

스마트 FA를 위한 음성인식 지능로봇제어에 관한 연구

A Study On Intelligent Robot Control Based On Voice Recognition For Smart FA

심현석^{1*}, 김민성¹, 최민혁¹, 배호영¹, 김희진¹, 김두범¹, 한성현²

H.S. Sim¹, M.S. Kim¹, M.H. Choi¹, H.Y. Bae¹, H.J. Kim¹, D.B. Kim¹, S.H. Han²

〈Abstract〉

This Study Propose A New Approach To Impliment A Intelligent Robot Control Based on Voice Recognition For Smart Factory Automation Since human usually communicate each other by voices, it is very convenient if voice is used to command humanoid robots or the other type robot system. A lot of researches has been performed about voice recognition systems for this purpose.

Hidden Markov Model is a robust statistical methodology for efficient voice recognition in noise environments. It has being tested in a wide range of applications. A prediction approach traditionally applied for the text compression and coding, Prediction by Partial Matching which is a finite-context statistical modeling technique and can predict the next characters based on the context, has shown a great potential in developing novel solutions to several language modeling problems in speech recognition.

It was illustrated the reliability of voice recognition by experiments for humanoid robot with 26 joints as the purpose of application to the manufacturing process.

Keywords : Smart FA, Intelligent Robot Control, Real Time Impliment, Voice Command

1* 정회원, 교신저자, 경남대학교 대학원 석사과정
E-mail: hs8388@dscast.co.kr
2 정회원, 경남대학교 기계공학부 교수

1* M.S. Course, Graduate School Kyungnam University.
2 Dept. of Mechanical Engineering, Kyungnam University.

1. 서론

첨단산업사회로 변화하면서 사람들은 어렵고, 힘들고, 더러운 환경에서의 근무를 기피하는 현상이 일어나고 있다. 그래서 사람을 대신하여 다양한 작업을 수행하는 무인화 기술이 매우 요구되고 있는 실정이다. 즉, 기계장치들이 발전함에 따라 보다 편하게 조작하기 위한 연구가 진행되고 있으며 주로 음성, 동작, 장애물회피 등을 통한 작업들이 많이 제안되고 있다.

음성인식 기술은 사람의 언어를 해석하여 적절한 행동을 수행할 수 있는 지능시스템을 만드는 것을 목적으로 한다. 이러한 기술은 정보 산업 분야, 디지털 통신분야, 가전분야, 멀티미디어 등에 적용 범위가 크게 확대되고 있으며, PCS, 장난감, 전화기, 인터넷, 컴퓨터, 자동차 주변 장치, navigation system 등 지속적으로 획기적인 제품들이 생산되고 있다[1,2].

따라서 로봇의 교시에 음성인식 기술을 활용한다면 사람에게 보다 친숙하면서 전문가가 아니더라도 로봇의 교시를 가능하게 하여, 전문 엔지니어 양성 및 프로그램 재개발에 따른 비용과 시간의 절감 효과를 가져 올 수 있다. 특히 열악한 환경에서의 로봇의 원격 조작 및 신체장애자를 위한 서비스를 제공할 수 있는 로봇으로도 이용될 수 있는 등 응용분야가 다양하다[3,4].

본 연구에서는 음성인식 기법을 이용하여 휴머노이드 로봇의 실시간 작업 동작제어에 대한 성능 실험을 수행하고자 한다. 먼저 음성인식알고리즘 개발을 위한 연구를 수행한다. 세부내용으로 백터 양자화, 훈련 백터의 분할, 백터 심볼 추출, 잡음 처리 방법, 끝점 검출 알고리즘, 특징 추출 알고리즘과 고성능 음성인식 알고리즘에 대하여 연구를 수행한다. 그 다음으로 음성인식시스템 구성에 관한 연구를 한다. 세부내용으로는 음성인식시스

템의 하드웨어 구성을 하여 음성인식시스템의 구현을 통한 훈련 인식단계를 분석하고 신경회로망 학습 인식단계를 분석에 관한 연구를 수행한다. 마지막으로 음성인식알고리즘 성능 실험을 수행하고 휴머노이드 로봇의 동작제어를 음성 인식을 통하여 작업명령을 위한 작업동작제어 실현에 대한 성능실험을 검증하였다.

