《研究論文》

Journal of the Human Engineering Society of Korea
Vol. 10 No. 1, June 1991.

산업용 로보트의 동작제어 명령어의 인식에 관한 연구

(A Study on the voice command recognition at the motion control in the industrial robot)

이순요†, 권규식††, 김홍태††

ABSTRACT

The teach pendant and keyboard have been used as an input device of control command in human-robot system. But, many problems occur in case that the user is a novice. So, speech recognition system is required to communicate between a human and the robot.

In this study, Korean voice commands, eight robot commands, and ten digits based on the broad phonetic analysis are described. Applying broad phonetic analysis, phonemes of voice commands are divided into phoneme groups, such as plosive, fricative, affricative, nasal, and glide sound, having similar features.

And then, the feature parameters and their ranges to detect phoneme groups are found by minimax method. Classification rules are consisted of combination of the feature parameters, such as zero crossing rate(ZCR), log energy(LE), up and down(UD), formant frequency, and their ranges. Voice commands were recognized by the classification rules.

The recognition rate was over 90 percent in this experiment. Also, this experiment showed that the recognition rate about digits was better than that about robot commands.

I. 서 론

음성은 인간의 의사소통에 없어서는 안될 수단이다. 인간은 자신이 생각하고 있는 바를 음성으로 다른 사람에게 전하는 경우가 대단히 많다. 특히 일상생활에서는 거의 모든 경우에 자신의 의사를 음성으로 전하고 있으므로 인간-로보트 시스템에서 로보트와의 의사소통에도 음성을 이용하는 것이 바람직할 것이다.

인간-로보트 시스템에서 인간조작자가 로보트를 제어함에 있어서는 기존의 정보 입력수단으로 가장 보편적인 Teach Pendant나 키보드에 의할 경우, 사용자가 초보자일 경우에는 음성을 사용하는 것보다 에러를 많이 범하게 되기 때문에 자연스러운 정보전달수단인 음성을 통하여 의사소통을 하는 것이 효율적일 것이다[1].

음성을 통한 의사소통을 위해서는 인간과 로보트

† 고려대학교 산업공학과

†† 고려대학교 산업공학과 대학원

의 관계에 관한 Human Interface 기술 중에서 인간과 로보트의 대화기능을 갖추어야 하며, 이에는 음성입력을 위한 음성인식과 음성출력을 위한 음성합성이 있다[2].

이러한 Human Interface 기술에서 음성입력의 장점을 정리하면 다음과 같다[3].

① 가장 자연스럽고 신속한 입력수단이다.
② 전화 등의 통신수단을 이용해 정보를 주고 받을 수 있다.
③ 입력자가 정보를 입력하면서 동시에 자신이 입력한 내용을 별다른 절차없이 확인할 수 있다.
④ 입력하는 과정에서 시각이나 촉각에 점유되지 않으므로 계속해서 작업장을 감시하면서 정보를 입력할 수 있다.

이러한 음성인식기능을 인간-로보트 시스템에 적용하게 되면, 인간과 로보트사이의 협조작업시 작업수행의 효율화를 기할 수 있을 뿐만 아니라 인간조작자에게 가해지는 작업부담을 경감시켜 주는 잇점을 가져오게 된다[4].

음성입력을 위한 음성인식기술은 대상으로 하는 화자(話者)의 범위에 따라 화자종속방식(Speaker Dependent System)과 화자독립방식(Speaker Independent System)으로 나누어질 수 있다[5], [6].

본 연구에서는 화자종속방식에서처럼 패턴을 따로 저장해야 되는 기억장소와 이의 관리 및 처리를 위한 하드웨어가 필요없이 음성을 입력받아 수치데이터로 만드는 기본의 회로만 있으면 입력신호의 분석과 음성의 특성을 뽑아내고, 식별하는 부분은 소프트웨어로 할 수 있는 음소분석방법에 의한 화자독립방식에 대해 연구하고자 한다.

음소분석방법은 음성신호의 음특성에 대한 구별과 분석에 기초한 방법으로서 초기에 음성신호를 분석하여 음소를 분리하는 상세 음소분석(Detailed Phonetic Analysis) 방식으로부터 음성신호를 모음, 비음, 마찰음 등과 같은 유사한 음들의 집단으로 나누어 구분하는 개략적 음소분석(Broad Phonetic Analysis) 방식으로 연구가 진행되고 있다[7], [8].

