

論文99-36S-7-9'

# 인터넷상에 3차원 모델을 이용한 한-일간 실시간 수화 통신 시스템의 구축을 위한 기초적인 검토

## (A Study on the Construction of a Real-time Sign-language Communication System between Korean and Japanese Using 3D Model on the Internet)

金商雲\*, 吳芝英\*, 青木由直\*\*

(Sang-Woon Kim, Ji-Young Oh, and Yoshinao Aoki)

### 요 약

수화 통신은 이중 언어간의 통신 수단으로 사용될 수 있다. 이 논문에서는 3차원 모델을 이용하여 한-일간 수화 통신 시스템을 구현하여 그 가능성을 실험하였다. 실시간 통신을 위하여 통신 시스템을 클라이언트/서버 구조로 하였으며, 지적 통신방식을 도입하였다. 각 클라이언트에 3차원 모델을 준비하여 놓고, 실제의 수화영상 대신에 애니메이션 생성을 위한 파라미터 만을 전송하였다. 클라이언트에서 입력된 문장은 서버로 전송되어 한국 또는 일본 수화 파라미터로 변환한 다음 다시 클라이언트로 전송되어 수화 애니메이션으로 재생된다. 또한 자연스러운 수화 애니메이션을 위하여 감정 표현과 가변 프레임 방식 및 3차 스플라인 보간식을 이용하였다. 실험을 위한 통신 시스템은 윈도우 플랫폼에서 Visual C++와 Open Inventor 라이브러리를 이용하여 구현하였다. 실험 결과 제안 시스템이 언어의 장벽을 넘을 수 있는 비언어 통신수단으로 이용될 수 있는 가능성을 보였다.

### Abstract

Sign-language communication can be a useful way of exchanging message between people who using different languages. In this paper, we report an experimental survey on the construction of a Korean-Japanese sign-language communication system using 3D model. For real-time communication, we introduced an intelligent communication method and built the system as a client-server architecture on the Internet. A character model is stored previously in the clients and a series of animation parameters are sent instead of real image data. The input-sentence is converted into a series of parameters of Korean sign language or Japanese sign language at server. The parameters are transmitted to clients and used for generating the animation. We also employ the emotional expressions, variable frames allocation method, and a cubic spline interpolation for the purpose of enhancing the reality of animation. The proposed system is implemented with Visual C++ and Open Inventor library on Windows platform. Experimental results show a possibility that the system could be used as a non-verbal communication means beyond the linguistic barrier.

\* 正會員, 明知大學校 컴퓨터學部  
(Division of Computer Science & Engineering,  
Myongji University)

\*\* 北海道大學 大學院 工學研究科  
(School of Engineering, Hokkaido University)

※ 이 研究는 韓國學術振興財團의 自由公募(1998-001-E00979)로 수행된 것으로, 日本 通信·放送機構 (TAO)의 國際共同研究로부터 실험기기 일부의 지원이 있었습니다.

接受日字:1999年3月16日, 수정완료일:1999年5月27日

## I. 서론

컴퓨터를 이용한 인터넷 통신은 사회를 통합하는 중요한 매체가 되었다. 그러나 농아자나 고령자와 같은 정보 취약자들이 손쉽게 사용할 수 있는 미디어는 아직까지 미흡한 실정이다. 따라서 제스처나 표정 등으로 의사를 전달할 수 있는 언어인 수화를 통신 시스템상에 구현하여 자유롭게 인터넷 통신을 할 수 있도록 하고, 인터넷을 통해 저렴한 비용으로 수화를 배울 수 있도록 하는 연구가 기업 및 대학에서 진행되고 있다<sup>[1, 4]</sup>.

또한 수화는 자연어와 달리 조사를 쓰지 않으며 동사는 기본형만을 사용하기 때문에 구문법이 간단하며 서로 다른 언어간 변환이 용이하다는 특징이 있다. 따라서 수화는 언어가 서로 달라 대화가 단절된 상황에서 간단한 의사소통을 할 수 있는 또 다른 통신수단이 될 수 있다<sup>[5]</sup>. 그리고 WWW 웹상의 사이버 스페이스에서 자신을 상징화시키는 아바타(avatar)로 네비게이션을 수행한다든지 아바타끼리 제스처나 표정으로 대화하도록 하는 시스템의 구현도 생각할 수 있다<sup>[6]</sup>. 수화는 이러한 아바타 통신에 사용할 수 있는 통신언어 구성의 기본이 될 수 있다.

이러한 점에 착안하여, 수화를 이용한 이언어간 통신 방식에 대한 연구가 진행되고 있다<sup>[5, 7, 8, 9]</sup>. 지금까지의 수화통신 시스템에서는 수화 영상을 실시간으로 전송할 수 있도록 하기 위하여 2차원 그래픽 모델을 이용하여 수화 애니메이션을 구현하였다<sup>[5, 7, 8]</sup>. 그러나 수화 영상이 2차원이었기 때문에 입체감이 결여되는 등 수화 동작을 정확하게 전달할 수 없었다. 또한 기존의 연구에서는 실시간 통신을 위하여 애니메이션 방식으로 선형 보간방식과 키프레임 사이를 일정한 수의 프레임으로 채우는 방식을 사용하였다. 그러나 선형 보간 방식은 각 키프레임을 직선으로 연결하기 때문에 곡선이 많은 수화 애니메이션은 어색하였다. 또한 키프레임 사이를 일정한 수의 프레임으로 채우기 때문에 키프레임 사이가 가까운 것은 상대적으로 느리게 디스플레이 되는 단점이 있었다.

이러한 단점을 보완하기 위하여 이 논문에서는 3차원 캐릭터 모델<sup>[11]</sup>을 이용하여 입체감을 느낄 수 있도록 하고, 입의 각도에서도 수화 제스처를 조망할 수 있도록 한다. 또한 일정한 수의 프레임에 의한 선형 보간의 문제점을 해결하기 위하여 프레임 수를 키

프레임간의 거리에 따라 달리하고, 완곡한 표현을 위하여 3차 스프라인법으로 보간 하는 방법을 이용한다<sup>[12]</sup>. 그리고 3차원 모델에 광원을 추가하여 공간감을 느낄 수 있도록 한다.

