보이는 행동패턴을 조사하고 어떤 요소가 경험적 만족도에 영향을 주는지 알아보았다. 그 결과 보이지 않는 영역에 대한 ‘호기심’과 ‘불확실함’이 영상을 탐색하는데 원동력이 되었으며 이는 새로운 장면이 시작되는 것이나 맥락적으로 정보를 암시하는 야외 배경에서 자주 나타나는 것을 확인하였다. 눈에 띄는 주요 피사체나 카메라의 이동과 같은 흥미로운 요소가 있을 땐 탐색을 멈추고 집중한다는 것 역시 실험을 통해 검증하였다. 인터뷰 결과 탐색을 통한 유의미한 정보 습득은 유저로 하여금 실존감을 느끼고 더욱 몰입하게 만든다는 것을 알아냈다.

참고문헌

[1] Yu-Chuan Su and Kristen Grauman. 2017. Making 360° Video Watchable in 2D: Learning Videography for Click Free Viewing. In *The IEEE Conference on Computer Vision and Patter Recognition (CVPR)*, July 2017.

[2] Carol C Kuhlthau. 1991. Inside the serach process: Information seeking from the user’s perspective. *Journal of the American society for information science* 42, 5 (1991), 361.

[3] Nicholas J Belkin. 1980. Anomalous states of knowledge as a basis for information retrieval. *Canadian journal of information science* 5, 1 (1980), 133–143.

가상 현실을 이용한 일상생활활동 과제: Virtual Supermarket을 통한 인지기능 훈련 도구 개발

김석현⁰, 진주희, 김광욱
한양대학교 컴퓨터소프트웨어학과

ksh4086@hanyang.ac.kr, jun.joohee@gmail.com, kenny@hanyang.ac.kr

Activities of Daily Living Task Using Virtual Reality: Training of Cognitive Function through Virtual Supermarket

Suckhyun Kim⁰, Joohee Jun, and Kwanguk (Kenny) Kim
Dept. of Computer and Software, Hanyang University, Seoul, Korea

요약

경도인지장애는 치매 증상은 없지만 인지기능이 저하된 상태를 말한다. 때문에 일상생활에서의 활동능력이 일반인들보다 떨어진다. 본 연구에서는 가상현실을 사용하여 경도인지장애 환자들이 훈련할 수 있는 일상생활 활동 과제인 Virtual Supermarket을 개발하고, 이를 정상군을 대상으로 검증하였다. 본 실험에서 주어진 과제의 요소들 간에 일정한 상관관계를 발견하였으며, 정상군에서 유의미한 패턴을 확인하였다. 이러한 결과를 통해 개발된 Virtual Supermarket이 신뢰할 수 있는 도구를임을 보였다.

1. 서론

경도인지장애(Mild cognitive impairment)는 치매증상 없는 인지기능저하 상태를 나타내며 일상적인 노화과정에서 치매로 전환하며 보고되는 현상으로 간주된다[1]. 급속도로 고령화되는 현대 사회에서 경도인지장애는 매우 중요한 사회적 이슈이다. 인지기능저하는 치매로 발전할 수 있고 이는 사회의 경제적 비용이 높아짐을 의미한다. 우리나라는 OECD 국가 중 고령화가 가장 급속도로[2]로 진행되고 있기 때문에 인지기능저하에 대해 더욱 주의를 기울여야한다. 경도인지장애는 사고력, 언어능력, 기억력, 판단력 등 광범위한 영역에 나타날 수 있다. 일상생활활동이 인지기능과 직접적인 관련이 있기에 일상생활활동을 문제없이 해결하는 것은 중요하다[3]. 일상생활과 관련된 행동들을 가상현실기술을 이용하여 환자를 테스트하고 평가할 수 있다. 가상현실은 현실과 유사하지만, 현실에서 일어날 수 있는 위험을 제거한 안전한 상황을 제공할 수 있다. 또한 현실에서 경험할 수 없는 상황들을 체험할 수 있다는 장점을 지니고 있다.

일상생활활동 과제에는 요리, 운전 등 여러가지 과제가 있다. 위의 과제들 같은 경우 문화의 차이에 따른 어려움이 있을 수 있다. 이에 본 연구에서는 쇼핑을 과제로 한 Virtual Supermarket을 개발하였다. 일반인들을 대상으로

로 Virtual Supermarket을 진행, 본 실험이 일상생활활동 과제로서 신뢰할 수 있는 실험임을 보이고자 한다.

2. 실험

2.1 피험자

피험자는 총 32명이 참가하였다(여성; N=14). 피험자의 평균나이는 25.6세이다. Korean-Mini Mental State Examination(K-MMSE)[4] 검사결과, 평균 28.5점이며, 표준편차는 1.1점으로 인지기능에 이상이 없었다.

2.2 장비 & 개발환경

가상현실환경은 Unity 2017.1.1f1를 통해 개발되었다. 실험은 Head-Mounted Display 시스템들(HMD; VIVE, HTC)를 사용하여 진행하였다.



그림 1: Virtual Supermarket

2.3 Virtual Supermarket 과제

피험자는 HMD를 통해 Virtual Supermarket을 탐색하고, 조이스틱을 사용해 제품을 이동시킬 수 있다. 실험은 총 16개의 과제로 이루어져 있으며 총 3블록으로 나누어 진행했다.

1번 블록은 6개의 과제로 이루어져 있다. 1번부터 5번과제는 특정 지점 찾아가기, 물건 집어 보기 등 VR환경에 적응하는 과제이다. 6번은 주어진 목록과 지도를 통해 상품 3가지를 카트에 넣는 시행을 3번 반복한다. 2번 블록은 7번부터 12번으로 이루어져 있으며, 주어진 목록과 지도를 통해 특정 상품 5가지를 기억한 후 상품들을 카트에 넣는 것이다. 7번은 목록과 지도를, 8번은

* 포스터 발표논문

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW 중심대학지원사업의 연구결과로 수행되었음(2016-0-00023),

Correspondence to K. Kim (kenny@hanyang.ac.kr)

지도만, 9번은 목록만 다시 볼 수 있다. 10번은 가격을 카운터에 가서 결제까지 완료해야 한다. 11번은 쇼핑 중에 상품 목록이 1개 변하고, 12번은 상품 목록이 2개 변한다. 3번 블록은 4개의 과제로 이루어져 있다. 13번은 보라색 상품 5개를 카트에 넣는 것이다. 14번은 바나나 우유를 만들기 위한 재료 3가지를, 15번은 볶음밥을 만들기 위한 재료 3가지를 창의적으로 카트에 넣어야 한다. 16번은 서로 다른 5개의 품목에서 먹을 수 없는 물건 한 개씩, 총 5개를 카트에 넣는 것이다.

2.4 실험절차

피험자는 실험에 대한 설명을 들은 후 동의서를 작성하였다. 그리고, 나이, 성별 등의 신상정보를 받고 K-MMSE 를 작성하였다. 본 실험은 3개의 블록으로 이루어졌으며, 1번째 블록 후엔 5분 2번째 블록후엔 10분의 휴식이 주어졌다. 마지막 블록이 끝난 후 실험환경에 대한 설문을 진행한 후 실험을 종료하였다.

2.5 측정 요소

Trajectory 는 피험자가 가상공간에서 움직인 거리, Total Time 은 과제를 완료하는데 걸린 시간, Hand Trajectory 는 가상공간에서 조이스틱을 움직인 거리, Map Count 는 도움 지도를 실행시킨 횟수, Map Duration 도움 지도를 실행시켜 놓은 시간, List Count 는 도움 목록을 실행시킨 횟수, List Duration 은 도움 목록을 실행시켜 놓은 시간이다.

3. 결과

결과는 SPSS 를 사용하여 상관분석(Correlation Analysis)을 수행하였다. 표 1의 값은 각 Task 에서 유의미한 상관관계가 나온 횟수를 의미한다. Total Time 과 Trajectory 와의 상관관계를 분석한 결과 17개의 과제 모두에서 유의미한 상관관계가 나타났고, Total Time 과 Hand Trajectory 사이에도 17개의 과제 모두에서 유의미한 상관관계가 나타났다. Map Count 와 Map Duration 의 상관관계를 분석한 결과 도움 지도를 실행시킬 수 있는 과제 3개 모두에서 유의미한 상관관계가 나타났으며, List Count 와 List Duration 의 상관관계를 분석한 결과 도움 목록을 실행시킬 수 있는 과제 5개 모두 유의미한 상관관계가 나타났다.

정상 패턴의 예로 6번과제 Trajectory 의 analysis of variance(ANOVA) 분석을 진행한 결과 세개의 컨디션

사이에서 유의미한 차이를 보였다. $F(2, 62) = 39.89, p < 0.001$. 또한 Hand Trajectory 에서도 유의미한 차이를 보였으며, $F(2, 62) = 5.71, p < 0.005$. Total Time 에서도 유의미한 차이를 보였다. $F(2, 62) = 73.81, p < 0.001$. 위 결과는 그림2에서 볼 수 있다.

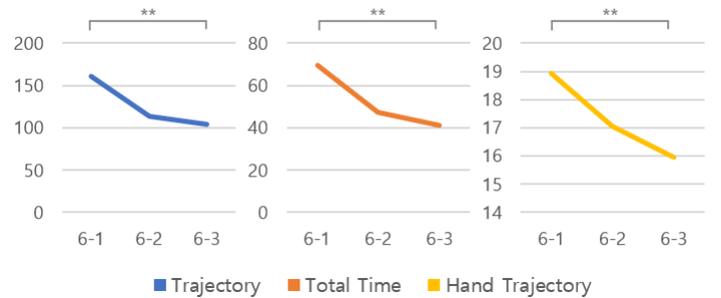


그림2: 6번과제 측정요소 그래프

4. 논의

본 연구는 Virtual Supermarket에서 Trajectory가 증가하면, 그 과제를 해결하는데 Total Time 이 증가하고, Hand Trajectory가 증가하면 그 과제를 해결하는데 Total Time이 증가한다는 사실을 실험을 통해 증명하였다. 현실의 슈퍼마켓에서 나타날 것이라고 가정한 요소들 간의 상관관계가, Virtual Supermarket의 측정 요소들 간의 상관관계에서 나타나기 때문에 본 실험은 유효한 일상생활활동 과제이다. 이러한 결과는 본 연구에서 개발한 일상생활활동 과제의 Ecological validity가 높다는 것을 반영한다. 또한, 6번 과제의 3번의 시행을 통해 반복적인 훈련이 유의미한 훈련 효과가 있다는 것을 보였다. 이러한 Ecological validity와 반복 훈련의 효과는 일상생활활동 과제로서의 Virtual Supermarket이 신뢰할 수 있는 과제임을 보였다. 따라서 향후, 일반인이 아닌 경도인지 장애 환자들을 대상으로 진행할 인지기능 훈련 결과에도 긍정적인 영향을 미칠 것이다. 그러나 본 연구는 분명한 한계점 역시 포함하고 있다. 실제 경도인지장애 환자를 대상으로 추가 연구를 수행하여야 하며, 평가 도구로써 뿐만 아니라 훈련 도구로서의 가치에 대하여 연구되어야 한다. 이러한 후속 연구를 통해 본 연구는 경도 인지 장애의 일상생활 연구 및 훈련에 도움을 줄 수 있을 것이다.

참고문헌

[1] Traykov L, Raoux N, Latour F, Gallo L, Hanon O, Baudic S, et al: Executive functions deficit in mild cognitive impairment. *Cog Behav Neurol* 2007;20:219-224

[2] 최용욱: 급속한 기대수명 증가의 함의. *KDI FOCUS* 2016, http://www.kdi.re.kr/research/subjects_view.jsp?pub_no=14826

[3] Martyr, A., Clare, L. Executive function and activities of daily living in Alzheimer's disease: a correlational meta-analysis. *Dement Geriatr Cogn Disord.* 2012;33:189-203.

[4] Kang Y, Na DL, Hahn S. A validity study on the Korean Mini-Mental State Examination (K-MMSE) in dementia patients. *J Korean Neurol Assoc* 1997;15:300-8

표 1: 상관관계가 나타난 과제의 개수

	Trajectory	Map Count	Map Duration	List Count	List Duration	Total Time	Hand Trajectory
최대 과제 개수	17	3	3	5	5	17	17
Trajectory		1	1	0	0	17	12
Map Count			3	2	2	2	2
Map Duration				2	2	2	2
List Count					5	0	4
List Duration						4	4
Total Time							17
Hand Trajectory							

Leap Motion 센서를 이용한 드론 애니메이션 제어*

김하준⁰, 임승찬, 김한설, 황인성, 김시은, 팔룰라 데니스, 이원섭
 한동대학교 ICT창업학부

(21500197, 21300623, 21200192, 21600800, 21500098)@handong.edu, dennisphalula@gmail.com, w.lee@handong.edu

Controlling the Drone Animation using Leap Motion Sensor

Hajun Kim, Insung Hwang, Hanseol Kim, Seungchan Lim, Sieun Kim, Dennis Phalula, Wonsup Lee
 School of Global Entrepreneurship and Information Communication Technology, Handong Global University

요약

손의 자세와 동작을 측정하는 Leap Motion 센서는 다양한 컴퓨터 그래픽에 활용되고 있다. 본 연구는 Leap Motion 센서로 측정된 손의 동작을 이용하여 드론의 애니메이션을 제어하는 연구를 수행하였다.

