

음성인식에 있어서 특징 파라미터의 기여도에 기반한 상태별 특징 파라미터 가중

State-Dependent Feature-Parameter Weighting By the Contribution of the feature parameter to the performance of the Speech Recognition

최 환 진*
(Hwan Jin Choi*)

요 약

본 논문에서는 은닉 마르코프 모델에 기반한 음성인식에 있어서 특징 파라미터의 인식 성능에 미치는 영향의 차이를 인식 시스템에 반영하여 인식성능을 향상시키기 위한 방법을 제안하였다. 특징 파라미터별 가중치를 유도하기 위해서 우선 상태별 특징 파라미터의 인식율에 대한 기여도를 가중치로 변환하고, 이를 특징 파라미터 각각의 상태에서의 출력확률에 곱하여 상태별 출력확률을 재 추정하게 된다.

실험결과, “가변가중”방법이 “고정가중”방법에 비해서 단어 인식의 경우 3.3%, 그리고 문장 인식율의 경우 5.3%의 성능향상을 보임으로써 상태별 특징 파라미터의 가중이 인식 성능 향상에 유효함을 알 수 있었다.

ABSTRACT

In this paper, we proposed the new approach to improve the performance of the speech recognizer based on the hidden Markov model by reflecting of the contribution of each feature parameter to the recognition rate.

To derive the feature-parameter weighting factor, we first transformed the contribution of each feature parameter for the recognition rate into the weighting value for the corresponding feature parameter, and then re-estimated the output probability by multiplying the weighting factor and the original output probability at a state.

Through several experiments, we found the effectiveness of the proposed method in that the variable-weighting method has shown the improvement of 3.3% at the word accuracy and that of 5.3% at the sentence accuracy over the fixed-weighting method.

I. 서 론

인간에 의해서 발화된 음성은 크게 화자의 특성을 나타내는 화자 정보(speaker information)와 화자가 전달하고자 하는 원래의 의도에 해당되는 언어적인 정보(language information)가 담겨져 있다. 화자 정보로는 기본 주파수, 포먼트의 케직등이 있으며[1, 2, 3], 이러한 정보는 발화하는 사람의 특성을 나타내는 것으로 화자인식(speaker recognition)을 위해서 사용되는 정보이다. 반면, 언어정보는 화자정보와는 무관하게 발화 대상자로부터 전달되는 정보로, 화자가 발성한 음성으로부터 화자가 의도한 내용을 문자의 형태로 변환하는 음성인식을 위해서 사용된다.

음성인식을 효과적으로 수행하기 위해서는 다양한 지식원들(knowledges)이 필요하다. 음성인식에 유효한 인식 모델을 만드는 것을 포함해서, 인간의 청각 특성등을 반영하여 음성인식에 유효한 특징들을 추출하거나 신호에 포함된 잡음등을 제거하며, 음향적인 해독과정에도 언어 모델등을 사용해서 발성에 따른 오류를 줄이는 방법들이 필요하게 된다. 이러한 음성인식에 유효한 지식원들은 인식을 위한 전 과정에서 각기 유효하게 적용되어 인식 성능을 향상시키기 된다.

실제적으로 음성인식에 있어서 음성신호의 특성을 표시하는 특징 파라미터의 선정은 음성인식에 있어서 매우 중요하다[4, 5]. 음성인식을 위한 초기 연구의 대부분은 음성신호를 효과적으로 표현하기 위한 방법들과 인식에 유효한 특징 파라미터의 추출과 관련되었으며, 현재에도 이러한 문제들은 꾸준히 연구의 주제가 되고 있다. 음성인식을 위해서 현재 널리 사용되는 특징 파라미터로는 선

*자연어정보처리연구부 시스템공학연구소
접수일자: 1997년 10월 13일

형예측계수(LPC: linear predictive coefficient), 캡스트럼 계수, 멜 캡스트럼(mel cepstrum), 대수 에너지(log energy) 등이 사용되고 있으며, 이러한 특징 파라미터들 이외에 이들 특징 파라미터의 시간적인 변화특성을 나타내는 차분 멜 캡스트럼(differential mel cepstrum), 차분 에너지(differential energy) 등이 사용된다[6, 7, 8]. 인식에 사용되는 특징 파라미터는 각기 고유의 특성을 가지고 있으며, 그러한 특성은 인식성능의 향상에 유효한 특성을 가지고 있어야 한다.