따라서 본 연구에서는 한글 음성입력에 대한 화자독립방식에서의 개략적 음소분석방법으로 음성인식을 하고자 한다. 본 연구의 특징을 기술하면 다음과 같다.

첫째, 8개의 음성명령어를 주명령으로 설정하여 이를 인식하게 한 후, 원하는 제어량을 나타내는 부명령을 인식하도록 하는 계층적 구조를 구성하고자 한다.

둘째, 음성신호의 음특성에 대한 구별과 분석에 기초한 음소분석방법을 이용한다.

세째, 음소분석에 의해 구성된 인식대상어휘로부터 음성식별을 하기 위해 영교차율(Zero Crossing Rate, ZCR), 대수 에너지(Log Energy, LE), 미분신호 영교차율(Up & Down, UD), 포만트 주파수(Formant Frequency) 등의 특징 파라미터를 추출한다.

네째, 음소분석에 의해 구분된 음집단들 사이의 특성을 식별할 수 있도록 특징 파라미터들의 범위의 조합으로 분류규칙(Classification rules)을 구성하고, 이를 통해 음성인식을 구현한다.

Ⅱ. 본 론

2. 1 음성인식의 과정

음성인식에 있어서 화자의 음성신호를 입력받아 이를 식별할 때까지의 과정을 단계별로 정리하면 그림 1과 같이 4단계로 나눌 수 있다.

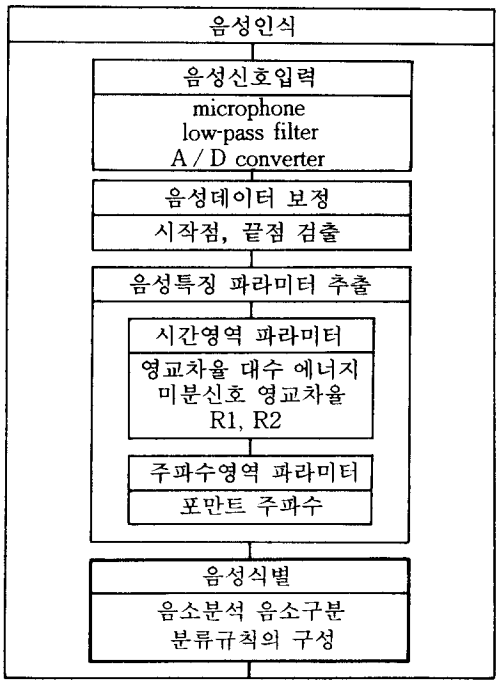

그림 1. 음성인식의 전체적 흐름도

첫번째 단계는 음성신호의 입력단계로서 화자가 마이크로폰에 음성을 입력하면 물리적 특성인 진동이 전기적인 회로에 의해 수치데이터 값으로 변하게 된다[9]. 두번째 단계는 음성데이터의 보정단계로서 얻어진 수치데이터로부터 순수한 음성신호만을 추출해내는 시작점, 끝점 검출을 행한다.

세번째 단계는 보정된 음성 수치데이터로부터 음성인식에 필요한 특징 파라미터를 추출하는 단계이다. 본 연구에서는 이러한 특징 파라미터로서 영교차율(ZCR), 대수 에너지(LE), 미분신호 영교차율(UD), R1, R2, 포만트 주파수 등을 사용한다. 네번째 단계는 추출된 음성 특징 파라미터들을 이용하여 인식대상어휘를 구분하기 위해 음소분석과 음소구분을 통해 분류규칙을 구성하게 된다. 이중선으로 되어 있는 부분이 최종적으로 음성식별을 하는 단계로서 가장 중요한 부분이다.

2. 1. 1 음성데이터 보정

음성입력자는 마이크로폰에 대고 발성을 함으로써 음성입력을 하는데 이 마이크로폰은 음성과 같은 소리를 전기신호의 파형으로 바꾸어 준다. 마이크로폰에 입력되어진 입력음성신호 중에서 음성분석에 필요하지 않는 환경잡음을 없애기 위해 저역통과여과기(Low-pass Filter)를 거치게 된다.