이하 제 2장에서는 이종언어간 수화 영상통신의 개요와 한-일 간 수화통신을 위해 필요한 사항을 검토하고, 제 3장에서는 수화 애니메이션에 필요한 3차원 모델의 구조 및 이로부터 추출되는 수화 파라미터에 대하여 설명한다. 제 4장에서는 보다 자연스러운 수화 제스처를 생성하기 위하여 도입한 애니메이션 방식을 살펴본 다음, 제 5장에서는 제안한 통신 시스템에 대한 실험 방식과 실험결과를 고찰한다. 끝으로 제 6장에서 결론을 맺는다.

## II. 이종 언어간의 수화통신

이 논문에서 이용하는 수화통신은 지적통신 방식으로, 송수신측에 각각 수화 애니메이션을 위한 3차원 모델을 준비하고, 실제의 수화영상을 전송하는 대신에 애니메이션 생성을 위한 파라미터만을 전송하는 방법이다<sup>[8]</sup>. 즉 송신측에서는 수화단어를 입력받아 애니메이션 파라미터를 생성한 다음 이를 텍스트 데이터로 전송하면, 수신측에서는 파라미터를 수신하여 수화 애니메이션을 재생하는 방법이다.

이 방법의 장점은 수화단어를 애니메이션 파라미터로 변환하는 단계에서 기계번역 등의 처리를 하여 서로 다른 언어의 수화로 바꿀 수 있다는 점이다. 따라서 이 논문에서는 이종언어간의 본격적인 수화통신을 위한 예비 단계로서, 언어 체계가 비슷한 한-일 간 수화통신을 검토한다. 먼저 송신 클라이언트측에서 키보드를 통해 입력한 문장을 서버로 전송하여 한국 또는 일본 수화 파라미터로 변환한 다음 이를 수신 클라이언트로 전송하여 해당 언어의 수화 애니메이션으로 재생하여 출력한다.

### 1. 한-일 수화의 비교

한국과 일본은 지리적으로 이웃하고 있는 나라로서 풍속이나 관습에 유사한 점이 많으며 언어에도 많은 특징을 공유하고 있다. 문장의 어순이나 구문구조가 거의 같고 명사를 나타내는 한자 의미도 거의 같다. 다만 술부나 용언의 활용에서 차이가 있으며 조사나 조동사 등의 의미 대응관계가 다르다는 특징이 있다<sup>[5]</sup>. 그런

데 한국수화나 일본수화 모두 조사를 쓰지 않으며 동사는 기본형만을 사용하기 때문에 수화번역은 자연어의 기계번역 보다 비교적 간단하게 수행할 수 있다.

한 일 양국의 표준적인 수화의 특징은 다음과 같다. 첫째, 두 나라의 표준수화는 모두 자연어의 어순에 따른다. 둘째, 양국의 수화에서는 모두 조사를 거의 생략하며, 동사나 형용사의 품사를 구분하지 않는다. 셋째, 존경어를 사용하지 않고 얼굴표정이나 신체의 자세로 나타낸다. 넷째, 상세한 표현에서는 차이점이 존재한다.

이러한 특징의 비교에서도 알 수 있듯이 두 나라의 수화번역이나 통역은 수화단어만을 교체하는 처리만으로 상당부분을 해결할 수 있게 된다.

따라서 이 논문에서는 직접번역 방식에 의한 한-일 수화통신을 전제로하여 수화번역 사전을 작성한다. 즉 각각의 수화단어에 대하여 두나라의 수화 파라미터를 쌍으로 결합하여 저장하는 방법을 이용한다. 한 나라의 수화단어를 입력하면 이를 이용하여 수화사전을 탐색하여 자신과 상대편 수화 파라미터를 검색하여 입력 단어에 해당하는 수화 애니메이션을 생성한다. 다음 절에서 수화 번역사전의 구조와 검색방법에 대하여 설명한다.

2. 한-일 수화 번역 사전의 구조

그림 1은 한글 수화(Korean Sign Language: KSL)와 일본 수화(Japanese Sign Language: JSL)로 수화 통신을 하기 위해 구축한 수화 번역 사전의 구조와 그 검색 과정을 도시한 것이다<sup>[10]</sup>. 여기서 수화 사전은 크게 한글 수화사전(KSL Dictionary), 변환사전(Translation Dictionary), 일본 수화사전(JSL Dictionary)로 구성된다. KSL, JSL 사전은 다시 단어사전(Word Dictionary)과 파라미터 사전(Parameter Array: PA)으로 구성된다.

수화 번역 사전을 이와 같은 구조로 구축한 이유는 수화 단어 정보와 수화 파라미터를 분리·저장하여, 수화 단어 검색의 속도를 향상시키기 위한 것이다. 여기서 단어사전은 단어 저장 구조인 트라이(trie) 구조와 각 단어에 대한 PA 사전 내의 위치(포인터) 및 동작 수에 대한 정보를 가지고 있다. 그리고 PA 사전은 수화 애니메이션에 필요한 상반신 관절각을 나타내는 54개 동작 파라미터 열로 구성된다.

변환사전은 한글 수화 단어와 일본 수화 단어를 일대일로 대응시키는 사전이다. 즉 입력된 한글 문장을

일본 수화로 번역하고자 할 때, 이 사전을 사용하게 된다. 그리고 JSL 사전은 KSL 사전과 같은 구조로 되어 있다. 예를 들어, “학교”의 한글 수화 단어를 일본 수화 “KakKyo”로 번역하는 과정은 다음과 같다:

- (1) KSL Dictionary 안에 있는 Word Dictionary의 트라이 구조를 검색하여 Translation Dictionary에 있는 한글 수화단어 “학교”의 위치(포인터)를 검색한다.
- (2) Translation Dictionary 에서 한글 수화단어 “학교”에 대응하는 일본어 수화단어가 “KakKyo”임을 알아낸다.
- (3) “KakKyo”로 JSL Dictionary 내의 Word Dictionary에서 트라이 구조를 검색하여 PA 사전 안의 “KakKyo”의 수화 파라미터에 대한 시작 위치(ID)와 동작 수(K)를 알아낸다.
- (4) PA사전의 주소 ID에서 시작해서 저장되어 있는 파라미터 열을 K개 만큼 읽어들이어 “KakKyo”의 수화 애니메이션을 생성한다.