음성인식을 위한 유효한 특징 파라미터의 선정과 그러한 특징 파라미터를 인식 시스템에서 어떻게 표현할 것인가에 따라 인식기의 성능이 크게 좌우된다. 앞서 설명하였듯이, 특징 파라미터가 음성신호의 다른 측면을 나타냄으로 가급적 중복된 정보를 나타내는 특징 파라미터를 여러개 사용하는 것보다는 인식에 유효한 특성을 다른 측면에서 표현하는 특징 파라미터를 다수 사용하는 것이 바람직하다. 하나의 특징 파라미터가 음성신호의 특성을 완전히 표현할 수 없으므로, 음성신호가 가지는 특성을 폭 넓고 상세히 표현할 수 있는 특징 파라미터들을 효과적으로 선택해야 한다. 이러한 특징 파라미터의 선택의 문제 이외에, 인식에 사용될 특징 파라미터들을 어떻게 인식을 위해서 표현할 것인가가 하는 표현의 문제가 남는다. 일반적으로 특징 파라미터는 일정 차원 수를 갖는 벡터의 형태로 표현된다. 따라서, 다수의 특징 파라미터를 사용하는 경우, 다수의 특징 파라미터들을 하나의 특징 벡터로 만들어서 사용할 것인가, 아니면 다수의 특징 파라미터들을 독립적으로 취급할 것인가 하는 표현의 문제가 있게 된다. 일반적으로, 다수의 특징 파라미터들을 독립적으로 사용하는 것이 하나의 단일벡터 형태로 표현하는 것보다 좋은 성능을 나타내는 것으로 알려져 있다[9].

음성인식을 위해서 사용될 특징 파라미터와 관련된 또 다른 고려점은 인식을 위해서 선택된 특징 파라미터의 중요성(weighting)이다. 음성신호의 특성을 나타내는 특징 파라미터들은 음성신호의 각기 다른 특성을 표현하고 있으며, 그러한 특성은 음성인식에 사용될 인식단위에 따라 중요성이 달라지므로, 그러한 차이가 특징 파라미터의 출력확률 계산 시에 반영된다면, 인식모델의 변별력을 보다 향상시킬 수 있을 것이다. 본 연구에서는 특징 파라미터의 가중을 위해서, 은닉 마르코프 모델을 구성하는 상태단위로 특징 파라미터의 중요도를 계산하고 이를 이용하여 상태별 출력확률의 계산 시 특징 파라미터별로 가중치를 반영하는 방법을 제안하고자 한다.

일반적인 은닉 마르코프 모델을 이용한 음성인식에 있어서, 다수의 특징 파라미터들이 사용되고 있으며, 인식모델의 한 상태에서의 출력확률은 특징 파라미터 각각의 입력벡터에 대한 해당 상태에서의 출현확률의 곱의 형태로 표현된다. 이때, 각 특징 파라미터의 중요도는 동일한 가중치로 출력확률의 계산과정 시에 반영됨으로써 은닉

마르코프 모델에서의 상태단위로 특징 파라미터의 중요도가 반영되지 않게 된다. 본 연구에서는 이러한 문제를 해결하기 위해서 퍼지 목적 함수를 이용해서 상태별 특징 파라미터의 가중치를 유도하고, 그러한 가중치를 입력의 시점에 입력된 각 특징 벡터의 현 상태에서의 출력확률값에 대한 로그 함수값에 선형 결합하는 방식으로, 특징 파라미터의 출력확률을 가중하는 방법인 상태 의존 특징 파라미터 가중(SD-FPW: state-dependent feature-parameter weighting)방식을 제안하고자 한다. 제안된 방법을 이용해서 실험한 결과, 기존의 동일한 가중치를 가지고 상태별 출력확률을 추정하는 방법에 비해서 보다 향상된 인식율을 보였으며, 음소별로 서로 다른 상태별 가중치를 나타낼 수 있었다.