2. 고성능 음성인식 알고리즘 개발

2.1 음성인식알고리즘 구성요소

HMM은 초기 상태 확률(initial state probability), 상태 천이 확률(state transition probability), 상태 관찰 확률(state observation probability)인 세 가지 파라미터로 구성된다. 이들을 간단히 모델 λ 로 표시하면 식 (2.1)과 같이 표현된다[5,6].

$$\lambda = (\Pi, A, B) \quad (2.1)$$

1) 초기상태 확률 $\Pi = \{\pi_i\}$

1부터 N까지 각각의 상태에서 t=1인 첫 번째 관찰이 일어날 확률을 나타낸다.

$$\pi_i = P(s_1 = i) \quad , \quad 1 \leq i \leq N \quad (2.2)$$

2) 상태천이 확률 행렬 $A = \{a_{ij}\}$

각 상태간의 천이할 확률을 나타낸다. 시간 t에서 상태 S_t 가 i일 때, 시간 t+1에서의 상태 S_{t+1} 이 j로 천이할 확률을 a_{ij} 로 나타내며, 이것은 행렬 A의 원소를 이룬다. 상태가 N개인 모델의 경우 $N \times N$ 행렬이 된다[7,8].

$$a_{ij} = \Pr(S_{t+1} = j | S_t = i) \quad (1 \leq i, j \leq N) \quad (2.3)$$

3) 상태관찰 확률 $B = \{b_i(O_t)\}$

각 상태에서 특정한 관찰이 일어날 확률들의 집합이다. 시간 t에서 상태 j일 때, O_t 가 관찰될 확률을 $b_i(O_t)$ 로 나타낸다.

$$b_i(O_t) = \Pr(O_t | S_t = j) \quad (1 \leq j \leq N) \quad (2.4)$$

2.2 음성인식 학습 알고리즘 분석

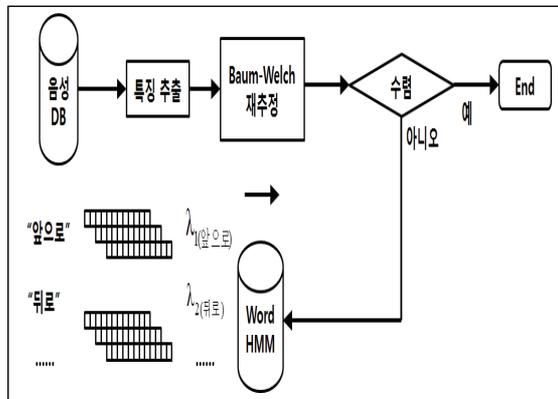


Fig. 1 Learning algorithm of HMM

HMM에 대하여 학습은 해당 음성 데이터가 발생할 확률을 최대로 하는 모델 파라미터 λ 를 추정하는 것이다. 이러한 문제는 가장 널리 사용되고 있는 Baum-Welch 알고리즘을 이용하여 해결할 수 있다[9,10].

2.3 음성인식 학습제어알고리즘 개발

음성구간 감지는 구현된 음성인식시스템에서 실시간으로 들어오는 음성을 유성음 및 무성음을 구별하는 단계로써, 만약 유성음이 감지되면 그 유

성음 각 프레임의 데이터 값을 절대에너지 방식으로 계산하여 설정된 크기보다 크면 음성을 간주하여 이 구간을 실제 음성으로 사용한다[11,12].

음성의 처리단위는 단구간 분석을 하는데, DMA (Direct Memory Access) 채널로 입력된 음성을 임시 저장하기 위해 버퍼 2개를 할당하고, 첫 번째 버퍼로 음성을 입력하는 중에 다른 버퍼에서는 에너지 계산을 하도록 하여 실시간으로 음성의 구간을 검출하도록 하였다[12,13].