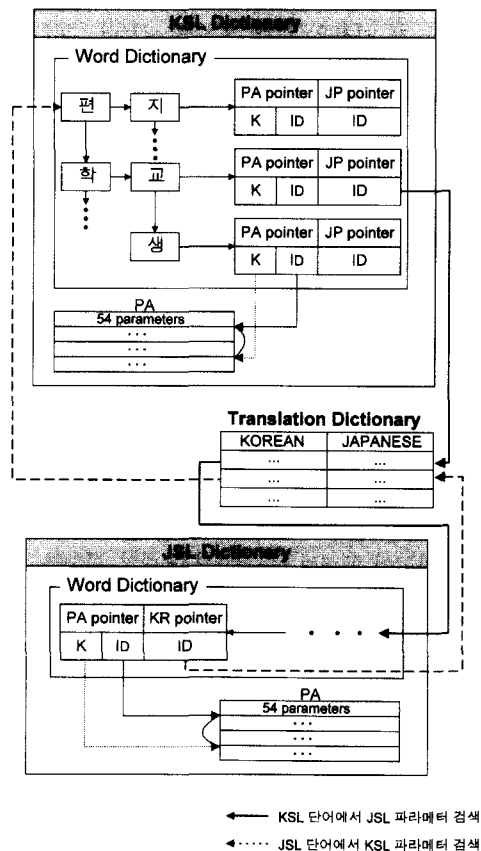


그림 1. 수화 사전의 구조와 검색 과정  
Fig. 1. The architecture of sign-language dictionary and its retrieval.

### III. 3차원 모델과 수화 파라미터

이 논문에서는 입체감 있는 수화 애니메이션을 생성하기 위하여 3차원 캐릭터 모델을 이용하고, 자연스러운 수화 통신을 위해서 얼굴 표정을 합성하여 사용한다.

#### 1. 3차원 모델의 구조

수화 애니메이션을 위한 3차원 캐릭터 모델의 구조는 그림 2와 같다<sup>[11]</sup>. 그림 2(a)는 3차원 모델의 와이어 프레임으로 1,071개의 제어점을 세 개씩 연결한 1,840개의 삼각형으로 구성되어 있다. 기존 연구<sup>[13]</sup>에서는 사진으로 실제 사람의 영상을 입력받아 3차원 모델에 텍스처 매핑(texture mapping)을 실행하는 방법을 이용하였다. 그러나 이 방법은 실제로 사람이 수화를 하는 것처럼 보일 수 있으나 텍스처 매핑에 많은 시간이 소요되므로 여기서는 이용하지 않았다. 그림 2(b)는 상반신 모델에서의 각 관절의 상관관계이다. 모델의 각 관절은 계층 구조를 이루며, 근육에 대한 점 데이터와 회전각 등의 데이터를 가진다.

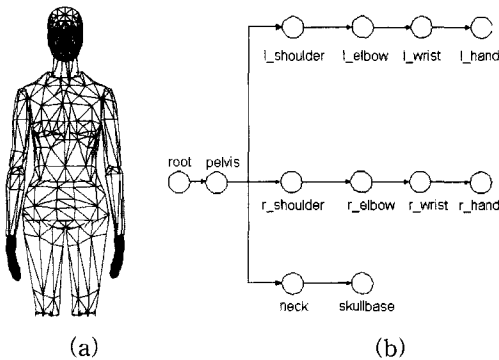


그림 2. 3차원 수화 애니메이션 모델 (a) 3차원 모델의 와이어 프레임 (b) 3차원 모델의 계층구조  
 Fig. 2. The 3D sign-language animation model (a) The wire frame of 3D model (b) The hierarchy of 3D model.

수화 애니메이션은 관절 회전각(joint angles)을 변화시켜가면서 수행된다. 예를 들어, 왼쪽 어깨(l\_shoulder)가 30° 만큼 회전하면, 하위 노드인 왼쪽 팔꿈치(l\_elbow)는 자동적으로 왼쪽어깨의 중심축을 기준으로 해서 30° 만큼 회전한다. 이러한 회전 관계는 계층 구조의 단말 노드까지 전달된다. 수화 애니메이션을 생성하기 위해서는 계층 구조와 함께 모델이

움직일 수 있는 축과 움직임의 범위를 정의하여야 한다. 수화 애니메이션을 위한 각 관절의 자유도(Degree Of Freedom: DOF)는 표 1과 같다.

표 1. 각 관절의 자유도  
 Table 1. The DOF(Degree of Freedom) of each joint.

관절 이름	DOF	좌표계
Shoulder	3	(x, y, z)
Elbow	2	(x, y)
Wrist	2	(x, z)
Metacarpophalangeal(finger)	2	(x, z)
Interphalangeal(finger)	1	(x)
Carpometacarpal(thumb)	2	(x, z)

여기서 자유도란 각 관절이 움직일 수 있는 축의 개수로써 한쪽 팔에 7개가 존재하고, 손가락 뿌리 관절(metacarpophalangeal)에 2개, 손가락 마디 관절(interphalangeal)에 각 1개씩 한 손가락에 4개가 존재하며, 한 손에는 20 자유도(=4×5)가 존재한다. 따라서 수화의 손과 팔 동작을 디스플레이하기 위해서는 한쪽 팔에 27개씩 전체 54개의 관절각 파라미터가 필요하다.

#### 2. 얼굴 표정 제어

수화통신에서는 정보의 많은 부분을 얼굴표정으로 전달하기 때문에 자연스러운 수화 통신을 위해서는 얼굴 표정을 사용해야 한다.

그러나 같은 단어를 표현하는 데에도 수화자는 그 정도나 경우에 따라 서로 다른 표정을 보인다. 따라서 이 논문에서는 “기쁨”, “슬픔”, “꺼림”, “화남”, “놀람”, “두려움”의 6가지 감정을 정의하고, 송신측에서 하나를 선택하여 전송한다.