본 논문의 구성은 다음과 같다. 2장에서는 연구의 동기와 기존의 방법이 갖는 문제점들에 대해서 살펴보고, 3장에서는 본 논문에서 제안한 퍼지 목적 함수를 이용한 상태별 특징 파라미터 가중 방법을 중심으로 설명한다. 4장에서는 제안된 방법과 기존 방법과의 성능 평가를 수행하고, 아울러 결과에 대해서 기술한다. 마지막 5장인 결론에서는 상태별 특징 파라미터의 가중에 따른 의의와 향후 연구방향에 대해서 기술하고자 한다.

II. 음성인식을 위한 특징 파라미터의 특성

일반적으로, 음성인식의 과정에서 사용되는 특징 파라미터를 인식기의 성능이라는 측면에서 보면 크게 두 가지 - "특징 파라미터의 표현 (representation of a feature parameter)"과 "특징 파라미터의 중요도 (degree of importance of a feature parameter)" - 측면에서 고려해 볼 수 있다[10]. "특징 파라미터의 표현"이라는 것은 음성인식에 사용될 특징 파라미터를 단일벡터 형태로 할 것인가, 아니면 독립적인 다수의 특징 파라미터들을 각기 독립적으로 취급할 것인가를 의미한다. 다수의 실험결과에 의하면, 다수의 특징 파라미터들을 독립적으로 취급하는 것이 단일 형태로 만들어 사용하는 경우에 비해서 보다 우수한 성능을 나타내는 것으로 알려졌다[11].

다음으로, "특징 파라미터의 중요도"는 특징 파라미터가 인식기의 성능 향상에 얼마나 영향을 미치는가를 나타낸다. 이러한 특징 파라미터의 중요도는 다시 크게 두 가지로 나뉘는데, 특징 파라미터가 음성인식을 위해서 적합한가를 나타내는 "적합성(feasibility)"과 이 특징 파라미터가 상대적으로 다른 특징 파라미터에 비해서 얼마나 유효한지를 나타내는 "유효성(effectiveness)"으로 나누어 볼 수 있다[10]. "적합성"은 "특징 파라미터의 선택(selection of a feature parameter)"이라는 문제와 연관된 것으로, 음성인식에 유효한 다수의 특징 파라미터들로부터 좋은 성능을 보이는 특징 파라미터를 선택하는 것과 관련되어 있다. 음성인식의 초기 연구는 대부분 음성 인식기에 대한 특징 파라미터의 적합성에 관한 연구들로

음성을 표현하는 많은 특징 파라미터들 중에서 음성인식에 유효한 특징 파라미터를 찾자 하는데 그 목적이 있다.

특징 파라미터의 중요도에 따른 두번째 측면인 “유효성”은 인식기에서 사용될 유효한 특징 파라미터들이 결정되었다는 가정 하에서 선택된 특징 파라미터들이 인식에 얼마나 유효한지를 나타낸다. 음성인식에 사용되는 특징 파라미터 각각은 인식에 유효하나, 하나의 특징 파라미터가 모든 음성신호의 특성을 대표할 수 없으므로 각기 고유한 특성을 가지게 된다. 이러한 특성으로 인해 신호성분을 해석하는데 있어서 한 특징 파라미터가 다른 특징 파라미터보다 신호를 더 잘 표현할 수도 있게 된다. 즉, 특징 파라미터가 동일한 신호에 대해 서로 다른 “유효성”을 가질 수 있음을 나타낸다고 할 수 있다.

