

MFCC-HMM-GMM을 이용한 근전도(EMG)신호 패턴인식의 성능 개선

최흥호^{1,2}, 김정호³, 권장우³

¹인제대학교 의용공학과 의용초음파 연구실

²인제대학교 BK21 바이오 조직 재생 사업팀

³동명대학교 컴퓨터공학과 인공지능 연구실

(Received August 3, 2006. October 16, 2006)

Performance Improvement of EMG Pattern Recognition Using MFCC-HMM-GMM

Heung-ho Choi^{1,2}, Jung-Ho Kim³, Jang-Woo Kwon³

¹Dept. of Biomedical Engineering, College of Bioscience Engineering, Inje University, Obang-dong, Gimhae, Korea

²BK21 Bio-Organ Tissue Regeneration Project Team, Inje University, Gimhae 621-749, Korea (Rep)

³Dept. of Computer Engineering, College of IT, Tongmyong University, Youngdang-dong, Busan, Korea

Abstract

This study proposes an approach to the performance improvement of EMG(Electromyogram) pattern recognition. MFCC(Mel-Frequency Cepstral Coefficients)'s approach is molded after the characteristics of the human hearing organ. While it supplies the most typical feature in frequency domain, it should be reorganized to detect the features in EMG signal. And the dynamic aspects of EMG are important for a task, such as a continuous prosthetic control or various time length EMG signal recognition, which have not been successfully mastered by the most approaches. Thus, this paper proposes reorganized MFCC and HMM-GMM, which is adaptable for the dynamic features of the signal. Moreover, it requires an analysis on the most suitable system setting for EMG pattern recognition. To meet the requirement, this study balanced the recognition-rate against the error-rates produced by the various settings when learning based on the EMG data for each motion.

Key words : EMG(Electromyogram), MFCC(Mel-Frequency Cepstral Coefficients), HMM(Hidden Markov Models), GMM(Gaussian Mixture Models), pattern recognition

1. 서 론

의수(Prosthetic Arm)는 결손된 손이나 팔을 대신해 실생활에서의 그 기능을 보조하는 역할을 수행한다.

의수 제어를 위해서 근전신호(Electromyogram, EMG)를 분석해 의수의 제어에 이용하는 것이 가장 보편적이며, 일반적이다. 표면 근전도(sEMG)는 피부표면에 전극을 부착하는 무통증, 비침습적인 방식이다. 하지만 실제 팔이나 손의 동작 패턴에 대한 근전신호의 내부에는 측정기기의 기계적 잡음, 근육수축시 표면전극에서 입력되는 외부잡음, 주위환경에서 발생된 잡음 등이 존재한다[1].

그러한 잡음을 실시간으로 예측하여 제거하기가 어렵다는 문제점이 있고, 만약 제거를 수행하여도 원 신호를 손상시켜서 근전신호의 고유성을 저해하는 역할이 될 수 있다. 따라서 각 동작 패턴마다 고유한 특징을 검출함에 있어 상당한 어려움이 따른다. 이는 인식 오차율 증가의 주요 원인이 된다. 또한, 각 개인의 근전신호 형태가 매우 다양하고, 동일인이라도 근의 분포상태, 전극의 부착 위치에 따라서 근전신호가 다양하게 출력된다. 따라서 정확한 패턴 분류에도 상당한 악영향을 미친다.

기존의 근전도 패턴 인식을 위한 특징 추출 과정에서는 주로 LPCC(Linear Predictive Cepstral Coefficients), 영교차율(Zero-Crossing), 절대적분치(Integral of Absolute Values, IAV), AR(Auto-regressive), Spike(극과), 근육의 피로도 등이 다양하게 적용되고 있으며 이들 비선형 시스템의 시간영역에서 근전신호를 주로 분석하기 때문에 잡음이 내포될 경우 인식률의 저하를 초래한다[2]. 이러한 문제 때문에 주파수 영역에서 특징을 검출하는 방

본 연구는 산업자원부(경상남도) 지역산업기술개발사업 (과제번호:11024335) 지원으로 수행되었음.

Corresponding Author : 최 흥 호
(621-749) 경남 김해시 어방동 607, 인제대학교 의용초음파 연구실
Tel : 055-320-3294 / Fax : 055-329-3294
E-mail : hhchoi@bme.inje.ac.kr

법과 관련해서 본 논문은 근전신호의 시간적 변화와 잡음에도 높은 인식을 향상을 보증하면서 동적인 의수제어를 위해 시간영역의 특징 추출기법보다 상대적으로 잡음에 강한 MFCC(Mel-Frequency Cepstral Coefficients) 특징 추출 기법을 제안한다. 음성 특징 추출에 주로 적용되는 이 기법은 인간의 청각 특성을 모델링한 것으로 주파수 별로 필터 बैं크를 구성해서 인식하는 방법이다. 음성 인식 특징 추출에 월등한 성능을 나타내고 있으며, 따라서 음성 인식에 쓰이는 이 기법을 근전도 특징 추출을 위해 재구성 할 필요가 있다.

