확산망을 이용한 음성인식

許 萬 澤*

* 부경대학교

(The Speech Recognition Using the Diffusion Network)

Huh, Man Tak

Pu-Kyung Nat'l Univ.

Abstract

In this paper, the pre-processing method for the recognition of single vowels by use of spectrum envelope[1] is presented. we use new method of an extrating spectrum envelope using the diffusion filter bank.[2,3] We reduced the total processing time, and got higher enhancement of discrimination. By getting 88.3 % of average recognition rate for single vowels of real voice through computer simulation, we confirmed it to be useful for speech recognition which use spectrum analysis for voice signal to have many frequency components.[4]

1. 서 론

음성인식의 최종목표에 다다르기 위해서는 인간음성인식, 화자독립, 무제한 어휘인식 등의 기술이 필수적이다.[5] 그리기 위해서는 궁극적으로 인간의 뉴런이론과 특성이 유사한 음성 처리기의 개발이 필수적이며 현재 공학의 각분야에서 각광을 받고 있는 신경회로망을 이용한 접근이 최근 많이 연구되고 있다.[6-8]이는 컴퓨터가 하나의 처리소자를 이용하여 하나의 정보만을 처리하는 반면 동물의 뇌 구조는 여러개의 세포가 복잡하게 연결되어 정보처리를 병렬로 하기 때문에 방대한 데이터를 신속히 처리할 수 있기 때문이다.

하만타 등은 신경회로망의 일종인 확산망이 공간주파수를 추출해 내는 대역통과 특성이 있다는 점과 신경세포간의 연결이 매우 적어 단순한 구조의 하드웨어 구현이 용이하다는 점에 착안하여 이를 이용한 대역통과필터의 구성을 제안하였다.[2] 종래의 확산망에서는 두 가우스함수의 차에 의해 DOG 연산을 하였지만 하만탁 등은 가우스함수를 독립적으로 차분하는 $g^{(1)}(t)$ 연산을 제안하였고, 또 차분회수를 증가시켜 필터의 선택도를 높일 수 있는 $g^{(2n)}(t)$ 연산자를 구현하였다.[9] 이에 의하면, 확산 필터의 중심 주파수는 확산회수와 차분회수에 의해 결정 되어지며 선택도는 차분회수에 의해 결정된다고 하였다. 또 이를 이용하여 확산회수와 차분회수를 순차적으로 적절히 조절해줌에 의한 순차필터뱅크(sequential filter bank)의 특성을 가지는 확산필터뱅크를 구현하였다. 이의 특성은 기존의 아날로그나 디지탈 필터뱅크가 가지고 있는 과다한 소자 사용에 따른 구조의 복잡성과 실제의 어려움[10,11]을 개선하였다.

하만탁은 확산필터뱅크를 이용하여 한국어 6 개 단모음의 스펙트럼 추출을 시뮬레이션한 결과 확산필터뱅크의 출력 스펙트럼은 주파수 성분을 잘 반영할 뿐만 아니라 포락선이 완만하여 바로 포먼트(formant) 추출이 가능하므로 다음 단계인 음성인식의 알고리즘과 하드웨어 구조를 단순화 시킬 수 있음을 예측하였다.[3]

본 논문에서는 확산필터뱅크의 스펙트럼 추출기능을 다음 효율적으로 개선하여 이를 설계하였고 또, 이를 이용하여 실제 음성신호의 스펙트럼을 추출하였으며 추출된 단모음들의 스펙트럼 포락선 형태를 분석하여 인식을 위한 알고리즘과 인식파라메터를 제안한다. 이들을 이용하여 단음절에 포함되어있는 단모음의 인식 시뮬레이션을 한 결과 인식율이 88.3 [%]에 이르러 조성에 관계없이 후속되는 모음을 인식할 수 있음을 알 수 있었다. 이로써 완만한 포락선을 가지면서 많은 주파수 성분을 추출할 수 있는 전처리기가 요구되는 음성신호의 인식에 대하여 확산필터뱅크가 대단히 효과적이며 음성인식 분야에 유용하게 사용될 수 있음을 확인하였다.