1. 서론

Leap Motion은 2013에 정식 출시된 3D motion control 기기로, 손가락 2개 정도 크기의 작은 PC의 input 장치이다. 인간-컴퓨터 상호작용 분야에서 Leap Motion은 다양한 컴퓨터 그래픽스 인터랙션을 구현하는데 사용되어 왔다. 예를 들어, Leap Motion을 이용하여 키보드나 마우스 같은 input 장치를 대신하기도 하고, 손이나 손가락 동작을 이용하는 게임들에 사용되기도 하고 있다. 뿐만 아니라 Leap Motion의 손 동작 제어 방식은 컴퓨터 그래픽스 인터랙션뿐 아니라 물리적인 장치들도 조작할 수 있는데, 손동작을 이용하여 드론을 제어하는 다양한 사례들이 YouTube에 소개되고 있다. 본 연구는 Scratch와 Go를 각각 이용하여 Leap Motion을 이용하여 DJI의 Tello라는 소형 드론을 효과적으로 제어하기 위한 동작의 적정 sensitivity를 연구하였다.

2. Leap Motion을 이용한 드론 제어 구현

2.1. Scratch

먼저, MIT에서 개발한 visual programming 언어인 Scratch를 이용하여 Leap Motion과 드론을 연동하였다. Scratch의 간단한 코딩을 이용하여 Leap Motion의 입력값을 받아 물리적 기기인 드론을 조작하고자 하였으나 Scratch에서는 Tello API를 제한적으로

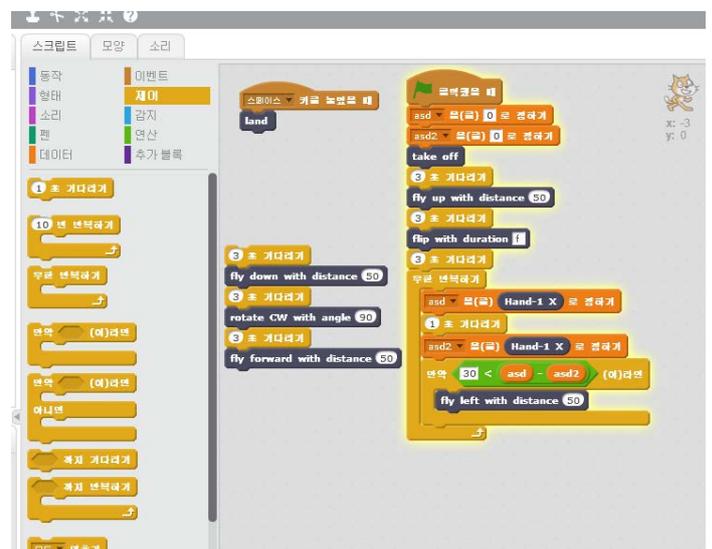


그림 1: Leap Motion으로 Tello를 조작하기 위한 Scratch 코딩 예

만 이용할 수 있고, 자연스러운 손동작에 따른 드론의 움직임을 조작하기에는 Scratch 프로그래밍에 많은 제약이 있었다. 즉, Scratch로 입력받은 손 동작에 따라 드론이 실시간으로 즉각 반응하지 않았기 때문에 인체 동작 sensitivity를 연구하기 어려웠다.

2.2. Go

Scratch의 제한적인 기능으로 인해 자연스러운 손동작을 이용한 컴퓨터 인터랙션을 구현하지 못하였기 때문에 고급 programming 언어를 이용하였으며, Google에서 제작한 Go 언어를 활용하는 GoBot library [1]에서 Leap Motion과 Tello를 제어할 수 있는 기능들을 제공하고 있으므로 이를 적용하였다. Leap Motion으로 손 및 손가락 관절들의 3차원 위치 정보를 받아서 이를 이용하여 드론의 방향을 제어하였다. Leap Motion 앞에서 손을 상하전후좌우 3축으로 움직이면 물리적 장치인 드론이 해당 방향으로 이동한다. 이 때 Leap Motion 앞에서의 손의 초기 위치를 기준으로 손의 상하전후좌우 상대적 위치 변화를 X, Y, Z 값으로 입력받는다(단위: cm). 입력 받은 X, Y, Z 값은 실제 드론이 이동하는 거리가 되는데, Leap Motion으로부터 매 0.1초마다 손의 위치

* 포스터 발표논문

* 학부생 주저자 논문임

* 본 논문은 요약논문(extended abstract)으로서, 본 논문의 원본 논문은 현재 타 학술대회 및 학술지에 제출 중

를 입력받아 드론이 이동할 거리를 드론에게 전달하기 때문에 드론에 전달되는 값은 결국 속도(=거리/시간)가 된다. 이때 Leap Motion 앞에서의 손이 움직이는 범위가 좁은 데에 비해 실제 드론이 움직이는 범위는 넓으므로 X, Y, Z 값에 일정 수치를 곱해줄 필요가 있다. 한편, 정지 상태에서 출발하거나, 진행 중에 방향 전환에 있어 가속과 감속 시간이 필요하다. 따라서 손의 움직임에 따라 드론의 움직임이 자연스럽게 synchronize되기 위해서는 Leap Motion으로 입력 받은 X, Y, Z 값에 sensitivity에 관련된 함수를 적용할 필요가 있다. 본 연구는 사용자가 가장 조작하기 용이한 sensitivity 정도 및 조작 방법에 대해 연구하고 있다. Go 언어 및 GoBot library를 이용한 Leap Motion 및 Tello 드론 제어 code는 다음 Github 사이트를 통해 공유되고 있다.

<https://github.com/HandongHCI/HandongHCI.github.io/tree/master/HCI2018S/Drone>

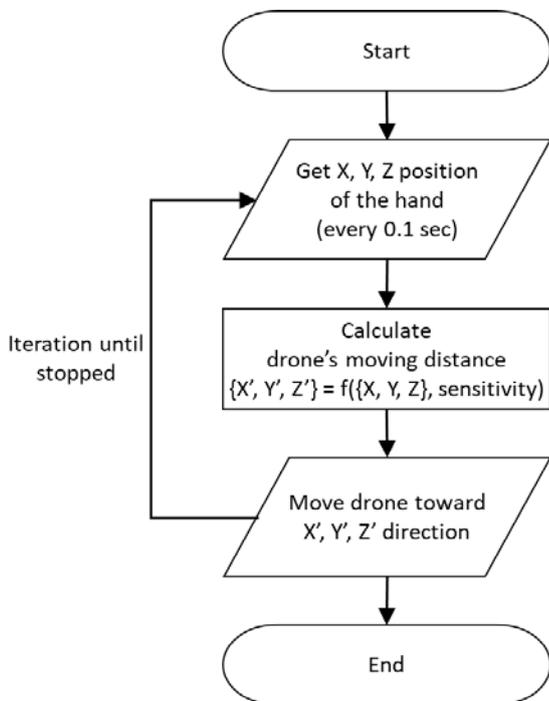


그림 2: Leap Motion을 이용한 드론 제어 순서도

시연 동영상은 아래와 같다. 두 영상은 동시에 촬영되었다.

1. Leap Motion 조작

<https://www.youtube.com/watch?v=UXO2BKxEk-Y>

2. Drone 애니메이션 제어

<https://www.youtube.com/watch?v=vtulx286ZTY>

3. 토의 및 추후 연구

손의 자연스러운 동작을 이용하여 드론 또는 로봇과 같은 물리적인 기기를 조작하기 위해서는 사람의 동작과 기기의 동작 간에 자연스러운 synchronization이 중요하다. 이는 컴퓨터의 마우스를 이용하여 모니터 상의 커

서를 조작하는 것과 비슷한 이치이지만, 물리적인 기기는 모터를 제어함에 있어 가속과 감속이 필요하고 이에 따라 손 동작과 기기의 움직임 간의 시간차가 발생한다. 이러한 시간차를 줄여 자연스러운 드론 애니메이션을 구현하는 알고리즘이 필요하다.

기존 Leap Motion을 이용한 드론 제어 사례들에서는 기술의 개발 및 적용에 관한 연구만 진행하였으며, 사용자와 드론 간의 자연스러운 인터랙션 메커니즘에 대해서는 제시하지 않았다. 본 연구는 자연스러운 드론 애니메이션 제어를 위해 Leap Motion을 이용한 드론 제어 알고리즘을 개발하고, 여러 사람들을 대상으로 서로 다른 드론 제어 sensitivity에 따른 드론 제어 용이성 및 감성적인 만족도 등을 평가하는 실험 연구를 진행하고 있다. 실험을 통해 사람들이 드론의 sensitivity를 변경해 가며 드론을 조작하기에 가장 적합한 손 움직임 정도와 이에 따른 드론의 반응 정도를 분석할 예정이다.

참고문헌

[1] GoBot, <https://gobot.io/>

회전 테이블의 기하적 변환을 이용한 점군 정합*

강현민⁰, 박용문, 김병준, 서용덕
 서강대학교 영상대학원
 (khm8206, stabilator, arrbem, yndk)@daehan.ac.kr

Point Cloud Registration Using Geometric Transformation of Turntable

Hyun-Min Kang ⁰, Yong-Moon Park, Byung-Jun Kim, Yong-Duek Seo
 Graduate School of Media, Sogang University

Abstract

In this paper, we propose a method of acquiring point cloud registration by using the geometric transformation of a turntable. ICP algorithm, which is one of the most well-known registration methods, produces weak results when used on symmetric objects. Therefore, in this paper, we propose a fast and accurate registration method that overcomes the shortcomings of ICP algorithm by converting point clouds from multi viewpoints into one turntable.

1. Introduction

In order to reconstruct a real object into a 3D model of the final object, an optimal posture alignment for 3D point clouds from several viewpoints is required. In the 3D registration method, the ICP (Iterative Closet Point) algorithm uses the least square method in order to estimate six rigid motion parameters between two different three-dimensional point clouds. ICP, however, may result in unsatisfactory registration results if it requires appropriate initial values for the convergence of the objective function in the 6-dimensional search space, or if the objective function is confined to the local minimum during the search. It is also nearly impossible to estimate the six parameters for a symmetrical object.

In this paper, we propose a method of physically reducing the search space by using a turntable and the proposed geometric correction below. This paper also shows that the registration of multiple point cloud can be performed quickly by estimating the geometric transformation between the 3D reconstruction system and the turntable, and using this estimated geometric transformation to search the one-dimensional rotation space.

2. Relative Position Estimation of the Turntable

If we know the rotation angle of the turntable in the system shown in Fig. 1 (a), we can estimate three rotation parameters among six rigid body motion parameters. The parameters can convert the multiple point cloud data into estimated positions of the turntable and the movement conversion between the respective point cloud data would infinitely become close to 0. The transformation matrix of the turntable for each pattern in Figure 1 can be defined as:

$$\begin{aligned} X_c^1 &= \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} X_w^1 \\ X_c^2 &= \begin{bmatrix} B & b \\ 0 & 1 \end{bmatrix} X_w^2 \\ X_T^1 &= \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} X_w^1 \\ X_T^2 &= \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} B & b \\ 0 & 1 \end{bmatrix} X_w^2 \end{aligned}$$

There is only rotation in the rigid transformation within a turntable. Thus, we can see that the transformation matrix equation between two turntable viewpoints is the same as $X_T^1 = R X_T^2$. R can be defined as $\begin{bmatrix} Z & 0 \\ 0 & 1 \end{bmatrix}$ because there is only rotation transformation on the Z axis, which has no

* 포스터발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 일본 논문은 3DSA 2018에 발표예정 되었음.

* 본 연구는 미래창조과학부 ‘범부처 Giga KOREA 사업’의 일환으로 수행하였음. [GK17P0300, 초실감 서비스를 위한 동적 객체의 실시간 4D 복원 기술 개발]

transformation. Using two correlations, we can estimate translation and rotation transformation of the turntable that we want:

$$\begin{aligned}
 X_T^1 &= R X_T^2 \\
 \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} Z & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} B & b \\ 0 & 1 \end{bmatrix} \\
 \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B & b \\ 0 & 1 \end{bmatrix}^{-1} &= \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} \\
 \begin{bmatrix} D & d \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} R Z R^T & -R Z R^T T + T \\ 0 & 1 \end{bmatrix} \\
 D &= R Z R^T, \quad d = -R Z R^T T + T
 \end{aligned}$$

In order to perform the registration of the point clouds acquired through the different viewpoints from the rotation of the turntable, all the acquired 3D point clouds are first converted into camera coordinates. Then, all 3D point clouds obtained through the camera coordinate system are converted into geometric correction pattern coordinates. Then, the point clouds that are converted into geometric correction pattern coordinates are reconverted into the turntable coordinates. Thus, the registration for the point clouds obtained through the turntable coordinates are affected only by the rotation of the turntable in the Z axis. Therefore, if we know the actual rotation angle of the turntable, it is possible to intuitively know the conversion between multiple point clouds, and to perform the registration for symmetric objects.

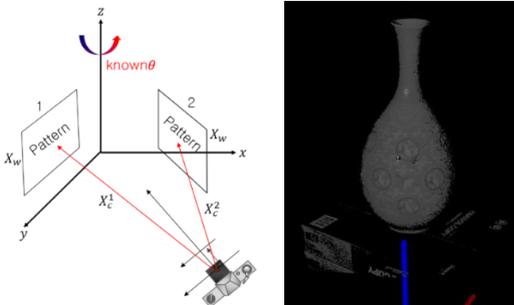


Fig 1 : Geometric correction estimation of a turntable and point cloud acquisition using structured light system

3. Experimental Results and Analysis

The actual object of the experiment was a symmetrically shaped porcelain, and the point cloud data was acquired (in 3D form) by using a structured light system. Also, four point clouds were obtained by rotating the turntable at 0, 90, 180, and 270 angles and determining each positions. The relative positional relationship between the structured light system and the turntable is estimated for the real data, and the 3D point cloud acquired from the structured light system is converted into turntable coordinates. It was found that registration of 3D point

cloud is possible by simply exploring the one-dimensional rotational space. The conversion between each point cloud using the four point clouds obtained at each point exists only in the rotation conversion to the Z axis. Since we know the angle of rotation, we can easily estimate the rotation transformation formula, and we can obtain the final point cloud registration by performing the geometric correction on the estimated rotation transformation point cloud.