이러한 근전신호의 특징을 분류하기 위한 분류기로 인공신경회로망인 MLP(Multilayer Perceptron), SOFM (Self-Organizing Feature Map)이다. 이들 분류기는 정적인 신호의 인식에는 뛰어나지만 근전 신호나 음성 신호처럼 시간영역에서 시간적으로 변하는 동적인 신호의 처리에는 적합하지 못하다는 단점을 가지고 있다. 근전도 신호들은 시변성이 강한 일종의 불규칙 신호라 볼 수 있으므로 확률 모델의 적용이 적합하다. 확률 모델(Stochastic Model)은 시변적 신호의 패턴 인식에 있어 효율적이며, 이러한 확률 모델들 중 대표적인 예는 마르코브 상태 천이 함수에 그 수학적 기초를 두고 있는 Markov model과 HMM(Hidden Markov Model)이다. 통계적 확률분포에 기인하는 HMM(Hidden Markov Models) 등을 단일 또는 혼합(Hybrid)하여 패턴 인식시스템을 구성한다. 일반적으로 HMM은 음성 연구 분야에서 활발한 연구가 진행되어 오고 있으며[3] 신경회로망과의 결합을 통하여 상호간의 장점을 증대시키기 위한 시도가 이루어져 오고 있다[4]. 그러나 생체 신호가 가지고 있는 특성이 다른 분야에서 제안하고 있는 합성 인식기에 적합하지 못하므로, 생체 신호의 특성을 잘 반영하는 인식기의 구성이 필요하다.

따라서 본 논문은 기존의 특징 추출 기법을 재구성한 MFCC를 이용해서 동작마다 근전도 신호의 특징을 추출하여 HMM-GMM으로 학습된 근전도 패턴 인식 시스템을 제안한다.

II. 본 론

A. EMG 특징 추출

연속적인 근전신호 분석을 수행하기 위해 먼저 근전도의 구간 검출이 먼저 수행되어야 한다. 각 프레임 별로 IAV 또는 Zero-Crossing을 통해 얻어진 근전신호의 에너지 값을 사전에 정의한 적절한 임계치를 적용하여 연속신호 내의 근전신호를 검출한다 [5][6].

$$IAV = \frac{1}{N} \sum_{i=1}^N |x1_i| + \frac{1}{M} \sum_{i=1}^M |x2_i| \quad (1)$$

근전신호의 특징 추출을 위해서 음성 특징 추출에 주로 적용되는 MFCC를 적용한다. 이 기법은 인간의 청각 특성을 모델링한 것으로 주파수 별로 필터 बैं크를 구성해서 인식하는 방법이다[7][8].

일반적인 MFCC 추출 방법은 먼저 해밍 윈도우를 씌워서 블록

단위의 프레임으로 나눈 뒤에 이후부터의 처리는 모두 프레임 단위로 이루어진다.

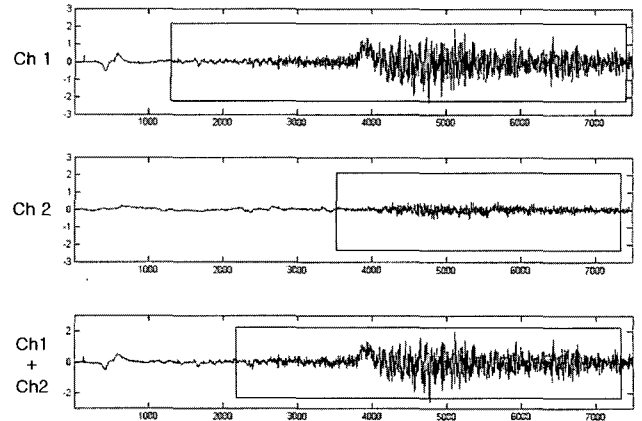


그림 1. 근전신호 구간 추출
Fig. 1. Detection the section of EMG

$$x[n], \quad n = 0, 1, \dots, N-1 \quad (2)$$

한 프레임의 음성신호는 FFT를 이용하여 주파수 영역으로 변환된다.

$$X_a[k] = \sum_{n=0}^{N-1} x[n] e^{-\frac{j2\pi nk}{N}} \quad (3)$$

주파수 대역을 여러개의 필터 बैं크로 나누고 각 बैं크에서의 에너지를 구한다.

$$H_m[k] = \begin{cases} \frac{2(k-f[m-1])}{(f[m+1]-f[m-1])(f[m]-f[m-1])} & k < f[m-1] \\ \frac{2(f[m+1]-k)}{(f[m+1]-f[m-1])(f[m+1]-f[m])} & \text{for } f[m-1] \leq k \leq f[m] \\ 0 & k > f[m+1] \end{cases} \quad (4)$$

다음과 같이 밴드 에너지에 로그를 취한 후

$$S[m] = \ln \left[\sum_{k=0}^{N-1} |X_a[k]|^2 H_m[k] \right] \quad (5)$$

이산 코사인 변환(Discrete Cosine Transform, DCT) 과정을 다음과 같이 수행한다.

$$c[n] = \sum_{m=0}^{M-1} S[m] \cos(\pi n(m-1/2)/M) \quad (6)$$

최종적으로 MFCC의 특징 계수 $c[n]$ 이 얻어진다. 이러한 계수 데이터를 다시 Delta MFCC($\Delta c[n]$) 또는 Delta-Delata MFCC

$(\Delta\Delta c[n])$ 를 얻기 위해 식(7)을 수행한다.

$$\begin{aligned} \Delta c[n] &= c[n+1] - c[n] \\ \Delta\Delta c[n] &= \Delta c[n+1] - \Delta c[n] \end{aligned} \quad (7)$$

여기서 기존의 MFCC 필터 बैं크 타입을 직접적으로 적용하기에는 적절하지 못하다. 근전신호는 음성신호보다 상대적으로 낮은 주파수 대역[10Hz~500Hz]을 지니고 있다. 따라서 기존의 MFCC와는 다른 다양한 타입의 필터 बैं크로 재구성이 요구되며, 채널 갯수 또한 다양하게 변경하면서 근전도 인식에 가장 적합한 구성을 취할 필요가 있다.