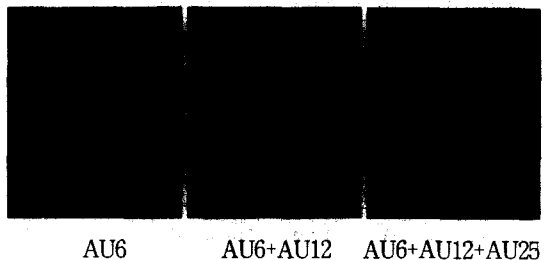


그림 3. “기쁨”에 대한 얼굴 표정의 생성 예  
 Fig. 3. The generation example of facial expression “Joy”.

감정을 표현하기 위하여는 FACS(Facial Action Coding System)의 AU(Action Unit)을 조합하여 이용한다. 예를 들어 “기쁨”의 경우 AU06(arg) + AU12(arg) + AU25(arg)로 구현한다<sup>[14]</sup>. 여기서 arg는 감정마다 정해진 강도를 보간하는 파라미터이다. 그림 3은 “기쁨”에 대한 얼굴 표정의 생성과정이다.

#### IV. 자연스러운 수화 애니메이션의 생성

이 논문의 수화 애니메이션 방식은 키 프레임 애니메이션 방식으로, 각 동작의 주요 장면들을 숙련된 애니메이터가 준비하고, 각 키 프레임의 사이는 보간 공식을 이용하여 채우는 방식이다. 수화 파라미터 열을 키 프레임으로 이용하여 수화 애니메이션을 생성하는 과정은 다음과 같다. 먼저 수화 파라미터 열로부터 동작의 크기 변화량을 계산하여 적정한 보간 프레임(in-between)의 수를 결정한 다음, 정해진 프레임 수에 맞추어 키 프레임 사이의 프레임을 생성하여 프레임 큐에 저장한다. 그리고 실제 애니메이션을 생성할 때에는 중간 프레임을 계산하지 않고 프레임 큐로부터 프레임을 읽어와 애니메이션을 수행한다<sup>[12]</sup>. 여기서 키 프레임은 수화 파라미터열에 해당한다.

##### 1. 가변 프레임 보간

가변 프레임 보간이란 키 프레임간의 프레임 수를 미리 고정시키지 않고 동작의 변화에 맞게 설정하는 보간법으로, 동작 변화가 클 경우에는 보간 프레임을 많이 이용하고 작은 동작에서는 소수의 프레임만으로 보간하는 방법이다. 이 논문에서는 동작 변화의 크기는 이전 동작 파라미터와 현 동작 사이의 파라미터 변화를 이용하며, 보간 프레임의 수는 식(1)과 같이 구한다.

$$F_i = \text{MAX}_j \left( \frac{11}{180} \times |P_{i-1} - P_j| \right) + 2 \quad (1 \leq j \leq 14) \quad (1)$$

여기서,  $i$  : 키 프레임 번호,

$j$  : 파라미터 번호,

$P_i$  :  $i$ 번째 키 프레임의  $j$ 번째 관절각 파라미터,

$F_i$  :  $i$ 번째 키 프레임과  $i-1$ 번째 키 프레임 사이의 보간 프레임 수.

식(1)에서  $P_j$ 는 팔의 관절각 54개로 구성되는 키 프레임 파라미터이고, 상수 2는 실험적으로 결정한 최소 보간 프레임 수이다. 보간 프레임 수는 키 프레임 사이에서 각도 변화가 가장 큰 관절각 파라미터를 기준으로 설정한다. 즉 팔이 움직일 수 있는 각도의 최대 범위를  $\pm 180^\circ$ 로 한 다음 이를 11등분으로 나누어 팔의 움직임을 근사 시키는 방법이다. 그런데 손가락의 각도 변화는 팔이나 어깨의 움직임에 비해 동작의 크기가 아주 작다. 따라서 식(1)은 14개 팔동작 파라미터의 변화에 대해서만 적용하였다.

##### 2. 3차 스플라인 보간

이 논문에서는 선형보간의 문제점을 극복하기 위하여 3차 스플라인 보간(cubic spline interpolation)법을 이용한다. 키 프레임 사이의 애니메이션 경로를 구하기 위한 3차 스플라인 보간식은 식(2)와 같다.

$$g(x_{ij}) = ax_{ij}^3 + bx_{ij}^2 + cx_{ij} + d; \quad (\forall j, 1 \leq j \leq 54) \quad (2)$$

$$x_{ij} = \frac{k}{F_i}; \quad (k=1, \dots, F_i)$$

여기서,  $i$  : 키 프레임 번호,

$j$  : 파라미터 번호,

$F_i$  :  $i-1$ 번째 키 프레임과  $i$ 번째 키 프레임 사이의 보간 프레임 수.

보간 함수  $g(x_{ij})$ 는  $i-1$ 번째 키 프레임과  $i$ 번째 키 프레임 사이에서  $j$ 번째 관절각 파라미터에 대한 3차 스플라인 보간식이다. 즉, 각 키 프레임 사이의 54개의 관절각 파라미터 각각에 대해 3차 스플라인 보간식을 적용하여 보간 프레임을 생성한다. 여기서  $g(x_{ij})$ 의 계수  $a, b, c, d$ 는 경계조건과 식(3)의 연립방정식으로 구한다.

$$\begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ 1 & 4 & 1 & 0 & \cdot & \cdot \\ 0 & 1 & 4 & 1 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & 1 & 4 & 1 \\ \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ \cdot \\ \cdot \\ P_{m-1} \end{bmatrix} = \begin{bmatrix} P_0 \\ 3(P_2 - P_0) \\ \cdot \\ \cdot \\ 3(P_{m-1} - P_{m-3}) \\ P_{m-1} \end{bmatrix} \quad (3)$$

여기서  $P_i$ 와  $P_j$ 는 제  $i$ 번째 키 프레임 및 이 프레임의 1차 미분이고,  $m$ 은 키 프레임의 수이다. 예를 들어, 한 번에 전송되어온 키 프레임 수가 5개일 경우  $m$ 의 값은 5가 되며 결정해야 할 파라미터 수는  $216(=4 \times 54)$ 이 된다.