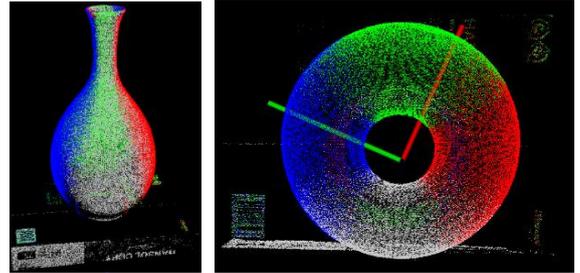


Fig 2 : Geometric correction estimation of turntable and point cloud acquisition using structured light system (Each color is a point cloud acquired at different points in view)

4. Conclusion

We propose a method of registration using the turntable geometric transformation in order to overcome ICP algorithm's shortcomings on symmetrical objects. As a result of implementing this method, we were able to perform the registration of the model of a symmetrical pottery and obtain the final integrated point cloud. Future work will be to make a mesh model using a point cloud registration from multiple view points, and to be mapping a texture to it.

참고문헌

[1] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell., 14(2):239–256, 1992.
 [2] S. Zhang and S.-T. Yau. “Generic non-sinusoidal phase error correction for three-dimensional shape measurement using a digital video projector.” Appl. Opt., 46(1):36-43, 2007. 10, 13, 14
 [3] C. J. Taylor, “Implementing high resolution structured light by exploiting projector blur.” Workshop on Applications of Computer Vision WACV, page 9-16. IEEE, 2012
 [4] Y. Park, J. Kim, Y. Seo, “Object Shape Restoration & Self-Localization Based on Multi-Structured Optical System”, 29th Workshop on Image Processing and Image Understanding, IPIU2017

Funny Room: 다중 사용자간 협업을 활용한 캐주얼 가상현실 콘텐츠 제작*

허하영⁰, 박원준, 박성준, 민성미, 김진모*
부산가톨릭대학교 소프트웨어학과

hayoung6177@hanmail.net⁰, phs6048@nate.com, zm0nn@naver.com, tjdal3627@naver.com, jmkim11@cup.ac.kr*

Funny Room: Casual Virtual Reality Contents Production using Multi-user Collaboration

Hayoung Heo⁰, Wonjun Park, Seongjun Park, Sungmi Min, Jinmo Kim*
Department of Software, Catholic University of Pusan

요약

본 연구는 두 명의 사용자가 가상현실 기반의 협업 환경에서 각각의 역할과 행동을 기반으로 함께 소통하고 협업함으로써 공통의 목표를 달성해가는 캐주얼 가상현실 콘텐츠 (Funny Room)를 제작한다. 기존의 1인 중심의 가상현실 콘텐츠에서 제공하는 현존감과 경험을 넘어 만족하는 체험을 제시할 수 있는 방향을 직접 콘텐츠 제작하는 과정을 통해 설명함으로써 누구나 쉽게 협업 기반 가상현실 콘텐츠를 제작할 수 있도록 한다.

1. 서론

가상현실 콘텐츠 제작과 관련된 대부분의 사례 또는 관련 연구들은 한 명의 사용자의 몰입과 경험에 초점을 맞추어 이루어지고 있다. 그리고 이 과정에서 사용자가 실제 환경 또는 상황에 존재하지 않음에도 마치 그 장소에 자신이 존재하고 있는 듯한 심리적 경험을 가질 수 있도록 가상환경 또는 객체와의 현실감 있는 상호작용 방법이나 햅틱 시스템을 활용한 촉각 전달, 모션 플랫폼을 활용한 사실적인 걷기 표현 등에 대한 현존감 연구들이 다양하게 이루어지고 있다[1]. 최근에는 다중 사용자가 함께 가상 환경에 공존하여 체험하는 가상협업환경에 가상현실 기술이 적용되어 경험을 공유함으로써 몰입형 상호작용에 사회적 상호작용이 결합되어 향상된 현존감을 제공하기도 한다[2]. 하지만 다중 사용자를 고려한 대부분의 가상현실 콘텐츠의 경우 기존의 1인 중심의 체험 방식을 다중 사용자로 확장 시키는 것 이상의 게임 요소나 상호작용 방식 등이 없어 여러 사용자가 함께 미션을 수행하거나 목표를 달성해나간다는 것 이외에 사회적 상호작용에 기반한 현존감 향상은 기대하기가 어렵다. 따라서 본 연구는 다중 사용자가 함께 소통하고 협업할 수 있는 환경적 또는 게임 요소들을 고려한 캐주얼 가상현실 콘텐츠인 Funny Room을 제작하고, 전체적인 제작 과정을 단계적으로 설명함으로써 협업과 소통을 고려한 몰입형 가상현실 콘텐츠를 쉽고 효율적으로 제작할 수 있도록 한다.

2. 가상협업환경에서의 가상현실 콘텐츠 제작

2.1. 캐주얼 가상현실 콘텐츠 기획

제안하는 가상현실 콘텐츠 Funny Room의 목표는 NPC가 임의로 제시하는 숫자를 제한된 시간 안에 다트와 다트 판을 사용하여 빠르고 정확하게 넘기면 성공하는 것이다. 이때 HMD를 착용한 두 명의 사용자는 독립된 역할과 행동 (다트 또는 다트 판)을 기반으로 서로 소통과 협업을 통해 함께 목적을 달성해 나가는 방식으로 1인 사용자 중심의 몰입형 가상현실 콘텐츠 이상의 현존감과 새로운 경험을 제공하는 것이 제안하는 콘텐츠 기획의 핵심이다. 그림 1은 이를 나타낸 것으로 다른 위치에 존재하는 사용자 중 한 명의 사용자는 다트를 다트 판에 던지고, 다른 한 명의 사용자는 날아오는 다트를 정확히 다트 판에 맞혀야 한다.



그림 1: 제안하는 캐주얼 가상현실 콘텐츠의 체험 과정의 예

2.2. 개발 환경 구축

본 연구의 콘텐츠는 유니티 (Unity) 3D 엔진을 기반으로 SteamVR 플러그인을 통합하여 개발환경을 구축하였고, 따라서 사용자는 HTC Vive HMD를 착용한 상태에서 전용 컨트롤러를 통해 체험을 하는 환경이다. 또한, 다중 사용자간 통신은 유니티 3D 엔진의 유넷 (U-Net)을 기반으로 P2P 방식에 적합한 호스트 기능 (호스트가 서버를 개설하고, 클라이언트는 해당 서버의 IP에 접속하는 방식)을 활용한다.

2.3. 다중 사용자간 협업을 통한 게임 상호작용

본 연구는 다중 사용자간 협업과 소통에 기반 하여 함

* 포스터 발표논문

* 본 연구는 2018년도 정부(교육부)의 재원으로 한국연구재단 (NRF) 지원으로 수행되었음. (No. 2017R1D1A1B03030286)

* 학부생 주저자 논문임

게 체험하는 콘텐츠를 제시함을 목표로 한다. 이를 위해 두 명의 사용자가 각각 가상환경 또는 객체와 상호작용함은 물론 서로간의 상호작용을 통해 정보를 교환하고 행동을 공유하는 구조를 가진다. 전체적인 구조를 바탕으로 콘텐츠가 진행되는 흐름은 다음과 같다.

1. NPC 캐릭터를 통한 설명, 조작법, 목표 점수 전달
2. 다트, 다트 판 사용자 결정
3. 사용자 각각의 아이템 랜덤 선택
4. 주어진 상호작용에 기반 하여 콘텐츠 실행
5. 목표 달성, 실패 유무 결정

우선, 사용자가 모두 콘텐츠에 접속하면 NPC가 진행과 관련된 설명(미션, 조작법 등)을 진행한다. 그리고 40~99 사이의 임의 값을 설정하여 목표 점수를 사용자에게 전달한다(그림 2(a)). 그런 다음 사용자가 의사소통을 하여 사용자 중 한 명은 다트를 나머지 한 명은 다트 판을 선택하고 해당되는 자리로 각자 이동한다. 이때 다트 판을 선택한 사용자는 게임 요소가 적용된 4개의 상자(기본 다트 판, 네모 다트 판, 별 다트 판, 배드민턴 채 다트 판) 중 하나를 선택하면 무작위로 하나의 다트 판이 결정된다(그림 2(b)). 다트 사용자는 다트 판의 숫자를 확인하고 다트를 던지고, 다트 판 사용자는 날아오는 다트를 정확히 다트 판에 맞힐 수 있도록 행동한다. 이 과정에서 게임적 요소를 추가하여 다트 판 중앙에 있는 물음표 과녁을 맞히게 되면 아이템 돌림판이 돌아가 미션에 도움이 되는 아이템(점수 두 배 다트, 두 배 큰 다트 판, 속도 감소 다트) 또는 방해가 되는 아이템(회전하는 다트 판, 시야 흐림 효과, 팽)을 획득할 수 있다(그림 2(c)). 두 명의 사용자가 함께 컨트롤러의 트리거 버튼을 누르게 되면 돌림판이 조금 더 움직이기 때문에 원하는 아이템을 획득하기 위하여 사용자가 함께 협업할 수 있는 구조로 제작하였다. 마지막으로 시간 제한을 두어 주어진 시간 안에 미션을 성공하지 못하면 처음으로 돌아가도록 설정하였다.

3. 구현 및 분석

다중 사용자의 협업에 기반한 캐주얼 가상현실 콘텐츠는 Unity 3D 2017.3.1.f1 엔진을 기반으로 SteamVR 플러그인을 통합하여 제작하였고, 콘텐츠에 사용된 그래픽 리소스 중 주 캐릭터, NPC 그리고 협업에 필요한 핵심 리소스들은 3ds Max 2017, Photoshop CS5, Illustrator CS5를 활용하여 직접 제작하였다. 그 밖에 배경을 포함한 기타 객체들을 유니티 3D 엔진에서 제공하는 에셋 스토어(asset store)의 리소스를 활용하였다. 두 명의 사용자가 그림 2와 같이 직접 제작한 캐주얼 가상현실 콘텐츠를 그림 1과 같이 HTC Vive HMD와 전용 컨트롤러를 사용하여 체험할 수 있다.

4. 결론

본 연구는 가상현실에 여러 사용자가 함께 참여하여 소통하고 협업할 수 있는 상호작용과 다양한 게임 요소를 추가함으로써 보다 향상된 현존감과 만족하는 경험을 제공하는 가상협업환경을 제시함을 목적으로 캐주얼 가상현실 콘텐츠 Funny Room을 제작하였다.

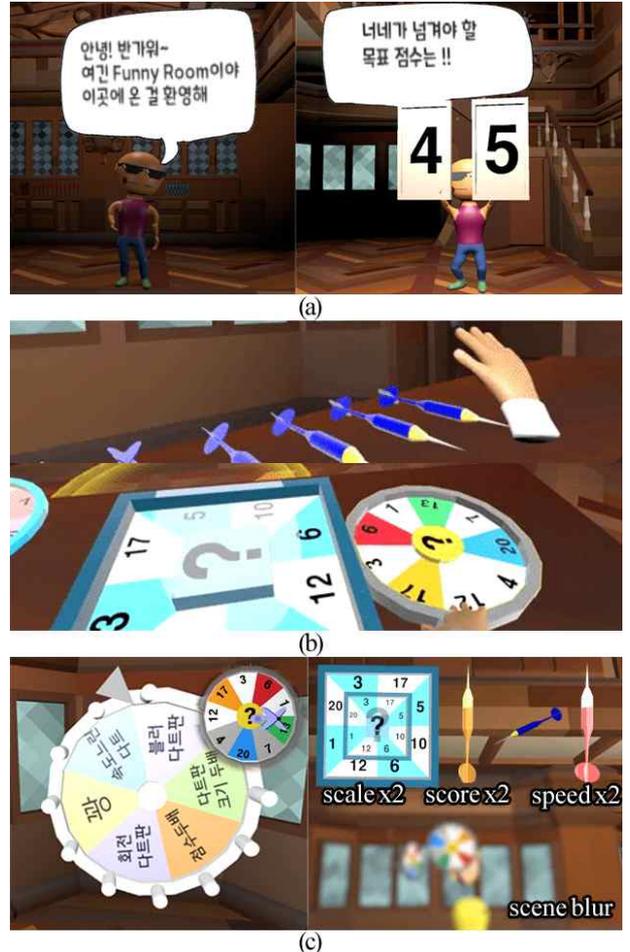


그림 2: 게임 요소와 다중 사용자간 상호작용이 포함된 콘텐츠 제작 결과: (a) NPC 캐릭터 정보전달, (b) 다트, 다트 판 선택, (c) 아이템 돌림판에서 제공하는 아이템들

이는 가상현실에 참여하는 HMD 사용자 각각에게 구분된 역할과 행동을 부여하고 함께 소통하고 협업하여야만 주어진 목표를 달성할 수 있도록 기획, 제작하였다. 여기에 여러 게임적 요소(다트, 다트 판, 아이템 돌림판 등)들을 활용하여 사용자가 함께 협업해 나감으로써 주어진 미션을 달성하도록 하였다.

현재는 두 명의 사용자가 함께 협업하는 것만을 고려하고 있기 때문에 더 많은 사용자가 함께 소통하고 협업할 수 있는 구조의 콘텐츠를 향후에 제작하고자 한다. 그리고 참가자를 대상으로 한 설문 실험을 통해 현존감과 경험에 대한 분석을 체계적으로 진행할 계획이다.

참고문헌

- [1] J. Lee, K. Jeong, and J. Kim, Mave: Maze-based immersive virtual environment for new presence and experience, *Computer Animation and Virtual Worlds*, 28(3-4), e1756, 2017.
- [2] K.-D. Le, M. Fjeld, A. Alavi and A. Kunz, Immersive environment for distributed creative collaboration, In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp.16:1-16:4, 2017.