여과기를 거친 입력음성신호는 물리적으로 아날로그의 성질을 가지므로 음성분석을 위한 수치데이터로 바꿔주기 위해 A / D Converter를 사용하여 1 / 10초 동안에 2000개의 음성신호 수치데이터를 입력받아 파일에 저장할 수 있다.

2. 1. 2 음성데이터 보정

음성신호의 처리이전의 과정에서 가장 중요한 것은 주위 환경의 잡음으로부터 정확한 음성신호만을 추출해내는 일로서 이것은 음성신호에서 시작점과 끝점의 정확한 검출을 의미한다. 시작점과 끝점이 정확히 검출되면 음성신호의 데이터 처리량을 최소로 줄일 수 있고 정확한 처리를 할 수 있다. 격리단어인식에 있어서 시작점과 끝점 검출을 위해서는 음성신호를 크게 무음(Silence)과 음성, 그리고 다시 무음의 세 부분으로 나누게 되는데 양 무음사이의 음성을 검출하는 것이다[9].

본 연구에서는 시작점과 끝점의 검출을 위한 방법으로 단구간 에너지를 이용한다. 단구간 에너지는 수치데이터의 값을 제곱하게 되므로 큰 신호값에 민감하여 음성부분과 무음부분을 구분하는데 유용하다.

2. 1. 3 특징 파라미터의 추출

음성신호는 유성음, 무성음, 무음의 3가지로 나눌 수 있는데 이러한 음성신호를 분석하는 방법에는 시간영역과 주파수영역상에서 분석하는 방법이 있다. 시간영역에서의 분석을 통한 것에는 영교차율(ZCR), 대수 에너지(LE), 미분신호 영교차율(UD) 등이 있고, 주파수영역에서의 분석을 통한 것에는 포만트 주파수 추출이 있다.

영교차율은 분석구간내에서 파형이 영점축과 교차하는 횟수를 말하며, 디지탈 신호처리에 있어서 유성음과 무성음을 판별하는데 유용한 파라미터이다[10].

$$ZCR = \sum_{m=n-N+1}^{n} \{1 - \mathrm{sgn}\{X(m+1)\} \times \mathrm{sgn}\{X(m)\}\} / 2,$$

$$\mathrm{sgn}\{X(m)\} = \begin{cases} 1, & M(m) \geq 0 \\ -1, & M(m) < 0 \end{cases}$$

대수 에너지는 샘플값의 제곱의 합에 대수를 취해 준 것으로서 무음과 무성음을 구분하는데 유용하다[11].

$$LE = 10 \ \mathrm{Log} \ \{\sum_{m=n-N+1}^{n} X^2(m)\}$$

음성신호의 값의 차를 다시 샘플값으로 하여 그것의 부호변환에 대한 횟수를 미분신호 영교차율이라고 한다[11].

$$UD = 1/2 \sum_{m=n-N+1}^{n} \{\mathrm{sgn}\{S(m+1)-S(m)\} - \mathrm{sgn}\{S(m)-S(m-1)\}\}$$

또한, 음절의 구성에 있어서 자음을 구분하는데 있어서 좀더 정확하게 구분을 하기 위해서 R1(= ZCR / LE)과 R2(UD / ZCR)를 사용한다.

포만트 주파수의 추출을 위해서 현재 신호처리연구에서 가장 보편적으로 쓰이고 있는 고속 퓨리어 변환(Fast Fourier Transform, FFT)를 사용하여 음성신호를 변화시킨다[12].

$$F(u) = 1/2 \{Feven(u) + Fodd(u) \ W_{2m}^{u}\}$$

여기서, $Feven(u) = 1/M \sum_{x=0}^{M-1} f(2x) \ W_M^u$,

$Fodd(u) = a/M \sum_{x=0}^{M-1} f(2x+1) \ W_M^{ux}$

2. 1. 4 음성식별을 위한 분류규칙의 구성

음성의 최종적인 인식을 위한 단계인 음성의 식별과정은 음소분석, 음소구분, 분류규칙의 구성 등의 단계를 포함하고 있다. 음소분석단계에서는 본 연구의 인식대상 어휘인 로보트 명령어와 제어량을 나타내는 숫자음의 계층적 구조와 이들에 대해 한국어의 음운학적 규칙에 따라 음소분석을 행하게 된다.