이러한 “유효성”은 특징 파라미터의 변별력을 나타내는 것으로 입의의 신호성분에 대해서 그러한 신호성분을 보다 더 잘 표현할 수 있는 특징 파라미터가 그렇지 못한 특징 파라미터에 비해서 보다 가중되어야 함을 나타낸다. 특징 파라미터의 “유효성”에 따른 특징 파라미터의 가중을 위한 제안된 방법에 대해서는 아래의 3장에서 상세히 다루고자 한다. 다음으로 앞서 설명한 “특징 파라미터의 표현”과 “특징 파라미터의 중요도”라는 특징 파라미터가 갖는 두 가지 측면에서 보다 상세히 알아보도록 하자.

2.1 특징 파라미터의 표현

온닉 마르코프 모델을 이용한 음성 인식모델의 생성시, 다수의 특징 파라미터를 사용하는 것이 단일 특징 파라미터를 사용하는 것에 비해서 인식율을 향상시키는 것으로 알려져 있다. 멜 캡스트럼 계수, 에너지등과 같은 정적인 특징 파라미터 이외에, $n(n > 1)$ 개의 프레임에 걸친 특징 파라미터의 변화특성을 나타내는 동적 특징 파라미터인 차분 에너지, 차분 멜 캡스트럼 계수등을 함께 사용하는 편이 보다 향상된 인식율을 나타내는 것으로 알려졌다.

이러한 다수의 특징 파라미터들을 사용하는 방식은 크게 두 가지로 나눌 수 있다. 첫 번째로, 다수의 파라미터들을 하나의 특징 파라미터로 병합하여 사용하는 경우와 각기 독립적인 형태로 사용하는 방법이 있다. 예를 들어서, 인식을 위해서 14차 멜 캡스트럼 계수와 14차 차분 멜 캡스트럼 계수, 그리고 1차 에너지 계수와 1차 차분 에너지 계수등 4개의 특징 파라미터를 사용하였다고 가정하자. 하나의 특징벡터로 병합하여 사용하는 경우에는 음성신호의 분석구간인 프레임마다 30차원의 특징 파라미터 벡터를 생성하는 경우에 해당된다. 이러한 방법과는 달리, 특징 파라미터 각각을 별도로 사용하는 방법이 있다. 이 방법은 Gupta에 의해서 제안된 방법으로[11], 성능적인 측면에서 다수의 특징 파라미터들을 하나의 특징 파라미터 벡터로 사용하는 방법에 비해서 독립적으로 사용하는 것이 보다 좋은 성능을 나타내는 것으로 알려져

있다. 다수의 특징 파라미터를 별도로 사용하는 방법은 다수의 특징 파라미터들을 하나의 특징벡터로 사용하는 경우에 비해서 다음과 같은 장점들을 갖는다.

- 복수의 특징 파라미터를 사용하는 것이 양자화에 따른 오류를 줄인다.
- 동적인 범위와 정밀도가 높다.

다수의 특징 파라미터들을 하나의 특징 벡터로 사용하는 경우에 각 특징 파라미터별로 오류를 최소화한 양자화된 값을 선택하는 것이 아니라, 하나의 특징 벡터로 묶인 모든 특징 파라미터들에 대해서 오류를 최소화하는 단일 특징벡터를 선택함으로써 양자화 오류가 커지게 된다. 이에 반해서, 복수의 특징 벡터들을 독립적으로 사용하는 경우에는 각각의 특징 파라미터별로 양자화 함으로 오류를 줄일 수 있게 된다. 두 번째로, 코드북의 크기와 사용되는 파라미터 수에 대한 조합 수만큼 복수 특징 벡터를 사용하는 경우에는 선택 가능함으로, 보다 상세한 변별력이 있게 됨으로써, 병합된 단일 특징벡터를 사용하는 것에 비해서 복수의 특징벡터를 사용하는 것이 바람직하다.

