

Application of Artificial Neural Network For Sign Language Translation

Jeong-Ran Cho*, Hyung-Hoon Kim**

Abstract

In the case of a hearing impaired person using sign language, there are many difficulties in communicating with a normal person who does not understand sign language. The sign language translation system is a system that enables communication between the hearing impaired person using sign language and the normal person who does not understand sign language in this situation. Previous studies on sign language translation systems for communication between normal people and hearing impaired people using sign language are classified into two types using video image system and shape input device. However, the existing sign language translation system does not solve such difficulties due to some problems. Existing sign language translation systems have some problems that they do not recognize various sign language expressions of sign language users and require special devices. Therefore, in this paper, a sign language translation system using an artificial neural network is devised to overcome the problems of the existing system.

▶ Keyword: Sign Language Translation, Artificial Neural Network, Machine Learning, Hearing Impaired Person

I. Introduction

컴퓨터과학기술 및 정보통신환경이 눈부시게 발전하면서 사람들 사이의 정보교환의 형태도 다양하게 발전하고 있다. 그러나 대부분의 상호간 소통을 위한 환경은 비장애인들을 대상으로 하고 있어 장애인의 경우에 대해서는 많은 어려움이 존재하고 있다. 청각 장애인들은 수화라는 소통 수단을 사용하여 상호간의 소통을 이루어가고 있지만, 비장애인들과의 소통에는 많은 어려움이 존재한다. 청각 장애인은 음성 문화권에서 생활하고 있지만 수화를 사용하여 비장애인과 소통해야 하는 어려움으로 인하여 일상생활 및 사회생활에 있어서 많은 불편함과 본인의 욕구와 다르게 여러 가지 불이익을 받고 있는 상황이다.

조사된 통계 자료에 의하면 청각장애인은 수화라는 소통 수단을 사용하는 특별함을 가지고 있으나 독립적으로 외출이 가능하고 사회 각 분야에 진출하여 역할을 수행하고 있는 것으로 파악되고 있다. 보건복지부 한국보건사회연구원에서 2015년 조사한 자료에 의하면 본인이 병의원에 가고 싶을 때 가지 못한 경험이 있는지 알아본 결과 19.1%가 있다고 응답하여 진료가 필요하나 받지 못한 경우가 있는 것으로 나타나고 있다. 이

가운데 뇌병변장애에 이어 청각장애인이 19.5%의 비율로 다른 장애 유형에 비해 높게 나타나 있다.[1] 상호간에 원활한 소통이 이루어지지 않았을 때 발생하는 여러 가지 문제적 상황 가운데 특히 병원진료와 같은 경우에는 비장애인과의 최소한의 소통이라도 반드시 이루어져야 하는 환경이라 할 수 있다.

전통적으로 언어번역 문제는 다른 분야의 문제와 다르게 정확한 함수적 관계식에 의해서 처리하기 어렵다는 특징을 가지고 있다. 이와 같은 언어번역에 대한 어려운 특징은 청각장애인이 대화 수단으로 사용하고 있는 수화에 대한 번역에 있어서도 동일하게 적용되는 특징이다. 수화번역시스템은 수화영상에 대한 인식 능력과 언어번역에 대한 기능을 모두 요구하는 기술로서 기존 언어번역에 대한 과제보다 더 많은 기술이 요구되고 복잡한 처리 과정을 요구하는 분야라 할 수 있다.

기존 수화번역시스템에 대한 일반적인 접근 방법은 크게 수화영상에 대한 영상처리 및 인식기법에 기반 한 수화번역시스템과 형상입력장치인 데이터 글로브와 같은 물리적센서를 사용하여 손가락의 구부림과 손의 방향에 대한 정보 입력장치를 기

• First Author: Jeong-Ran Cho, Corresponding Author: Hyung-Hoon Kim
*Jeong-Ran Cho (jrcho@kwu.ac.kr), Dept. of Health Administration, Kwangju Womens University
**Hyung-Hoon Kim (hhkim@kwu.ac.kr), Dept. of Cosmetic Science, Kwangju Womens University
• Received: 2019. 01. 11, Revised: 2019. 02. 12, Accepted: 2019. 02. 12.

반 한 수화 번역시스템이 연구되고 있다. 형상입력장치인 데이터 글로브를 착용한 수화 번역시스템은 손가락의 움직임을 감지하는 방식으로 수화형상을 인식하고 번역하는 시스템이다. 또한 기존의 수화영상에 대한 영상처리 및 인식기법은 표준 수화영상에 대한 결정된 특징정보를 담고 있는 수화영상데이터베이스를 근거로 수화영상을 인식하고 번역하는 시스템이다.

데이터 글로브와 같은 형상입력장치를 사용한 수화번역시스템은 수화를 인식하기에 필요한 상세한 움직임 감지가 어렵다는 기술적인 문제점과 함께 수화번역시스템 사용성 측면에 있어서 특별한 장치 등 이용 환경을 요구함으로써 불편함 등 현실화의 어려움이 존재한다. 또한 기존의 수화영상에 대한 영상처리 및 인식기법은 수화영상에 대한 몇 가지 결정된 특징정보를 담고 있는 수화영상데이터베이스를 근거로 하고 있기 때문에 수화 사용자의 다양한 수화 표현을 인식하기 어렵고, 변화되는 수화영상 및 기호에 대해 적응성이 떨어진다는 문제점이 있다.