음소구분단계에서는 음소분석에 의해 분류된 비음, 유음, 파열음, 마찰음, 파찰음 등을 구분할 수 있는 특징 파라미터들과 이들의 범위를 찾아내게 된다. 분류규칙의 구성에서는 앞의 단계에서 찾아진 특징 파라미터들의 범위를 규칙화하여 분류규칙을 구성하고, 입력음성으로부터 추출된 특징 파라미터의 값이 해당되는 범위를 찾아 해당 음성을 인식하게 된다. 여기에서는 본 연구에서 중요한 부분을 차지하고 있는 분류규칙의 구성에 대하여 살펴보고자 한다.

로보트의 기본적 동작을 표현하는 명령어와 그에 따르는 제어량을 나타내는 숫자를 인식하기 위한 분류규칙을 구성하는데는, 앞서 구해진 자음의 인식을 위해 구성된 파열음, 유음, 비음에 대한 특징 파라미터들의 범위가 활용된다. 또 자음의 구분을 통해 인식이 안될 경우에는 모음을 구분하기 위한 포만트 주파수를 사용한다.

먼저 인식대상으로 하는 음성의 첫 소리에 대한 구분을 행하고, 첫 음소로서 구분이 되지 않으면, 마지막 음소의 구분을 통하여 해당음을 인식하게 된다.