2.2 특징 파라미터의 중요도

특징 파라미터가 인식기의 성능 향상에 얼마나 영향을 미치는가를 나타내는 “특징 파라미터의 중요도”는 앞서 설명한 특징 파라미터의 인식기 내에서의 표현보다도 중요하다. 그 이유는 실제적으로 인식기의 성능에 영향을 미치는 특징 파라미터의 선택과 연관되어 있기 때문이다.

음성인식을 위한 초기의 연구는 음성인식에 유효한 특징 파라미터를 선정(selection of feature parameters)하는 문제가 대부분이었다. 다양한 특징 파라미터들이 제안되었으며, 인식방향에 대한 다양한 성능평가를 통해서 음성인식에 유효한 특징 파라미터들이 선택되었다. 현재는 주파수 특성을 나타내는 정적인 특징 파라미터와 동적인 특징 파라미터들이 혼합되어 음성인식에서 사용되고 있다[12, 13].

음성인식을 위해서 사용할 특징 파라미터들이 선정되었다고 하여도, 여전히 고려할 문제는 남는다. “어떠한 특징 파라미터가 얼마나 인식에 유효한 것인가?”. 음성인식을 위해서 사용되는 특징 파라미터는 각기 고유한 특성을 가지며, 이러한 특성은 하나의 특징 파라미터가 어떤 측면에서 다른 특징 파라미터보다 음성신호를 잘 표현할 수 있음을 나타내며, 이는 특징 파라미터가 동일한 신호에 대해서 서로 다른 “유효성”을 가질 수 있음을 나타낸다고 할 수 있다. “유효성”은 특징 파라미터의 변별력을 보다 향상시키기 위해서 사용할 수 있으며, 입의의 신호성분에 대해서 그러한 신호성분을 보다 더 잘 표현할 수 있는 특징 파라미터가 그렇지 못한 특징 파라미터에 비해서 보다 가중될 수 있음을 의미한다고 할 수 있다.

특징 파라미터의 각기 다른 유효성을 살펴보기 위해서, 20명의 화자가 각각 5분장씩 발성한 총 100분장용 대상으로 각 특징 파라미터의 유효성을 위한 실험을 수행하였다. 그 실험결과는 그림 1과 같다.

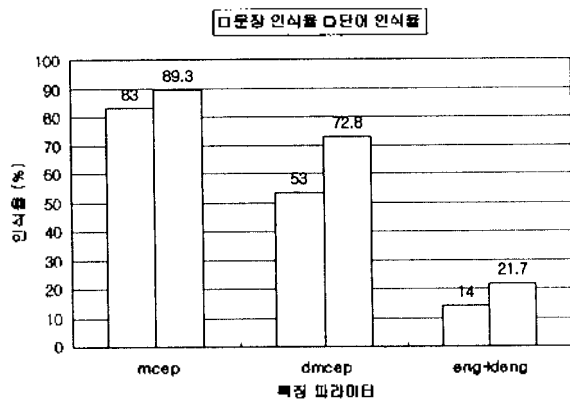


그림 1. 특징 파라미터별 유효성

위의 그림 1은 특징 파라미터 각각이 인식기의 성능에 미치는 영향을 측정하기 위해서 수행한 실험 결과이다. 본 연구에서 사용되는 특징 파라미터들 전부를 사용해서 100%의 인식율을 갖는 학습 자료들에 대해서 단일 특징 파라미터를 사용하였을 경우의 각각의 성능을 표시한다. 멜 캡스트림이 가장 좋은 성능을 보였으며, 차분 멜 캡스트림, 그리고 에너지 순으로 인식에 유효함을 보였다. 이러한 실험결과로부터, 특징 파라미터가 인식기의 성능에 미치는 영향이 다르며, 그러한 차이가 인식기의 성능에 반영되어야 함을 알 수 있다. 개별적인 특징 파라미터의 사용에 따른 인식기의 성능 이외에, 복합적인 특징 파라미터들의 사용에 따른 인식의 성능을 살펴보기로 하자. 아래의 그림 2는 다수의 특징 파라미터들을 함께 사용한 경우, 인식기의 인식 성능을 측정한 결과를 보이고 있다.