2. 확산필터뱅크

동물의 신경망은 신경세포가 자극을 받게 되면 이웃하는 신경세포 쪽으로 그 자극이 전달되어 가는데 이와 모방하여 확산망이 제안되었으며 확산망의 전달함수는 가우스함수의 특성을 이용한 신호처리에 아주 유용하게 적용될 수 있다. 본 장에서는 가우스함수를 이용한 확산필터[2,3]를 사용하여 만든 확산필터뱅크의 이론 및 확산필터뱅크에 의한 스펙트럼 추출원리를 논하고자 한다.

확산망의 수행회수 k와 차분망의 수행회수 n은 시간이 경과되어야 증가시킬 수 있으므로 추출 채널의 순서는 그 중심주파수에 해당하는 k와 n이 증가하는 방향으로만 설정할 수 있다. 또한 확산과 차분의 연산은 순서에 상관없이 회수에만 의존한다. 선택도 Q는 확산회수 k가 비교적 작을 때 근사적으로 차분회수 n의 제곱근에 비례하고 확산회수 k와는 관계가 없다. 그러나 k은 차분회수 n을 고정시켜 선택도를 변화하지 않고 중심주파수만 변화 시킬 수 있음을 보여준다. 그리고 또한 확산회수

와 차분회수가 같은 비율로 변화함에 따라 중심주파수는 변화하지 않고 선택도만 변화함을 보여준다. 주어진 확산망이 추출할 수 있는 중심주파수의 범위는 확산회수 k와 차분회수 n의 가변범위 및 샘플링 주파수에 의하여 정해진다. 확산계수를 고정하면 입력된 신호는 k와 n의 값, 그리고 샘플링 주파수에 의하여 결정된 중심주파수에 따라 여과처리된다.

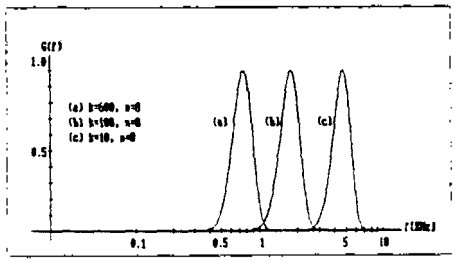


Fig. 1. The case varied with center freqency only.

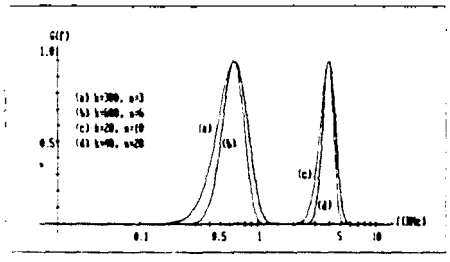


Fig. 2. The case varied with selectivity only.

이제 앞서 서술한 여파특성을 가진 각 필터 채널을 조합하여 목적하는 필터뱅크 시스템을 구성하기로 한다. 일반적인 확산필터뱅크 시스템의 전체 구성은 그림 3과 같이 하였다. 여기

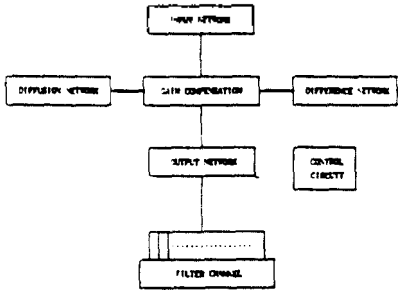


Fig. 3. General block diagram of diffusion filter bank.