기존 수화번역시스템이 공통적으로 가지고 있는 문제점은 수화 사용자의 다양한 수화 표현을 인식하면서 사용의 편리성을 좋게 하는 수화번역시스템으로 만들기 어렵다는 것이다. 기존 수화번역시스템은 수화에 대한 데이터베이스를 구성하여 비디오 영상 또는 형상입력장치로부터 인식된 수화를 저장된 데이터베이스의 내용을 근거로 번역하는 체계를 가지고 있다. 따라서 수화 사용자의 다양한 수화의 특징을 모두 반영한 시스템을 구현하기란 쉽지 않은 근본적인 문제점을 가지고 있다.

인공신경망은 오래전부터 과학자와 공학자에게 큰 관심을 받아 왔고, 지금까지 개발된 최고의 계산도구중 하나로 평가 받고 있다. 이와 같이 기대하는 큰 이유는 인공신경망이 경험적 데이터 또는 과거의 사례 데이터로부터 학습할 수 있는 사람의 능력을 시뮬레이션할 수 있기 때문이다. 이 인공신경망은 불완전한 정보로 학습 또는 훈련 되었을 때에도 적절한 결과 값을 결정하거나 추측할 수 있다는 장점을 가지고 있다. 또한 인공신경망은 훈련 데이터에 포함되어 있지 않은 전혀 새로운 데이터에 대해서도 이를 받아들여 타당한 결과를 추측함으로써 사람의 창의적 처리를 어느 정도 모방하고 있다고 평가할 수 있다. 이에 본 논문에서는 기존 수화번역시스템의 특징과 문제점을 검토하고 이를 보완하기 위한 방법으로 인공신경망을 이용한 수화번역시스템을 고안하였다.

II. Analysis on Translation System

2.1 Sign Language Concept

수화란 손의 움직임과 비수지 신호(얼굴표정과 몸짓)를 사용하여 표현하는 시각언어이다. 한국 수화의 정립은 1947년 국립맹아학교 초대 교장인 윤백원에 의하여 한국 수화 지문자가 만들어지면서 그 형태를 갖추기 시작하였고, 1982년 서울 농아학교 교직원들이 만든 약 6,000여 단어가 수록된 표준 수화 사전

이 발간되어 수화 보급에 획기적인 기틀이 마련되었다.[2, 3] 한국수화는 문장구성과 문법범주 각각의 측면에서 다음과 같은 통사론적 특징이 있다.[4]

(1) 문장구성 측면의 특징

한국수화의 문장구성상의 통사론적 특징은 어휘 형태소를 중심으로 의미 구조가 통합되어 문장을 생성해 낸다. 즉, 개체와 동작을 나타내는 의미형태소를 끝에 완료시제나 추측 등을 나타내는 의미형태소를 붙여 문장을 완성시키는 통사규칙을 갖는다. 예를 들어 “순이가 왔다”를 표현하는 수화문장구성 <순회><오다><끝> 에서 <순회><오다>는 개체와 동작을 나타내는 의미형태소들이고, <끝>은 완료시제를 나타내는 의미형태소이다.

(2) 문법범주 측면의 특징

한국수화의 문법범주상의 통사론적 특징은 어휘 형태소와 비수지적 보조언어가 통합되어 문장을 생성해 낸다. 수화에서의 비수지적 보조언어에는 눈 또는 눈동자의 움직임, 안면 근육의 움직임, 입술이나 코의 움직임, 머리, 턱, 목, 목의 움직임, 그리고 수화 어휘표현상의 크고 작음, 긴장도, 강약, 처짐, 빠름과 느림 등이 있다. 예를 들어 청유문은 ‘바라다’, ‘부탁하다’ 등의 어휘형태소와 간절한 얼굴표정, 눈빛, 고개 움직임 등을 통합하여 표현한다.

우리나라에서 상용되는 수화에는 한국수화와 국어 대응식 표준수화가 있다. 한국수화는 청각장애인들이 음성언어의 의미를 그들 나름으로 자연스럽게 표현해 온 신호체계이다. 이 자연수화를 농교육계에서는 공식적으로 한국수화(Korean Sign Language)라고 한다. 국어 대응식 표준수화는 한국어 문법과 어순에 따라 만들어진 수화이다. 이는 ‘고안된 한국수화’, ‘고안된 국어 대응식 수화’, ‘고안된 국어 대응식 표준한국수화’, ‘한글식 수화’, ‘한글식 표준수화’ 등으로 다양하게 부르는데, 이중 농교육계에서 공식적으로 쓰이고 있는 명칭은 ‘국어 대응식 표준수화’이다. 한국수화와 국어 대응식 표준수화의 언어학적 특징은 표 1과 같다.

Table 1. Characteristics of Korean Sign Language and Standard Sign Language for Korean

	Korean sign language	Standard Sign Language for Korean
Agreement with the Korean language	It has a somewhat unique grammar, which does not match the meaning of words and grammar with Korean.	Use of Termination, Expletive, Pronoun is correct.
Ease of understanding	Easy to understand Even if you do not understand Korean	Difficult to understand if you do not understand Korean
Use of A finger character	Use but few	Use aggressively
Combined with A Spoken word	Not Combined	Combined

2.2 Sign Language Translation System

지금까지 연구된 수화 번역 시스템 몇 가지를 소개하고자 한다.

(1) 자연수화문장생성을 위한 한국어-수화 형태소 변환

[Su-Hyun Park, Seok-Hoon Kang, Hyuk-Chul Kwon,1996]에서는 한국어-수화 번역 시스템을 위한 수화 형태소의 정의 및 분석과 한국어-수화의 형태소 변환 규칙을 제안하였다. 한국어를 수화로 번역하기 위한 한국어-수화 번역시스템(TeST)의 개요도를 그림 1과 같이 구성하였다. 한국어 텍스트를 입력받아 기존의 언어처리시스템과 같이 형태소 해석과 구문 및 의미해석을 행한다. 이 결과를 수화 형태소로 변환하여 동작 사전 및 연결규칙에 의해 수화 동영상 생성하게 된다.