본 연구에서는 사용된 프레임의 크기는 256샘플로 설정하였고, 프레임 중첩 간격은 80샘플로 설정하였다. 여기에서 음성인지 아닌지를 판단하는 절대 에너지 방식의 임계값은 실험 데이터를 통하여 획득된 값에 의해서 에너지 스펙트럼 값을 50000으로 설정하여 유성음 무성음을 판단하였다 [14,15].

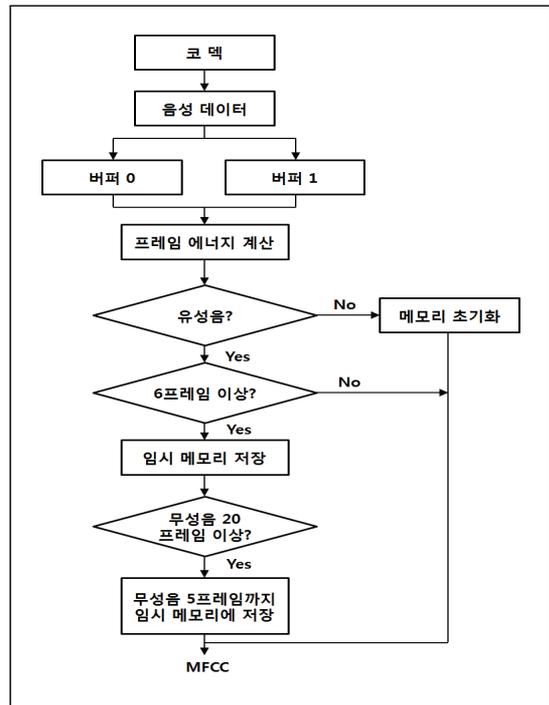


Fig. 2 Algorithm of endpoint detection

3. 성능실험

3.1 음성인식알고리즘

(1) 음성인식모델의 구조 및 특징

음성인식 방법은 Endpoint detection에서 음성의 시작점을 검출하면, 검출된 음성에 대해서 front end feature extraction에서 MFCC로 변환하고, MFCC를 이용하여 acoustic matcher에서 voice model과 acoustic score를 구한다. 그리고 그 score를 이용하여 인식 network을 parsing해서 인식된 결과가 나오면 rejection에서 음성의 신뢰도를 평가하게 된다. Fig. 3 Structure of Robot Model With 26 Joints 은 인식시스템의 개념적인 모델이다.



Fig. 3 Structure of Robot Model With 26 Joints

번호	시작	마지막	변위	포크	포트	인 이름	프레임수	실행시간[ms]
2	97	94	-3	0	PO.0	Center	8	120
3	155	156	1	0	PO.0	Center	8	130
4	92	95	3	0	PO.0	R_RoL	10	320
5	162	161	-1	0	PO.0	R_RoL	1	20
6	144	145	1	0	PO.0	LU	10	90
7	110	112	2	0	PO.0	Center	8	120
8	128	129	3	0	PO.0	Center	8	130
9	128	132	4	0	PO.0	L_RoL	10	320
10	128	128	0	4	PO.0	L_RoL	1	20
11	124	126	2	4	PO.0	RU	10	90
12	49	58	9	4	PO.0	Center	8	120
13	212	216	4	4	PO.0	Center	8	130
14	139	137	-2	4	PO.0	R_RoL	10	320
15	117	117	0	4	PO.0	R_RoL	1	20
16	127	127	0	4	PO.0	LU	10	90
17	127	127	0	4	PO.0	Center	8	120
18	53	39	-14	4	PO.0	Center	8	130
19	197	197	0	4	PO.0	L_RoL	10	320
20	127	127	0	4	PO.0	L_RoL	1	20