서 입력 신호의 스펙트럼만을 요구할 때, 여파출력의 실효치만 검출하면 되므로, 한 프레임의 길이를 t_f라 하고 샘플링 주파수를 f_s, 추출할 최저 주파수를 f_L, 한 프레임에 입력되는 최저 주파수 성분의 주기수를 T_L라 하면, 한 프레임에 필요한 신경 세포의 수 M은

$$M = \frac{T_L}{f_L} \cdot f_s \qquad (1)$$

와 같다. 식 (1)에 의해 각 확산망에서, 출력망을 제외한 입력망, 확산망, 차분망 및 이득 보상 회로의 신경 세포의 수는 여파할 프레임의 샘플링 수 M과 가장자리 문제[12]를 제거하기 위한 여분의 신경 세포수 m을 합한 크기이며 출력망의 크기는 M이다. 여기서 한 프레임의 길이 t_f는 최저 주파수에 따라 결정되므로 M은 최저 주파수와 샘플링 주파수에 비례하여 정해진다. 그림 3에 있는 확산필터뱅크의 동작은 다음과 같다. 입력망은 저역 필터를 통과한 후 샘플링된 외부신호를 직렬로 받아 들여 가면서 쉬프트 동작을 하는데, 샘플링된 외부 신호를 최초에는 M+m 개를 받아 들인 후 다음 부터는 M개씩 쉬프트한 후 이득보상회로로 병렬로 내 보낸다. 출력망은 여파특성을 결정하는 k와 n의 합이 작은 순으로 여파된 신호를 이득 보상회로로 부터 받아 k와 n의 합이 작은 필터 채널 순으로 차례대로 내 보낸다. 이득 보상 회로는 먼저 입력망으로부터 신호를 병렬로 받아 확산망, 또는 차분망으로 내 보내면서 제어 시스템으로부터 전송된 이득보상계수에 의하여 스케일링한다. 확산망과 차분망은 이득보상회로로부터 받은 신호를 제어 시스템으로부터 보내온 회수만큼 처리하여 다시 이득보상회로로 전송한다. 제어 시스템은 각 블럭에서의 입력과 출력의 출입을 사전 계획된 대로 제어하고 이득 보상 계수를 계산하여 이득보상회로로 전송하며 사전 프로그램된 확산 회수와 차분 회수를 확산망과 차분망으로 각각 전송한다. 필터 뱅크의 최종 출력은 출력망으로부터 보내온 여파된 프레임별 신호의 실효치를 구하여 각 필터채널로 전송된 값들이다.

3. 단모음의 스펙트럼 분석 및 인식실험

3-1. 음성의 분석

음성분석에 있어서 가장 중요한 것 중에 하나는 특성을 추출하는 것이다. 음성 특성은 크게 주파수영역 분석에 의한 것과 시간영역 분석에 의한 것, 이 두 가지를 조합한 것들이 있다.[2] 기존의 방법 중에는 시간영역상의 음성파형을 특성으로 많이 추출하였다. 그러나 음성파형은 시간에 따른 변화량이 많아 데이타 양이 많은 단점이 있어 이를 주파수 영역으로 변환시켜 추출하는 방식이 많이 사용된다. 음성이 성도(vocal tract)로부터 발성된다는 사실을 근거로 성도의 형태를 필터로 가정하고 그 필터계수를 음성의 특징으로 삼는 방법과 동물의 귀가 음성을 분석하는 방법을 이용한 auditory 분석도 있다. 최근에는 시간영역의 동적특성 (dynamic feature)을 주파수 영역의 특징(spectral feature)들과 함께 사용하기도 한다.[3]

말의 흐름 속에서 분별이 가능한 최소단위의 소리를 단음(phone) 또는 음소(phoneme)이라고 하며 단음은 제각기 소리값이 다르지만 유사성과 공통성에 따라 분류된다. 분류되는 구분은 분류기준에 따라 다양하게 분석될 수 있으나 크게 자음과 모음으로 분류된다. 우리나라의 언어적인 단음체는 음절구조의 관점에서 초성, 중성, 종성의 삼분법으로 나누며 초성

과 중성은 자음에 중성은 모음에 해당하며 반모음은 모음에 속한다. 단모음은 그 음을 길게 발음 하더라도 그 음이 변하지 아니하는 모음으로 /아/, /에/, /이/, /오/, /으/, /어/, /우/, /애/ 등이 있다.[14] 또한 단모음의 분석에는 포만트에 의한 방법이 주로 쓰인다. 포만트는 발성자에 따라서, 또 전후에 연결되는 음소의 영향 즉 조음결합에 의해 주파수가 변동한다. 포만트 주파수는 성도의 기하학적 형상이나 크기와의 상관관계가 크다. 그러나, 그림 4[13]에서 보는 바와 같이 각 단모음의 영역이 중복되는 부분이 있어 1, 2차 포만트만으로는 그 구별의 정확도가 다소 떨어진다.