그림 2와 그림 3은 재구성된 필터 बैं크 타입과 적절한 채널 갯수를 적용하여 구성한 필터 बैं크이다. 그림 2은 삼각 필터 बैं크이며, Type 1은 고주파 대역으로 갈수록 필터 बैं크의 크기가 증가되는 형태이며, 필터들마다 오버랩(overlap)형태를 유지하고 있다. Type 2는 Type 1과 비슷하며, 단지 오버랩형태가 아니다. Type 3은 주파수 대역이 일정한 필터 बैं크이며, 오버랩(overlap) 형태이다. Type 4는 Type 3과 비슷하지만 오버랩 형태가 아니다. 그리고 필터는 점점 증가하다가 다시 감소하는 형태를 유지하면서 최고일 때 가중치는 1, 최저일 때 가중치는 0으로 설정한다.

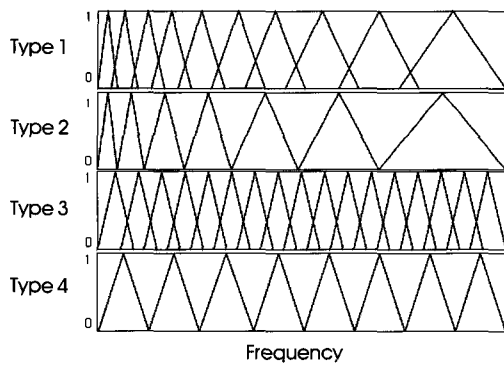


그림 2. 삼각 필터 बैं크
Fig. 2. Triangular filter banks

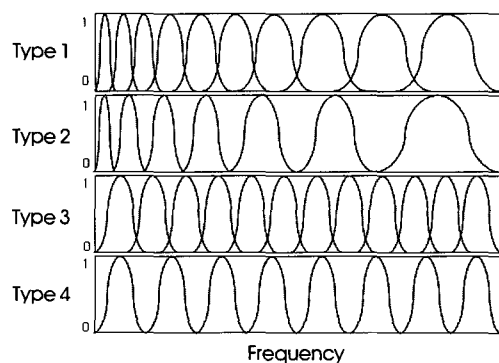


그림 3. Hanning 필터 बैं크
Fig. 3. Hanning filter banks

MFCC에서 추출된 특징 벡터들은 K-means와 LBG 벡터 양자화를 적용해서 클러스터로 구성된다.

MFCC 특징벡터 집합 $[x_1, \dots, x_N]$ 으로부터 K 개의 중심집합 $[y_1, \dots, y_K]$ 을 만들어서 x_n 이 y_k 에 가까울 경우 클러스터 X_k 에 라벨링하고, 전체 집합에 대한 K 개의 클러스터 $\{X_1, \dots, X_K\}$ 로 구성한다.

새로운 클러스터들의 각각의 중심을 다음처럼 갱신한다.

$$y_i = c(X_i), i = 1, \dots, K \quad (8)$$

다시 데이터와 가장 거리가 가까운 클러스터 중심과의 총 왜곡 (Distortion)을 구한다.

$$D = \sum_{n=1}^N d(x_n, y_{i(n)}), i(n) = k, \text{ if } x_n \in X_k \quad (9)$$

왜곡이 안정적인지를 확인하기 위해서는 다음과 같은 상대적인 감소치를 체크하면 된다. ΔD 는 왜곡율을 나타내며, 임계치 이하이면 종료한다.

$$\Delta D = \frac{D_{prev} - D_{curr}}{D_{prev}} < 10^{-4} \quad (10)$$

B. EMG 패턴인식 시스템 구성

근전신호를 분류하기 위한 분류기로 인공신경회로망인 MLP (Multilayer Perceptron), SOFM (Self-Organizing Feature Map) 이 주로 적용되어진다. 그림 4은 HMM의 동적인 신호에 대한 적용 특성과 학습속도의 향상을 위해 MLP의 탁월한 분류력을 결합시킨 형태이다. 이 구조는 1단의 HMM과 2단의 MLP를 직렬 연결한 형태를 갖는다. 별도의 MLP를 두고 이를 세분하여 훈련시킴으로써 오인식율을 줄이려는 의도이다[5][6].

훈련 과정에서는 채널 1과 2를 통해서 입력된 데이터에서 추출한 LPCC계수가 벡터 양자화된 후 기본 동작별로 6개의 HMM을 생성한다.

하지만 그림 4와 같은 분류기는 정적인 신호의 인식에는 뛰어나지만 근전 신호나 음성 신호처럼 시간영역에서 시간적으로 변하는 동적인 신호의 처리에는 상대적으로 적합하지 못하다는 단점을 가지고 있다.

그림 5와 같이 실시간으로 각 채널마다 입력으로 들어오는 근전 신호를 분석하기 위해 한 패턴 당 2채널을 동시에 분석한 후 추출된 MFCC 특징 벡터를 HMM 기법을 이용해 각 패턴에 대한 HMM을 구성한다[9]. 동일한 동작이라고 하여도 2개의 채널의 전극이 부착된 부위는 상이하므로 전혀 다른 신호가 입력된다[14]. 한 동작 패턴에 대해 2개의 채널에서 근전도 신호가 받으므로 1쌍의 HMMs이 구성되며, K-means와 LBG 알고리즘을 통해서 양자화를 거쳐 클러스터를 구성하고, Baum-Welch 알고리

층을 통해 총 6쌍(12 HMMs)의 모델이 구성된다. 이런 형태는 다른 분류기보다 상대적인 인식률 증가를 보여줄 것으로 판단된다.