Fig. 4 Motion control program of robot

(2) 성능 실험 결과

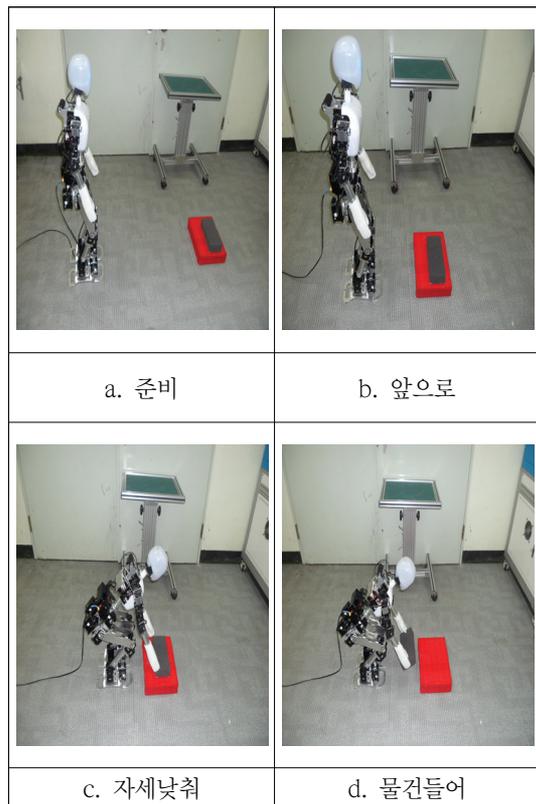


Fig. 5 Performance experiment scene of voice recognition I

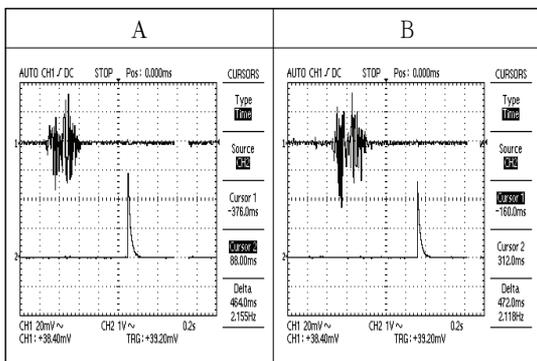


Fig. 6 Voice recognition result graph about "Junbi"

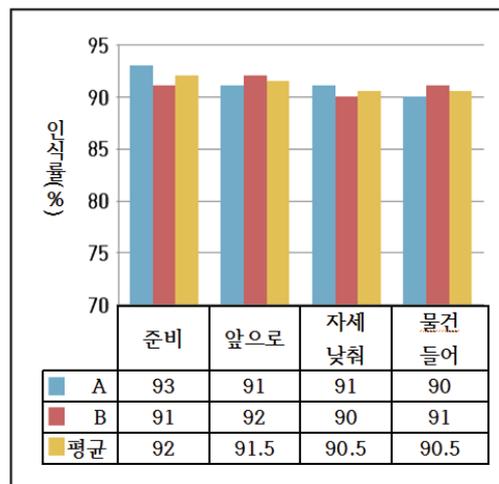


Fig. 9 Voice recognition experiment result

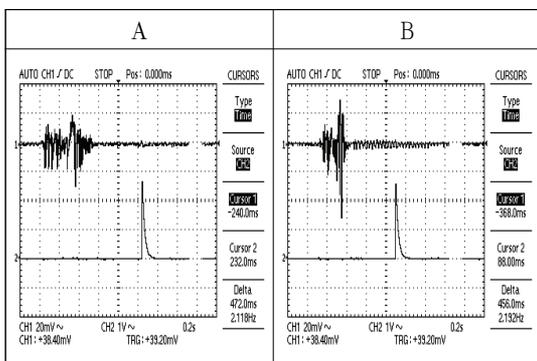


Fig. 7 Voice recognition result graph about "Apuro"

표 54

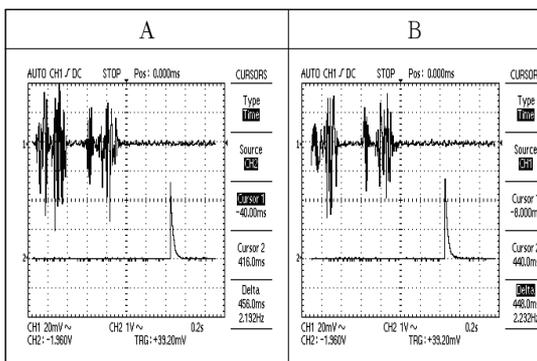


Fig. 8 Voice recognition result graph about "Jasaenatchou"

3. 결론

본 연구에서는 음성인식 기법을 이용한 휴머노이드 로봇의 실시간 작업 동작제어에 관한 연구를 수행하였다. 구현된 음성인식시스템은 통계적 확률방법을 이용하여 인식을 수행 할 경우의 오인식을 최소화하기 위한 후처리 과정으로 신경회로망의 일종인 방사 기저함수 신경회로망을 결합하여 보다 높은 인식률을 얻을 수 있는 통계적 확률방법과 방사 기저함수 신경회로망의 시스템을 적용하였다.