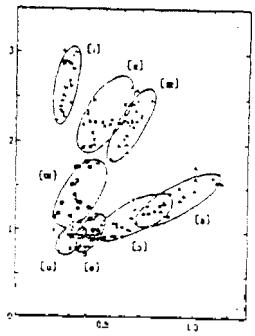


Fig. 4. The vowel triangle.

3-2. 스펙트럼 추출

본 논문에서는 확산필터뱅크의 스펙트럼 추출 기능이 더욱 효율적으로 되도록 개선하였으며 추출된 단모음들의 스펙트럼을 분석하는 데에도 종래의 포만트 분석법이 모음들 사이에 중복영역이 있음으로 인하여 분리가 잘되지 않는 것을 피하기 위해 스펙트럼 포락선의 형태를 분석하였고 분석된 결과에 의해서 인식 시뮬레이션을 행하였다.

최근의 확산필터뱅크는 분석대역을 이분하고 각 대역을 신호의 샘플링에서부터 독립적으로 처리하였으나 각 대역의 저주파 성분을 추출하기 위해서는 여전히 많은 확산회수가 요구되었다.[3] 이는 해당 대역의 최고주파수보다 낮은 주파수 성분을 검출해내기 위해서는 확산회수를 증가시키야 하는데 추출할 주파수가 아주 낮아지면 확산회수는 매우 커져야 한다. 확산회수가 크다는 것은 그만큼 필터링 시간이 많이 소요되어 실시간 처리에 애로를 줄 수 있고 가장자리를 위한 여분의 신경세포가 많아야 하고 이득보상계수의 변화도 커지므로 하드웨어 구현을 어렵게 한다. 최고확산회수를 줄이기 위해서는 각 분석대역의 최고주파수와 최저주파수의 비를 줄여야 하는데 전체 분석 대역폭을 줄일 수는 없으므로 대역폭을 여러개로 분할할 수 밖에 없다. 분할된 각 대역의 최고주파수는 샘플링 주파수의 조정에 의해서 결정할 수 있으므로 최고 샘플링 주파수로 샘플링된 신호를 각 분할대역의 최고주파수에 맞게 분주를 하면 되는데 이에 앞서 나이퀴스트 이론에 의해 고주파 성분을 제거해야 한다.

본 논문에서는 고역성분을 제거하기 위해 신호를 반복 확산처리하여 저역필터링을 하였다. 실제로 76 [Hz] ~ 4296 [Hz]의 음성대역을 6 개로 분할하여 최고주파수가 각각 5,000 [Hz], 2,500 [Hz], 1,000 [Hz], 500 [Hz], 250 [Hz] 및 200 [Hz]가 되도록 하였다. 이렇게 함으로써 이분하였을 때 최고확산회수가 800여 회[3] 요구되던 것을 40여 회로 감소시켰다. 이때 저역필터링에 의한 이득변화는 시뮬레이션에 의해 그 계수를 구하여 보상하였다. 또한 종래에는 확산필터뱅크의 각 차분회수를 8로 하였으나 스펙트럼의 분별력을 높이기 위해 20으로 설정하여 선택도를 높였다. 그림 5의 (a)와 (b)에 종래의 확산필터뱅크에 의한 스펙트럼과 본 논문에서 개선한 확산필터뱅크에 의한 스펙트럼을 보여 준다. 주파수 분석대역을 음성신호의 대역에 맞추어 76[Hz] ~ 4296[Hz]로 하고 검출주파수의 채널수는 162 개로 하였다. 선택도는 달팽이관의 선택도보다 높은 5.2 정도가 되도록 차분회수를 20으로 고정하였다. 추이진 확산형이 추출할 수 있는 중심주파수의 범위는 확산확수 k와 차분회수 n의 가변범위 및 샘플링주파수에 의해 결정된 중심주파수에 따라 여파된다. 또한 추출주파수 대역도 6 개로 분할하여 저주파 대역의 샘플링주파수를 10[KHz]로 하여 2,00[Hz] ~ 4,296[Hz], 샘플링주파수를 5[KHz]로 하여942[Hz] ~ 1,914[Hz], 샘플링주파수를 2[KHz]로 하여 380[Hz] ~ 859[Hz], 샘플링주파수를 1[KHz]로 188[Hz] ~ 366[Hz], 샘플링주파수를 500[Hz]로 하여 95[Hz] ~ 183[Hz], 샘플링주파수를 400[Hz]로 하여 76[Hz] ~ 94[Hz] 범위의 162 개 주파수 성분