Video see-through HMD 와 optical see-through HMD 간 색상 및 크기 차이를 위한 보정 방법

김 주 환¹
원광대학교 컴퓨터공학과¹
bbang4193769@naver.com

김 호 원²
한국전자통신연구원(ETRI)²
hw-kim@etri.re.kr

조 동 식*^o
원광대학교 디지털콘텐츠공학과*^o
dongsik1005@wku.ac.kr*^o

Exploring Scale and Color Difference among MR Devices for Improved MR Experience

JuHwan Kim¹
Wonkwang Univ.¹

Howon Kim²
ETRI²

Dongsik Jo*^o
Wonkwang Univ.*^o

요약

최근, 실제 세계에 가상의 물체를 합성하여 현실에서 가상의 물체를 보고 상호작용하는 경험을 제공하는 혼합현실(MR) 기술이 널리 사용하고 있다. 하지만, 이러한 혼합현실 기술은 기기별(예. video see-through HMD 혹은 optical see-through HMD)로 가상의 물체에 대한 크기, 칼라 등이 다르게 보여 사용자마다 동일한 경험을 제공하는데 한계가 있다. 예를 들면, 혼합현실 환경에서 온라인 쇼핑을 하기 위해 사용하는 3D 가상 모델은 혼합현실 기기마다의 특성에 의해 사용자별로 서로 다른 크기, 서로 다른 색상으로 판단될 수가 있다. 따라서, 본 논문에서는 이러한 한계점을 극복하기 위해 혼합현실 기기별로 서로 다르게 가시화되는 요소(색상, 크기)를 정의하고, 실험 가이드라인을 제시하여 향후 혼합현실 기반의 체험 효과를 극대화하고, 사용자에게 동일한 경험을 제공하기 위한 보정 방법론을 제시하고자 한다.

1. 서론

혼합현실(MR)은 실제 세계에 가상의 물체를 합성하여 가상의 물체가 마치 현실 세계에 원래 존재했던 것처럼 표현하고, 상호작용하는 기술이다[1]. 예를 들면, 온라인 쇼핑을 하기 위한 소비자들은 상품을 구매하기 위해서 직접 매장에 가서 상품을 보기 위해 가지 않아도 되고, 집에서 소비자가 구매하려는 제품을 가상의 물체로 표현하고, 혼합현실 환경으로 가시화하여 원하는 물건을 살펴보면서 구매에 대한 판단을 할 수 있을 것이다. 최근, 혼합현실 사용자 체험을 제공하기 위해 다양한 형태의 장치들이 글로벌 기업에서 출시되어 상품화되고 있다. 이러한 혼합현실 기기는 주로 머리에 착용하는 방식

을 통해 실 세계의 합성된 가상의 물체에 대한 사용자 경험을 제공하고 있다. 이를 위한 혼합현실 기기는 크게 비디오 영상 신호를 입력으로 하여 영상을 합성하는 방식인 video see-through 방식과 광학적으로 실제 세계를 보면서 합성된 가상의 물체를 보는 optical see-through 방식으로 나누어 진다 [2]. 또한, 전통적인 방식인 스마트폰의 카메라 영상을 이용하여 합성된 가상 물체를 보는 방법도 여전히 많이 사용되고 있다. 스마트폰을 이용하는 방법은 사용자의 눈의 위치에서 합성된 결과를 보지 않고, 사용자 손에 위치에서 혼합현실 환경을 보는 indirect한 방법이라고 알려져 있다. 이러한 AR/MR의 다양한 가시화 장치의 경우 표준화된 장치가 없고, 개별 회사마다 시야각(FOV), 색상, 해상도, 착용감, 무게, 센서 딜레이, 초점 거리, 렌더링 방법 등이 달라 가상 물체가 보여지는 정도, 사용자의 느낌 등 수많은 관련 요소가 관계 하고 있어서 아직 많은 연구가 진행 되어야 할 상황이다 [3]. 또한, 다양한 디바이스 환경에서 3D 모델의 일관된 사용자 경험을 제공하기 위한 파라미터가 무엇인지를 알기 위한 연구가 필요한 실정이다 [4].

한 가지 예로, 장치별로 합성된 가상의 물체가 색상의 차이 혹은 크기의 차이 등 사용자마다 다르게 느끼기 때문에 동일한 가상 물체를 가시화 하여도 장치마다 각각 서로 다르게 보이는 문제점이 발생하게 된다. 따라서, 사용자가 동일한 3D 가상 물체를 서로 다른 색상과 크기의 물체로 인식하는 경우가 발생할 수 있을 것이다. 이에 본 논문에서는 이러한 문제를 해결하기 위해서 가시화된 3D 가상 물체의 크기, 색감 등 요소에 대한 차이점을 사용자 인지 측면에서 측정 및 조사하기 위한 실험 방법론을 제시하고자 한다. 또한, 본 논문에서는 제시된 실험 방법론을 통해 3D 가상 물체의 경험 정보 일치를 위해 기기별로 보정(adjustment)할 수 있는 방법론을 제시하고자 한다.

2. 혼합현실(MR) 실험 방법

그림 1은 혼합현실 환경에서 피실험자 기반 측정을 위한 실험 방법에 대한 한 가지 예를 보여주고 있다. 피실험자는 혼합현실 기기를 착용하고 실제 물체 (예. 큐빅

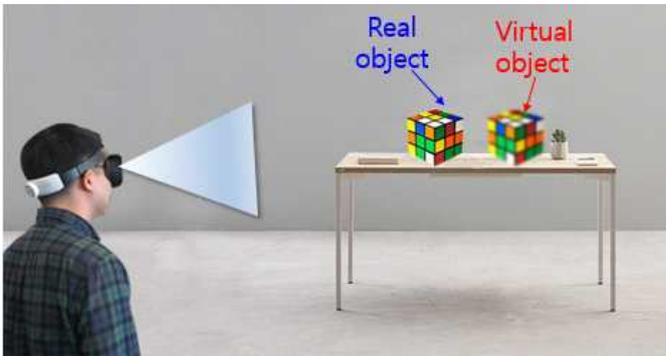
* 학부생 주저자 논문임 (포스터발표논문)

* 교신저자: 조동식(Dongsik Jo)*^o

* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2018년도 문화기술연구개발 지원사업으로 수행되었음

* 현재 타 학술대회(한국정보과학회 KCC 2018)에 발표함

퍼즐)와 가상 물체(예.가상 큐빅 퍼즐)을 동시에 보게 된다. 크기의 경우 실험의 방법은 다음과 같다. MR 장치를 착용한 피실험자는 실제 물체의 크기와 비교하여 가상의 물체의 크기를 맞추기 위해 조이스틱 장치를 이용하여 조절하는 작업을 하게 된다. 이를 통해 가상의 물체에 대한 크기가 MR 장치별로 사용자가 다르게 느끼는 정도를 측정할 수 있게 된다. 색상의 경우 실험 방법은 피실험자가 실제 물체와 비교할 수 있도록 색상이 조금씩 차이가 나는 가상의 물체를 객관식으로 선택할 수 있도록 한다. 이러한 피실험자를 대상으로 한 실험을 MR 장치별로 반복하여 수행하게 되면 장치별로 사용자가 서로 다르게 평균적으로 느끼는 크기와 색상의 차이에 대한 파라미터 값이 추출될 것으로 본다.



〈그림1. 피실험자 실험 방법: 실제 큐브, 가상 큐브 간 색상/크기 비교〉

3. 혼합현실(MR) 기기별 보정 파라미터 적용 방법

본 논문에서는 혼합현실 기기별로 3D 가상 모델이 서로 다르게 보이는 것에 관해 크기와 색상을 고려하였고, 아래와 같이 혼합현실 기기에서 콘텐츠 가시화를 초기화 보정을 한다면 3D 가상 모델에 관해 사용자에게 동일한 경험을 제공할 수 있을 것으로 본다.

$$AC_{new} = R_{original} \pm DC_{device} \pm DC_{rending} \dots\dots(1)$$

$$AS_{new} = RS_{original} \pm DS_{device} \pm DS_{rending} \dots\dots(2)$$

수식 1은 혼합현실 기기별 맞춤형으로 칼라 (색상)에 대한 파라미터 보정식이고, 수식 2는 크기에 대한 보정식이다. 보정을 하기 위해서는 혼합현실 장치에서 렌더링을 통해 발생하는 에러를 보정하는 방법으로 진행하게 된다. 수식 1에서 AC_{new} 는 혼합현실 기기에서 보정되는 칼라 값이고, $R_{original}$ 는 실제 물체 고유의 색상, DC_{device} 는 혼합현실 장치에서 실제 물체를 바라보았을 때 변화되는 색상 값, $DC_{rending}$ 는 혼합현실 장치에서 가상의 물체를 보았을 때의 렌더링에 의해 변화되는 값이다. 즉, 이러한 색상 보정 결과를 수행하면, 혼합현실 기기를 착용한 사용자는 가상의 3D 모델과 실제 물체를 동시에 투영하였을 경우 가상과 실제 물체가 색상에서는 동일하다고 판단하게 될 것이다. 실험에서는 수식 1에서 동일한 조명 조건에서 수행될 수 있도록 컬러미터를 이용하여 지금 현재의 실제 물체에 대한 값을 가상 물체에 적용하는 방법으로 수행하고자 한다. 수식 2에서는 AS_{new} 는 혼합현실 기기에서 보정되는 크기

(scale) 보정 값이고, $RS_{original}$ 는 실제 물체 고유의 크기 DS_{device} 혼합현실 장치에서 실제 물체를 바라보았을 경우 변화되는 차이 값, $DS_{rending}$ 는 혼합현실 장치에서 가상의 물체를 보았을 때 렌더링에 의해 변화되는 차이 값이다. 수식을 보정할 수 있는 크기 또한 색상과 동일한 방식으로 보정을 할 수가 있다. 피실험자를 대상으로 한 실험을 통해서 수행한 결과 수식(1),(2)와 같이 linear transform 형태로 나타나는지를 검토할 필요가 있고, 보정 작업을 위해서는 사용자에게 제공되는 MR 콘텐츠의 3D 가상 모델을 기기별로 서로 다르게 추출된 보정 파라미터 값을 적용하는 방법으로 수행하면 된다. 즉, 사용자가 착용하는 기기별로 색상 및 크기 등에 관한 파라미터를 확정하고, 이를 체험하려는 3D 가상 물체에 자동으로 피팅(fitting)할 수 있을 것이다.

4. 결론 및 향후 연구

본 논문에서는 다양한 혼합현실 장치별로 서로 다른 사용자 경험 (예. 크기가 다르다고 생각 등)을 동일한 3D 모델이 합성되는 느낌을 전달하기 위한 방법론을 제시하였다. 이를 위해 피실험자를 대상으로 한 실험 방법을 제시하였고, 또한, 실험의 결과를 통해 추출될 수 있는 있는 파라미터 결과를 이용하여 각 장치별로 가상의 3D 모델을 미리 보정(adjustment)할 수 있는 방법을 제시하였다. 이는 서로 다른 혼합현실 장치를 이용하는 사용자에게 가시화되는 가상의 물체에 대한 속성을 변경하여 동일한 3D 혼합현실 경험을 제공하는데 도움이 될 것으로 전망한다. 향후 본 논문에서 설계된 요소와 실험 방법을 통해 실제 피실험자 대상으로 혼합현실 실험을 수행하고, 실제 제품으로 출시된 다양한 혼합현실 기기에 적용하여 사용자 경험이 비슷한지 통계적으로 차이가 있는지 등에 대한 시스템 구성 및 실험을 직접적으로 진행하고자 한다. 또한, 사용자별로 다르게 느끼는 차이에 대한 파라미터를 보정할 수 있는 방법에 대한 연구도 수행하고자 한다. 본 논문의 결과를 통해 향후 AR/MR 장치에 따라 사용자가 인지하는 특성(예. 색상/크기)이 다른 정도를 찾고, 다양한 AR/MR을 지원하는 이종 디바이스 간 일관된 AR/MR 실감 체험을 제공할 수 있을 것으로 기대한다.

참고문헌

- [1] Stephan Beck et al., "Immersive group-to-group telepresence," in IEEE TVCG, 2013.
- [2] Dongsik Jo, Ki-Hong Kim, and Gerard Jounghyun Kim, "Effects of avatar and background types on users' co-presence and trust for mixed reality-based teleconference systems," in Proceedings CASA, 2017.
- [3] Andrew Maimone et al., "General-purpose telepresence with head-worn optical see-through displays and projector-based lighting," in Proceedings Virtual Reality (VR), 2013.
- [4] Sabah Boustila, Dominique Bechmann, and Antono Capobianco, "Evaluation of factors affecting distance perception in immersive virtual environments during virtual visits of houses," in Proceeding IMH'15, 2015.

가상현실 방향전환보행을 위한 경로예측 시스템*

조용훈⁰, 이동용, 이인권
연세대학교 컴퓨터과학과
cyhterry@naver.com, pflidy2850@gmail.com, iklee@yonsei.ac.kr

Path Prediction Using LSTM Network for Redirected Walking in Virtual Reality

Yong-Hun Cho⁰, Dong-Yong Lee, In-Kwon Lee
Dept. of Computer Science, Yonsei University

요약

가상현실을 사용하기 위해 좁은 물리공간의 한계를 극복하기 위해서 방향전환보행(Redirected walking)은 매우 좋은 대안이 될 수 있다. 본 논문에서는 방향전환보행을 더욱 효율적으로 하기 위한 사용자의 경로예측 알고리즘을 제안한다. 이를 활용하면 효율적인 공간 활용을 통해 직접 보행을 통한 몰입감 있는 가상현실이 체험이 가능하다.