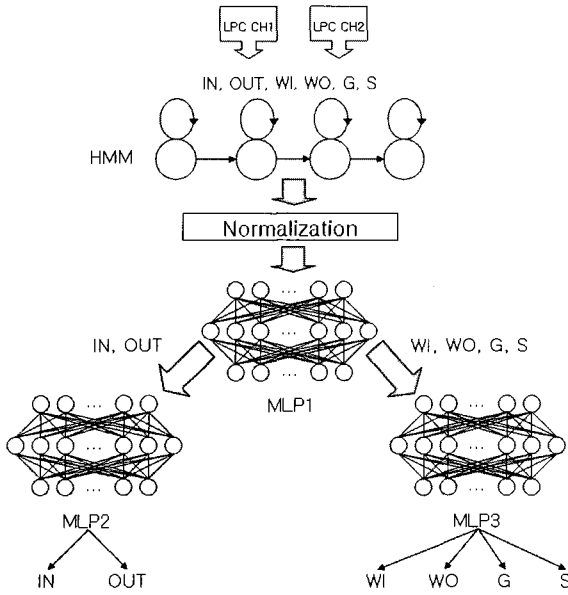


그림 4. MLP와 HMM 합성 분류기
Fig. 4. Composition Classifier of MLP and HMM

이렇게 학습된 HMMs 시스템은 다시 실시간으로 입력되어 들어오는 근전신호를 분석하게 되고, 특징을 추출하여 Viterbi 알고리즘을 통해서 결과치를 출력하게 되고, 정규화 과정을 수행하게

된다. 패턴당 1쌍에 출력되는 결과치를 결합하여 총 6개의 결과치를 출력하게 되고, 이들 값을 비교하여 최고치를 유지하는 모델을 인식모델로 선택하게 된다.

다음은 6개의 패턴에 대한 각 HMM 모델을 구성하는 과정이다. 모델 파라미터 λ 가 주어졌을 때 관측 심볼열 O 에 대한 확률, $P(O|\lambda)$ 는 전향 과정(Forward procedure)알고리즘을 이용한다.

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i) \tag{11}$$

여기서 식(12)는 초기 확률 π 와 처음 출력확률로 초기화를 한다.

$$\alpha_1(i) = \pi_i b_i(o_1) \quad 1 \leq i \leq N \tag{12}$$

시간 t 에서 $t+1$ 로 천이 할 때, $t+1$ 에서 상태 j 가 될 확률은 식(13)과 같다.

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(o_{t+1}) \tag{13}$$

$$1 \leq t \leq T-1, \quad 1 \leq j \leq N$$

N 개의 모델 각각에 대한 $P(O|\lambda_N)$ 는 다음과 같이 나타낼 수 있다.

$$B \triangleq [P(O|\lambda_1), P(O|\lambda_2), \dots, P(O|\lambda_N)] \tag{14}$$

$$a = \frac{\sum_{i=1}^N P(O|\lambda_i)}{N} \tag{15}$$

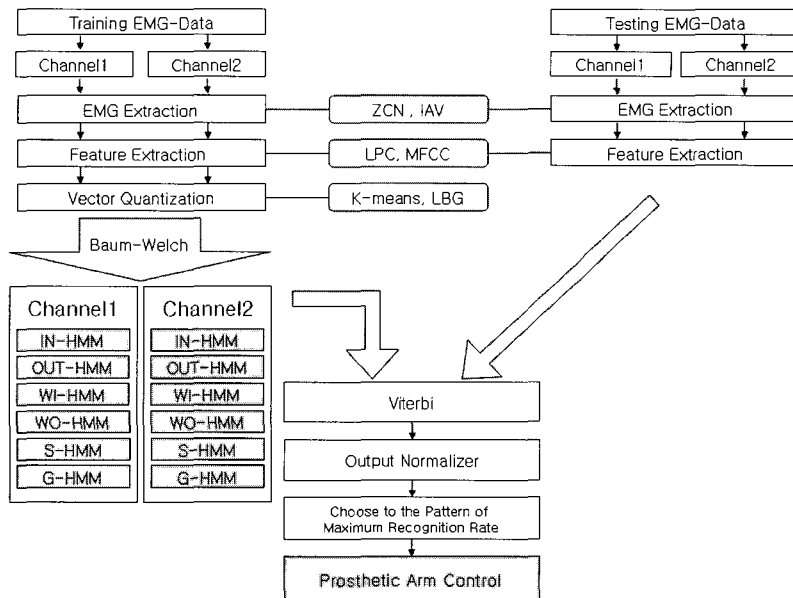


그림 5. HMM-GMM 인식 시스템 구성도
Fig. 5. Architecture of a HMM-GMM System

$$C = a/B = [a/P(O|\lambda_1), a/P(O|\lambda_2), \dots, a/P(O|\lambda_N)] \quad (16)$$

여기서 a 는 가중치이며, $C = c_1, c_2, \dots, c_N$ 이다.

가. HMM-MLP 구성

식(16)로 HMM의 출력치는 주어지고 C 를 입력으로 하는 MLP의 은닉층의 출력 값은 다음과 같이 주어지며

$$y_i = f(v_j^t c), \text{ for } j=1, 2, \dots, J \quad (17)$$

여기서 v_j 는 V 의 j 번째 행에 해당하는 열벡터이다. 출력층의 출력은 식(18)처럼 주어진다.

$$o_k = f(w_k^t y), \text{ for } k=1, 2, \dots, K \quad (18)$$

여기서 w_k 는 W 의 k 번째 행에 해당하는 열벡터이다. 따라서 쌍극성 연속 함수인 경우 출력층에서의 오차 신호항은 식(19)로 은닉층에서의 오차 신호항은 식(20)로 주어진다.

$$\delta_{ok} = \frac{1}{2}(d_k - o_k)(1 - o_k^2), \text{ for } k=1, 2, \dots, K \quad (19)$$

$$\delta_{yj} = \frac{1}{2}(1 - y_j^2) \sum_{k=1}^K \delta_{ok} w_{kj}, \text{ for } j=1, 2, \dots, J \quad (20)$$

나. HMM-GMM 구성

그림 6은 HMM의 각 모델은 좌우진행형(left-to right) 모델을 사용하고 근전도 신호에 의한 근육의 수축이 발화 상태(firing state)-수축의 과도기 상태(contraction transition state)-수축의 과도기 유지 상태(contraction transition state)-수축의 정상 상태(contraction steady state)유지의 4단계를 갖는다고 가정하고 각 모델에 4개의 상태를 할당하였다. GMM 기법을 기반으로 하여 각 상태(State)들마다 다양한 특징 벡터수를 구성해서 테스트 한다 [10][11].