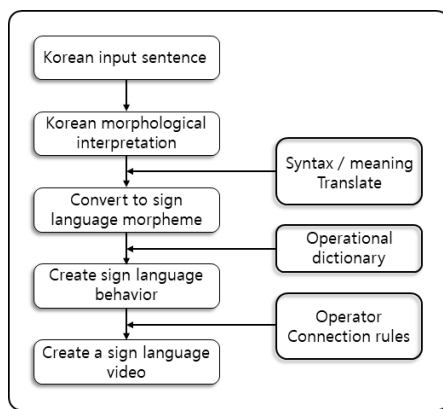


Fig. 1. Overview of the TeST System

한국어-수화 번역시스템은 다음과 같이 두 단계로 구성된다. 1단계에서는 입력된 한국어 텍스트를 언어 처리과정을 통해 해석하고, 한국어 형태소 해석결과를 수화문장 구조에 적합한 형태소로 분리한 후 재배열한다. 2단계에서는 수화사전에서 각각의 토큰에 해당하는 수화이미지를 검색하고 이를 이용하여 수화 동영상을 생성한다.

한국어를 수화로 바꿀 때 형태소분석 만으로는 그 의미를 정확히 번역해 낼 수 없다. 그것은 한국어가 지닌 의미의 모호함과 수화만의 독특한 규칙들에 기인한다. 즉, 문장에 따라 다른 의미를 지닌 수화, 주어의 입장에 따라 수화의 방향이 변하는 경우, 관형어구들, 시제의 표현, 비수지적 보조언어의 정리 등과 같은 문제의 해결이 뒤따라야 완전한 한국어-수화 번역 시스템이 가능하다.

(2) 한국어-수화 번역에 필요한 사전구성과 이를 이용한 번역시스템

번역의 과정은 한국어 문장이 입력되면 먼저 형태소분석기를 통해서 문장을 어근과 조사 등 기본단위로 분석한다. 한국어와 한국 수화는 기본적인 어순은 같지만 조사 및 접속사의 사용, 어순의 도치, 시제의 표현 등 표현의 차이가 있으므로 재배열/생성 단계에서 한국어와 수화 표현의 차이에서 발생하는 구성의 변경 작업을 처리한다. 재배열/생성 단계의 출력은 목적

어인 수화로 표현해야 할 단어로서 사전 검색의 키워드이다. 최종적으로 재배열된 단어들은 사전 검색을 통해 번역하고자 하는 목적어인 수화의 동작으로 표현된다.[2] 번역에 관련된 자연어 처리는 형태소 분석, 구문 분석, 사전 구성 등 원시언어의 해석과정과 목적어언어로의 변환 및 생성 과정 등 많은 작업을 수행해야 한다.[2, 4, 5]

[Kyung-Hyuk Kwon, Yo-Seop Woo, Hong-Ki Min,2000]에서는 문어체 한국어 문장들을 청각장애인이 이해 가능한 수화로 번역하였고, 또한 번역에 사용한 사전의 유용성도 입증하였다. 하지만 실제 수화의 표현은 단순한 손동작만으로는 표현되지 않고 얼굴의 표정과 동작의 크기 등을 고려해야 하며, 화자에 따라 수화의 방향이 바뀌는 문제 등 단순한 데이터베이스 검색만으로는 해결할 수 없는 문제들이 많이 있다. 또한 재배열/생성 과정에서 발생하는 많은 예외의 경우들에 대한 연구와 다양한 한국어에 대한 수화 사전 구성에 많은 시간과 노력이 필요하며 수화 GIF 생성을 위해 사용한 디지털 사진을 색상 수와 해상도가 적은 애니메이션으로 변환하여 출력 속도를 향상시키고 저장 공간을 줄여야 한다.

(3) 응급의료 상황을 위한 양방향 수화시스템

[Ryu M. W.,2015]에서는 한국수화와 국어 대응식 표준수화에서 사용하는 총수화 단어의 수가 너무 많아 빠른 치료를 요구하는 응급상황에서 쓰기에 적합하지 않기 때문에 응급실 문진 리스트를 바탕으로 한 응급수화를 제시하였고, 응급의료 상황을 위한 양방향 수화 시스템을 제안하였다.

의료 전문인과 청각장애인 환자 간의 의사소통을 위해 의료 전문인이 질문을 텍스트로 입력하면 응급수화 연산에 따라 진행된 결과가 언어set에 연결되어 수화 이미지를 보여준다. 수화 이미지를 보여 줌으로써 수화를 모르는 의료 전문인이 청각장애인 환자의 현재 상태를 물어볼 수 있도록 하였다. 그리고 스마트폰 카메라로 청각 장애인의 수화 모습을 촬영하면 어떤 뜻을 가진 수화인지 의료 전문인이 알 수 있도록 해석해주는 수화 인식 시스템을 구현하여 양방향 수화 시스템을 구현하였다.