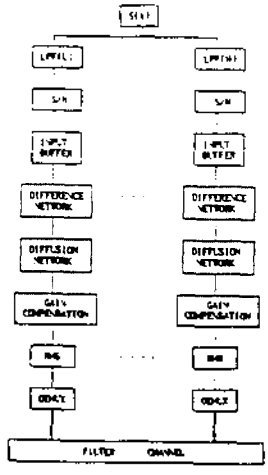
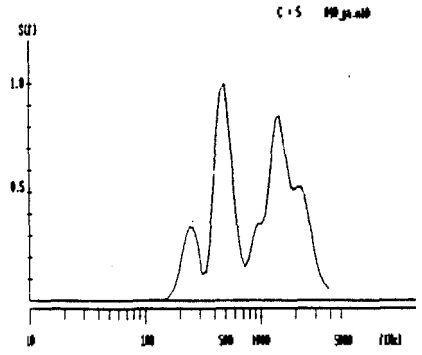
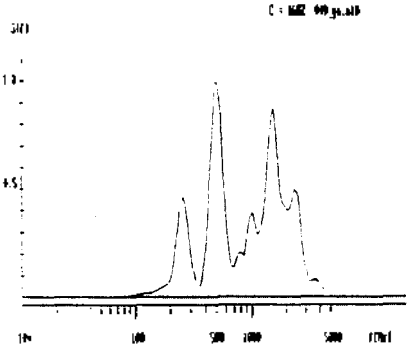


Fig. 5. The block diagram of the diffusion filter bank producing the spectrum of speech signals with n=20.

을 검출하였다. 전체 시스템의 구성은 선택도가 고정된 점을 고려하여 그림 6과 같이 하였으며 그 동작은 다음과 같다. 먼저 입력된 신호는 각 대역별로 저역통과필터를 거치게 되는데 각 밴드의 최고주파수 이상은 완전 차단되도록 하였다. 저역통과필터를 거쳐 나온 신호들은 각 밴드의 차단주파수에 맞도록



(a)



(b)

Fig. 6. The spectrum extracted from diffusion filter bank.
(a) Out from conventional diffusion filter bank.
(b) Out from advanced diffusion filter bank.