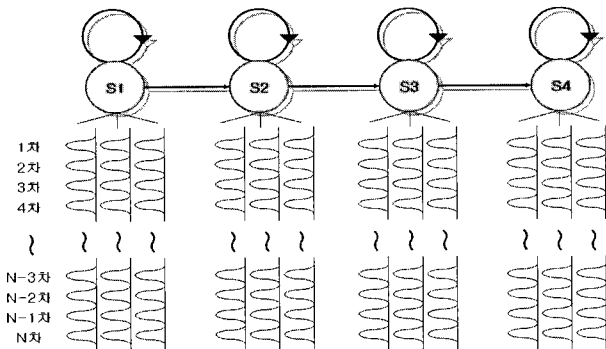


그림 6. HMM-GMM 구성 모듈
Fig. 6. Module of HMM-GMM

j 번째 영역에서, 상태열 T_j 개의 학습 벡터를 $Y_j = Y_{t_j=1}, \dots, x_T Y_{t_j=T_j}$ 라 하면, 가우시안 성분 밀도는 M_j 성분밀도의 가중된 합으로 표현할 수 있다.

$$p(y_{t_j}|\lambda) = \sum_{i=1}^{M_j} p_{j,i} \frac{1}{(2\pi)^{\frac{L}{2}} |\Sigma_{j,i}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(y_{t_j} - \mu_{j,i})^T \Sigma_{j,i}^{-1} (y_{t_j} - \mu_{j,i})\right\} \quad (21)$$

여기에서, $\mu_{j,i}$ 는 평균벡터이고, $\Sigma_{j,i}$ 는 분산행렬이다. 혼합 성분의 가중치는 $\sum_{j=1}^K \sum_{i=1}^{M_j} p_{j,i} = 1$ 를 만족한다.

$Y = \{Y_1, \dots, Y_K\}$ 가 주어지면, 모델을 위한 가우시안 성분 밀도 함수는 성분의 평균벡터(μ), 공분산 행렬(Σ), 가중치(p)로 나타낼 수 있다.

$$\lambda = \{p_{j,i}, \mu_{j,i}, \Sigma_{j,i}\} \quad \begin{matrix} i=1, \dots, M_j \\ j=1, \dots, K \end{matrix} \quad (22)$$

여기에서, j 번째 영역의 i 번째 클래스의 사후확률(A posterior probability)은

$$p(j, i|y_t, \lambda) = \frac{p_{j,i} b_i(y_{t_j})}{\sum_{i=1}^{M_j} p_{j,i} b_i(y_{t_j})} \quad (23)$$

이다. 근전도 패턴 식별을 위하여, 6개의 동작 패턴은 GMM의 $\theta_1, \theta_2, \dots, \theta_6$ 로 나타내고, 패턴의 주성분 벡터를 이용하여 GMM의 최대 사후확률 값을 갖는 패턴 모델 l 를 찾을 수 있다[12].

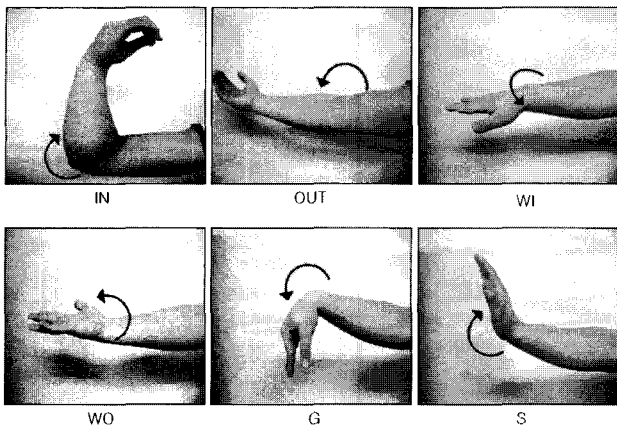
$$\hat{s} = \max \sum_{t=1}^T \log p(y_t|\theta_l) \quad (24)$$

III. 실험 및 결과

A. 데이터 획득

실험은 우선 의수 제어를 위한 기본 기능을 6가지의 기본 운동으로 제한한다. 그림 7에 설정한 팔의 기본 기능이 나타나 있다. 그림처럼 6가지 기본 운동에 대한 신호를 장기간에 걸쳐 개별적으로 수집하였다. 시험 시스템의 구성은 그림 5와 같이 구성하였으며, 20대 후반의 정상인을 실험 대상으로 선택하였다. 증폭기는 Nihon Kohden사의 생체 증폭기인 AB-621G를 사용하였으며, 시간 상수는 피부와 전극사이에 움직임으로 유기되는 포텐셜 및 근포텐셜 등이 원인인 저주파 잡음 제거 등을 고려하여 0.01초로 하였고, 감도는 0.2mV/div를 사용하여 A/D변환기의 입력 범위를 충분히 사용할 수 있게 하였다. 실제로 EMG 신호의 유효 성분을 가진 주파

수 영역은 참조하는 근육에 따라 다르다[1]. 전극은 Ag-Cl 일회용 전극을 사용하였으며, 이두박근과 삼두박근의 중심부와 끝부분에 각기 2개씩의 전극(쌍극법)을, 그 중앙에 접지전극을 부착하여 시험하였다. 표면 전극 2쌍을 실험자의 이두박근(Channel 1)과 삼두박근(Channel 2)에 각각 부착하였으며. 충분한 EMG 정보를 유지하도록 EMG 신호 샘플링률을 10000 samples/s로 설정하였다. 5-2000Hz의 Band Pass 필터를 적용했다. 그리고 연속적인 근전신호 분석을 수행하기 위해 데이터 윈도우 길이는 25.6 msec(256 points)로 설정한다. 또한 윈도우 이동증가분은 10 msec(100 points)로 설정한다.