1. 서론

가상현실은 현실에서 경험하기 힘든 다양한 콘텐츠를 제공하며 방대한 상상력의 결과물을 체험할 수 있다. 가상현실을 체험하기 위해 사용자는 컨트롤러를 사용하게 되는데, 컨트롤러 없이 실제 보행을 통해 탐험하게 되면 훨씬 더 큰 몰입감을 느낄 수 있다.

다만, 물리적 공간의 한계로 먼 거리를 보행하기가 어려운데, 이를 보완하기 위해 방향전환보행이 사용된다. 방향전환보행은 가상의 움직임과 실제 움직임 간의 괴리를 발생시켜 공간을 더 효율적으로 사용할 수 있게 한다. 예로, 사용자의 이동거리나 회전각도를 증감시켜 사용자가 가상현실 상에서 보는 뷰와 다른 뷰를 가능하게 한다. 또한, 사용자가 가상현실 상에서 직진을 한다고 생각하지만 실제로 조금씩 회전을 주어 원형을 돌게 만들 수 있어, 공간을 매우 효율적으로 사용할 수 있게 된다[그림 1].

이 때, 사용자가 움직일 방향을 미리 예측 할 수 있다면, 방향전환보행을 사전에 미리 적용하여 좀 더 효율적인 공간사용이 가능하다. 방향전환보행 방법중 FORCE[6]이나 MPCRed[3] 같은 알고리즘의 경우 방 구조 등을 고려하여 최적의 방향전환보행 파라미터를 계획 하는데, 이 때 사용자의 경로를 미리 예측하면 더 효율적이고 적은 계산으로 방향전환보행을 수행할 수 있다. 이전 연구에서는 사용자의 기존 경로를 연장하여 사용하거나 눈의 움직임을 추적[4], 혹은 사람의 보행

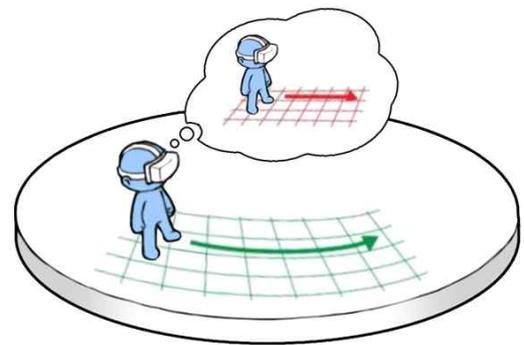


그림 1 방향전환보행

모델[5]을 사용하여 경로를 예측하였다.

본 논문에서는 미로형의 공간에서 사용자 보행 데이터를 수집하고, Long Short-Term Memory(LSTM)[2]를 이용하여 사용자 경로를 학습하여 사용자의 미래 경로를 예측할 수 있는 알고리즘을 개발하였다.

2. 데이터 수집

사용자는 Nescher 등의 연구에 쓰인 환경[3]과 유사한 20m x 20m 가상 공간[그림 2]을 총 10분동안 자유롭게 탐험하게 되며, 사용자의 움직임은 HTC VIVE와 트래커를 통해 추적되었다. 실제 공간의 크기는 4m x 7m이며, Azmandian 등의 연구[1]에서 구현된 바와 같이 방의 중심부로 사용자를 유도하는 steer-to-center 알고리즘이 사용되었다. 사용자가 추적되는 공간 이상으로 나가게 되면 2:1 리셋(사용자가 가상공간에서 한바퀴를 회전 할 동안 실제 공간에서는 반바퀴를 회전)을 사용하여 반대방향으로 나갈 수 있게 한다. 총 7명의 사용자가 2번씩 실험을 진행하였다. 사용자의 움직임은 0.01초 단위로 기록되었으며 지나친 중복을 막기 위해 두 프레임 마다 한 프레임을 생략하였다.

3. 경로 예측 모델

수집된 데이터는 사용자의 시간에 따른 연속적인 데이터이므로 Recurrent Neural Network(RNN)의 변형인 LSTM모델을 사용한다. RNN은 과거의 연속적인 데이

* 포스터 논문

* 본 논문은 요약논문 (Extended Abstract) 으로서 초기 연구 진행 결과를 포함하고 있음.

* 이 논문은 삼성전자 미래기술육성센터의 지원을 받아 수행된 연구임(과제번호 SRFC-IT1601-04).

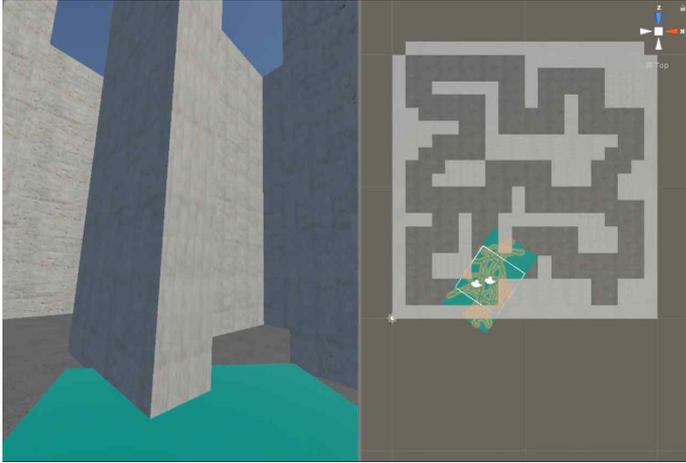


그림 2 렌더링된 가상환경(좌), 전체 맵의 구조(우)

터를 이용해 미래의 정보를 예측하는데, LSTM 모델은 RNN에 비해 이전 정보를 더 잘 기억하고 저장해낸다. 경로 예측 모델의 입력은 4차원의 100프레임 정보로 현재의 좌표와 방향을 나타낸다. 출력은 가상환경에서 100프레임 후의 예측 좌표가 된다. 128개의 은닉 계층(hidden layer)가 존재하며, 이 계층은 fully-connected layer와 연결되어 출력으로 연결된다.

이 경로예측 모델을 사용하면 사용자의 미래 위치를 예측할 수 있다. 방향전환보행 사용 시에, 사용자가 이동할 위치를 미리 알 수 있으므로, 한발 앞서 방향전환보행을 적용하여 더 효율적인 공간 활용이 가능해진다. 딥러닝의 특성상 학습 후 테스트 단계는 거의 실시간으로 이루어 질 수 있기 때문에, 빠른 연산이 가능하다.

3. 경로 예측 결과

경로예측 모델을 사용하여 사용자의 경로를 예측한 결과는 [그림 3]과 같다.

테스트 데이터는 실제 사용자가 걸어다니며 수집한 데이터 중 일부이다. 빨간색 선이 실제 사용자가 이동한 경로이며, 파란색 선은 과거 사용자 경로를 입력으로 하여 경로예측 모델이 출력해낸 예측 위치 결과이다. 빨간색과 파란색 선이 비슷하면 예측이 잘 되었다고 할 수 있다. Mean-Squared Error(MSE)를 측정한 결과 약 0.152의 수치가 나왔다. MSE는 예측 좌표와 실제 사용자 좌표 간의 차이를 측정한 것으로 20m x 20m 가상환경에서 평균적으로 0.4m 가량의 오차가 나온 것으로 이해할 수 있으며 실제 사용자 경로에서 크게 벗어나지 않는 것을 볼 수 있다.

[그림 3]의 (10,5) 부분처럼 몇몇 부분에서 예측 결과가 멍치는 결과가 종종 나오는 것을 볼 수 있다. 맵의 특정 부분에서 사용자들이 변칙적인 행동을 하거나 멈추는 등의 동작이 원인이 될 수 있으며 사용자의 속도와 같은 내재적인 요인도 영향을 끼칠 수 있는 것으로 보인다.

4. 결론

본 논문에서는 사용자의 이동 데이터를 토대로 가상현

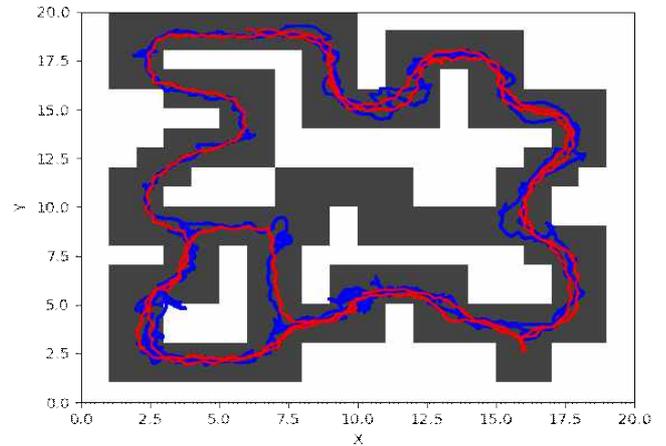


그림 3 사용자 경로 예측 결과

실 상에서의 경로예측 모델을 제안하였다. 결과는 경로 예측 결과가 실제 보행 경로와 상당히 유사한 모습을 보였다. 이를 사용하면 미리 효율적인 방향전환보행을 위한 판단을 할 수 있어, 좁은 물리적 공간을 더욱 효율적으로 사용할 수 있게 된다. 추후, 모델의 과인튜닝과 여러 예외처리 등을 통해 모델을 더욱 정확하게 개선시킬 예정이다.

참고문헌

- [1] M. Azmandian, T. Grechkin, M. Bolas, and E. Suma. The redirected walking toolkit: a unified development platform for exploring large virtual environments. In 2016 IEEE 2nd Workshop on Everyday Virtual Reality (WEVR), pp. 9–14, March 2016.
- [2] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, Nov. 1997.
- [3] T. Nescher, Y.-Y. Huang, and A. Kunz. Planning redirection techniques for optimal free walking experience using model predictive control. In 2014 IEEE Symposium on 3D User Interfaces (3DUI), pp. 111–118, March 2014.
- [4] M. Zank and A. Kunz. Eye tracking for locomotion prediction in redirected walking. In 2016 IEEE Symposium on 3D User Interfaces (3DUI), pp. 49–58, March 2016.
- [5] M. Zank and A. Kunz. Where are you going? using human locomotion models for target estimation. *Vis. Comput.*, 32(10):1323–1335, Oct. 2016.
- [6] M. A. Zmuda, J. L. Wonser, E. R. Bachmann, and E. Hodgson. Optimizing constrained-environment redirected walking instructions using search techniques. *IEEE Transactions on Visualization and Computer Graphics*, 19(11):1872–1884, Nov 2013.

가상현실환경에서 사운드를 이용한 시각적 멀미 감소*

박찬중⁰, 윤종철¹, 이인권²

^{0,2}연세대학교 컴퓨터과학과, ¹강원대학교 컴퓨터미디어산업공학부

⁰cjprist@yonsei.ac.kr, ¹media19@kangwon.ac.kr, ²iklee@yonsei.ac.kr,

Visual Sickness Reduction Using Sound in Virtual Reality

Chan Jong Park⁰, Jong-Chul Yoon¹, In-Kwon Lee²

^{0,2}Dept. of Computer Science, Yonsei University, ¹Division of Computer, Media and Industrial Engineering

요약

가상현실환경에서 시각적 멀미(Visual sickness)를 감소시키기 위한 많은 연구들이 수행되었다. 본 논문에서는 사운드의 입체감과 그 크기의 변화를 다르게 설정하였을 때, 시각적 멀미의 변화를 사용자 설문조사를 통하여 측정하였다. 2D와 3D 사운드를 삽입했을 때, 멀미감이 감소하였고, 가상현실에 대한 존재감(Presence)도 증가하는 것을 볼 수 있었다.

1. 서론

가상현실(Virtual Reality, VR) 환경의 한 단계 도약을 위해서 VR 디바이스가 유발시키는 시각적 멀미를 감소시키는 연구가 활발하게 수행되고 있다. 시각적 멀미는 가상현실환경 구성에 있어 사용자 편의성을 감소시키는 가장 큰 요소이다. (그림 1)의 (a)와 같이 현실에서 물체를 주시할 때, 초점거리와 주시거리는 동일한 거리를 가진다. 하지만 (b)와 같은 스테레오 환경에서, 주시거리는 (a)와 동일하지만 눈의 초점은 디스플레이에 위치하는 인지 불일치가 시각적 멀미를 유발하게 된다. 본 논문에서는, 이러한 문제를 해결하기 위해 시각적 자극에 의한 사이버 멀미 해결 방법이 아닌, 가상현실 환경에서 사운드 제어를 사용하여 인지의 불일치를 보정하는 연구를 수행하였다.

비디오게임에 대한 시각적 멀미와 존재감에서 사운드의 역할에 대한 연구들 [1, 2]에서는, 큰 역할을 하지는 않지만 사운드가 시각적 멀미를 감소시킨다는 결과가 보고되었다. 본 연구에서는 이러한 연구들을 바탕으로 가상현실환경에서 사운드를 이용한 시각적 멀미감소와 현실감 증가에 대한 연구를 수행하였다.

2. 가상현실환경에서의 시각적 멀미와 사운드 실험

2.1 사운드 게인(Sound Gain)

사운드 게인이란 물체의 운동에서 발생하는 사운드를 실제 물체보다 더 멀게 또는 가깝게 위치시켜 왜곡함으로써 사용자들의 시각적 멀미를 감소시키는 방법이다.