이는 의료인의 수화지식 부족으로 치료시기가 지체되어 사망으로까지 이어진 경우와 의사소통 문제로 청각 장애인들에게 잘못된 진료행위, 즉 자신의 질병과 다른 치료를 하고 있었던 경우(진단의 오류), 치료방법을 잘못 알아들은 경우(치료방법의 오류), 가족의 강요에 의한 치료행위와 같은 형태의 오류가 발생하는 경우를 방지할 수 있게 하였다. 또한, 사람들이 많이 사용하고 있는 스마트폰에서 사용할 수 있게 애플리케이션으로 구현함으로써 휴대성과 이동성, 범용성이 뛰어나고 비용에 대한 걱정도 줄었다.

더 나아가 수화인식 부분의 인식률을 높이기 위해서는 윤곽선 검출 후 특징점을 추출하는 방법과 많은 데이터베이스를 두어 Template 매칭을 진행하는 방법을 사용한다면 더 유용한 시스템이 될 것이다.

(4) 딥러닝 기반의 수화번역 시스템

[Seung-Seok Lee, Jeong-Hyun Heo, Seung-Woo No, Hyeon-Jin Yoon, So-Hyun Park, Chan-Kyu Kim, 2018]에서는 언어장애인들과 원활한 의사 소통을 위해 수화를 딥러닝 기법으로 번역하고 번역된 문장들을 음성으로 출력시키는 딥러닝 기반의 수화번역 시스템을 제안하고 있다.

그림 2는 딥러닝 기법이 적용된 수화번역 시스템의 설계도이다. 사용자 User가 수화 동작을 취하면 장갑에 장착되어 있는 Flex sensor와 MPU6050 sensor의 센서값들이 변화하게 되고 Arduino uno는 변화한 센서값들을 PC로 전송한다. Deep Learning으로 학습되어있던 PC는 센서값이 표현하는 동작을 감지하여 학습되어 있던 데이터에서 알맞은 문장을 찾아낸다. 그리고 다시 Arduino의 Speaker module(ELB080905)로 전송한다. 문장이 전송되면 Speaker module은 문장을 음성으로 출력하여 의사소통이 가능하게 설계하였다.

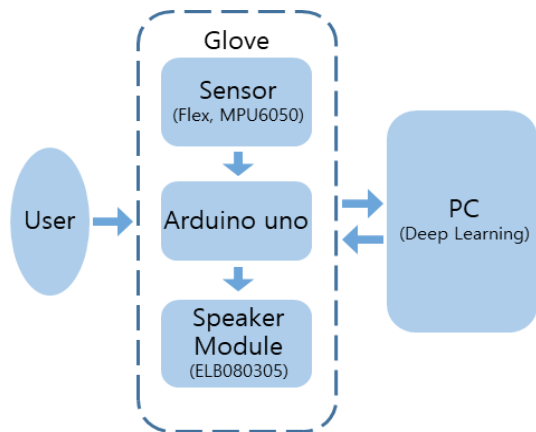


Fig. 2. System design with deep running technique

(5) 스마트폰 수화번역 어플리케이션 개발

[Haesun Paeng, Jihun Lee, Jaeyun Cho, 2012]에서는 스마트폰 모바일 3D엔진 기술을 활용하여 수화의 언어적 특성을 반영한 수화번역 어플리케이션을 개발하였다. 스마트폰에서 서비스가 가능한 이 어플리케이션은 입력된 문자데이터를 실시간으로 자원 사이의 관계-의미 정보를 파악하여 3D 모바일 환경에 출력하여 문자 정보의 수화 번역이 가능한 기술을 구현하였다.

주요 기술개발 내용은 다음과 같다.

① 모바일 디바이스를 통하여 입력된 문장 분석

모바일 디바이스를 통하여 입력된 문장은 문장단위로 인식되어 문장의 형태소 단위로 분석, 일반명사/복합명사/대명사/형용사/동사 등을 추출한다. 추출된 형태소를 분석하여 각 구성요소에 맞는 신호를 부여하고, 메시지를 매칭시킨다. 이때 수화의 언어적 특성인 조사 생략, 의문문의 표현, 시제표현을 고려하여야 한다.

② 수화 동작 표현이 가능한 3D 캐릭터 제작

모바일 환경에서 자연스러운 동작표현이 가능하며 수화 통역 시스템에서 최적화된 수화동작이 표현될 수 있는 캐릭터를 설계

한다. 캐릭터의 디테일한 면이 살아날 수 있도록 Normal-map 기술을 적용하여 사실감 있는 캐릭터를 구현한다.

③ 3D 모바일 환경 구축

모바일 3D 엔진을 활용하여 개발된 3D 캐릭터가 수화동작을 자유자재로 표현 할 수 있고, 실시간으로 사용자가 조작할 수 있는 모바일 환경을 구축한다.

④ 3D 데이터의 실시간 모바일 재생기술 구축

입력된 문자 데이터를 실시간으로 자원 사이의 관계-의미 정보를 파악하여 특정 신호를 호출, 3D 데이터는 이 특정 신호를 받아서 3D 모바일 viewer창에 해당 신호의 동작을 디스플레이 한다.

전문 수화통역사가 아닌 일반인의 수화번역이 가능하고 문화복지 산업에서 사용가능한 어플리케이션 시장은 현재까지 미개척 CT기술 개발 분야로서 향후 고부가가치 산업으로서 성장할 가능성이 매우 높은 분야이다.