번호	심볼	기본 운동
1	IN	굴곡(屈曲) 운동
2	OUT	신장(伸張) 운동
3	WI	내전(內轉) 운동
4	WO	외전(外轉) 운동
5	G	폐수(閉手) 운동
6	S	개수(開手) 운동

그림 7. 7가지 동작
Fig. 7. 7 Motions

팔과 손동작의 6가지 패턴에 대해서 데이터를 분석하여, 동작마다 각각 30개의 데이터를 반복해서 획득하였다. 이들 기본동작들은 IN(굴곡 운동), OUT(신장 운동), WI(내전 운동), WO(외전 운동), G(폐수 운동), S(개수 운동)으로 구성한다. 이두박근과 삼두박근에 부착한 표면전극(surface electrode)에서 동시에 연속적인 데이터를 취득한 후 절대 적분치 알고리즘을 적용하여 에너지 분포가 높은 구간만을 학습에 적용한다. 2개의 채널에서 입력 받은 근전신호를 절대적분치 식(1)을 적용하여 설정된 임계치 내의 두 채널에 대한 근신호 구간을 추출한다.

B. 실험결과

본 실험은 2개의 Channel에서 입력된 근전도 신호에서 6개 패턴에 대한 인식률을 비교하기 위해서는 다음과 같은 사항을 고려한다[13].

- 특징 벡터 차수의 수와 인식률
- 다른 특징 벡터간의 결합을 통한 인식률
- MFCC의 필터타입에 따른 인식률
- MFCC의 필터 개수 변화에 따른 인식률
- 각 분류기에 따른 인식률 차이

표1 은 특징 벡터의 차수를 변화시키면서 그 인식률 차이를 비교한 결과이다. LPCC는 시간영역에서 검출된 타 특징들에 비해 상대적으로 높은 인식률을 나타내므로 MFCC의 특징과 비교 분석이 적합하다. 따라서 다음 실험들을 위해서 제안된 시스템과의 비교대상으로 선정했다.

인식률 결과를 도출하기 위해 각 패턴당 15개의 학습용 샘플과 15개의 테스트용 샘플을 수집하여, 분류기의 패턴 학습에 이용된 특징 데이터는 패턴당 학습용 샘플 15개와 테스트용 샘플 15개를 혼합해서 총 30개 샘플을 이용했다. 다음 실험들은 모션 당 30번의 테스트를 통해서 에러율을 검출했다.

표 1. LPCC와 MFCC 차수별 에러율
Table 1. Error Rates of Changes of LPCC and MFCC oders.

Motion Order	IN	OUT	WI	WO	G	S
LPCC(8)	0/30	1/30	3/30	1/30	2/30	0/30
LPCC(9)	0/30	1/30	4/30	0/30	2/30	0/30
LPCC(10)	0/30	0/30	5/30	1/30	2/30	0/30
LPCC(11)	0/30	1/30	6/30	1/30	1/30	0/30
LPCC(12)	0/30	1/30	6/30	1/30	2/30	0/30
LPCC(13)	0/30	0/30	5/30	1/30	3/30	0/30
LPCC(14)	0/30	1/30	5/30	1/30	4/30	0/30
LPCC(15)	0/30	1/30	5/30	1/30	3/30	0/30

Motion Order	IN	OUT	WI	WO	G	S
MFCC(8)	0/30	0/30	1/30	1/30	1/30	0/30
MFCC(9)	0/30	1/30	1/30	0/30	1/30	0/30
MFCC(10)	0/30	0/30	1/30	1/30	1/30	0/30
MFCC(11)	0/30	1/30	1/30	1/30	1/30	0/30
MFCC(12)	0/30	1/30	1/30	1/30	1/30	0/30
MFCC(13)	0/30	1/30	0/30	1/30	1/30	0/30
MFCC(14)	0/30	1/30	1/30	1/30	2/30	0/30
MFCC(15)	0/30	1/30	2/30	1/30	2/30	0/30

실험 결과에서 LPCC와 MFCC의 인식 에러율이 상당히 차이가 있음을 보여준다. 상대적으로 MFCC의 전체적 에러율이 낮음을 확인할 수 있다. 특히 WI, G는 인식률이 상대적으로 낮았다. 또한 LPCC와 MFCC 모두 차수가 증가함에 따라 차츰 인식률이 떨어 짐을 보여준다. 따라서 가장 적절한 차수의 수는 8-10으로 나타났다.

표2와 표3은 재구성한 필터 뱅크 타입에 대한 에러율을 나타낸다. 실험을 위해 표1의 실험결과에서 도출한 결과를 반영한 것이다. 차수의 수를 적절한 9로 설정하여 테스트 하였다. 뱅크의 갯수

는 차수의 수와 비슷하게 10개로 구성하였다. 삼각 필터 बैं크 타입들과 해닝 필터 बैं크 타입들을 비교를 통해 삼각 필터 타입이 대체로 인식률이 개선된 결과를 보여준다. 특히 삼각 필터 타입 1이 가장 어려움이 적은 결과를 도출하였다. 이는 본래의 음성인식에서 적용되는 MFCC의 타입과 유사하며, 따라서 음성이나 근전도는 고주파보다는 저주파에 유효한 특징에 충분히 포함되어 있음을 상기시켜 주는 부분이다.