### V. 실험 및 결과 고찰

#### 1. 통신 시스템의 구성

본 논문의 수화 통신 시스템은 중앙 집중적인 서버-클라이언트 구조로, 전체 구조는 그림 4와 같다<sup>[8]</sup>. 이 시스템은 크게 서버 응용 프로그램과 클라이언트 응용 프로그램으로 구성된다. 서버 응용 프로그램은 통신모듈(Communication)과 변환모듈(Translation)로 구성된다. 통신모듈은 각 클라이언트의 연결을 관리하며, 클라이언트로부터 들어오는 요청을 해석하여 필요한 데이터를 보내준다. 클라이언트가 보내는 요청에는 연결 시작, 대화 그룹 생성, 대화 그룹 참가, 문장 변환·전송, 연결 종료 등이 있다. 예를 들어, 문장 변환 요청이 들어오면, 이는 변환모듈이 입력문장을 수화 파라미터로 변환하여 보내준다. 이를 위해 변환모듈은 수화 단어 사전에 가진다. 수화 단어 사전은 문장 분석을 위한 언어 정보와 함께 수화 단어의 파라미터를 저장하고 있다.

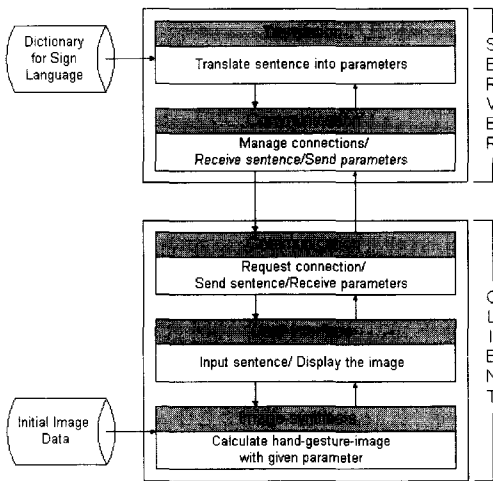


그림 4. 수화 통신 시스템의 구조  
Fig. 4. The architecture of sign-language communication system.

클라이언트 응용 프로그램은 인터페이스 모듈(User-Interface), 통신모듈(Communication)과 영상합성모듈(Image-synthesis)로 구성된다. 인터페이스 모듈은 사용자의 명령 입력 또는 문장 입력을 받아 이를 통신 모듈로 전달한다. 통신모듈은 서버로부터 오는 메시지의 수신, 사용자 인터페이스로부터 전달된 명령의 송신을 담당한다. 서버에서 수화 파라미터가 전달되면, 통신모듈이 이를 영상합성모듈 모듈로 전달한다. 영상합성모듈 모듈은 클라이언트의 파일 시스템에 저장되어 있는 수화 모델 데이터에 수화 파라미터를 합성하여 수화 애니메이션을 생성한다.

이 시스템은 윈도우즈 플랫폼에서 Visual-C++ 5.0을 이용하여 구현했다. 대화자는 각자의 IP 주소를 이용하여 서버에 접속할 수 있으며, 간단한 수화문장을 입력하면 이 문장에 해당하는 수화 애니메이션을 볼 수 있고 동시에 상대방 대화자들에서도 재생된다. 또한, 3차원 모델에서는 Open Inventor 라이브러리를 도입했으며, 인체 모델은 "http://www.ywd.com"에서 VRML 파일 포맷으로 받은 것을 Open Inventor 2.1 파일 포맷으로 변환시켜 사용했다.

이 논문에서는 키보드를 통하여 입력한 한글 또는 일본어 문장을 KSL 또는 JSL파라미터로 변환하여 사용하는 방법으로 수화 입력장치를 대체하였다<sup>[15]</sup>. 예를 들어, 한글 문장 "나는학교에 갑니다(I go to school)"가 입력되었을 경우 먼저 이문장을 스페이스를 기준으로 구문 분석하여 세 개의 KSL단어 "나(I)", "학교(school)", "가다(go)"로 바꾼다. 그리고 이 세 KSL 단어를 이용하여 수화사전을 검색하여 KSL 또는 JSL 애니메이션 파라미터를 얻는다.

#### 2. 키 프레임 에디터

그림 5는 3차원 수화 애니메이션을 생성하기 위하여 필요한 키 프레임 에디터의 인터페이스이다. 왼쪽은 수화 애니메이션을 위한 아바타이고 오른쪽은 각 관절각 값을 입력하는 패널이다. 오른쪽에 각 관절각

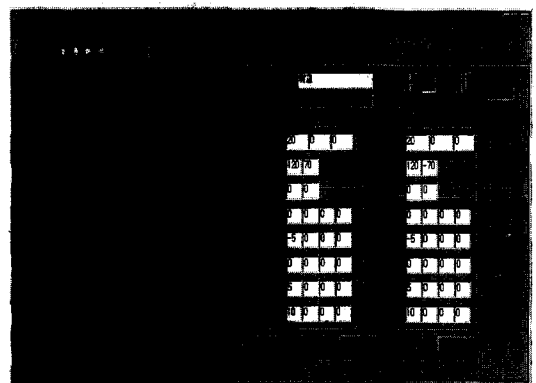


그림 5. 3차원 수화 애니메이션 생성을 위한 키 프레임 에디터  
Fig. 5. Key-frame editor for generating 3D sign-language animation.

값을 입력하면 이에 따라 왼쪽 아바타의 팔이 움직이게 된다. 따라서 아바타 팔의 움직임을 관찰하면서 적당한 값을 찾아내는 방법으로 원하는 키 프레임을 작성한다.

3. 실험 결과 고찰

먼저 키보드로 입력한 문장을 분석하여 3차원 수화 애니메이션을 생성하는 실험을 하였다. 기존의 2차원 모델을 사용했을 경우<sup>[7]</sup>와 비교한 결과 2차원 모델의 경우 전면의 한 방향에서만 관측이 가능한 반면, 3차원 모델의 경우 시점을 돌려가면서 볼 수 있기 때문에 보다 입체감 있는 수화 동작을 관측할 수 있었다. 또한 3차원 모델에 의한 수화 애니메이션의 정확도를 측정하는 실험을 하였다. 표 2는 초급 수화자 6명이 주어진 수화단어에 대한 수화 애니메이션을 인식하는 실험 결과이다. 평균 인식률은 84%로서 2차원의 평균 인식률(71%)보다 향상되었음을 보인다.