III. Proposed Translation System

3.1 Concept of Artificial Neural Network

인공신경망은 사람의 신경학적 처리 능력을 시뮬레이션할 수 있다. 사람의 뇌는 평균적으로 약 100억개의 뉴런을 포함하고 있고, 각 뉴런은 다른 뉴런과 10,000개 정도의 연결을 갖고 있다. 그림 3에서 각 뉴런은 신호를 입력 받는 수상돌기와 정보를 수신하여 처리하는 세포 몸체, 그리고 다음 뉴런에 정보를 전달하기 위한 축삭돌기인 3개의 주요 부분으로 구성된다.[1]

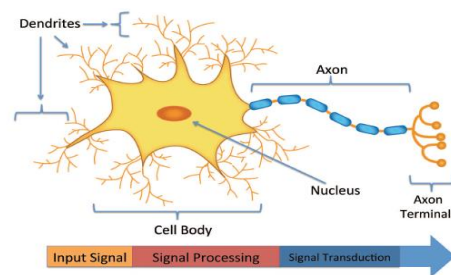


Fig. 3. Conceptual structure of biological neurons

사람의 뇌는 뇌세포간의 연결을 통해 학습하고, 패턴을 인식하고, 결과를 예측하게 된다. 이와 같은 생물학적 뇌의 특성을 모방하여 만들어진 인공신경망은 많은 노드들로 구성되며, 각 노드는 사람의 뇌세포에 해당하는 인공신경세포이다. 모든 노드들 사이의 연결은 서로 다른 가중치를 갖는다. 인공신경망은 사람 두뇌의 동작을 흉내 내며, 잠재적으로 사람처럼 생각할 수 있는 기계를 개발하려는 과학자들의 소중한 꿈을 실현하게 한다. 인공신경망은 복잡한 다차원적 문제에 대한 데이터 모델링과 패턴 인식을 통하여 사람 두뇌의 학습과 일반화 동작을 시뮬레이션 한다.

인공신경망 모델과 일반적인 통계적 모델 사이의 중요한 차이점은 인공신경망은 특별한 수학적 함수 없이 독립변수와 종속변수 사이의 관계성을 일반화할 수 있다는 것이다. 따라서 인공신경망은 수화번역시스템과 같은 수화자의 다양한 수화 표현에 대한 인식과 언어번역과 같은 비함수적 관계성에 대한 문제를 해결하는 좋은 도구로 분석되어 진다.

새로운 인공신경망이 계속적으로 개발되면서 많은 유형이 존재하지만, 모든 인공신경망은 전이함수, 학습 규칙, 연결 형식에 의하여 종류가 분류될 수 있다. 가장 일반적으로 사용되는 인공신경망의 종류는 오차역전파로 훈련된 전방향 전과 인공신경망이다. 그림 4에 보인 것과 같이 전방향 전과 인공신경망은 입력층, 하나 이상의 은닉층 그리고 출력층으로 구성된다.

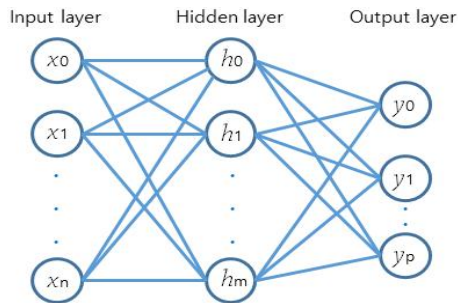


Fig. 4. Structure of artificial neural network

입력층은 외부 소스로부터 데이터를 공급 받으며, 은닉층은 입력층에서 입력된 데이터를 연결된 링크의 가중치에 근거하여 다음 계층에 전달한다. 최종 결과 값은 출력층에 생성된다. 각 계층의 노드는 다음 계층에 대한 출력을 계산하기 위해 이전 계층으로부터 입력을 받아 합계를 계산하고 시그모이드 전이함수에 의하여 출력을 식(1), (2)와 같이 계산한다.

$$y_j = b + \sum w_{ij}x_i \quad (1)$$

$$f(y_j) = \frac{1}{1 + \exp(-\alpha y_j)} \quad (2)$$

식 (1)에서, w_{ij} 는 이전 계층에 있는 노드 i 와 현재 계층에 있는 노드 j 사이의 링크에 대한 가중치이고, x_i 는 이전 계층으로부터의 출력 값이고 b 는 편향 값이다. 식 (2)는 전이함수로 사용된 시그모이드함수로서 $f(y_j)$ 는 출력 값으로 다음 계층에 전달되는 값이고, α 는 시그모이드 함수의 형태와 관계된 파라미터이다. 이 시그모이드 함수의 장점은 과도한 감쇠 없이 작은 신호를 통과시킬 수 있고, 포화 없이 큰 신호를 수용할 수 있다는 것이다. 시그모이드 함수의 비선형성은 α 가 증가함에 따라 강화된다.

인공신경망에서 은닉층의 개수와 은닉층에 포함되는 노드의 개수는 인공신경망의 기능과 성능에 영향을 준다. 최적의 은닉층의 개수와 노드의 개수에 대한 선택은 일반적으로 시행착오의 방법으로 결정되어 진다. 인공신경망에 대한 은닉층 노드의

개수에 대한 상한선에 대해 식 (3)을 참고하여 결정할 수 있다.

$$N_{hidden} = \frac{N_{trn}}{[R + (N_{inp} + N_{out})]} \quad (3)$$

식 (3)에서, N_{hidden} 은 은닉층에 있는 노드의 개수이고, N_{trn} 은 훈련 데이터의 개수, R 은 5에서 10까지 범위 안에 있는 상수 값, N_{inp} 는 입력층 노드의 개수 그리고 N_{out} 는 출력층 노드의 개수이다.