음성인식알고리즘의 인식실험 및 실제 로봇에 적용하여 성능평가를 통해 다음과 같은 결과를 얻었다.

첫 번째로 임의의 환경에서 음성에 의한 보행 로봇의 경로생성이 가능한 통계적 확률방법을 이용한 음성인식알고리즘을 이용하여, 인식률 향상을 위한 방안으로 통계적 확률방법에 의한 인식을 수행한 후 그 신뢰도를 적절히 판단하여 신뢰도 낮은 경우 신경회로망에 의한 재인식 과정을 후처

리에 도입하였다.

두 번째로 로봇적용 단어를 대상으로 실제 보행로봇에 적용하여 음성교시에 의한 경로생성 실험을 수행하여 성능평가를 하였다. 설정된 환경에서 로봇은 전체 인식단어와 로봇적용 단어에 대해 80%이상의 인식률을 얻을 수 있었다.

후 기

본 연구는 로봇융합원천기술실용화개발사업의 일환으로 연구되었음.

참고문헌

- [1] H. J. M. Steeneken and D. A. van Leenwen, "MultiLingual Assessment of speaker independent large vocabulary speech-recognition systems: SQUALE Project," Proc. EUROSPEECH-95, pp. 1271-1274
- [2] T. Matsuoka "Large-vocabulary continuous-speech recognition using Japanese business newspaper (Nikkei)" DARPA Speech Recognition Workshop, Feb. 1997, pp. 137 - 142.
- [3] 구명완, "음성인식기술 현황과 전망," 전자공학회지, 제 20권 제 5호, pp.548-555, 1993.
- [4] J. T. Spooner and K. M. Passino, "Stable adaptive control using fuzzy systems and neural networks," IEEE Trans. Fuzzy Syst., vol. 4, pp. 339-359, Aug. 1996.
- [5] 이건상, 양성일, 권영현 "음성인식(Speech Recognition)", 한양대학교 출판부, 2001
- [6] L.-X. Wang, A Course in Fuzzy Systems and Control. Upper Saddle River, NJ: Prentice-Hall, 1997.
- [7] Jang, Sun, and Mizutani, Neuro-Fuzzy and Soft Computing, Prentice-Hall, 1997.
- [8] 김정훈, "음성인식처리용 임베디드 시스템의 설계 및 구현에 관한 연구", 한국해양대학교 공학 석사학위 논문, 2000
- [9] 김성규, "음성명령에 의한 로봇의 구동," 충남대학교 석사학위논문, 2011
- [10] L. R. Rabinar and B. H. Juang, "An Introduction to Hidden Markov Models," IEEE ASSP Magazine, Vol. 3, No. 1, pp 4-16, January 1986
- [11] M.T Jhonson, "Capacity and Complexity of HMM Duration Modeling Techniques," IEEE Signal Processing Letters, Vol. 12. 2, pp. 407-410, 2005
- [12] 양준석, 구영목, 조상영, 원종범, 원종대, 한성현, "A Study on Design and Control of Flexible Gripper System for Handling Working of the Forging Process in Heat Resisting Environment", 한국산업융합학회 논문집 제 18 권 제 4호, pp.216-223, 2015
- [13] Seong-Su Lee, Yong-Wook Kim, Hun Oh, Wal-Seo Park "Implementation of Self-adaptive System using the Algorithm of Neural Network Learning Gain," International Journal of Control, Automation, and Systems, vol.6, no.3, pp. 453-459, June 2008. Specification,
- [14] D.Psalitis, A. sideris, and A. Yamamura, "AMultilayered Neural Network Controller", IEEE Control Systems Magazine, pp.17-21 1988
- [15] 정양근, 조상영, 양준석, 한성현, "화자독립방식에 의한 음성인식 알고리즘 개발 및 실시간 실현에 관한 연구", 한국산업융합학회 논문집 제 18권 제 4호, pp.250-258, 2015
- [16] 정양근, 심현석, 이우송, 한성현, "영상 피드백에 의한 4축 스카라 로봇의 실시간 궤적추적제어에 관한 연구", 한국산업융합학회 논문집 제 17권 제 3호, pp.136-144, 2014
- [17] Lawrence Rabiner, "A Tutorial on Hidden Markov Models and Selected Application in Speech Recognition", Proc. IEEE, Vol 77, No. 2, february 1989.
- [18] 조현수 "HMM을 기반으로 한 자율이동로봇의 음성명령 인식시스템의 개발", 부산대학교 공학

