

수화 통역을 위한 VR 콘텐츠 개발

Development of Virtual Reality Contents for Korean Sign Language Interpretation

나길항, Kilhang Na*, 이병호, Byungho Lee**, 김종훈, Jonghun Kim**, 김종남, Jongnam Kim**, 정영기, Youngkee Jung***

요약 ~ 본 논문은 영화, 방송, 애니메이션 등의 다양한 동영상 콘텐츠에 수화 애니메이션을 합성하여 동영상 콘텐츠를 청각 및 언어장애인들에게 이해시키기 위한 수화 통역 VR 콘텐츠 시스템을 제안하고자 한다. 제안된 시스템은 수화 사전에 있는 수화들을 3D 애니메이션으로 DB화하기 위해, 모션 캡처 시스템과 데이터 글러브를 사용하여 실제 사람처럼 자연스러운 애니메이션을 생성하였다. 최종적으로 동영상 콘텐츠의 자막이나 대본의 구문분석을 한 후, 이를 수화용 단어자막을 통해 수화 애니메이션을 DB에서 검색한 후, 실시간적으로 기존 동영상 콘텐츠와 동기합성을 하여 수화 통역 콘텐츠를 제공하는 VR 콘텐츠 시스템을 구현하였고 이 시스템을 동화용 애니메이션에 적용하였다.

핵심어: Virtual Reality, Korean Sign Language, Motion Capture, Data Glove

1. 서론

사회가 발전함에 따라 사람들의 생활수준과 함께 의식수준이 향상되면서 장애인의 복지에 대한 사회적 관심이 점차 높아져 가고 있다. 국가차원에서 일반국민과 장애인 계층 간 정보격차를 줄이기 위해 많은 노력을 기울이고 있다. 이에 따라 청각장애인의 모국어라 할 수 있는 수화도 이전 어느 때보다 많은 주목을 받고 있다[1].

최근 텔레비전에서는 일반국민에 비해 정보 획득에 많은 어려움을 겪고 있는 청각장애인의 이해를 돕기 위해 일부 프로그램에서 수화 통역사를 통한 콘텐츠를 제공하거나 자막 방송을 내보내는 등 여러 의미 있는 시도가 이루어지고 있다. 뿐만 아니라 문자중계서비스, 영상중계서비스 등과 같은 통신중계서비스 (Telecommunication Relay Service:TRS)를 제공하고 있다. 하지만 대부분의 청각장애인들은 문자보다는 수화를 훨씬 빨리 습득하고 이해하기 때문에 수화 통역사를 통한 콘텐츠를 제공받기를 원한다[2-4].

인터넷을 통한 많은 사이버강좌에서 청각장애인들을 위해 수화통역 콘텐츠를 활용한 교육을 시도 하고 있다. 수화 통역 콘텐츠는 수화 통역된 교과목의 동영상 선택하면, <그

림1>과 같이 교수의 강좌와 함께 오른쪽 하단부분에 수화 통역사가 교과내용을 통역하는 화면이 나타나 학습이 용이하다. 또한 수화 통역 콘텐츠는 강의 동영상뿐만 아니라 영화, 방송, 애니메이션 등의 콘텐츠 등에도 활용될 예정이다.



그림 1. 사이버대학의 수화 통역 콘텐츠

이처럼 수화 통역 콘텐츠의 수요가 늘어나는 시점에서 수화통역사가 많이 필요하다. 현재 전국 122개의 수화통역센터에는 약 3백여 명의 수화통역사가 청각장애인의 교육상담,

*주저자 : 호남대학교 컴퓨터공학과 석사과정 e-mail: nallang3637@gmail.com

**공동저자 : 호남대학교 컴퓨터공학과

***교신저자 : 호남대학교 컴퓨터공학과 조교수; e-mail: ykjung@honam.ac.kr

직업상담, 의료상담 등을 위한 수화통역 서비스를 제공하지만, 매우 부족한 실정이다. 따라서 전문 수화통역사를 대신 해줄 수화 애니메이션 생성 시스템에 대한 연구들이 다양하게 이루어지고 있다.

하지만 기존의 수화 애니메이션은 <그림2>와 같이 OpenGL 및 DirectX를 기반으로 한 모델링에 지나지 않아 그 형태가 부자연스럽고 동작에 대해 파라미터 형태의 데이터베이스 구축으로 재사용이 어렵고, 동작이 부드럽지 못하였다. 또한 인식에 있어서의 어려움으로 인해 약 100여개에 지나지 않는 데이터베이스를 구축하고 있는 실정이다[5-6].