또한 훈련 데이터의 크기 및 훈련의 양은 인공신경망의 사용에 있어서 매우 중요하다. 덜 훈련된 인공신경망은 큰 오류를 가지게 되며, 너무 과하게 훈련된 인공신경망은 패턴인식 및 일반화에 대한 능력을 오히려 떨어뜨리는 현상을 만들기 때문이다. 식 (4)는 입력층 노드의 개수, 은닉층 노드의 개수, 출력층 노드의 개수와 관련하여 훈련 데이터의 양을 합리적으로 예측하는데 사용된다.

$$n_s = \beta(n_h(n_i + 1) + n_o(n_h + 1)) \quad (4)$$

식 (4)에서, n_h 은 은닉층 노드의 개수, n_i 은 입력층 노드의 개수, n_o 는 출력층 노드의 개수, 그리고 n_s 는 훈련 데이터 양의 개수이다. 상수 β 는 과적합의 정도에 관련된 파라미터이다.

3.2 Sign Language Translation System using Artificial Neural Network

본 논문에서는 기존 수화번역시스템의 단점을 보완하여 청각 장애인과 비장애인과의 수화를 사용한 원활한 소통을 위하여 인공신경망을 사용한 수화번역시스템을 고안하였다. 수화번역시스템은 수화를 한국어로 번역하는 수화-한국어번역시스템과 한국어를 수화로 번역하는 한국어-수화번역시스템으로 구성된다.

3.2.1 Sign Language-Korean Translation System

수화-한국어번역시스템은 수화영상을 한국어로 번역하는 시스템으로서 청각 장애인의 수화 내용을 비장애인이 한국어로 번역된 내용으로 이해하기 위하여 사용되는 시스템이다.

그림 5는 일반적인 절차적 프로그램과 인공신경망을 통한 기계학습을 사용한 수화-한국어번역시스템에 대한 개념도이다. 수화-한국어번역시스템은 크게 수화영상전처리단계, 인공신경망수화영상인식단계, 수화문장처리단계, 인공신경망한국어번역단계의 4단계로 구성된다.

수화영상전처리단계는 스마트폰 또는 영상입력장치로부터 입력된 수화자의 수화영상을 동영상으로 입력받아 수화영상인식에 적절한 상태로 이미지처리를 수행하는 단계이다. 수화영상전처리단계를 통하여 수화자의 수화영상은 동영상으로부터 주요한 정지영상인 수화영상이 선별되어진다. 선별된 수화영상으로부터 수화자의 손 영역을 찾아 손 영역에 대한 수화영역을

추출하고 다음 단계에서 요구하는 이미지 색상 수, 이미지 크기, 노이즈 제거와 같은 수화영상의 인식에 필요한 이미지 처리를 수행한다.

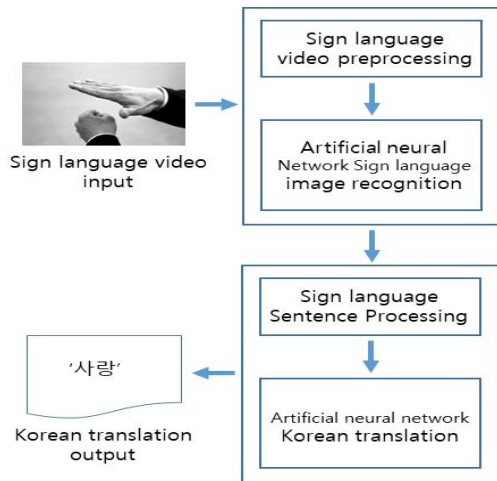


Fig. 5. Concept of sign language-korean translation system using artificial neural network

인공신경망수화영상인식단계는 인공신경망으로 구축된 시스템으로서 수화영상전처리단계로부터 전달된 수화영상을 인식하는 단계이다. 인공신경망수화영상인식단계는 사전에 수화영상과 정답 기호 라벨의 세트로 구성된 충분한 훈련 데이터를 사용하여 학습된 인공신경망 수화영상인식시스템이다.

그림 6은 수화영상전처리단계로부터 전달된 수화영상을 인식하기 위한 인공신경망으로 구성된 인공신경망수화영상인식시스템의 개념도를 나타내며 여기에서는 숫자 수화에 대한 내용으로 제한하여 표현하였다. 인공신경망수화영상인식단계의 인식 과정은 입력층에서 수화영상전처리단계로부터 전달된 수화영상을 입력 받아 이미 학습된 인공신경망의 은닉층을 통하여 매핑이 이루어져 출력층에서 인식된 결과를 예측하게 된다.

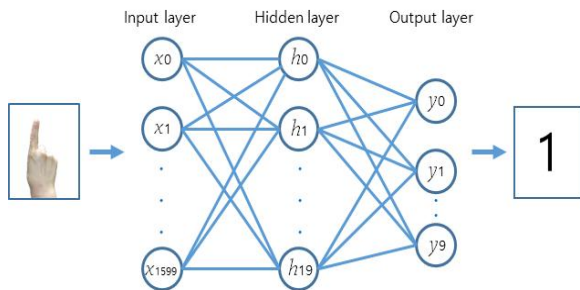


Fig. 6. Artificial neural network concept of sign language video recognition system

수화문장처리단계는 인공신경망수화영상인식단계에서 인식된 수화 인식 결과를 사용하여 수화 문장형태에 대하여 분석하여 수화에서 한국어로 번역하기 위한 준비 처리를 수행한다.

마지막으로 인공신경망한국어번역단계는 수화문장처리단계에서 전달된 인식된 각 수화 문장을 이미 학습된 인공신경망한

국어번역단계를 통하여 한국어 문장으로 번역한다. 인공신경망 한국어번역단계는 사전에 수화 언어와 정답 한국어 라벨 세트 로 구성된 충분한 훈련데이터를 사용하여 훈련된 인공신경망한 국어번역시스템이다.