주파수가 분주되게 하였다. 이들 신호는 230 샘플씩 직렬로 입력망에 입력된 후 200 샘플씩 쉬프트한 후 입력버퍼를 거쳐 차분망으로 병렬출력한다. 차분망에서 2차 차분을 20회 수행한 후 확산망으로 출력하고 확산망에서는 차분망으로 부터 받은 신호와 주어진 각 중심주파수에 해당하는 확산회수 만큼 반복한 후 이득보상회로로 출력한다. 이득보상회로는 확산망으로 부터 여분의 신경세포수 30개를 제외한 200 개의 신경세포로 부터 입력받은 신호에 스케일계수를 곱한 후 이를 실효치 산출회로로 보내며 여기에서 실효치를 구하여 그 결과를 해당 대역로 보낸다. 여기에서 이득보상회로의 출력이 완전한 정현파가 되지 않은 경우와 순시잡음에 의한 오차를 제거하기 위해 실효치를 취하였다. 분석 시료의 처리과정은 다음과 같다. 먼저 10 [KHz]로 샘플링된 음성 시료로 부터 신폭의 크기에 의하여 만들된 모음구간 600 개의 샘플을 분리하고 이를 반복 입력하여 3,000 개의 샘플을 마련한 다음, 각 대역에 알맞는 시료로 처리한다. 초기에 많은 수의 샘플을 마련한 것은 분주되어 감에 따라 샘플의 수가 작아지기 때문에 이를 미리 고려하고자 함이다. 위에서 언급한 각 대역에 알맞는 시료를 마련하기 위한 처리과정은 저역필터링과 샘플링 주파수와 분주로 나누어 지며 대역필터링은 21 대역의 최고주파수 이상의 주파수 성분이 완전히 제거 되도록 결정된 확산회수 만큼 확산을 반복함으로써 수행되고, 샘플링 주파수 분주는 확산처리된 시료에서 가장자리 제거를 한 다음 일정수의 샘플을 건너 뛰어 발췌하는 방법으로 수행된다. 여기서 샘플링 주파수의 분주율은 1/2 또는 1/5로 하였으며 이들을 반복 사용하여 6 개의 분석대역에 맞는 시료로 처리하였다. 이상과 같이 처리하여 마련된 시료는 각 해당 대역 확산필터뱅크에서 대역필터링이 되어 각 주파수 성분이 검출되어 시료 신호의 스펙트럼이 구하여진다.

3-3. 스펙트럼의 특징파라미터 분석 및 인식 시뮬레이션

분석대상이 된 단모음은 20세 중반의 세 남자가 발성한 1초성 + 중성1으로 구성된 단음절에서 단모음만 분리한 /아/, /어/, /에/, /오/, /우/ 및 /이/의 6 개이며 초성에는 파열음, 마찰음, 파찰음 및 비음등 초성에 올 수 있는 모든 자음을 포함시켰다. 이렇게 함으로써 보다 더 자연스럽게 발성될 뿐만 아니라 음소단위 분석에 상세적인 시료를 얻을 수 있으며 초성에 관계없이 후속되는 모음을 검정할 수 있음을 보여 줄 수 있다. 스펙트럼 분석은 스펙트럼의 포락선 형태를 모음에 따라 구분한 것으로써 주로 극대점과 극소점의 주파수축상의 상대적 위치를 기준하였다. 이는 모음의 주파수 조합이 발성자의 피치 주파수에 비례하기 때문에 주파수축을 대수눈금으로 하면 피치 주파수에 관계없이 극대점이나 극소점의 상대거리를 결정할 수 있다. 여기서 선정한 특징 파라미터는 낮은 주파수로부터 스캔하여 기준치 이상의 값을 가지면서 주변보다 높은 첫번째 극대점(P1), 극대점보다 일정치 이하의 값을 가지면서 주변보다 낮은 첫번째 극소점(Z1), Z1보다 일정치 이상의 값을 가지면서 주변보다 높은 첫번째 극대점(P2), P2보다 일정치 이하의 값을 가지면서 주변보다 낮은 첫번째 극소점(Z2), Z2보다 일정치 이상의 값을 가지면서 주변보다 높은 첫번째 극대점(P3), P3보다 일정치 이하의 값을 가지면서 주변보다 낮은 첫번째 극소점(Z3), P3로 부터 기준치 이상의 값을 가진 극대점(P3m), P2 와 P3 사이의 극대점 (P231, P232), Z3로 부터 일정 기준치 이상의 값을 가지면서 일정 주파수 이내 및 이외인 극대점(P3i 및 P3o) , P3 이후의 또 다른 기준치 이하의 극소점(Z0) 이다. 이들을 그림 7에 나타내었다. 이들은 분석작업에 사

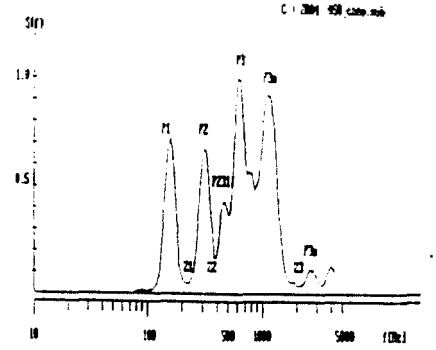


Fig. 7. The example of characteristic parameters.