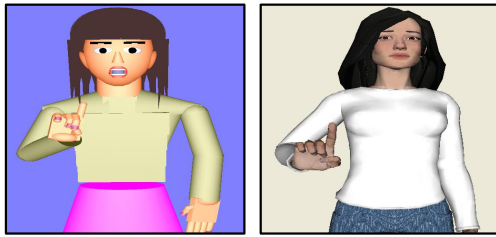


그림 2. OpenGL 기반의 수화 모델

본 논문에서는 청각 및 언어장애인에게 다양한 동영상 콘텐츠를 제공하기 위하여 수화 통역을 위한 VR 콘텐츠 시스템을 제안한다. 자연스러운 수화 동작을 추출하기 위해 모션 캡처 시스템과 정확한 손동작 표현을 위해 데이터 글러브를 이용하여 실제 사람과 같은 캐릭터 애니메이션을 생성하고자 하였다. 또한 자막/대본을 이용하여 자동으로 수화 통역 콘텐츠를 제작하기 위한 시스템을 개발 하였다.

2. 전체 시스템 구성

본 논문의 전체 시스템 구성은 <그림3>과 같다. 수화로 통역할 동영상 콘텐츠와 수화용 자막 파일을 불러오는 모듈과 수화 사전 데이터베이스와 수화 애니메이션 데이터베이스가 존재하는 수화 데이터베이스 모듈, 가상현실 공간에서 수화 애니메이션의 움직임을 실시간 3D 렌더링하여 화면에 보여주는 수화 뷰어 모듈, 마지막으로 기존 동영상과 최종 렌더링 화면을 동영상 파일로 저장하는 모듈로 구성된다.

먼저, 수화 통역 콘텐츠로 제작할 동영상 콘텐츠를 준비하고, 해당 동영상의 수화 애니메이션을 실행하기 위한 수화용 자막을 만든다. 이렇게 만든 수화용 자막은 해당 단어에 맞는 수화 애니메이션을 재생하기 위해 만든다. 이렇게 만든 수화용 자막과 준비한 동영상을 불러오고, 앞에서 만든 수화용 자막에 맞게 실시간으로 해당 수화 단어 데이터베이스와 수화 애니메이션 데이터베이스를 검색하여 해당하는 수화 단어를 찾는다. 찾은 수화 단어의 애니메이션을 실시간으로 화면에 렌더링하게 된다. 그렇게 하여 생성된 수화 애니메이션을 기존의 동영상과 병합하여 하나의 최종 수화 통역 동영상 콘텐츠가 제작된다.

3. 수화 애니메이션 생성

수화 애니메이션 생성에서는 자연스럽게 정확한 수화 애니메이션을 보여주기 위해 바디 모션 획득 파트와 손동작 모션 획득 파트로 분리하였다. 바디 모션 획득에서는 모션 캡처 시스템을 이용하였고, 손동작 모션 획득에서는 데이터 글러브를 이용하여 정확한 수화를 표현하고자 한다.