문헌 [7]의 2차원 수화 애니메이션에 대한 인지율을 분석하면, “가지다”, “보다”, “시간”, “고맙다”, “반갑다”와 같이 수화 동작이 크고 쉬운 동작일 경우에는 모두 정확하게 인지했다. 그러나, “손”은 정확하게 오른손이 가리키는 곳이 손의 어느 부분인지 인지하기가 힘들었다. “먹다”, “너”의 경우 손과 신체의 거리를 제대로 보기 힘들거나, 손이 정확히 가리키는 방향이 어느 쪽을 향하는지 알기가 힘들었다. 이와 같은 것들은 모두 앞쪽 영상만을 보여주는 2차원 모델에 기인한 문제점이다.

또한 “몸”의 경우 신체의 앞부분에서 손가락으로 원을 그리는 동작을 한다. 그러나, 구현된 동영상을 재생하기 위해 사용하는 보간법에서는 원을 그리는 동작을 제대로 표현할 수 없었다. 그러므로, 원 동작이 들어간 수화는 수화 애니메이션이 영성해진다.

이와 비교하여, 표2에서 3차원 수화 애니메이션에 대한 인지율을 분석하면 다음과 같은 몇가지 특징을 알 수 있다. 첫째, 실험 대상자 6명 간에 인식률의 차이나 알아보기 못하는 단어에 대한 차이가 별로 없다는 것이다. 이것은 이 시스템을 사용하는 사용자 상호간의 통신 중의 오해를 줄일 수 있고, 인식률이 저조한 애니메이션에 대해서는 개선 방식을 쉽게 찾을 수 있다는 장점이 있다.

둘째, 인식률이 저조한 “흰색”, “손”, “너” 외에는 전부 83 ~ 100%의 인식률을 나타냈다. “흰색”은 “하얗다”의 의미와 “색”이라는 동작이 합쳐져서 이루

어진다. 여기서, “하얗다”라는 애니메이션을 위해서는 모델이 치아를 가리키는 동작을 해야 하지만, 수화 파라미터의 제약상, 치아를 가리키는 동작이 불가능했다. 또한, “하얗다”라는 애니메이션을 할 충분한 시간이 없이 다음 동작으로 바로 넘어가 버린 것도 인식률을 0%로 만든 원인이다. “손”과 “너”의 수화 애니메이션은 모델의 수화 동작은 비교적 정확했지만, 다른 단어에 비해 제스처가 단순하고 애니메이션시간이 짧기 때문에, 문맥에 의한 이해가 요구되었던 것으로 생각된다.

표 2. 3차원 수화 애니메이션의 인지율

Table 2. The recognition rate of 3D sign-language animation.

수화단어	P1	P2	P3	P4	P5	P6	인식률*
흰색	×	×	×	×	×	×	0(83)
손	×	×	×	×	×	○	17(17)
사람	○	○	○	○	○	○	100(83)
편지	○	○	○	○	○	○	100(66)
작다	○	○	○	○	×	○	83(50)
먹다	○	○	×	○	○	○	83(0)
읽다	○	○	○	○	○	○	100(66)
가지다	○	○	×	○	○	○	83(100)
들어가다	○	○	○	○	×	○	83(66)
가다	○	○	○	○	○	○	100(83)
너	×	×	○	○	×	○	50(50)
교사	○	○	○	○	○	○	100(100)
보다	○	○	○	○	○	○	100(100)
시간	○	○	○	○	○	○	100(100)
뜻	○	○	○	○	○	○	100(66)
몸	○	○	×	○	○	○	83(50)
부탁	○	○	○	○	○	○	100(100)
열심히	○	○	○	○	○	○	100(83)
고맙다	○	○	○	○	○	○	100(100)
반갑다	○	○	○	○	○	○	100(100)
인지율**	85	85	75	90	75	95	84(71)

○:인지, ×:인지 못함,

Pi : 인지 실험을 수행한 대화자

( ) : 2D 수화 애니메이션의 인지율

인지율\* : 인지한 사람수/전체 사람수 ×100 (%)

인지율\*\* : 인지된 수화 단어수/전체 수화 단어수 ×100 (%)

마지막으로, 2차원의 인식률과 비교해서는 3차원의

수화 모델이 2차원 수화 모델의 고정된 시야를 극복할 수 있는 장점 때문에, 3차원 수화 애니메이션이 대체적으로 인식률이 높았다. 또한, 스플라인 보간 공식을 이용하여 원 동작이 들어가는 “몸” 동작을 개선하여 인식률이 높아진 것을 확인할 수 있었다. “몸”의 경우 가슴 앞에서 손가락으로 크게 원을 그리는 동작을 한다. 그러나, 선형 보간 방식에서는 원을 그리는 동작을 제대로 표현할 수 없었기 때문에 원 동작이 들어간 수화는 수화 애니메이션이 영성해진다. 한편 “나쁘다”의 경우 수화 동작 초반의 속도는 빠르고 후반에는 느려지는 강·약의 성분이 들어있다. 그러나, 구현된 시스템은 한 동작 단위 내에서 일정한 속도와 일정한 각도 변화로 움직이도록 되어 있다. 따라서 동작의 강·약이 수화의 의미 전달에 중요한 역할을 할 경우에는 그 인지도가 떨어지는 것을 확인하였다.

가변 프레임 방법에서 생성되는 보간 프레임의 수를 관찰하였다. 표 3은 4개의 문장을 입력했을 때 고정 프레임 방법과 가변 프레임 방법에서 생성되는 보간 프레임의 수를 비교한 것이다. 고정 프레임 방법에서는 키 프레임 당 7개가 소요된 반면에 가변 프레임의 경우에는 평균 5.45개이었으며 식(4)와 같이 시간 감소율은 22.14%이었다.

$$R = \frac{A-B}{A} \times 100 \approx 22.14\% \quad (4)$$

여기서,  $A = 7 \times 0.3 = 2.1$

(고정 프레임 방식의 평균 디스플레이 시간)

$B = 5.45 \times 0.3 = 1.64$

(가변 프레임 방식의 평균 디스플레이 시간)

가변 프레임 방식으로 고정 프레임에서 생기는 불필요한 프레임을 줄여 애니메이션 속도를 향상시킬 수 있으며 애니메이션 동작도 보다 자연스러웠다.