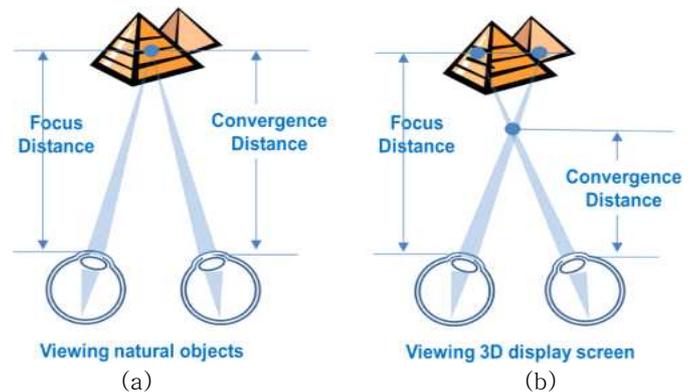


그림 1: 초점과 주시 운동 파괴에 의한 사이버멀미 유발

사운드 게인을 본 논문의 실험에 도입하여 초점과 주시 운동파괴에 의한 시각적 인지불일치를 청각적 왜곡을 통한 청각적 자극으로 시각적 인지불일치를 보정하는 가능성을 테스트하고자 한다. 예를 들어 사운드 게인이 2.0인 경우 사용자와의 거리가 1인 물체가 사용자를 향해 출발할 때, 물체의 사운드는 거리가 물체보다 두 배인 2의 위치에서 출발하게 된다. 이때 사용자에게 도달되는 시간은 동일해야 하므로 사운드는 물체보다 2배의 속도로 사용자에게 접근한다(그림 2 참조).

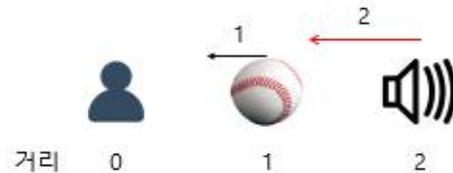


그림 2: 2.0의 사운드 게인 예시

2.2 실험 설계

본 연구의 실험은 가상현실 야구게임 환경에서 피칭머신에서 발사하는 공을 스트라이크와 볼로 구분하는 것이다 (그림 3). 실험에서는 사운드를 사운드가 없는 환경, 사운드의 크기가 일정한 2D 사운드 환경, 사용자와의 거리에 따라 사운드의 크기가 달라지는 3D 사운드 환경의 세 가지로 나누어서 실험을 진행했으며, 3D 사운드 환경에서는 사운드 게인(Sound Gain)을 0.5, 1.0, 2.0의 세 가지 경우로 나누어 총 다섯 가지 환경에서 실험을 진행했다. 이 때 피칭머신에서 불타는 공을 발사하여 사운드의 효과를 증대하고자 했다.

* 요약 논문 (Extended Abstract)

* 포스터 발표논문 (Poster Presentation Paper)

* 본 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임.

(No. NRF-2017R1A2B4005469)

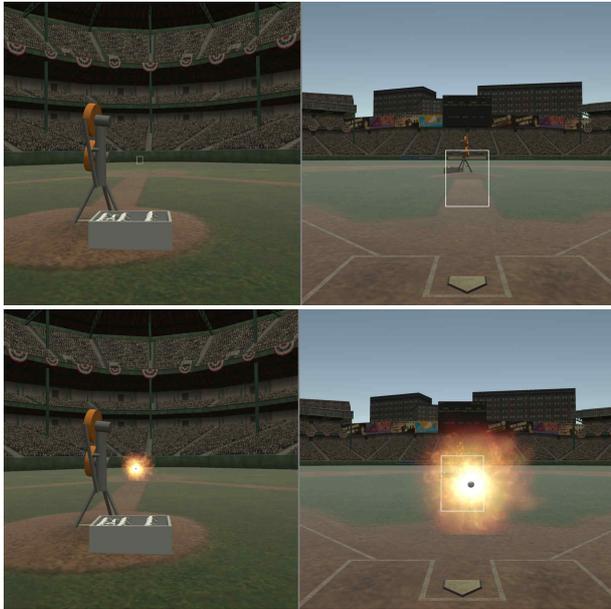


그림 3: 불타는 공을 발사하는 실험 환경의 예

실험에 앞서서 사용자는 훈련세션을 거치며 실험 환경에 적응하는 시간을 가진다. 훈련세션이 끝나면 각 실험 환경마다 10개의 공을 스트라이크와 볼로 판정하는 실험을 진행한다.

2.3 실험의 설문과 실험 참가자

시각적 멀미를 측정하기 위한 설문지는 SSQ (Simulator Sickness Questionnaire)[3]을 사용하였고, 현실감 측정에는 PQ (Presence Questionnaire)[4]를 사용하였다. SSQ가 클수록 멀미가 크다는 것을 뜻하고, PQ의 경우에는 클수록 존재감이 크다는 것을 뜻한다. 실험 참가자는 총 다섯 명이며, 평균 26세이고, 참가자 모두 가상환경 디바이스에 대한 경험이 있었다.

2.4 실험의 가설

- 가설 1: 시각적 멀미는 no sound > 2D sound > 3D sound의 순서가 될 것이다.
- 가설 2: 사운드 계인이 커 질수록 시각적 멀미가 감소할 것이다.
- 가설 3: 현실감은 no sound < 2D sound < 3D sound가 될 것이다.
- 가설 4: 사운드 계인이 커 질수록 존재감이 증가할 것이다.

3. 결과

(그림 4)의 그래프에서 SSQ 값은 각 환경에서의 실험 후의 SSQ에서 실험 전의 SSQ 점수를 뺀 값의 평균이고, PQ 값은 각 환경에서의 PQ점수의 평균이다. 전체 환경에서의 SSQ 값과 PQ 값의 평균은 각각 2.47, 44.04 이고 두 값의 표준편차는 각각 2.41, 3.58이다. (3D, 2.0Gain) 경우를 제외하고 오히려 멀미가 증가한 결과가 나왔다. PQ의 경우, 사운드가 없는 경우 빼고는 모두 비슷하게 현실감이 높아진 결과를 보이고 있다.

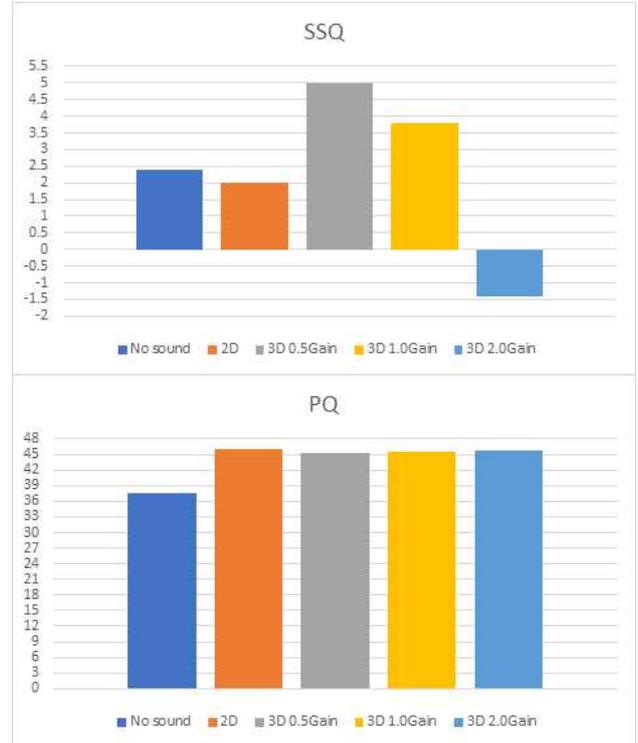


그림 4: SSQ, PQ 점수의 실험환경에 따른 평균 그래프

사운드의 입체감이 변해도 멀미감은 비슷하거나 오히려 사운드가 없을 때 보다 멀미감이 심해졌으므로 가설 1은 틀렸다고 말할 수 있다. 계인과 멀미감의 관계로 보아 가설 2는 부분적으로 맞았다. 사운드가 없을 때보다 있을 때 현실감이 증가하고 있으므로 가설 3은 부분적으로 맞았으나, 계인과 현실감의 관계는 모호하므로 가설 4는 맞았다고 볼 수 없다.

4. 결론

본 연구에서는 이렇게 사운드의 입체감과 계인을 변화시켰을 때 사용자가 느끼는 멀미감의 변화를 측정해 보고자 하였다. 사운드는 사용자의 멀미감을 감소시키고 현실감을 높이는데 기여할 수 있다는 것을 일부 확인할 수 있었다. 그러나, 가설과 일치하지 않는 결과들이 나온 가장 큰 요인은 실험 참여자의 수가 많지 않았기 때문으로 추측된다. 사운드의 유무나 현실감 측정에 의미 있는 데이터를 얻기 위해서 더 많은 실험 참여자를 대상으로 계속 실험을 진행 하고 있으며, 야구 게임 이외에도 다른 형태의 실험을 더 추가하여 실험할 계획이다.

참고문헌

[1] B. Keshavarz and H. Hecht, "Visually induced motion sickness and presence in videogames: The role of sound", Proc. of Human Fact. and Ergo. Soc., 56(1), 2012.
 [2] J. Dalman, et al., "Could sound be used as a strategy for reducing symptoms of perceived motion sickness?", Journal of Neuro Engineering and Rehabilitation. 2008.
 [3] R. Kennedy, et al., "Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness", Int.J.of Avi.Psychology, 3(3), 1993.
 [4] B. Witmer, et al., "Measuring Presence in Virtual Environments: A Presence Questionnaire", Presence: Teleoperators and Virtual Environments, 7(3), 1998.

증강현실을 이용한 아동 대상 영어단어 학습 콘텐츠 구현*

김은경⁰, 유영재, 정지연^{1,2}, 송창근, 이정, 김선정
 한림대학교 융합소프트웨어학과, 심리학과¹, 인터랙션디자인 대학원²
 {dossiba9603, YYJ0716}@hallym.ac.kr, qkdua45@naver.com, {cgsong, airjung, sunkim}@hallym.ac.kr

Implementation of English Word Learning Contents for Children using Augmented Reality

Eun-Kyoung Kim⁰, Young-Je Yu, Ji-Yeon Jeong¹, Chang-Geun Song, Jung Lee, Sun-Jeong Kim
 Dept. Convergence Software, Dept. of Psychology¹, Graduate School of Interaction Design², Hallym University

요약

아동을 대상으로 하는 다양한 영어 교육 콘텐츠 중 효과적으로 아동들이 영어 단어를 학습할 수 있도록 증강현실을 기반으로 하는 콘텐츠를 구현한다. 학습에 대한 흥미 유발을 통해 학습자에게 자율적인 학습 환경을 구축하기 위해 콘텐츠는 영어 단어에 함께 대응되는 3D 가상물체를 애니메이션으로 디스플레이 한다. 영어 단어에 대응되는 3D 가상물체가 영어 단어를 중심으로 원을 그리는 움직임을 통해 학습자들이 영어 단어에 집중하여 보다 효과적으로 암기할 수 있도록 한다.

1. 서론

현재 수 많은 증강현실 콘텐츠가 존재한다. 특히 증강현실은 교육, 건설디자인, 게임 등 실시간 상호작용이 중요한 분야에 많이 사용되고 있다. 증강현실이 이와 같은 분야에서 주목 받고 있는 이유는 실시간으로 사용자와 가상물체간의 상호작용이 용이한 기술 중 하나이기 때문이다.

교육 분야에서 사용자, 즉 학습자와 가상물체간의 상호작용이 중요한 이슈 중 하나인 이유는 학습자에게 즉각적인 궁금증을 해소시킬 수 있고 단순히 Text형식이 아닌 시각적, 청각적 요소 등을 해소시켜 주어 학습에 많은 도움을 주기 때문이다. 본 논문에서는 구성주의 기반의 교육에 입각하여 유용한 콘텐츠를 제공한다[1]. Shelton은 증강현실과 교육에 있어서 구성주의적인 관점으로 바라보고 있는데 구성주의의 핵심 의견은 학습은 학습자에 의해 구성된다는 것이다[2]. 학습이 학습자에 의해 구성되기 위해서는 능동적인 학습을 유도해야 한다. 증강현실을 통해 가상물체와 상호작용을 하며 흥미 유발을 통해 자율적인 학습 환경을 구축해 주는데 도움을 준다.

또한 주의집중은 아동 학습에 있어서 새로운 정보를 습득하는데 반드시 필요하고 특히 학습을 증진시킬 수 있는 주요 기초 능력이다[4]. 한 가지 특정 대상을 제시하면 한 활동에 집중하는 주의 집중도가 높아진다. 따라서 학습을 증진시켜 줄 뿐만 아니라 사용자 기억에도 도움을 주기 때문에 영어 단어 학습에 큰 기여를 한다[5]. 본 논문에서 주의 집중도를 높일 수 있는 방법을 모색해 보았다.

2. 본론

2.1. 개발 도구

본 논문에서는 증강현실을 이용하여 구현하였다. 증강현실에는 마커(Marker) 기반 방식과 비마커(Marker-less) 기반 방식 두 가지로 나뉜다.

마커 기반 방식은 약속된 특정 형태의 무늬를 추적하는 방식이다. 이 방식은 비마커 기반 방식에 비해 알고리즘이 복잡하지 않고 쉽게 인식이 된다는 장점이 있다. 따라서 시간 지연 없이 실시간으로 인식이 가능하다.

본 논문에서 제안하는 콘텐츠는 영어 단어에 대응하는 3D 가상물체를 제공한다. 즉, 영어 단어를 마커로 사용한다. 이에 맞는 개발 도구인 ‘웹캠’ 사에서 제작한 증강현실 소프트웨어 플랫폼 Vuforia를 사용하였다.

2.2. 주요 개발 과정

Vuforia에서는 10만 가지의 단어까지 영어 단어를 인식하는 것이 가능하다. 단, 마커로 쓰이는 영어 단어는 영어사전에 게재된 단어 또는 일상에서 자주 쓰이는 명사가 가능하다. 그림 1처럼 다채로운 색상이나 기울인 글꼴에도 인식률이 좋다.



그림 1: 마커로 사용된 영어 단어

* 포스터 발표논문

* 학부생 주저자 논문임

* 본 연구는 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2017R1D1A1B03035576).