3.1 모션 캡처 시스템을 이용한 바디 모션 획득

기존 OpenGL 방식의 3D 모델링은 너무 딱딱하고 부자연

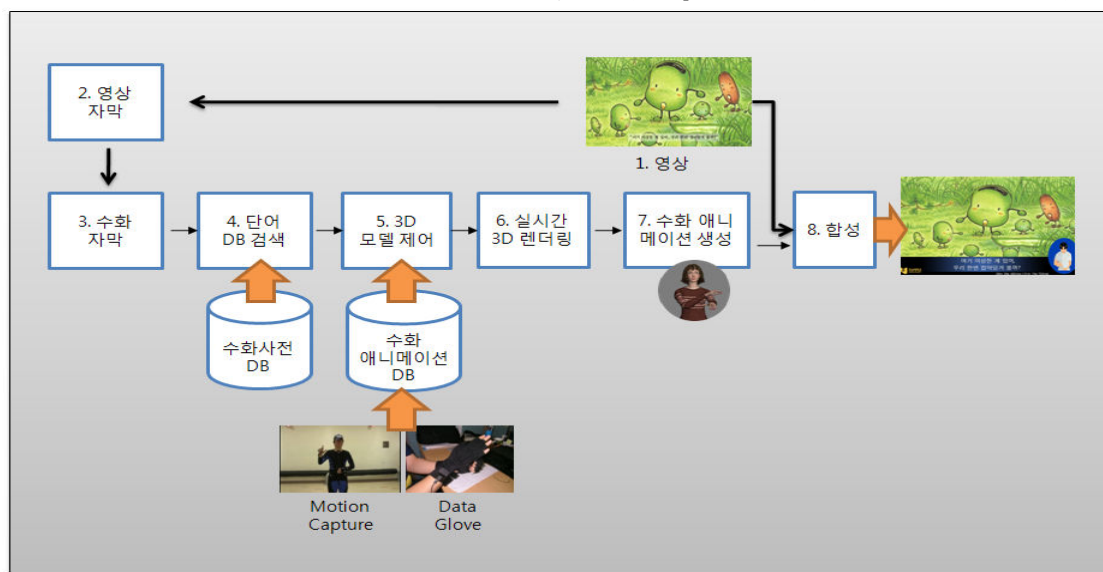


그림 3. 전체 시스템 구성도

스러우며, 거북한 감이 없지 않아 있었다. 여기서는 보다 자연스럽고 친근한 캐릭터 애니메이션을 위해 모션 캡처 시스템을 이용하였다.

모션 캡처는 사람이나 물체(Object)의 움직임을 컴퓨터가 사용 가능한 형태로 기록하는 것을 의미한다. 넓은 의미로는 사람이나 동물 등과 같은 다관절체의 움직임을 추적하고 움직임을 분석 가능한 데이터를 만드는 과정이고, 좁은 의미로는 움직임을 녹화하고 데이터화 하여 컴퓨터로 보내는 일련의 과정을 의미한다. 또한 물체의 움직임을 추적하여 얻어진 데이터를 모델링 된 캐릭터에 적용하여 같은 움직임을 하도록 하는 기술이다.

모션 캡처를 이용하면 자연스러운 동작의 위치값(X, Y, Z)과 회전값($\theta_x, \theta_y, \theta_z$)을 얻을 수 있고, 이를 보다 쉽게 데이터베이스화 할 수 있다. 모션 캡처의 종류에는 기계적 방식, 자기장을 이용하는 방식, 광학적 방식 등이 있지만, 그 중에서도 가장 일반적으로 많이 알려져 있는 광학적 방식은 적외선 카메라로 사람이나 사물에 부착된 광학적인 센서(반사구)인 마커(Marker)를 인식하는 방식이다. 광학적 방식의 장점으로서는 마커들의 위치를 조정하기 쉬우며, 고속으로 촬영할 수 있어 유실되는 동작이 거의 없고, 동작에 제한이 없어 액터(Actor, 배우)가 자유로운 동작을 할 수 있어서 매우 디테일한 동작을 캡처해낼 수 있다. 또한 거의 실시간의 움직임을 포착하는 것이 가능하다. 단점으로는 장비 자체가 고가라는 점도 있지만 액터의 각 관절에 부착하는 마커가 카메라로부터 숨어 데이터의 손실의 우려가 있고, 빛에 민감하여 데이터의 오류가 발생할 우려가 있다.

여기서는 MotionAnalysis사 Eagle Digital Motion Capture System을 이용하였으며, 총 8대의 카메라를 사용하여 캡처 볼륨(Volume)을 형성하였다. 보통 한 사람의 액터의 모션캡처를 받기 위해 기본적으로 28개의 마커를 사용하지만, 보다 자연스럽고 정확한 모션캡처를 위해 각 관절에 33개의 마커를 부착하였다.

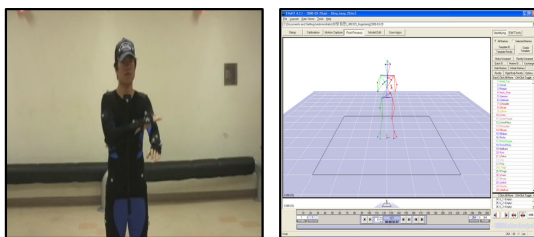


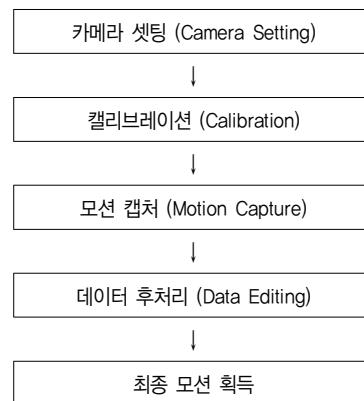
그림 4. 모션 캡처 장면과 후처리 장면

모션 캡처를 하기 위해서는 <표2>처럼 여러 과정의 작업을 필요로 한다. 먼저 수화 연기자 관절에 적외선에 반응하는 마커를 부착하고, 모션을 촬영할 8대의 카메라를 설정