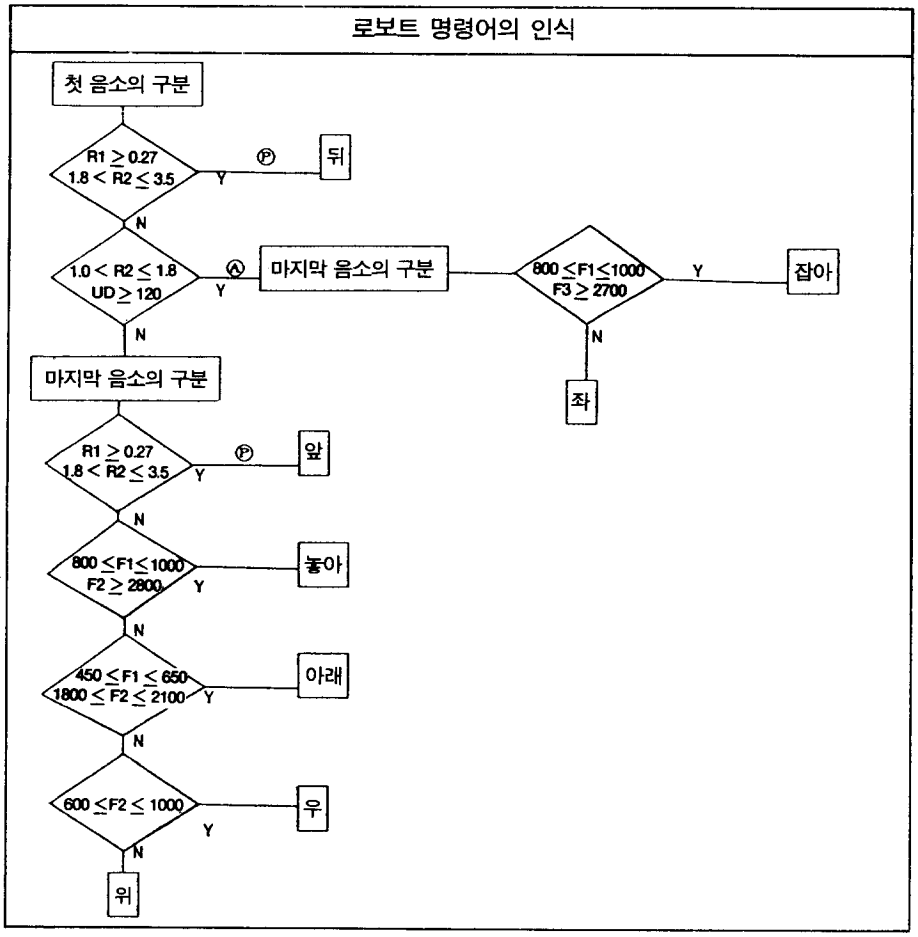

그림 2. 로보트 명령어의 인식 흐름도

(1) 로보트 명령어의 인식

그림 2에서와 같이 먼저 입력되는 음성의 첫 음소에 의한 구분을 행한다. 8개의 로보트 명령어 중에서 첫 음소가 파열음인 것은『뒤』하나 밖에 없으므로, 입력되는 음성의 첫 음소가 파열음의 특성을 가지면 곧바로『뒤』로 인식하게 된다.

첫 음소가 파열음이 아닐 때에는 파찰음의 특성을 가지게 되는가를 알아본다. 그런데 파찰음의 특성을 가지는 것은『좌』와『잡아』두개가 있으므로, 이들의 구분을 위해 포만트 주파수에 의한 마지막 음소의 구분을 행한다.

나머지 5개의 명령어에서는 첫 음소가 모두 비음이므로, 마지막 음소에 의한 구분을 행한다. 즉, 마지막 음소가 파열음의 특성을 가질 경우에는『앞』으로 인식하게 된다. 그리고『놓아』,『아래』,『우』,『위』는 마지막 음소가 모음으로 끝나게 되므로, 포만트 주파수를 통해 구분을 행하게 된다.

(2) 숫자음의 인식

그림 3의 인식과정을 보면 로보트 명령어의 인식과정에서와 같이 먼저 입력되는 음성의 첫 음소에 의한 구분을 행한다. 10개의 숫자음 중에서 첫 음소가 파찰음인 것은『칠』하나 밖에 없으므로, 입력되는 음성의 첫음소가 파찰음의 특성을 가지면『뒤』로 인식하게 된다.

첫 음소가 파찰음이 아닐 때에는 파열음의 특성을 가지게 되는 가를 알아본다. 그런데 파열음의 특성을 가지는 것은『구』와『팔』두 개가 있으므로, 이들의 구분을 위해 마지막 음소의 구분을 행한다. 여기서『팔』은 마지막 음소가 유음이므로, 유음의 특성을 가진 음은『팔』로 인식하게 되고 그렇지 않