마커로 해당 단어가 인식되면 아래 그림 2와 같이 그 단어를 Text로 출력한 뒤 단어에 대응하는 3D 가상물체를 찾아 띄운다. 가상물체에는 간단한 애니메이션이 포함되어 있다. 뿐만 아니라 주의 집중도를 높이기 위해 3D 가상 물체가 영어 단어를 중심으로 원운동 애니메이션을 하며 움직인다. 안드로이드로 빌드 하여 휴대성이 높아 어디에서나 실행이 가능하다.



그림 2: 실행 화면

3. 구현결과

본 논문에서 제시한 애플리케이션을 동작하면 일반 카메라 애플리케이션과 같은 화면(이하 카메라 화면)이 나타난다. 일반 사진이나 숫자는 인식하지 못하므로 상호작용 없이 카메라 화면이 출력 된다. 하지만 영어 단어를 인식하면 인식된 영어 단어가 마커의 하단에 나타난다. 만약 애플리케이션 내에 등록된 3D 가상 물체와 그에 부합하는 영어 단어가 인식되면 등록되어 있는 3D 가상 물체를 영어 단어 주변을 원을 그리며 움직인다.

4. 결론 및 향후 연구

본 논문에서는 마커 기반 증강현실을 이용한 영어 단어 학습 콘텐츠를 개발했다. 개발된 콘텐츠를 통해 다음과 같은 기대효과를 가질 수 있다.

첫째, 대상이 아동이므로 시각적인 자극에 약하다. 따라서 애니메이션을 통해 시각적 자극은 학습자에게 적극적인 학습 자세를 불러 일으킬 수 있다. 즉, 능동적인 학습을 유도한다.

둘째, 3D 가상물체를 여러 각도에서 관찰이 가능하다. 만약 평면 상의 그림이었다면 단면만 관찰하게 된다. 그에 비해 본 논문에서 개발한 학습 콘텐츠는 다각도에서 관찰이 가능하기 때문에 경험하기 힘든 물체 등 직접 경험하지 않아도 가상으로 간접 경험이 가능하다.

셋째, 기존의 증강현실을 이용한 아동 대상 영어단어 학습 콘텐츠는 3D 가상물체의 움직임이 정적이었다. 반

면 본 연구로 개발된 콘텐츠는 3D 가상물체가 원을 그리며 영어 단어 주변을 움직여 기존의 콘텐츠에서 제공하는 3D 가상물체의 움직임에 비해 동적이다. 이는 성인에 비해 집중력이 약한 아동의 시선을 사로잡아 이목을 집중시킨다. 따라서 사용자는 자연스럽게 3D 가상물체의 움직임을 따라 대상을 주의 집중하며 주시하는 시간이 늘어나고 주의집중도가 높아진다. 주의집중도가 높아지면 새로운 정보를 습득하는데 도움을 주는데 주시한 3D 가상물체는 단어와 대응되는 이미지를 가진 물체이다. 이는 주의집중된 물체의 이미지와 시야에 근접해 있는 영어단어를 보다 효과적으로 기억하게 되며 대응되는 물체의 이미지와 영어단어가 동시에 연상되는 학습효과가 있다. 따라서 본 논문에서 제시한 콘텐츠는 기존의 콘텐츠와 비교했을 때 3D 가상물체의 동적인 움직임을 통해 더욱 쉽게 영어 단어를 기억할 수 있도록 학습에 도움을 준다.

향후에는 사용자가 일상생활에서 빈도수가 높은 명사 단어 약 1000개에 관련된 가상 물체를 추가할 수 있을 뿐만 아니라 사용자가 직접 해당 단어에 대응하는 물체를 스캐닝 하여 클라우드 서버에 올릴 수 있게 하여 가상 물체의 다양성을 넓히는 콘텐츠를 추가하여 보완할 계획이다.

그리고 본 연구로 개발된 콘텐츠의 교육효과를 측정하기 위해, 텍스트, 시각, 청각 등을 조합하여 멀티미디어 콘텐츠가 아동들의 암기 학습에 효과적인지를 검증할 계획이다.

참고문헌

- [1] 서진원, 구성주의 교육방법의 구현요소로서의 학교도서관 활용수업에 관한 연구, *한국도서관·정보학회지*, 제 42권, 제 2호, 215-236 (22 pages), 2011.6.
- [2] B. E. Shelton, How augmented reality helps students learn dynamic spatial relationships, Unpublished doctoral dissertation, University of Washington, Washington, 2003.
- [3] 백승만, 디지털 콘텐츠에서 애니메이션을 이용한 학습용 교재 개발 연구, *디지털디자인학연구*, 제 3권, 13-20 (8 pages), 2002.2.
- [4] Rizzo A. A., & Buckwalter, U., Psycho-neu-ro-physiological assessment and rehabilitation in virtual environments: Cognition, clinical, and human factors in advanced human computer interactions, Amsterdam: *IOS Press*, 1997.
- [5] 최계영, 백화점 공간의 연속 주시에 나타난 주의집중 특성, *한국실내디자인학회 논문집*, 제 24권, 제 6호, 128-136(9 pages), 2015.12.

마커 기반 증강 현실을 이용한 네일 아트 어플리케이션*

한예진⁰, 서수종^{1,2}, 송창근, 이정, 김선정
 한림대학교 융합소프트웨어학과, 심리학과¹, 인터랙션디자인 대학원²
 {nekoto96, westwaterbell, cgsong, airjung, sunkim}@hallym.ac.kr

A Nail Art Application Using Markers Based on Augmented Reality

Yejin Han⁰, Sujong Seo^{1,2}, Chang-Geun Song, Jung Lee, Sun-Jeong Kim

Dept. of Convergence Software, Dept. of Psychology¹, Graduate School of Interaction Design², Hallym University

요약

기술의 발전에 따라 소비자들은 재화의 구매 전 성공적인 소비를 위해 간접적으로 대상을 체험해 보는 방식으로 소비패턴이 바뀌었다. 본 논문에서는 증강 현실(Augmented Reality)를 활용한 콘텐츠들 중 서비스 혹은 상품의 구매 전 체험을 돕는 어플리케이션을 제안한다. 네일 아트 서비스를 받고자 하는 소비자가 직접 손에 서비스를 받기 전 자신에게 어울리는 디자인을 증강 현실 어플리케이션을 통해 체험해 볼 수 있다. 사전 체험을 통해 시행착오 없이 마음에 드는 디자인을 선택할 수 있도록 도움 줄 것을 기대한다.

1. 서론

디지털 기술이 발전함에 따라 가상 현실(VR, Virtual Reality)에 대한 세간의 관심이 증가하였다. 그에 따라 가상 현실 기술을 활용한 게임, 마케팅, 의료용, 쇼핑 프로그램 등이 활발하게 제작 및 사용되고 있다[1]. 특히 증강 현실(AR, Augmented Reality)은 현실과 가상 물체를 이음매 없이(seamlessly) 합쳐 실시간으로 보여줘 사용자에게 몰입감을 준다[2]. 오늘날에는 매우 상용화 되어 우리 주변에서 흔히 볼 수 있는 기술이다. 게임 '포켓몬 GO'와 마이크로소프트(Microsoft)사의 홀로렌즈(Hololens)의 사례들을 통해 인기를 실감할 수 있다.

특히 쇼핑분야에서도 높은 활용을 들 수 있다. 이케아(Ikea)의 이케아 플레이스(IKEA Place)는 구매 전 가구를 집에 미리 배치해 볼 수 있어 구매를 돕는다[3].

또 FX기어(FXGear)사의 가상현실 피팅 기기 FX미러(FX Mirror)는 3D가상 피팅 서비스를 제공해 기존 피팅의 번거로움을 줄여준다.

본 논문에서는 네일 아트 서비스에 AR을 접목하여 사용자가 쉽게 간접적으로 서비스를 체험할 수 있는 어플리케이션을 제안한다. 통계청에 따르면, 2017년 네일 아트 창업 규모는 3조원에 달한다. 그 수요와 사업체는 최근 몇 년간 크게 발전했다. 또한 최근 네일 서비스에 대한 인식이 변화하여 서비스를 받는 대상은 20~50대로 다양한 연령층을 보이고, 남성들의 활발한 참여로 네일 케어는 더이상 여성의 전유물이 아니게 되었다[4].

시장의 규모가 점점 커질 수로 소비자가 받을 수 있는 서비스의 종류도 다양해진다. 본 어플리케이션을 통해 미리 다양한 디자인들을 손에 적용시켜본다면 단 한번의 시도로 원하는 디자인을 고를 수 있어 시술 비용과 시간을 단축시키는 데에 효과적일 것이라 예상된다.

2. 본론

2.1. 관련 연구

본 논문에서는 증강 현실 구현이 가장 핵심적인 기술이다. 증강 현실에는 위치와 이미지 자체를 인식하는 마커리스(markerless) 방식과 컴퓨터가 인식하기 좋은 특정 패턴, 예를 들어 QR코드 같은 것을 인식하는 기술을 marker방식이라고 한다.

2011년 전형배와 홍성용은 증강 현실 네일샵 적용 방안으로 마커리스 방식을 채택하고, 색상을 통해 손가락 끝을 검출하여 해당 위치에 매니큐어 색상을 보여주는 기술을 제안하였다[5]. 하지만 본 논문에서는 marker방식을 사용해 3차원 좌표를 정확히 받아들여 현실감을 줄 것이다. 또, 마커(marker) 방식이 마커리스 방식보다 계산 속도가 빠르기 때문에 소비자들이 어플리케이션을 통해 디자인이 적용된 자신의 손을 실시간으로 볼 수 있을 것이다[6].

* 포스터 발표논문

* 학부생 주저자 논문임

* 본 연구는 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2017R1D1A1B03035576).

2.2. 설계

본 연구에서 제안하는 시스템의 전체적인 구조는 다음과 같다. 증강 현실은 퀄컴(Qualcomm)에서 제작한 증강 현실 소프트웨어 뷰포리아(Vuforia)와 Unity3D를 사용하여 구현한다. 사용자는 디자인을 입혀보고 싶은 손톱에 마커가 그려진 인조 네일 틱을 부착한다. 이때 마커는 뷰포리아의 데이터 베이스에 저장된 것으로 모든 손가락에 네일틱을 부착할 수 있도록 10가지의 종류를 가진다. 뷰포리아에서 인식하기 좋은 마커는 뚜렷한 명암 등의 특징이 필요하므로 그림 1과 같은 마커를 디자인하였다. 어플리케이션 실행 후 마커가 인식이 될 때 디자인이 증강되도록 설계했다. 어플리케이션을 실행한 뒤 원하는 디자인을 선택하고 자신의 손을 휴대폰 카메라로 비추면 증강된 네일 디자인을 그림2와 같이 볼 수 있다.



그림 1 어플리케이션에 사용된 마커 디자인과 인조 네일틱

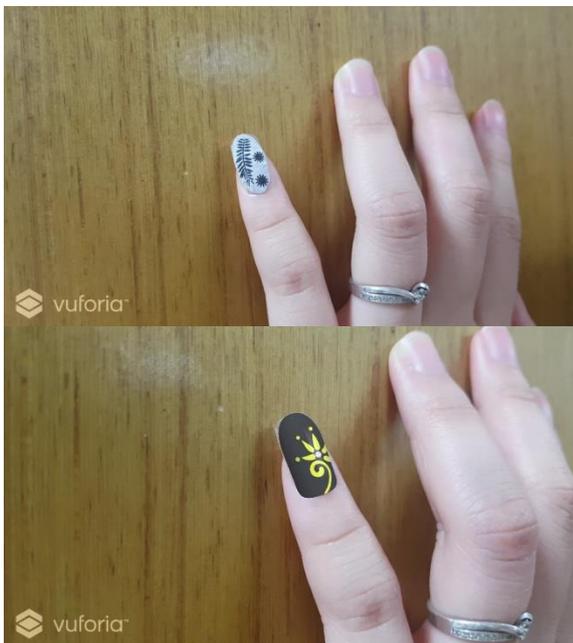


그림 2 네일 디자인이 증강되기 전(상)과 후(하)의 모습

사용자는 네일 디자인을 바꿔가며 자신의 손과 어울리는 혹은 자신의 마음에 드는 디자인을 고른 뒤 실제 네일 서비스를 받으면 된다.

3. 결론

본 논문에서는 4차 산업혁명시대의 대표적인 차세대

기술로 손꼽히는 마커 기반 증강 현실 네일 아트 어플리케이션을 제안하였다.

본 연구에서 기대되는 점은 네일 아트를 받는 소비자들이 보다 편리하게 다양한 디자인을 미리 적용해보고 시행착오를 겪지 않는 것이다. 스마트폰 어플리케이션을 사용하기 때문에 간편하게 사용할 수 있다.

다만 마커 기반이기 때문에 항상 마커가 그려진 인조 네일 틱을 가지고 수반해야 하는 번거로움이 생긴다. 그렇기 때문에 마커리스 방식이면서도 손끝 검출 계산 및 3차원 좌표 계산을 빠르고 정확하게 하는 알고리즘의 고안이 필요하다.

그리고 현재 어플리케이션에서는 미리 등록된 디자인만 사용해 볼 수 있고 디자인 자체는 2차원 이미지이기 때문에 생동감이 떨어진다. 인터넷에서 디자인을 받아오는 방식 혹은 사용자가 직접 촬영하고 등록된 디자인을 증강할 수 있는 어플리케이션으로 발전시킬 수 있다. 또 손톱과 유사한 3차원 모델을 만들어 디자인의 무늬 뿐만 아니라 파츠 혹은 큐빅 같은 악세서리들까지 더욱 현실성 있는 어플리케이션을 만들 수 있도록 지속적으로 확장할 계획이다.