한다. 카메라는 많으면 많을수록 좋지만, 얻고자 하는 볼륨에 따라서 카메라는 조정 가능하다. 그리고 캡처 공간의 중심을 정의하고 카메라의 위치와 방향을 설정하고 각 카메라의 렌즈 왜곡률을 보정하는 과정의 캘리브레이션(Calibration)을 수행한 후 모션 캡처를 실시한다.

모션 캡처 과정을 통해 얻은 모션 데이터는 몇 가지 문제점을 가지고 있다. 첫 번째는 마커가 바뀌거나 혹은 데이터는 존재하여도 마커가 부여되지 않은 경우, 두 번째는 데이터가 없어서 끊어진 경우, 마지막으로 데이터가 튀는 경우이다. 마커가 바뀐 경우에는 정확한 마커를 지정해 주고, 마커가 사라진 경우에는 찾아서 제거하며 튀는 경우에는 찾아서 제거해주는 작업을 한다. 이처럼 문제점들을 찾아서 해결하는 과정이 데이터 후처리 작업이다. 후처리 작업에 따라 모션 캡처 데이터가 부드럽고 자연스러움이 결정된다. 이렇게 후처리된 모션 데이터는 Tracked ASCII(.trc) 파일로 저장된다.

표 2. 모션 캡처 작업 과정



3.2 데이터 글러브를 이용한 손동작 모션 획득

여기서는 손가락 관절 데이터 즉, 수화에서 가장 중요하게 표현되어야 할 손동작 인식을 위해 데이터 글러브를 사용하였다. 데이터 글러브를 통해 손의 모양, 손의 방향, 그리고 운동 방향의 변화를 검출하였다.

데이터 글러브는 손에 끼고 사용하는 컴퓨터용 멀티미디어 입력장치로 손에 끼고서 손을 움직이면, 글러브 안에 장착된 센서가 이를 감지하여 컴퓨터에 위치를 입력하는 장치이다. 데이터 글러브는 기본적인 방법으로 손가락의 구부러짐과 펴는 동작에 따라 여러 가지 명령을 정의할 수 있다.

여기서는 5DT사의 Data Glove를 사용하였다. Data Glove는 사용자 손목의 움직임과 5개의 손가락 골굴 움직임을 센싱한다. <그림5>는 여기서 사용된 데이터 글러브 장비이다[7].



그림 5. 5DT 사의 Data Gloves

데이터 글러브 데이터는 모션 캡처 데이터와 함께 동시에 획득 할 수 있으나, 모션 캡처 과정에서 후처리 과정을 거치지 않고 획득하기 때문에 부자연스럽고 부정확한 데이터가 획득된다. 따라서 본 논문에서는 모션 캡처 데이터를 먼저 획득한 후에 데이터 글러브를 이용해 손가락 관절 데이터를 획득하였다.

데이터 글러브 데이터를 레코딩(Recording) 하기 위해서도 몇 가지 일련의 과정이 필요하게 되는데, 이는 앞에서 설명한 모션 캡처 작업과 거의 비슷하다. 가장 먼저 데이터 글러브를 인식하기 위한 설정을 한다. 데이터 글러브의 데이터는 Autodesk사의 MotionBuilder S/W를 통해 획득하게 되는데, 데이터 글러브 제작사에서 제공하는 Plug-In을 이용한다. 데이터 글러브의 정확한 데이터 획득을 위해 캘리브레이션(Calibration) 과정을 수행한다. 데이터 글러브의 캘리브레이션 방법은 각 손가락 관절을 오므렸다 폈다하는 과정을 2~3회 반복하여 설정한다. 캘리브레이션 과정은 연결된 데이터 글러브의 터미를 생성하여, 그 터미가 데이터 글러브의 움직임과 일치하는지를 판단하고, 미세한 부분까지 교정한다. <그림6>은 데이터 글러브로 획득한 손가락 관절에 대한 계층구조를 보여준다.