표 2. 삼각 필터 बैं크에 대한 어려움

Table 2. Error Rates of Triangular filter banks

Motion Type	IN	OUT	WI	WO	G	S
Type 1	0/30	0/30	1/30	1/30	1/30	0/30
Type 2	1/30	0/30	1/30	1/30	1/30	0/30
Type 3	0/30	0/30	4/30	2/30	1/30	0/30
Type 4	1/30	1/30	5/30	2/30	1/30	0/30

표 3. 해닝 필터 बैं크에 대한 어려움

Table 3. Error Rates of Hanning filter banks

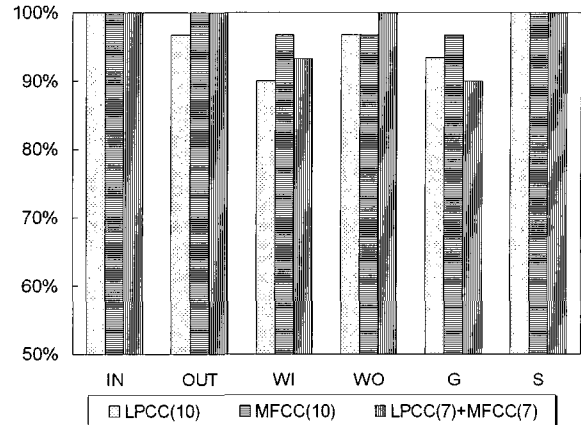
Motion Type	IN	OUT	WI	WO	G	S
Type 1	0/30	0/30	3/30	1/30	1/30	0/30
Type 2	1/30	0/30	3/30	1/30	1/30	0/30
Type 3	0/30	1/30	3/30	2/30	1/30	0/30
Type 4	1/30	2/30	5/30	2/30	2/30	0/30

학습을 통해서 구성된 6쌍의 HMMs은 테스트 데이터에서 추출된 특징 벡터들과 Viterbi 알고리즘을 통해 비교 되어져서 최대값을 출력하는 모델을 인식결과로 결정한다. 그림 7은 LPCC와 MFCC를 적용해서 동작별 인식률을 비교 분석한 결과이다.

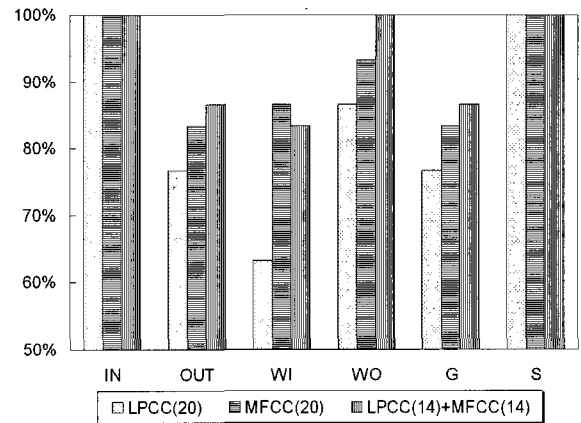
실험 결과를 통해 LPCC 보다 잡음요소에 강한 MFCC가 OUT, WI, G 동작에 대해서 인식률이 개선되었음을 보여준다. LPCC의 경우는 전체 약 95%의 인식률을 보였지만 MFCC는 실험을 통해서 약 98%의 인식률을 나타내면서 약 3%의 인식률 개선을 보여준다. 그리고 LPCC와 MFCC의 특징을 혼합한 결과는 LPCC보다 OUT, WI, WO 동작에 대해서 개선된 인식률을 보여주지만, 이는 연산량이 더 많기 때문에 실시간 인식시스템 구현에는 적당치 않다. 오히려 MFCC만으로 구성된 결과와 비교해보더라도 큰 인식률의 차이는 없다. 따라서 MFCC만 적용하는 것이 적절하다.

그림 7의 (a)와 (b)는 특징 차수가 인식률에 상당한 중요한 요소를 보여준다. 과도한 특징 차수의 증가는 오히려 인식률을 낮추는 결과를 보여준다.

그림 8은 각 모션에 대한 근전도 신호에서 검출한 특징 벡터들로 학습된 분류기(Classifier)들의 인식결과를 보여준다. DHMM, Cascaded DHMM, DHMM-MLP, 제안한 시스템인 HMM-GMM에 대해서 실험하였다. HMM-GMM 분류기의 경우 MFCC 특징 벡터 차수를 10, 필터뱅크수를 10으로 설정하였고, 다른 분류기들은 LPCC 특징 차수를 10으로 설정하여 실험을 수행하였다.



(a) 10차



(b) 20차

그림 8. 각 동작에 대한 인식률

Fig. 8. Recognition rates of each motions

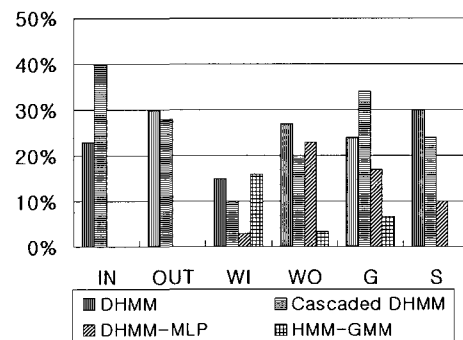


그림 9. 동작 패턴 분류기들에 대한 어려움(%)

Fig. 9. Error Rates of Motion Classifiers(%)

Type	Error Rates	총 어려움(%)
DHMM		25.3
Cascaded DHMM		26.5
DHMM-MLP		9.2
HMM-GMM		4.5

IV. 결 론

본 논문에서는 의수의 연속 제어시 인식을 향상을 위해서 MFCC 특징 검출 기법과 HMM-GMM 패턴 분류기를 도입하였다. 잡음이 내포된 근전신호에서 고유 특징을 검출하기 위해 필터의 재구성하였다.