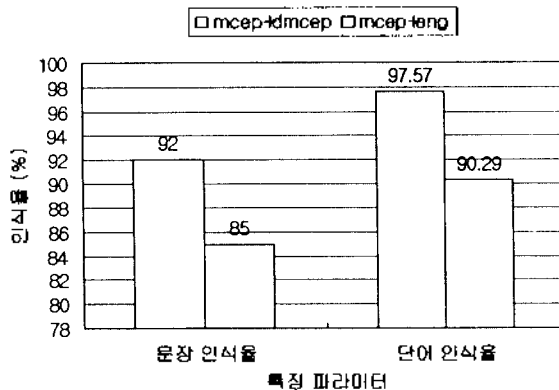


그림 2. 특징 파라미터들의 결합에 따른 인식성능

멜 캡스트림과 차분 멜 캡스트림을 복합적으로 사용하는 경우에 97.57%의 단어 인식율을 나타냄으로, 단순히 멜 캡스트림 특징 파라미터만을 사용하는 경우에 비해서 8.3%의 성능 향상을 보였다. 이러한 결과로부터, 멜 캡스트림과 차분 멜 캡스트림을 함께 사용한 경우, 차분 멜 캡스트림의 사용에 따른 유효성이 8.3%임을 알 수 있다. 아울러, 이러한 두가지 특징 파라미터를 복합적으로 사용하는 경우 이외에, 에너지를 함께 사용하는 경우의 인식율이 100%라고 할 때, 에너지 파라미터에 따른 성능 향상율은 약 2.5%정도임을 알 수 있다. 두 번째로, 멜 캡스트림과 에너지를 복합적으로 사용하는 경우에 90.29%를 나타냄으로써 단순히 멜 캡스트림만을 사용하는 경우에 비해서, 1%정도의 향상이 있음을 알 수 있다. 세가지 특징 파라미터를 복합적으로 사용하는 경우의 인식율이 100%라고 할 때, 차분 멜 캡스트림의 성능 향상율은 약 9.7% 정도라고 생각할 수 있다. 두 개의 실험결과로부터, 멜 캡스트림과 차분 멜 캡스트림을 복합적으로 사용하는 경우가 멜 캡스트림과 에너지를 복합적으로 사용하는 경우에 비해서 문장 인식율에서 7%의 향상율을 보이며, 단어 인식율적인 측면에서도 7.3%의 성능 향상을 보임을 알 수 있다. 이러한 결과를 종합해보면, 음성 인식에 미치는 특징 파라미터의 평균적인 기여도는 대략 5:3.5:1.5의 비율을 나타내는 것을 알 수 있다. 특징 파라미터의 인식기에 미치는 기여도의 차이가 서로 다르기 때문에 이러한 차이를 인식기에 반영하는 것이 필요하다. 본 연구에서는 이러한 예비 실험결과를 통해서, 인식 모델에서 특징 파라미터 각각의 "유효성"에 기반하여 인식 모델 내에서 특징 파라미터를 가중하는 방법을 제안하고자 한다.

III. 퍼지 기여도에 의한 특징 파라미터 가중

앞서 설명한 바와 같이, 특징 파라미터의 인식성능에 미치는 영향을 인식기에 반영하기 위해서 어산형 은닉 마르코프 모델을 구성하는 상태별로 특징 파라미터 각각의 상태에서의 중요도를 특징 파라미터별 출력확률에 가중하여 최종적인 상태별 출력확률을 추정하는 방법을 제안하고자 한다. 먼저, 상태별 특징 파라미터의 가중치를 적용하여 상태별 출력확률을 추정하는 추정식에 대해서 설명하고, 다음으로 특징 파라미터의 가중치를 유도하는 방법에 대해서 기술하고자 한다.