3.2.2 Korean-Sign Language Translation System

한국어-수화번역시스템은 한국어를 입력받아 수화영상으로 번역하는 시스템으로서 비장애인이 표현한 한국어를 청각 장애 인이 사용하는 수화언어로 번역하여 제공함으로써 비장애인의 한국어를 청각 장애인이 수화로 이해할 수 있도록 하기 위하여 사용되는 시스템이다.

그림 7은 일반적인 절차적 프로그램과 인공신경망을 통한 기계학습을 사용한 한국어-수화번역시스템에 대한 개념도이다. 한국어-수화번역시스템은 크게 문장형태소분석단계, 인공신경 망수화번역단계, 수화영상처리단계의 3단계로 구성된다.

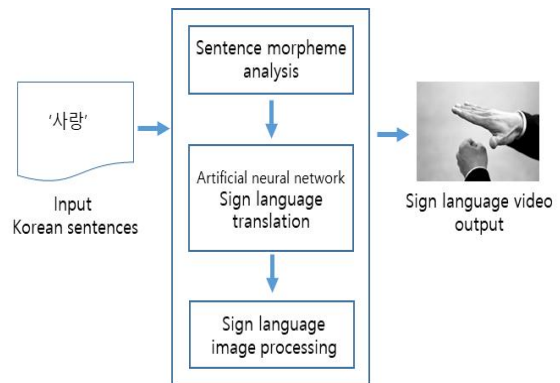


Fig. 7. Concept of korean-sign language translation system using artificial neural network

문장형태소분석단계는 비장애인이 입력한 한국어 문장에 대 해 형태소를 분석하여 한국어에서 수화로 번역하기 위한 준비 처리를 수행한다. 인공신경망수화번역단계는 인공신경망으로 구축된 시스템으로서 문장형태소분석단계로부터 전달된 한국 어 문장을 수화로 번역하는 단계이다. 인공신경망수화번역시스 템은 사전에 한국어문장과 정답 수화기호 라벨의 세트로 구성 된 충분한 훈련 데이터를 사용하여 학습된 인공신경망수화번역 시스템이다.

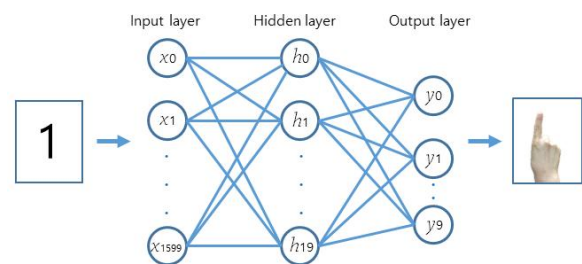


Fig. 8. Artificial neural network concept of sign language translation system

그림 8은 문장형태소분석단계로부터 전달된 한국어를 수화로 번역하기 위한 인공신경망으로 구성된 인공신경망수화번역시스템의 개념도를 나타내며 여기에서는 숫자에 대한 내용으로 제한하여 표현하였다.

인공신경망수화번역단계의 번역과정은 입력층에서 문장형태소분석단계로부터 전달된 한국어를 입력 받아 이미 학습된 인공신경망의 은닉층을 통하여 매핑이 이루어져 출력층에서 인식된 수화 결과를 예측하게 된다. 마지막으로 수화영상처리단계는 전단계에 의하여 한국어에서 수화로 번역된 결과에 대한 수화영상을 생성 및 처리하는 단계이다.

IV. Conclusion

정보통신기술이 발전함에 따라서 사람들 사이의 정보교환의 형태가 다양하게 발전하고 있다. 그러나 수화를 사용하고 있는 청각 장애인의 경우에 대해서는 수화를 이해하지 못하는 비장애인과 소통을 하는데 있어서 아직도 많은 어려움이 존재하고 있는 상황이다. 특히 수화를 이해하지 못하는 음성 문화권에서 생활하고 있는 청각 장애인은 수화를 사용해야 하는 이유로 일상생활 및 사회생활에 있어서 많은 불편함과 본인의 욕구와 다르게 여러 가지 불이익을 받고 있는 상황이다.

따라서 본 논문에서는 수화를 사용하는 청각 장애인의 비장애인과 소통에 대한 문제를 해결하기 위하여 수화번역시스템에 대해 검토하였다. 기존 수화번역시스템에 대한 검토 결과에 의하면 기존의 방법은 크게 수화영상에 대한 영상처리 및 인식기법에 기반한 수화번역시스템과 형상입력장치와 같은 물리적 센서를 사용한 수화번역시스템의 두 가지로 나누어 진행되고 있음을 알 수 있었다.

기존 수화번역시스템이 공통적으로 가지고 있는 문제점은 수화 사용자의 다양한 수화 표현을 인식하기 어렵고, 수화 사용자에게 특별한 장치를 요구함으로써 편리성에 대한 문제점이 존재한다. 기존 수화번역시스템은 수화에 대한 수화영상특징정보 또는 수화자의 움직임 특징정보에 대한 데이터베이스를 근거로 번역하는 체계를 사용하고 있어서 수화 사용자의 다양한 수화 표현을 인식하기가 근본적으로 어려운 문제점을 가지고 있다.

전통적으로 언어번역 문제는 정확한 함수적 관계식 또는 몇 가지 특징정보에 기반한 선형적 매핑 관계로 처리하기 어렵다는 특징을 가지는 분야이다. 이와 같은 어려움은 수화번역에 있어서도 동일하게 적용되며, 수화영상인식에 대한 어려움까지 존재하여 더욱더 많은 기술과 복잡한 처리과정이 요구되는 분야라 할 수 있다.