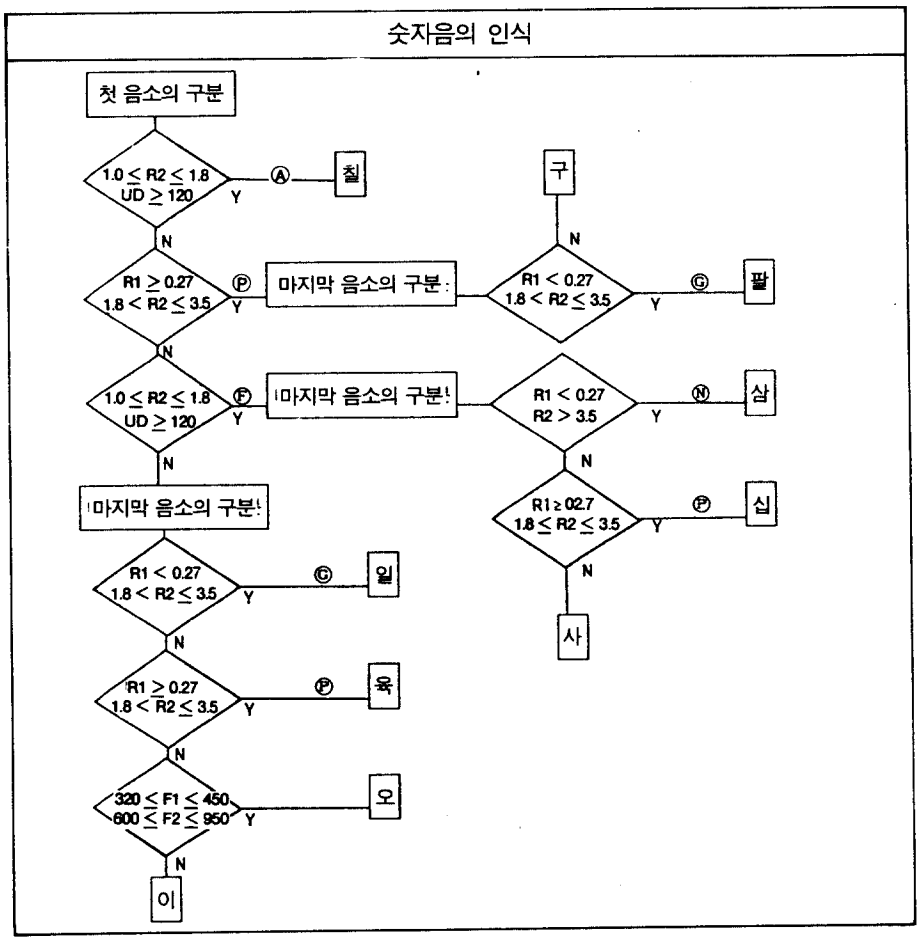

그림 3. 숫자음의 인식 흐름도

으면 『구』로 인식하게 된다.

입력음성의 첫 음소가 파찰음도 아니고 파열음도 아닐 때에는 마찰음의 특성을 가지는 가를 알아본다. 그런데 마찰음의 특성을 가지는 것에는 『삼』, 『사』가 있으므로, 이들의 구분을 위해 마지막 음소의 구분을 행한다. 여기서 『삼』은 마지막 음소가 비음이므로, 비음의 특성을 가진 음은 『삼』으로 인식하게 된다. 또 『십』은 마지막 음소가 파열음이므로 파열음의 특성을 가진 음은 『십』으로 인식하게 되고, 그렇지 않으면 『사』로 인식하게 된다. 나머지 『일』, 『이』, 『오』, 『육』 4개의 숫자음에서는 첫 음소가 모두 비음이므로, 마지막 음소에 의한 구분을 행한다.

즉, 마지막 음소가 유음의 특성을 가진 경우에는 『일』로 인식하게 되고, 마지막 음소가 파열음의 특성을 가질 경우에는 『육』으로 인식하게 된다. 그리고, 『오』와 『이』는 마지막 음소가 모음으로 끝나게 되므로 포만트 주파수를 통해 구분을 행하게 된다.

Ⅲ. 실 험

3. 1 실험방법 및 절차

본 연구에서 음성인식을 위한 하드웨어적인 장치로는 인간조작자의 음성을 입력하기 위한 마이크로폰과 환경잡음을 없애기 위한 저역통과여과기, 음성데이터를 수치데이터로 바꾸어 주는 A / D Converter, 그리고 수치데이터를 파일에 저장할 수 있는 IBM / PC-AT가 사용된다.

본 연구의 실험으로는 본 연구에서 제시한 인식방법을 이용하여, 이를 로보트 제어시스템에 활용하기 위해 로보트 명령어와 숫자음의 인식율에 관해 알아보기로 한다. 인식대상으로 하는 음성은 로보트의 기본적인 동작을 표현할 수 있는 로보트 명령어와 제어량을 나타내는 숫자음이다.

로보트 명령어로는 로보트의 미시적 동작제어시 움직이는 방향을 나타내는 『앞』, 『뒤』, 『위』, 『아래』, 『좌』, 『우』의 6가지와 로보트가 작업대상물체를 잡고 놓는 동작을 위한 『잡아』와 『놓아』의 명령을 합하여 8가지 명령을 인식대상으로 하였다. 숫자음으로는 『일』, 『이』, 『삼』, 『사』, 『오』, 『육』, 『칠』, 『팔』, 『구』, 『십』의 단독 숫자음을 인식대상으로 하였다.

음성인식의 실험을 위한 실험대상자로는 본 연구의 음성인식에 대해 기초지식이 있는 3인의 대학원생으로 선정하였는데 이는 피실험자간의 변동을 줄이기 위해 결정한 것이다. 각 피실험자는 실험환경에 익숙해질 때까지 예비실험을 거친 후에, 인식 대상음성을 통계적으로 안정된 값을 얻을 수 있도록 각각 10회씩 입력하여 실험한다.