원(circle) 애니메이션(키 프레임 수: 4)을 위하여 선형 보간식과 3차 스플라인 보간식을 적용한 결과는 그림 6과 같다. 그림 6(a)는 선형 보간의 결과로서 전체 28프레임이 소요되었고, (b)는 3차 스플라인 보간의 결과로 16 프레임이 소요되었다. 즉 선형 보간의 애니메이션은 많은 프레임을 사용하고도 다각형을 생성한 반면 스플라인 보간은 비교적 적은 수의 프레임으로 원에 근사한 동작을 생성한다. 또한 선형 보간에서는 곡선 동작의 수화 애니메이션을 위해서 키 프레임을 6~8개 만들어 주어야 했지만, 스플라인 보간을

이용하면 3~4개의 키 프레임으로 보다 자연스러운 애니메이션을 만들 수 있었다.

표 3. 고정 프레임법과 가변 프레임법의 보간 프레임 수 비교

Table 3. Comparison of the frame numbers by fixed-frame and variable-frame

수화 애니메이션	키 프레임 수	보간 프레임 수	
		선형 프레임법	가변 프레임법
나는 학교에 갑니다	7	49	48
선생님 안녕하세요	9	63	41
당신은 누구십니까	6	42	26
아침에 모두 기뻐하다	9	63	54
평균	8	54	42

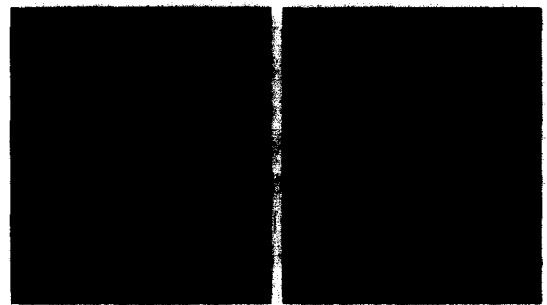


그림 6. 원 동작의 비교 (a) 선형 보간식을 적용한 원 동작 (b) 스플라인 보간식을 적용한 원 동작  
Fig. 6. The comparison of circular motions (a) The circular motion with linear interpolation (b) The circular motion with spline interpolation.

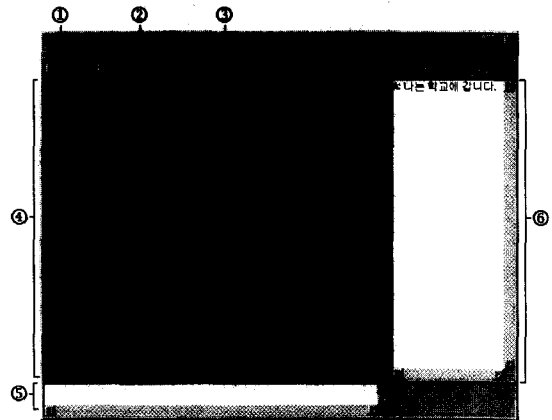


그림 7. 3차원 수화 통신용 윈도우  
Fig. 7. The window for 3D sign-language communication.



끝으로 한-일간 수화영상 통신을 위하여 인터넷 상에 클라이언트-서버 시스템을 구축하여 실험하였다. 그림 7은 통신 시스템에 연결된 클라이언트 "A"와 클라이언트 "B"의 화면에 나타나는 통신용 윈도우이다. 여기서, ①은 연결·종료 버튼, ②는 얼굴 표정 선택 버튼, ③은 한글수화, 일본수화, 한-일 수화 번역, 일-한 수화 번역을 선택하는 버튼, ④는 3차원 모델의 수화 채팅 창, ⑤는 사용자의 문장 입력 창, ⑥은 송수신 메시지를 보여주는 창이다.

대화자가 한 사람씩 서버에 연결될 때마다 그 대화자 분신에 해당하는 3차원 캐릭터 모델이 윈도우에 하나씩 추가된다. 연결된 대화자들이 각자 수화문장을 입력하면, 이 문장은 서버로 전송되어 수화 파라미터로 변환된 후 다시 클라이언트들로 전송된다. 이때 수화 번역은 서버가 입력 문장을 수화 파라미터로 바꾸는 과정에서 일어난다. 즉 대화자의 선택에 따라 서버가 한글 또는 일본어 수화 파라미터로 변환하여 클라이언트에 보내면, 클라이언트들은 이 수화 파라미터를 이용하여 애니메이션을 생성하여 디스플레이하게 된다.

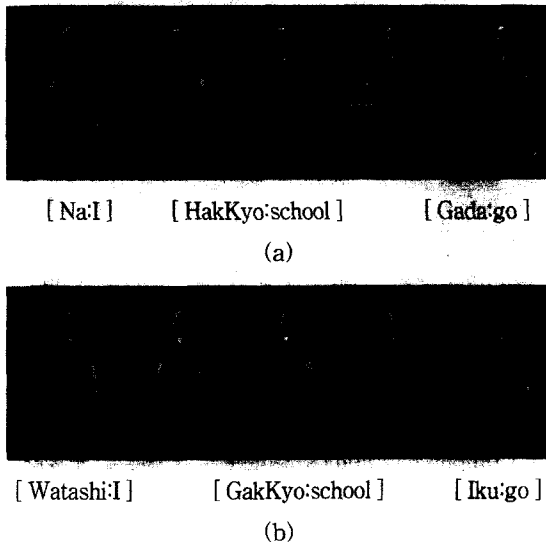


그림 8. 3차원 수화 번역의 예 (a) 한글 수화 "NaNeun HakKyoEa KamNiDa" (b) 일본어로 번역된 수화 "WatashiHa KakKyoE Ikimasu"

Fig. 8. Example of 3D sign-language translation (a) The KSL "NaNeun HakKyoEa KamNiDa" (b) The translated JSL "WatashiHa KakKyoE Ikimasu".

그림 8은 클라이언트 "A"와 "B"사이에 전송되는 "나는 학교에 갑니다(I go to school)"에 대한 3차원

수화 애니메이션의 예이다. 그림 8(a)는 한글 수화 "NaNeun HakKyoEa KamNiDa"가 클라이언트 "A"에 디스플레이되는 모습이고, 그림 8(b)는 전송되어온 한글수화가 클라이언트 "B"에서 일본 수화 "WatashiHa KakKyoE Ikimasu"로 변환되어 디스플레이되는 모습이다. 또한 3차원 모델에서는 음영이 중요한 역할을 한다. 따라서 이 논문에서는 공간감을 더하기 위해 광원을 추가하였다. 광원을 추가하지 않은 모델인 그림 7에 비하여 광원을 추가한 애니메이션인 그림 8(a), (b)가 더욱 입체감이 있음을 보이고 있다.