인공신경망은 경험적 데이터 또는 과거의 사례 데이터로부터 학습할 수 있는 사람의 능력을 시뮬레이션하기 위하여 사용되는 계산도구이다. 인공신경망은 언어번역과 같이 비합수적이고 수화 사용자의 다양한 수화 표현과 같은 다양한 변수를 훈련 데이터로 받아들여 타당성 있고 합리적인 결과를 예측할 수

있는 컴퓨터 계산 모델 가운데 하나이다. 또한 인공신경망은 전혀 새로운 데이터에 대해서도 이를 받아들여 적절한 결과를 추측함으로써 사람의 창의적 능력을 어느 정도 모방할 수 있다고 평가되는 도구이다.

본 논문에서는 기존 수화번역시스템의 특징과 문제점을 검토하였고 이를 보완하기 위한 방법으로 인공신경망을 이용한 수화번역시스템을 고안하였다. 본 논문에서 고안된 인공신경망을 사용한 수화번역시스템에서는 기존의 다양한 절차적 프로세스인 영상처리기술과 병합하여 구성함으로써 수화번역시스템의 정확성과 실용성을 더욱더 높이는 방향이 되도록 고안하였다. 추후 연구에서는 본 논문에서 고안된 인공신경망 수화번역시스템의 실질적 가능성을 검토하기 위한 다양한 실험과 분석을 진행하고, 여기에 필요한 훈련 데이터의 준비, 훈련 방법, 인공신경망 설계기술 등에 대해 연구하고자 한다.

REFERENCES

- [1] Vijaykumar Sutariya, Anastasia Groshev, Prabodh Sadana, Deepak Bhatia, Yashwant Pathak, "Artificial Neural Network in Drug Delivery and Pharmaceutical Research", The Open Bioinformatics Journal, Jul. 2013.
- [2] Kyung-Hyuk Kwon, Yo-Seop Woo, Hong-Ki Min, "Design and Implementation of a Korean Text to Sign Language Translation System", The KIPS Transactions, Vol 7, No 3, pp. 756-765, March 2000.
- [3] Korea deaf children's neighbors love society, "Open sign language", Book published blue ears, pp. 231-232, Jan. 1998.
- [4] Su-Hyun Park, Seok-Hoon Kang, Hyuk-Chul Kwon, "Morpheme Conversion for Korean Text-to-sign Language Translation System", The KIPS Transactions, Vol.5, No.3, pp. 688-702, March 1996.
- [5] Mi-Sook Eum, "Analysis of Syntactic Characteristics of Korean Sign Language", Daegu university Graduate School M.S., 1996.
- [6] Ryu M. W., "Design and Implementation of Bi-directional Sign Language System for Emergency Medical Situation", Pusan National University M.S, Feb. 2015.
- [7] Seung-Seok Lee, Jeong-Hyun Heo, Seung-Woo No, Hyeon-Jin Yoon, So-Hyun Park, Chan-Kyu Kim, "Sign Language Translation System based on Deep Learning for speech disorders", Proceedings of Symposium of the Korean Institute of communications and Information Sciences, pp. 951-952, Jun. 2018.
- [8] Haesun Paeng, Jihun Lee, Jaeyun Cho, "Research for Sign Language Translation Application Development Base on Smartphone Environment", Proceedings of

- Symposium of The HCI Society of Korea, pp. 62-65.
Jan. 2012.
- [9] Hyo-Seob, Se-Jin Chang, Byung-Ju Shin, Yun-Mo Yang, "Sign-Language Recognition using the Information of Hand Shape and Moving Direction", Journal of KISS(B): Software and Applications, Vol.26, No.6, pp. 804-810, Jun. 1999.
- [10] Hoon Jang, Young-Jun Oh, "Sign language interpretation system design and implementation", Proceedings of The KISS Conference, Vol.29, No.22, pp. 691-693, 2002.
- [11] F. Liu, M. J. Er, "A novel efficient learning algorithm for self-generating fuzzy neural network with applications", International Journal of Neural Systems, Vol. 22, pp21-35, Feb. 2012.
- [12] D. K. Kim, "C++ API OpenCV Programming", The Publish Company of KAME, May 2015.
- [13] Lee Hyun-Suk, Kim Seung-Pil, Chung Wan-Young, "Development of Sign Language Translation System using Motion Recognition of Kinect", Journal of the institute of signal processing and systems, Vol. 14, No. 4, pp.235-242, Oct. 2013.

Authors



Jeong-Ran Cho received the B.S., M.S. and Ph.D. degrees in Computer Science from Chonnam National University, Korea, in 1987, 1989 and 1999, respectively. Dr. Cho joined the faculty of Kwangju Womens University, Gwangju, Korea, in 1994.

She is currently a Professor in the Department of Health Administration at Kwangju Womens University. She is interested in database, parallel computing, internet and mobile computing, and multimedia contents service.



Hyung-Hoon Kim received the B.S. degrees in Computer Science from Chonnam National University, Korea, in 1986. He received the M.S. degrees in Computer Science from KAIST, Korea, in 1988. He received the Ph.D. degrees in Computer

Science from Hanyang University, Korea, in 2007. Dr. Kim joined the faculty of Kwangju Womens University, Gwangju, Korea, in 1994. He is currently a Professor in the Department of Cosmetic Science at Kwangju Womens University. He is interested in web programming, medical information system, information security, and AI.