3.1 특징 파라미터 가중의 정의

상태별 특징 파라미터의 출력확률에 가중된 가중치는 상태에서의 특징 파라미터의 출력확률과 밀접한 관련을 가진다. 상태에서 높은 출력확률을 가진 특징 파라미터가 그렇지 못한 특징 파라미터에 비해서 높은 출력확률을 가지는 것이 바람직하다. 따라서, 특징 파라미터에 가중될 가중치는 상태별 특징 파라미터의 출력확률에 기반한다. 상태에서 주어진 특징 파라미터들에 대한 출력확

률의 일반적인 추정식은 다음과 같다.

$$P_s(t) = \prod_{p=1}^P \Pr(\vec{x}_p(t)|s) \quad (1)$$

이산형 은닉 마르코프 모델에서 다중 특징 파라미터들을 사용하는 경우에 상태 s 에서 각 특징 파라미터별로 상태에서의 출력확률을 구하고, 이들의 곱의 형태로 정의된다. 이러한 식에서 적은 출력값을 갖는 특징 파라미터의 출력확률에 대해서 곱 연산을 수행하는 경우, 그 값이 컴퓨터에서 처리할 수 있는 값의 범위 이하로 작아짐으로써 계산상의 오류를 유발할 수 있게 된다. 이러한 문제를 해결하기 위해서 일반적으로 특징 파라미터의 출력확률값에 로그함수를 적용해서 확률값의 곱의 형태가 아닌 확률에 대한 로그값의 합의 형태로 표현한다. 위의 식에 대한 로그함수를 적용한 식은 아래와 같다.

$$\log P_s(t) = \sum_{p=1}^P \log \Pr(\vec{x}_p(t)|s) \quad (2)$$

본 연구에서는 상태별 출력확률의 추정에서 각 특징 파라미터의 상태에서의 중요도를 각 특징 파라미터별 출력확률의 로그형태에 대한 곱의 형태로 가중하도록 하였다. 변형된 수식은 아래의 식 3과 같다.

$$\log P_s(t) = \sum_{p=1}^P [\log P(\vec{x}_p(t)|s) \times w_s(p)] \quad (3)$$

여기서, 상태별 특징 파라미터의 가중치를 나타내는 $w_s(p)$ 는 상태 s 에서 p 번째 특징 파라미터의 가중치를 나타내며, 다음과 같은 제약조건을 가진다.

$$\sum_p w_s(p) = 1 \quad (4)$$

위의 조건에 의해서, 수식 3에서 한 상태 s 에서 상태에서의 출력확률은 각 특징 파라미터의 출력확률에 대한 가중된 평균값의 형태를 취하게 된다. 수식 3을 원래의 특징 파라미터의 확률의 곱의 형태로 표시하면, 다음과 같은 형태가 된다.

$$P_s(t) = \prod_{p=1}^P P(\vec{x}_p(t)|s)^{w_s(p)} \quad (5)$$

위의 수식에서 알 수 있듯이, 상태에서 특징 파라미터에 대한 가중치는 특징 파라미터의 출력확률의 지수의 형태로 표시됨으로써, 중요도가 높을수록 특징 파라미터 확률의 지수배로 가중된다. 다음으로, 특징 파라미터에 대한 상태별 가중치를 유도하는 방법에 대해서 알아보도록 하자.