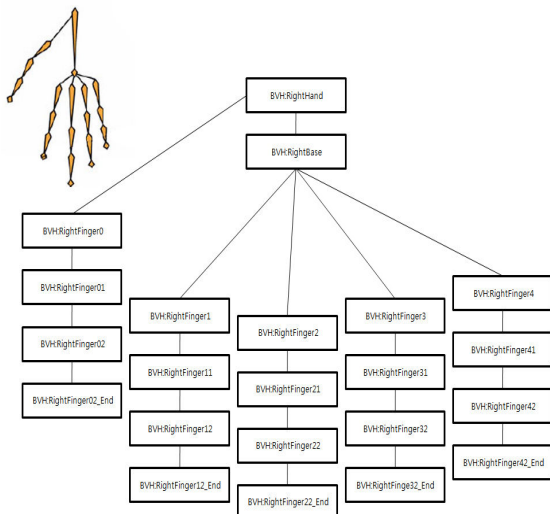
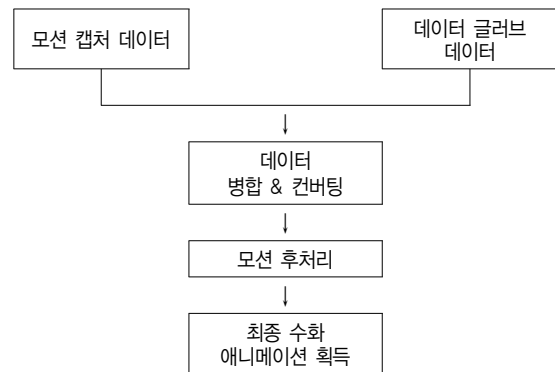


그림 6. 오른손에 대한 계층 구조

3.3 최종 수화 애니메이션 생성

지금까지 모션 캡처 시스템을 이용하여 바다 모션에 대한 데이터와 데이터 글러브를 이용하여 손가락 관절에 대한 데이터를 획득하였다. 여기서는 <표3>처럼 앞에서 획득한 데이터들을 하나의 데이터로 병합하여 하나의 수화 애니메이션을 생성하고, 모션 후처리 과정을 통하여 최종 수화 애니메이션을 생성하였다.

표 3. 수화 애니메이션 생성 구성도



모션 캡처를 이용하여 획득된 Tracked ASCII(.trc) 파일은 바다에 대한 단순한 마커의 위치 값만을 가지고 있으므로 이를 데이터 글러브를 통해 획득한 손동작 데이터와 병합하여, 뼈대 구조를 가지는 모션 데이터로 변환시키는 작업이 필요하다. 이렇게 변환한 후, Cycling과 Blending, Retargetting 등의 작업을 통해 원하는 모션으로 변형한다.

<그림7>은 앞에서 획득한 모션 캡처 데이터를 가져와 MotionBuilder S/W 내에서 터미를 생성한 후 각 마커를 터미에 매칭 시키는 장면이다. 이렇게 매칭을 시키고, 기존의 뼈대 데이터를 입력받아 마커 데이터와 뼈대 데이터를 링크시킨다. 그렇게 생성된 뼈대 데이터는 최종 애니메이션 툴(Tool)인 3ds Max로 작업하게 된다.

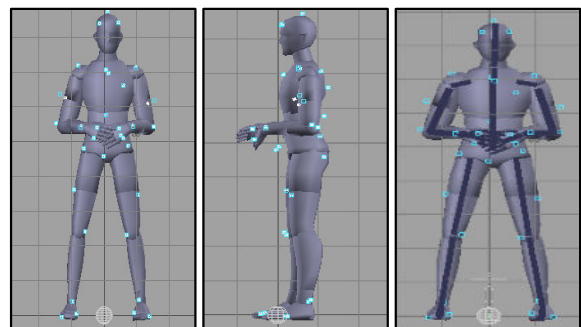


그림 7. 마커와 터미의 매칭

이렇게 처리된 데이터는 *.BIP 파일(Biped File)로 전환되고, 전환된 모션을 3D 캐릭터에 적용시키면 <그림8>과 같이 최종 모션이 획득된다.