위와 같은 방법으로 각 입력음에 대하여 인식을 제대로 한 경우와 제대로 하지 못한 경우를 조사하여 인식율을 계산한다. 여기에서 인식율은 음성입력횟수에 대해 정확히 인식한 횟수의 비를 말한다.

3. 2 실험결과 및 분석

본 연구에서는 음소분석에 의한 분류규칙의 구성을 통해 음성인식을 실험하였다. 이와 같은 방법으로 실시한 8개의 로보트 명령어에 대한 인식실험의 결과가 표 1에, 숫자음에 대한 인식실험 결과가 표

표 1. 로보트 명령어의 인식율 실험결과

입력음성 \ 출력음성	앞	뒤	위	아래	좌	우	잡아	놓아	인식율 (%)
앞	27		1	1			1		90.0
뒤		29						1	96.7
위		1	28			1			93.3
아 래	2			25			1	2	83.3
좌					28		1	1	93.3
우		1	1			27		1	90.0
잡 아	1				2		26	1	86.7
놓 아				2	1			27	90.0

표 2. 숫자음의 인식율 실험결과

입력음성 \ 출력음성	일	이	삼	사	오	육	칠	팔	구	십	인식율 (%)
일	27	2					1				90.0
이	1	28			1						93.3
삼			29	1							96.7
사			1	28						1	93.3
오		1			28	1					93.3
육					3	36			1		86.7
칠			1				27	1		1	90.0
팔							1	27	2		90.0
구					1			3	26		86.7
십			1	1						28	93.3

2에 나타나 있다.

표 1의 결과를 살펴 보면 8개의 로보트 명령어에 대해 평균 90% 이상의 인식율을 얻었다. 각 명령어에 대해 비교적 고른 인식율을 얻었지만, 『아래』, 『잡아』, 『놓아』와 같은 2음절의 명령어에서 낮은 인식율이 나왔다.

『아래』가 『앞』과 『놓아』로 잘못 인식된 것은 마지막 음소에서 포만트 주파수를 구하는 데, 모음 『ㅐ』의 추출이 용이하지 않았고, 『ㅐ』와 『ㅏ』의 구분이 미흡했기 때문이다. 『잡아』와 『놓아』의 경우에도 각각 모음 『ㅏ』와 『ㅘ』, 『ㅏ』와 『ㅐ』를 구분하기 위해서 사용된 포만트 주파수의 추출이 미흡했기 때문이다.

또한, 표 2의 결과를 살펴보면 평균 인식율이 90% 이상으로 각 숫자음의 인식율이 고르게 나타났으나, 『일』과 『이』, 『오』와 『육』, 그리고 『팔』과 『구』의 구분이 제대로 안되어 서로 잘못 인식을 하여 인식율이 낮아졌다. 『일』과 『이』의 구분에서는 『일』에서 유음인 『ㄹ』이 제대로 추출되지 못하여, 『이』와 비슷한 포만트 주파수를 가지게 되어 잘못 인식이 되었다.

그리고 『육』이 『오』로 잘못 인식된 것은 『육』의 종성인 『ㄱ』이 거의 묵음에 가깝고, 『오』와 『육』의 포만트 주파수의 범위가 비슷하기 때문이라고 생각된다. 『팔』과 『구』의 인식에서는 발음하는 사람에 따라 초성인 『ㅍ』과 『ㄱ』이 짧게 발음되므로 이를 구분하기가 힘들었기 때문이다.

이상의 로보트 명령어와 숫자음에 대한 실험결과에서 보는 바와 같이 숫자음의 인식율보다 로보트 명령어의 인식율이 떨어짐을 알 수 있다. 이것은 로보트 명령어의 인식에서 연결음으로 된 음성을 인식하는 것이 단음절로 된 음성을 인식할 때보다 음소추출이 용이하지 않았기 때문이다.

Ⅳ. 결 론

본 연구에서는 인간-로보트 시스템에서 로보트를 제어하기 위한 입력수단으로써 기존의 Teach Pendant나 키보드 대신에 가장 자연스러운 입력수단인 음성을 이용하였다. 로보트 제어시 필요로 하는 8개의 명령어와 10개의 숫자음을 인식대상어휘로 하고, 이들을 계층적으로 구성하였다. 즉, 로보트 명령어를 입력하여 인식하게 되면, 다시 숫자음을 입력하여 인식하게 되는 것이다.