### VI. 결론

이 논문에서는 3차원 수화 애니메이션을 이용한 이중 언어간 비언어 통신 시스템을 인터넷상에 구축하기 위한 기초적인 내용을 검토하였다. 우선 실시간 통신을 위해서는 지적 영상 통신 방식을 도입하여 수화 애니메이션에 필요한 파라미터만을 전송하는 방식을 이용하였다. 또한 자연스러운 수화 애니메이션을 생성하기 위해서 감정표정을 위한 파라미터를 추가하는 방식과 가변 프레임방식 및 3차 스프라인 보간식을 도입하였다. 그리고 수화 입력은 키보드를 통하여 입력한 문장을 수화 파라미터로 변환하는 방식을 사용하였다. 이 때 한-일간 수화 파라미터를 직접변환하여 수화를 번역하는 방법을 검토하였다.

실험 결과, 제안한 시스템을 이용하여 실시간의 이중 언어간 비언어 통신이 가능하다는 것을 확인하였다. 또한 3차원 모델과 애니메이션 방식을 개선하여 2차원 모델의 경우보다 인식률을 높였다. 그러나 수화 동작이 모두 등속도를 가정했기 때문에 여전히 몇가지 동작은 이해할 수 없는 경우가 발생하였다. 이를 해결하기 위해서는 단순한 수화 파라미터열 뿐 아니라 속도 정보를 추가하는 방식과 이를 어떻게 애니메이션으로 생성해야 하는지에 대한 연구가 필요하다.

### 참 고 문 헌

[ 1 ] <http://www.hitachi.co.jp/Prod/comp/app/shuwa/index.html>.  
 [ 2 ] <http://www.nttdata.co.jp/profile/philan/finger-lang/default.html>.

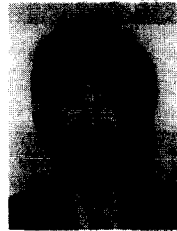
- [3] [http://chihara.aist-nara.ac.jp/new-home/people/people\\_j.html](http://chihara.aist-nara.ac.jp/new-home/people/people_j.html).
- [4] <http://www.media.eng.hokudai.ac.jp/~tana/study/link.html>.
- [5] 李今燦, 青木由直, 金商雲, “日韓手話通信のためのCGアニメーションシステム構築の基礎的検討,” 電子情報通信學會論文誌, vol. J82-D-I, no. 5, pp. 613 - 621, 1999. 05.
- [6] B. Damer, Avatar, Peachpit Press, 1998.
- [7] J.-Y. Oh, S.-W. Kim, S. Tanahashi and Y. Aoki, “A sign language chatting System for non-verbal communication between different languages,” Proc. of ITC-CSCC'98, Sokcho, Korea, pp. 693 - 696, July 1998.
- [8] 吳芝英, 金商雲, 青木由直, “지적 영상통신에 의한 수화채팅시스템의 구현,” 大韓電子工學會秋季綜合學術大會論文集, vol. 20, no. 2, pp. 1221-1224, 1997. 11
- [9] 吳芝英, 金商雲, 青木由直, “3차원 모델을 이용한 한-일 수화 영상통신 시스템의 구현,” 大韓電子工學會夏季綜合學術大會論文集, vol. 21, no. 1, pp. 925-928, 1998. 06
- [10] S.-W. Kim, S.-H. Shin, J. Li and Y. Aoki, “A dictionary structure for sign language translation between Korean and Japanese”, Proc. of the third International Media Symposium'98, Sapporo, pp. 128 - 132, Feb. 1998.
- [11] <http://ece.uwaterloo.ca/~h-anim/spec.html>
- [12] 李種雨, 吳芝英, 金商雲, 青木由直, “스플라인 보간법을 이용한 3차원 수화 애니메이션의 생성,” 大韓電子工學會秋季綜合學術大會論文集, vol. 21, no. 2, pp. 931-934, 1998. 11
- [13] 최원영, “상반신 3차원 모델을 이용한 수화 동영상 생성 시스템,” 공학석사 학위논문, 명지대학교, 1996
- [14] 愼晟孝, 金商雲, 青木由直, “한글 수화 영상통신을 위한 감정과 입모양 표현방법,” 大韓電子工學會秋季綜合學術大會論文集, vol. 20, no. 2, pp. 1217-1221, 1997. 11
- [15] 오지영, 신성효, 김상운, 棚橋 眞, 青木由直, “한글 텍스트의 수화영상 변환 시스템,” 大韓電子工學會夏季綜合學術大會論文集, vol. 20, no. 1, pp. 311-314, 1997. 06

저 자 소 개



金商雲(正會員)

1978年 2月 韓國航空大學 通信情報工學科卒. 1980年 2月 延世大學校 大學院 碩士課程卒. 1988年 2月 同 大學院 博士課程卒(工學博士). 1989年-現 明知大學校 Computer 工學科 助教授, 副教授, 教授. 1992年 - 1993年 北海道大學 情報工學科 客員研究員. 主關心研究分野: Pattern 認識 및 學習, Media工學 등.



青木 由直

1964年 北海道大學 工學部 電子工學科卒. 1966年 同大學 大學院 修士課程卒. 1969年- 1971年 Canada Laval 大學 留學. 1966年 - 現在 北海道大學 講師, 助教授, 教授(情報 Media工學 講座). 中國 瀋陽大學 및 黑龍江大學 客員教授. 主關心研究分野는 信號圖像處理, Computer Graphics, 衛星通信. 工學博士. 主著書는 「波動信號處理」(北森出版), 「オペレータ法ディジタル信號處理」(コロナ社), 「コンピュータグラフィックス講義」(コロナ社).



吳芝英(學生會員)

1993年 3月 - 1997年 2月 明知大學校 Computer工學科 卒(工學士). 1997年 3月 - 1999年 2月 明知大學校 大學院 Computer工學科卒(工學碩士). 主關心研究分野: Pattern 認識, HCI