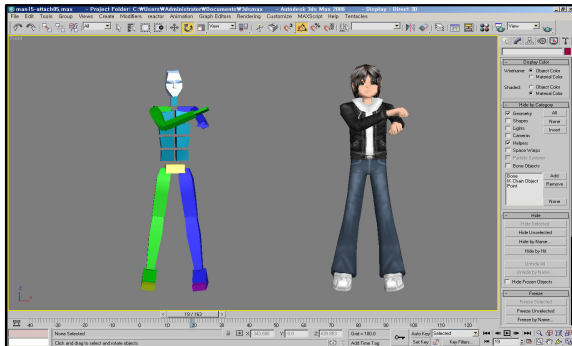


그림 8. Biped와 3D 캐릭터에 적용

최종 수화 캐릭터 애니메이션을 생성하기 위한 전체 흐름도는 <그림9>과 같다.

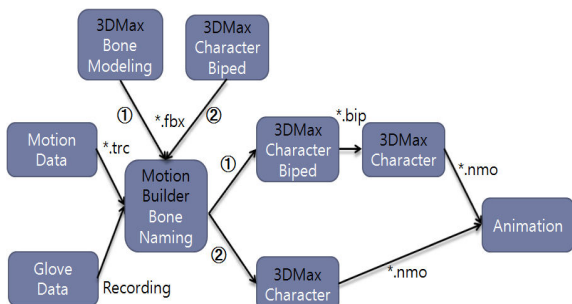


그림 9. 수화 애니메이션 생성 흐름도

4. 수화 자막 생성 및 DB 구축

수화 통역 콘텐츠 제작 시스템에는 자연스러운 동작과 정확한 손동작도 중요하지만, 정확한 수화를 제때에 표현하는 것과 많은 양의 수화 단어 및 애니메이션을 구축하였는지 또한 중요하다.

4.1 수화용 자막 생성

수화는 국어와 기본적인 어순은 같다. 하지만 조사 및 접속사의 사용, 어순의 도치, 시제의 표현 등에서 차이를 보이므로 한국어 문장을 수화와 지문자로 구분하여야 한다. 따라서 여기서는 자막/대본으로 입력받은 텍스트 파일을 시간에 맞도록 수화용 자막으로 변환한다[8].

4.2 수화 단어 데이터베이스 구축

수화 데이터베이스는 수화 단어 데이터베이스와 수화 애니메이션 데이터베이스로 구성되어 있다. 수화 단어 데이터베이스는 한국수화사전에 실린 표준 수화 단어와 문장을 데이터베이스화하고 수화 애니메이션이 해당단어에 맞도록 연동시켰다.

수화의 구성 요소는 크게 수화 어휘, 지문자, 그리고 몸짓의 세 가지로 볼 수 있다. 한국 수화는 약 6800여 개의 수화 단어와 40개의 지문자(指文字), 그리고 36개의 지숫자로 이루어져 있다[8].

수화는 국어와 기본적인 어순은 같다. 하지만 조사 및 접속사의 사용, 어순의 도치, 시제의 표현 등에서 차이를 보이므로 한국어 문장을 수화와 지문자로 구분하여야 한다. 예를 들자면 “나는 당신을 사랑합니다.”의 수화는 아래와 같이 분석된다.

주 어 : 나는 => 나(수화) + 는(지문자)
목적어 : 당신을 => 당신(수화) +을(지문자)
서술어 : 사랑합니다 => 사랑(수화) + 하다(수화)

여기서는 일상생활에 가장 많이 쓰이는 기본 수화 단어 약 1,000개의 데이터베이스를 구축하였다. <그림10>은 수화 단어 데이터베이스의 수화 인덱싱 샘플을 보여주고 있다.

index	category	word_kr	word_cn	word_en	part	behavior	code
1	gi	ㄱ	NULL	NULL	NULL	NULL	gi_a_001
2	gi	가감	NULL	NULL	NULL	NULL	gi_a_002
3	gi	가건물	NULL	NULL	NULL	NULL	gi_a_003

그림 10. 수화 단어 데이터베이스의 수화 인덱싱 샘플

4.3 수화 애니메이션 데이터베이스 구축

수화 애니메이션 데이터베이스는 이 시스템의 핵심인 모션 캡처와 데이터 글러브를 통해 전문 인력 수준의 정확한 수화 애니메이션을 구축하고 연동되는 수화 단어에 알맞은 애니메이션을 검색하여 적절한 애니메이션을 재생한다.