이러한 로보트 명령어와 숫자음에 대한 음성식별을 실현하기 위해서, 음소분석에 의해 인식대상음성의 음소를 파열음, 마찰음, 파찰음, 비음, 유음, 모음 등의 비슷한 특성을 가지는 집단으로 나누었다. 이렇게 나누어진 음집단을 구분하기 위해 시간영역에서의 파라미터와 주파수영역에서의 파라미터를 추출하였고, 이를 통해 각각의 집단을 구분할 수 있는 파라미터와 이들의 범위를 찾아내었다. 여기서 찾아낸 파라미터들을 기초로 하여, 음성명령을 식별할 수 있는 분류규칙을 구성함으로써, 입력되는 음성명령어의 첫 음소와 마지막 음소만을 구분

하게 되어 음성식별이 가능하게 되었다.

실험결과 평균 90% 이상의 인식율을 얻을 수 있었는데, 연결음에서의 인식율이 단음절에서의 인식율보다 낮음을 알 수 있었다. 이는 연결음에서의 음소추출이 용이하지 않았기 때문이다.

본 연구에서는 작업자가 이해하기 쉬운 로보트 명령어와 숫자음을 통하여 로보트를 제어할 수 있는 한국어 음성명령어의 인식을 가능하게 함으로써, 본 연구에서의 결과가 앞으로 공장자동화를 위한 음성인식 시스템의 구축에 도움이 될 수 있을 것으로 본다. 앞으로의 연구방향으로는 연결단어나 연속어의 인식방법에 관한 것이 필요하며, 이를 실제 로보트와 연결하여 활용함으로써 로보트 동작제어를 위한 정보전달 수단에 효율적으로 적용하기 위한 방법의 연구도 진행되어야 할 것이다.

참 고 문 헌

〔1〕 Reising, J. M. and Curry, D. G., "A Comparison of Voice and Mutifunction Controls : Logic Design is the Key", Ergonomics, Vol. 30, No. 7, pp. 1063-1077, 1987.

〔2〕 오영환, 음화상 인간, 도서출판 기다리, 1988.

〔3〕 長町三生, 現代の人間工學わ朝倉書店, 1986.

〔4〕 이순요, 이창민, 박세권, "Human Interface 기술의 일환으로서 ASR과 Fuzzy Set Theory를 이용한 Developed-ERES / WCS 설계에 관한 연구", 한국 자동제어 학술회의 논문집, pp. 76-81, 1988.

〔5〕 Staugaard, A. C., Jr., Robotics and AI : An Introduction to Applied Machine Intelligence, Prentice-Hall, pp. 128-195, 1987.

〔6〕 Furui, S., "A Training Procedure for Isolated Word Recognition Systems", IEEE Trans. Acoust. Speech Signal Processing, vol. 28, No. 2, pp. 129-136, 1980.

〔7〕 Cole, R. A., "Phonetic Classification in New Generation Speech Recognition Systems", SPEECH TECH '86, International Space Business Review, pp. 43-46, 1986.

〔8〕 Helander, M., "Systems Design for Automated Speech Recognition" in Handbook of Human-Computer Interaction, Elsevier Science Publishers, pp. 301-319, 1988.

〔9〕 Lamel, L. F. and Rabiner, L. R., "An Improved Endpoint Detector for Isolated Word Recognition", IEEE Trans. Acoust. Speech Signal Processing, Vol. 29, No. 4, pp. 777-785, 1981.

〔10〕 Robort Ito, M. and Donaldson, R. W., "Zero-Crossing Measurement for Analysis and Recognition of Speech Sounds", IEEE Trans. Audio and Electroacoustics, Vol. 19, No. 3, pp. 235-242, 1971.

〔11〕 Wiitten, I. H., Principles of Computer Speech, Academic Press, 1985.

〔12〕 Oran Brigham, E., The Fast Fourier Transform and Its Application, Prentice-Hall, pp. 131-166, 1988.