용된 시료들의 스펙트럼 포락선을 관찰하면서 구분하기에 적당한 것들을 선정한 것이다. 이에 다음 분석과정은 이들 특징 파라미터들을 사용하여 그 포락선의 특징을 이진화하는 것이다. 이에 예외적으로 P2 와 P3 의 주파수 비는 수치로서 나타내었는데 그 값 G2 는 식 (2)와 같다.

$$G2 = 10\log\left(\frac{P3}{P2}\right) \qquad (2)$$

여기서 P1과 P2의 비인 G1은 변별력이 전혀 없어서 무시하였다. 다음 P3m의 존재 유무를 EP3m, 이와 같이 하여 EP23l, EP23z, EP3i, EP3o, 및 EZ0 등을 설정하고 또 P3 의 값이 P2 나 P1 보다 기준치 이상 큰지(P3max)를 조사하였다. 이들 특징들을 모음별로 분석한 결과를 표 1에 보여준다. 표 1에서 알 수 있듯이 분석한 시료들의 스펙트럼 포락선 형태는 대부분 이들로써 구분이 되었으며 그 구분이 중복되는 경우도 소수 있었다. 이는 순모음 구간의 분리 알고리듬이 너무 단순하고 또 감음의 영향을 크게 고려하지 않았기 때문으로 추정된다.

Table 1. The analysis table of characteristic parameter of the vowels.

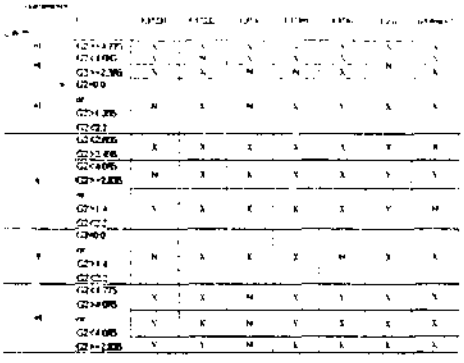


Table 2. The experimental data of the recognition simulation of /a/.

입력 / 인식	아	어	오	우	이	에	무음	시료수	인식율 (%)
아	40	2						42	95.2
어		30	3	1		2	1	37	81.1
오			29	6	2	1		38	76.3
우			2	35				37	94.6
이					38			38	100.0
에		3		4		31		38	82.6

총시료수	총인식수	오인식수	동정불능수	인식율(%)
230	203	26	1	88.3

본 논문에서 분석에 사용한 시료는 /아/가 42 개, /어/ 및 /우/가 각각 37 개, /에/, /이/ 및 /오/가 각각 38 개이다. 이들을 분석한 결과를 이용하여 인식 시뮬레이션을 한 결과를 표 2에 나타내었고 또 모음별 인식율 분석표를 표 2에 나타내었다. 표 2와 살펴보면 /아/ 및 /이/의 인식율이 다른 것에 비해 상당히 높고 /오/의 인식율이 상당히 낮다는 것을 알 수 있다. 이는 그림 4의 모음 삼각도에서 알 수 있듯이 /아/ 및 /이/는 중첩된 영역이 거의 없어 타모음과의 구별이 용이하고 /오/는 타모음과 중첩된 영역이 많아 인식률이 저조한 것으로 생각된다.

4. 결 론

본 논문에서는 화산필터뱅크를 사용한 한국어 단모음 인식 시뮬레이션을 행하였다. 스펙트럼 포락선의 변별력을 높이기 위해 화산필터의 선택도를 높였으며, 필터의 채널수를 증가시키 주파수 분해능을 높였다. 필터뱅크의 처리시간을 단축시키기 위하여 분석대역을 여러개로 분할하여 각 분할대역에서의 최대 화산회수를 줄였다.