수화 애니메이션 데이터베이스는 이 시스템의 핵심인 모션 캡처와 데이터 글러브를 통해 획득한 수화 애니메이션이 구축되어 있다. 획득한 애니메이션은 앞서 설명한 것처럼 3ds Max의 바이페드(Biped) 형식의 애니메이션으로 저장된다. 이를 실시간 렌더링할 Virtools 엔진에서 사용되어지는 *.nmo 파일 형식으로 파일 포맷을 변경한다. 변경된 *.nmo 파일들의 집합이 수화 애니메이션 데이터베이스가 되는 것이다.

수화 애니메이션 데이터베이스에는 일상생활에서 가장 많이 쓰이는 수화 단어 중 50여 개의 애니메이션을 선별하여 구축하였다.



5. 수화 통역 콘텐츠 구현 및 실험

수화 통역 콘텐츠의 유저 인터페이스는 <그림11>과 같다. 손말새랑은 손 우리말로 수화를 지칭하는 손말과 ‘새롭고 순수하게 사랑을 베풀며 받자’ 라는 사랑의 합성어이다. 수화 통역 콘텐츠인 손말새랑을 실행한 후에 수화 콘텐츠 생성 버튼을 클릭하여 실행하면, 수화 통역 콘텐츠로 수정할 원본 동영상상을 불러오고, 그에 해당하는 자막파일을 불러온다. 그리고는 실시간으로 수화 통역된 콘텐츠가 최종 동영상 파일로 생성이 된다. 일시정지와 정지 버튼으로 중간에 잘못 저장될 경우에는 중간에 정지가 가능하다.



그림 11. 수화 통역 VR 콘텐츠 시작 화면

본 논문에서는 어린이를 대상으로 제작된 동화 “누에공과 콩알 친구들” 동영상 콘텐츠를 수화 통역 콘텐츠로 제작하였다.



그림 12. 수화 통역 VR 콘텐츠를 통해 생성된 최종 수화 통역 콘텐츠

6. 결론

본 논문에서는 청각 및 언어장애인들에게 있어서 기본 의사소통인 수화로 표현이 가능한 동영상 콘텐츠를 제공하기 위해 수화 통역을 위한 VR 콘텐츠 시스템을 제안하였다. 자연스러운 수화 동작을 추출하기 위하여 모션 캡처 시스템과 정확한 손동작 표현을 위해 데이터 글러브를 사용하여 실제 사람과 같은 캐릭터 애니메이션을 구축하고, 동영상의 자막이나 대본을 통해 생성한 수화용 자막을 이용하여 해당 자막에 맞는 수화 애니메이션을 실시간 생성하였다. 이를 기존 동영상과 생성된 수화 애니메이션을 합성하여 최종 수화 통역 콘텐츠를 제작하는 시스템을 개발할 수 있었다.

앞으로 지금의 시스템을 기반으로 하여 수화 통역을 원하는 동영상 콘텐츠의 수화용 자막을 자동으로 생성하는 모듈과 실감나는 애니메이션을 위해서는 손동작뿐만 아니라 캐릭터의 얼굴 표정도 제어될 필요가 있다. 이에 대해서는 향후 연구에서 해결하고자 한다.



감사의 글

본 연구는 호남대학교 지역혁신센터(RIC(T)) 지원에 의해 수행되었음.



참고문헌

- [1] 한국정보문화진흥원, 2007 장애인 정보격차 실태조사 보고서, 서울, 한국, 2007.
- [2] 한국정보문화진흥원, "2006-2007 정보격차 해소 백서", 서울, 한국, 2007, pp. 70-79.
- [3] 통신중계서비스, <http://www.relaycall.or.kr/>
- [4] 홍경순, "정보통신기술을 활용한 청각장애인의 의사소통 지원방안", 한국정보문화진흥원 KADO 이슈 리포트, 서울, 한국, 2006, pp. 87-99.
- [5] 오영준, 박광현, 정성훈, 장효영, 변증남, "수화 아바타의 동작 데이터베이스를 위한 그래픽 사용자 인터페이스의 개선", HCI2007 학술대회, 한국, 2007.
- [6] 배성조, "한글 수화 애니메이션 시스템", 명지대학교 정보통신공학과, 석사학위논문, 한국, 2004.
- [7] 5DT, <http://5dt.com/>
- [8] 문화관광부 국립국어원, 한국농아인협회, "한국수화사전", 한국표준수화규범 제정 추진위원회, 2007.