참고문헌

- [1] 최병록, “4차 산업혁명시대의 소비자이슈와 소비자정책”, 한국기술혁신학회 학술대회, pp. 471~479, Nov, 2017.
- [2] R. Azuma, “A Survey of Augmented Reality,” *PRESENCE: Teleoperators and Virtual Environment*, vol.6, no.4, pp. 355-385. Aug. 1997.
- [3] 조희경, 김성훈, “증강현실 어플리케이션 사용자의 참여와 몰입도 증진을 위한 콘텐츠 디자인 연구”, 한국디자인문화학회지, 제 23 권, 제 4 호, pp.733-742, 2017.
- [4] 김은영, “유형별 네일샵의 서비스품질이 고객만족도에 미치는 영향”, 건국대학교 석사학위논문, pp. 15-62, 2011.
- [5] 전형배, 홍성용, “증강현실을 이용한 네일 샵 적용 방안 연구”, 한국정보과학회 학술발표논문집, 제 38 권, 제 1C 호, pp.362-364, 2011.
- [6] 이상진, 류남훈, 오경숙, 이혜미, 김응곤, “증강현실 기반의 3D 어류 백과사전 시스템 설계”, 한국통신학회 학술대회논문집, pp.709-710, 2010.

시뮬레이터 멀미의 요인 조사 및 분석*

전한재, 송창근, 이정, 김선정
 한림대학교 융합소프트웨어 학과
 {jhj7872, cgsong, airjung, sunkim}@hallym.ac.kr

Investigation and Analysis of Simulator Sickness Factors

Hanjae Jeon, Chang-Geun Song, Jung Lee, Sun-Jeong Kim
 Dept. of Convergence Software, Hallym University

요약

시뮬레이터 멀미는 시각을 통해 얻은 움직임 정보와 전정 기관의 정보가 일치하지 않는 전정 불일치 현상, 그리고 VR을 체험하기 위한 기기의 성능이 사람과의 성능차를 만족하지 못해 실제 세계에서 체험하지 못한 경험이 원인으로 꼽힌다. 악영향을 끼칠 것이라 예상되는 요인을 조사하고 측정 방법에 대해 정리하였고, 이를 사용자가 느낀 불쾌감과 비교해 화면 패턴의 어떤 움직임이 시뮬레이터 멀미에 큰 영향을 끼치고, 어느 정도의 상관관계를 가지는지를 조사할 수 있을 것으로 생각된다.

1. 서론

시뮬레이터 멀미라고 불리는 이 현상을 해소하기 위해 기존 멀미의 이론과 시뮬레이터 멀미만의 이론이 제기되고 원인을 분석하는 연구가 진행되어왔다. 본 논문에서는 시뮬레이터 멀미에 영향을 끼칠 것으로 예상되는 특징적인 요인을 선정하고 조사 방법을 정리한다.

2. 시뮬레이터 멀미의 원인

2.1. 멀미

멀미는 시각적으로 감지된 움직임과 전정 기관으로 감지된 움직임 사이에 모순이 존재할 때 발생한다는 감각 충돌 이론[1]이 존재한다. 차에서 책을 읽으면 멀미가 나는 것은 전정 기관은 차가 주행한다고 뇌로 정보를 보내지만 시각으로는 차의 움직임과는 관련 없는 책을 보고 있기에 시스템 사이에 충돌이 일어나는 것을 이유로 들 수 있다. 같은 이유로 VR을 체험할 때, 디스플레이를 통해 움직임에 대한 정보를 뇌로 보내나 실제로 움직이지는 않으므로 시스템간 충돌이 일어나는 것을 이유로 들 수 있다.

2.2. Vection

시뮬레이터 멀미의 연구에서 보이는 공통적인 결론은 참가자의 물리적 움직임에 부적절한 시각으로 유도되는

Vection에 기인한다는 것이다. Vection이란 사람 주변의 환경이 움직이는 것을 보면 자신의 몸이 움직이고 있다고 착각을 느끼게 되는 현상이다. 세로 줄무늬 패턴을 가지는 원기둥 안에 사람이 들어가고 원기둥을 회전시키면 사람이 회전하고있다고 착각을 느끼게 된다. 움직이는 장면에서 휘도 변화, 색차 변화, 모양 및 패턴의 변화 중에서 Vection의 레벨에 가장 큰 영향을 미치는 것은 휘도 변화라는 연구가 있다[2][3].

Richard H. Y. So[4]의 연구에서 사용자가 보는 화면의 power spectrum density를 계산해 현재 화면의 복잡도를 계산하고 VR내 카메라의 각속도와 속도를 곱하는 방법을 제안했다. 그러나 이 방법으로는 상기한 예시의 경우 값을 구하지 못할 수 있으며 또 현재 화면만을 사용해 복잡도를 계산하므로 연속적인 화면의 시간에 대한 요소가 빠져 있을 것으로 생각된다. 이에 본 연구에서는 이전 화면과 현재 화면을 이용해 Optical Flow를 계산하고 특정 마스크와의 내적을 통해 Vection을 계산하는 방법을 확인하고 검토한다.

Vection은 3축에 따른 회전 및 이동을 각각 나누어 총 6가지의 타입을 가지게 된다[5]. 사용자가 바라보는 풍경이 이전과 비교해서 어떻게 변화했는가를 Optical Flow를 통해 표현할 수 있으며, 6 타입의 Vection은 각기 다른 Optical Flow의 특징을 가진다.

2.3. HMD의 특성

감각 충돌 이론 외에, HMD를 통해 경험하는 공간이 실제 공간과 차이 나기때문에 평소에 전혀 경험하지 못한 체험으로 불쾌감을 얻는다는 연구가 있다[6]. 이 연구에서는 VR을 체험하기 위해 지금은 널리 쓰이는 HMD라는 기기의 성능이 사람의 눈이 가질 수 있는 화면 해상도나 화면 시야 각, 화면 갱신을 등등을 만족하지 못해 불쾌감을 얻는다고 한다. 또, 사용자 머리의 움직임을 파악할 수 있도록 기기 주변의 풍경을 볼 수 있도록 실험을 했을 때 멀미가 현저하게 줄어드는 것을 확인했다고 한다.

2.4. 특징적인 요인의 선정

특징적인 요인은 크게 HW, SW 두 그룹으로 나눠서 생각한다. HW측 요인은 HMD의 화면 해상도, 화면 시야 각, 지연시간, 주변시야 차단 여부를 조사한다. 이들

* 포스터 발표논문

* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 제12회 한국정보과학회/한국빅데이터학회 공동 학술 심포지엄에 발표되었음.

* 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF -2017R1D1A1B 03035576).

은 기술 발전에 따라 눈과의 성능차를 극복할 수 있을 것이다. SW측 요인은 Optical Flow를 통해 계산된 6 타입의 Vection을 조사한다. 이 값과 사용자가 느끼는 불쾌감과 비교해 화면 구성물이 어느 방향으로 회전/이동하면 제일 불쾌감이 발생하는지, 어느 빈도로 회전/이동 패턴이 나타나면 불쾌감이 증가하는지를 분석할 수 있을 것이다. 이를 위해, Vection 데이터를 계산하는 방법을 정하고 예상하는 데이터가 나오는지 확인한다.

2.5. 실험 방법

회전, 가속, 감속이 포함되는 터널 주행을 시뮬레이션 대상으로 삼는다. 도로가 아닌 터널을 배경으로 삼은 이유는 시선 집중할 물체를 최대한 배제함과 VR 환경을 꾸미기 위한 물체를 최소화해 계산 요구량을 줄이기 위함이다. 터널 내부에는 먼 간격으로 전등을 배치해 터널 내부는 밝은 곳과 어두운 곳이 일정 간격으로 존재하며, 주행 시 반복되는 빛의 자극을 받도록 되어 있다. 사용자는 앞방향으로 천천히 가속했다가 일정 속도로 주행 후 천천히 감속하는 것을 체험한다. 각 20초, 10초, 20초 동안 수행하며 10번 반복한다. 체험 중 사용자가 느끼는 불쾌감을 0부터 10 사이의 숫자로 말하도록 주기적으로 질문을 한다. 실험 전후로 SSQ[7]라는 시뮬레이터 멀미를 정량화 하는 설문지를 작성하게 해, 주관적인 불쾌감 외에 보다 상세한 데이터를 구하도록 한다. Vection이 가지는 Optical Flow의 특징을 추출하기 위해 화면과 동일 크기의 행렬을 사전에 계산을 마친 뒤 터널 주행 시뮬레이션을 시작한다. 행렬 요소 하나 하나는 각 픽셀의 Vector가 되며 Vection의 타입과 좌표에 따라 다른 방향과 길이를 가지는 Vector이다. 주행 중, 현재 Optical Flow와 사전에 정의한 행렬 사이에 요소별 내적 수행 후 데이터를 더해 Vection을 구한다. 또, 이렇게 구한 값이 의도한대로 구해졌는지 확인하기 위해 사용자의 머리 회전 속도와 사용자 좌표 기준 머리 이동 속도를 계산한다. 원래 VR공간의 절대 좌표계를 기준으로 계산되었던 기차의 속도를 사용자의 머리 좌표계를 기준으로 계산하도록 변환시켜주고 기록한다. 이를 통해 VR 공간 안에서의 사용자의 이동 방향과 화면에 보이는 Optical Flow를 비교하도록 한다. 정면 이동하는 기차 방향과 화면이 동일한 방향이라면 Forward Vection의 변화가 머리 이동 속도 상승에 따라 유사하게 증가할 것이다. 그림 1의 그래프는 Forward Vecti-

on과 정면 이동 속도의 양상을 비교한 그래프이다. 모든 데이터는 초당 60번 기록한다.

2.6. 토론

실험 중 얻은 6 타입의 Vection 데이터를 구하고 예상과 같은 지 확인하기 위해 3 종류의 실험을 진행했다. 머리 방향을 기차 이동 방향과 일치시켜 사용자 기준으로 앞으로 이동할 때, 머리를 왼쪽으로 돌려 사용자 기준으로 오른쪽으로 이동할 때, 머리를 뒤로 젖혀 사용자 기준으로 아래로 이동할 때의 데이터를 확인한다. 전체적으로 회전 Vection의 변화폭이 선형 속도 Vection의 변화폭보다 큰 것을 확인했다. 이는 각 Vection의 계산법의 차이로 인한 것으로 생각된다. 회전 Vection 계산법은 모든 픽셀의 가중치를 동일하게 가지거나(Pitch, Yaw) 중앙 기준으로 방사 방향은 같은 가중치를 가지는데 반해 선형 속도 Vection 계산법은 원근에 따른 가중치를 다르게 지정했기 때문인 것으로 생각된다. 이 차이를 줄이기 위해 두 계산법에 다른 방식의 정규화를 적용할 것이다.

3. 결론

시뮬레이터 멀미에 영향을 끼칠 것으로 예상되는 요인을 조사하고 측정 방법에 대해 정리했다. Vection은 3축 회전/이동에 따라 6 타입으로 나뉠 수 있다. VR 내에서의 회전 및 이동 시 나타나는 화면의 움직임과 Vection의 패턴을 통해 값을 계산했다. 이 값을 예상대로 정확하게 구할 수 있다면, 추후 사용자가 느끼는 불쾌감과 비교해 화면 패턴의 어떤 움직임이 시뮬레이터 멀미에 큰 영향을 끼치고, 어느 정도의 상관관계를 가지는지를 조사할 수 있을 것으로 생각된다.

참고문헌

[1] J. T. Reason; J. J. Brand, Motion sickness, Motion sickness, 310, 1975.
 [2] Previc F. H.; Varner D. C., & Gillingham K. K, Visual scene effects on the somatogravic illusion, Aviation, Space, and Environmental Medicine, 1060-1063, 1992.
 [3] Livingstone M. & Hubel D, Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science, 240, 740-749, 1988.
 [4] Richard H. Y. So; Andy Ho; W. T. Lo, A Metric to Quantify Virtual Scene Movement for the Study of Cybersickness: Definition, Implementation, and Verification, Presence Volume 10, 193-215, 2001.
 [5] Steven M. LaValle, VIRTUAL REALITY, NPTEL and IIT Madras, 2016.
 [6] Jason D. Moss; Eric R. Muth, Characteristics of head-mounted displays and their effects on simulator sickness, Human Factors, 2011.
 [7] Robert S. Kennedy; Norman E. Lane; Kevin S. Berbaum & Michael G. Lilienthal, Simulator Sickness Questionnaire(SSQ) : An Enhanced Method for Quantifying Simulator Sickness, The International Journal of Aviation Psychology Volume 3, 203-220, 2009.

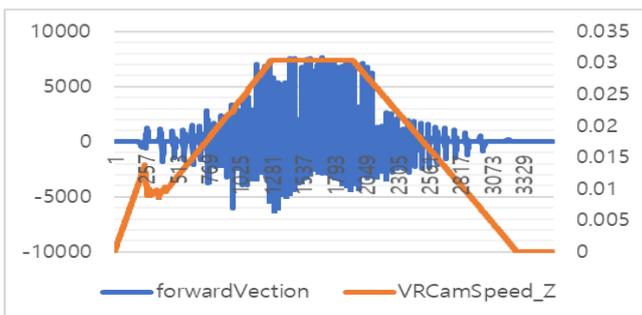


그림 1 정면 이동 시, z축 이동 Vection인 Forward Vection과 사용자 기준 z축 이동 속도를 비교한 그래프.

한국컴퓨터그래픽스학회 2018 학술대회 논문집

2018년 7월 10일 인쇄

2018년 7월 10일 발행

발행인: 김동호

편집인: 이지은, 최민규

발행처: 사단법인 한국컴퓨터그래픽스학회

주소: 서울특별시 성북구 삼선교로16길 116

한성대학교 연구관 911호(삼선동2가)

전화: 02-821-2889

FAX: 02-822-3622

URL: <http://cg-korea.org>

X GRAPHICS

From
To X

주최 | 한국컴퓨터그래픽스 학회

후원 |