- 석사 학위논문, 2001
- [19] 황원준, 박인만, 강연욱, 한성현, “초음파센서 기반 2휠구동로봇의 실시간 자율 주행제어에 관한 연구”, 한국산업융합학회 논문집 제 17권 제 3호, pp.151-169, 2014
- [20] 심병균, 정양근, 박인만, 황원준, 강연욱, 한성현, “로봇 핸드핑거의 파지제어에 관한 연구”, 한국산업융합학회 논문집 제 16권 제 4호, pp.141-146, 2013
- [21] P. Renevey, R. Vetter, and J. Kraus, “Robust speech recognition using missing feature theory and vector quantization,” in Proc. Of Eurospeech-2001. ESCA, 2001, pp. 1107-1110.
- [22] 임창욱, “연속 HMM을 이용한 우리말 음성 인식에 관한 연구”, 호서대학교 공학석사학위논문, 2005
- [23] Seong-Su, Lee, “Hybrid Neural Network Controller Design of DSP for 3-Phase Induction Motor of Real-time Control,” Wonkwang University, Master’s thesis, pp.20-27, 2004
- [24] S. Yamamoto, K. Nakadai, H. Tsujino, and H. Okuno, “Assessment of general applicability of robot audition system by recognizing three simultaneous speeches.” in Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-2004). IEEE and RSJ, 2001, p. to appear.
- [25] 정우근, “음성인식을 이용한 이동로봇의 제어”, 성균관대학교 석사학위논문, 2002
- [26] 정양근, 김민성, 조상영, 원종범, 원종대, 한성현, “고열 단조부품 핸들링을 위한 로봇 그리퍼의 방위 및 포지션 정밀제어에 관한 연구”, 한국산업융합학회 논문집 제 17권 제 3호, pp.151-169, 2014
- [27] L.R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Application in Speech Recognition," Proc. of the IEEE, Vol. 77, No.2, pp. 257-285, February 1989.
- [28] 강성인, “하이브리드 음성인식 기법을 이용한 인공지능형 이동로봇 시스템의 설계 및 구현,” 해양대학교 공학박사 학위논문, 2004.
- [29] 박진영, “HMM을 이용한 화자인증 시스템의 구현”, 동아대학교 공학석사 학위논문, 2004
- [30] X. D. Huang, Y. Ariki, and M. A. Jack, hidden markov models for speech Recognition, Edinburgh University Press, Edinburgh, England, 1990.
- [31] Myers C. S. and Rabiner L. R., "A comparative study of several Dynamic Time-Warping Algorithm for Connected-Word Recognition,"Bell system Tech. J., 60(7) pp 1389-1409, September 1981.
- [32] Siverman H. F., and Morgan D. P, "The application of dynamic programming to connected speech recognition," IEEE Transaction on Acoustics, Speech, and Signal Processing Magazine, Vol. 7, pp. 6-25, July 1990.
- [33] L.R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Application in Speech Recognition," Proc. of the IEEE, Vol. 77, No. 2, pp. 257-285, February 1989.

(접수: 2018.01.18. 수정: 2018.02.20. 게재확정: 2018.03.07.)