각각의 특징 차수를 증가시킨 실험의 경우 오히려 전체적으로 인식률이 낮아지는 결과가 발생했다. 이는 특징 차수가 인식률에 상당한 중요한 요소임을 보여준다. 실험에서 과도한 특징 차수의 증가는 오히려 인식률을 낮추는 결과를 보여준다. 이는 패턴마다 고유한 특징 분포가 평균화되면서 전체적인 인식률이 저하된다.

LPCC의 경우는 전체 약 95%의 인식률을 보였지만 MFCC는 실험을 통해서 약 98%의 인식률을 나타내면서 약 3%의 인식률 개선을 보여준다. LPCC 보다 잡음요소에 강한 MFCC가 OUT, WI, G 동작에 대해서 인식률이 개선되었다. 그리고 LPCC와 MFCC의 특징을 혼합한 결과는 LPCC보다 OUT, WI, WO 동작에 대해서 개선된 인식률을 보여주지만, 이는 연산량이 더 많기 때문에 실시간 인식시스템 구현에는 적당치 않다. 오히려 MFCC만으로 구성된 결과와 비교해서 큰 인식률의 차이는 없다. 따라서 MFCC만 적용하는 것이 적절하며, MFCC만으로 구성된 실험에서는 상대적으로 어느 한 패턴에 편파적으로 인식률이 높지 않고 고루 인식률이 높게 나타났다.

인식 시스템의 경우 DHMM 분류기의 경우는 신호의 동적인 특성에 대한 적응성이 뛰어나지만 적절한 특징의 분류 기능을 내포하고 있지 못하다. Cascaded DHMM의 경우는 단계적인 분류 과정 수행을 통해서 이를 개선하기 위한 시도였으나 오히려 DHMM보다 더 낮은 결과가 나타났다. 이에 비해 DHMM-MLP의 경우는 신경회로망을 통해서 특징 분류 기능을 추가함으로써 성능을 개선하였다. 하지만 신경회로망은 오히려 근전도 신호의 동적인 특성을 적절히 반영하지 못하였다. 이들 DHMM, Cascaded DHMM, DHMM-MLP의 분류기와 비교해서 HMM-GMM 분류기는 시간적으로 변하는 동적인 신호의 처리와 적절한 특징 분류를 통해서 개선된 성능을 보여주었다.

향후 안정적인 시스템의 개선을 위해 적절한 특징차수의 분석에 대한 재검토가 필요하며, 학습 및 인식 속도 개선 위한 분석도 요구된다.

참고문헌

- [1] Moshe Solomonow et al., "The EMG-force model of electrically stimulated muscle: Dependence on control strategy and predominant fiber composition," *IEEE Trans. Biomed. Eng.*, vol. 34, no. 9, pp. 692-703, 1987.
- [2] Willis J. Tompkins, *Biomedical Digital Signal Processing*, Prentice Hall International Editions, 1993, pp. 43-44.
- [3] G. Zavaliagkos, et al., "A hybrid segmental neural net/hidden Markov model system for continuous speech recognition," *IEEE Trans. Speech and Audio Processing*, vol. 2, no. 1, part II, pp. 151-152, 1994.
- [4] Lalit R.Bahl, Peter F.Brown, Peter V. de Souza, and Robert L.Mercer, "Estimating hidden Markov model parameters so as to maximize speech recognition accuracy," *IEEE Trans. Speech and Audio Processing*, vol. 1, no. 1, pp. 77-78, 1993.
- [5] Jangwoo Kwon et al., "Probabilistic-neural pattern classifier and muscle force estimation," in *Proc. the 15th annual international Conference of the IEEE EMBS*, Sandiego, U.S.A., Oct. 1993, pp.1145-1146.
- [6] Jangwoo Kwon, et al., "The application of counterpropagation neural networks for EMG pattern classification," in *Proc. 94 World Congress on Medical Physics and Biomedical Engineering*, Rio-dejanerio, Brasil, Aug. 1994, pp. 919.
- [7] L.R.Rabiner, R.W.Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Inc., 1978, pp. 359-362.
- [8] Lawrence Rabiner, Riing-Hwang Juang, *Fundamentals of Speech Recognition*, Prentice-Hall International, Inc., 1993, pp. 333-357.
- [9] Kai-Fu Lee, and Hsiao-Wuen Hon, "Speaker-independent phone recognition using hidden Markov models," *IEEE Trans on Acoustics, Signal processing*, vol. 37, no. 11, Nov. 1989.
- [10] Masakiyo Fujimoto, Yasuo Ariki, "Robust speech recognition in additive and channel noise environments using GMM and EM algorithm," in *Proc. (ICASSP '04). IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2004, I-941-4 vol.1.
- [11] Seiichi Nakagawa, Wei Zhang, Mitsuo Takahashi, "Text-independent speaker recognition by combining speaker-specific GMM with speaker adapted syllable-based HMM," in *Proc. (ICASSP '04). IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2004, I-81-4 vol.1.
- [12] Ming-Tao Gan, Hanmandlu, M., Ai Hui Tan, "From a gaussian mixture model to additive fuzzy systems," *IEEE Transactions on Fuzzy Systems*, vol. 13, pp. 303-316, 2005.
- [13] Bernard Hudgins and Philip Parker, and Robert N.Scott, "A new strategy for multifunction myoelectric control," *IEEE Trans on Biomedical Engineering*, vol. 40, no.1. 1993.
- [14] Ronald J.Triolo and Gordon D.Moskowitz, "The theoretical development of a multichannel time-Series myoelectric processor for simultaneous limb function detection and muscle force estimation," *IEEE Trans. Biomed. Eng.*, vol. 36, no. 10, 1989.