개선된 화산필터뱅크를 사용하여 한국어 단모음 6 개의 인식 시뮬레이션을 시행한 결과, 평균 88.3 [%]의 좋은 인식율을 얻었다. 인식을 위한 특징파라미터들은 스펙트럼 포락선의 형태를 나타내는 것으로서 주로 포락선의 극대점과 극소점의 상대적 주파수 위치및 특정 극점의 존재 유무로써 각 단모음이 간단히 구분되도록 하였다. 인식 시뮬레이션의 대상이 된 6 개 단모음 중 /아/와 /이/의 인식율은 매우 높았으나 /오/의 인식율은 상대적인 저조를 보였다. 이것은 음성의 특성상 음소 영역의 분리도에 기인함을 알 수 있었다. 본 시뮬레이션에서 인식 알고리듬이 간단함에도 불구하고 좋은 인식율을 보여줌으로써 화산필터뱅크가 스펙트럼 분석에 의한 음성인식에 아주 유용함을 확인하였다. 앞으로의 과제는 스펙트럼 포락선 패턴의 인식 알고리듬을 더욱 개선시키고, 인식 알고리듬에 더욱 효율성 있도록 화산필터뱅크의 성능을 개선하여 단모음뿐만 아니라 다른 음소들까지 좋은 인식율로 인식하게 하는 음소단위 음성인식기를 개발하는 일이다.

參 考 文 獻

1) 최영배, 양진우, 이형순, 김순협, "한국어 음소인식을 위한 신경회로망에 관한 연구", 한국음향학회지, Vol. 13, No 1, pp. 5-13, 1994.

2) 이세성, 허만타, 이종하, 남기곤, 김재창, 박의열, "화산신경망과 차분연산에 의한 대역통과 필터의 구현," 전자공학회논문지-B, 제32권 B편, 제7호, pp. 94-102, 1995.

3) 허만타, "화산필터뱅크를 이용한 음성신호의 스펙트럼 추출", 부산공업대학교 논문집, 37집, pp. 453-470, 1995.

4) 이종하, 정용근, 유태훈, 김재창, 박의열, 이양성, "∇^2G연산자를 이용한 음성신호의 주파수분석", 전자공학회논문지-B, 제28권, B편, 제4호, pp. 24-32, 1991.

5) L. Rabiner and B-H. Juang, *Fundamentals of Speech Recognition*, P-H, New Jersey, 1993.

6) 이종하, 정용근, 남기곤, 유태훈, 김재창, 박의열, 이양성, "∇^2G연산자의 신호분석 특성을 이용한 음성인식 신경회로

망에 관한 연구", 전자공학회논문지-B, 제29권 B편, 제10호, pp. 90-98, 1992.

7) 허만탁, 이종혁, 윤태훈, 남기곤, 김재창, 이양성, "음소 복정 파라미터를 이용한 부성자음 인식", 전자공학회논문지-B, 제31권 B편, 제4호, pp. 525-532, 1994.

8) 김순협, "음성인식 기술현황및 실용화 전망", 한국음향학회지, Vol. 13, No. 2, pp. 86-95, 1994.

9) 허만탁,남기곤,김재창,이종혁,김실중,윤태훈,박의열, "차분망을 이용한 확산필터 알고리듬의 개선 및 대역통과특성의 양자화 분석", 전자공학회논문지-B, 제33권, B편, 제7호, pp. 163-172, 1996.

10) H. Y-F. Lam, *Analog and Digital Filters*, Prentice Hall, 1979.

11) M. E. Van Valkenburg, *Analog Filter Design*, Holt Saunders, 1982.

12) 허만탁, *확산망과 차분연산에 의한 대역통과필터 구현과 특성해석*, 부산대학교, 공학박사학위논문, 1994.

13) 오영환, *음성언어론*, 형학사, pp. 169-264, 1994.

14) 유창돈, *현대 국어음운론*, 형설출판사, 1993.