3.2 상태 가중을 위한 가중치의 유도

앞서 설명하였듯이, 특징 파라미터의 가중형태는 특징 파라미터의 출력확률의 지수형태로 최종적인 출력확률

로 반영된다. 특징 파라미터의 가중치를 구하기 위해서, 본 연구에서는 상태에서 특징 파라미터의 출력확률에 기반하여 특징 파라미터의 가중치를 구하게 된다. 본 연구에서는 특징 파라미터의 가중치를 상태에서 출력확률의 추정과정에 반영하는 두가지 방법-상태별로 고정된 특징 파라미터별 가중치를 적용하는 방법과 인식 시 매시점마다 동적으로 특징 파라미터의 가중치를 출력확률에 적용하는 방법-을 제안하고자 한다. 방법별 특징 파라미터의 가중치 추정방법은 다음과 같다.

3.2.1 고정형 특징 파라미터 가중치

상태별로 고정된 특징 파라미터별 가중치를 구하기 위해서, 임의의 시점에서 상태별로 모든 특징 파라미터들 중에서 가장 큰 출력확률을 갖는 특징 파라미터의 출력확률과 가중치를 구한다. 그 식은 다음과 같다.

$$p_s^*(t) = \operatorname{argmax}_p P(\vec{x}_p(t)|s), 1 \leq t \leq N(s) \quad (6)$$

여기서, $p_s^*(t)$ 는 상태 s 에서 총 $N(s)$ 개의 특징 파라미터 벡터가 주었다고 할 때, 그 가운데에서 t 번째 특징 벡터들 중에서 최고의 출력확률을 나타내는 특징 파라미터의 색인을 나타낸다. 상태 s 에서 p 번째 특징 파라미터가 가장 큰 출력확률을 가진 빈도수는 $N_p(s)$ 라고 표시하고, 누적 확률값은 $N_p(s)$ 라고 표기한다. 상태별로 각 특징 파라미터의 평균적인 최고 출력 확률값은 다음과 같다.

$$\bar{N}_p(s) = \frac{N_p(s)}{N(s)}, 1 \leq p \leq P \quad (7)$$

이렇게 구한 평균적인 상태별 특징 파라미터의 출력확률들을 상태에서의 가중치로 사용하기 위해서 퍼지 목적 함수를 사용한다. 퍼지 목적 함수는 퍼지 벡터 양자화를 위해서 사용되는 방법으로, T 개의 특징벡터에 대해서 주어진 유한개의 대표적인 특징 파라미터들과의 거리에 대한 가중치의 최적값을 구하기 위해서 사용된다. 일반적으로 사용되는 퍼지 목적함수의 식은 아래와 같다.

$$D = \sum_{i=1}^T \sum_{j=1}^M m_{r,i}^f \cdot d(\vec{x}_i, \vec{c}_j) \quad (8)$$

여기서, M 은 총 코드워드의 수, F 는 퍼지의 정도(degree of fuzzy)를 나타내며, $F > 1$ 의 값을 가진다. 퍼지 목적 함수 D 를 최소화하는 입력패턴 \vec{x}_i 의 i 번째 코드워드 벡터 \vec{c}_j 에 대한 기여도인 $m_{r,i}$ 는 다음과 같은 제약조건을 가진다.

$$\sum_{i=1}^M m_{r,i} = 1 \quad (9)$$

위의 제약조건을 만족하는 가정 하에서 식 8의 퍼지 목적함수를 최소화하는 최적의 $m_{r,i}$ 에 해당되는 값의 추정은 Lagrange multiplier를 이용해서 아래의 식에 의해서 구해

진다.

$$m_{t,i} = \left[\sum_{j=1}^M |d(\vec{x}_t, \vec{c}_i) / d(\vec{x}_t, \vec{c}_j)|^{MF} \right]^{-1} \quad (10)$$

본 연구에서는 상태에서 특징 파라미터별 가중치를 계산하기 위해서, 위의 식 8에 식 7에서 구한 상태에서의 특징 파라미터별 평균 출력값을 적용해서 특징 파라미터별 가중치 값을 구하게 된다. 그 유도식은 아래와 같다.

$$w_s(i) = \left[\sum_{j=1}^P [\log N_s(i) / \log N_s(j)]^{MF-1} \right]^{-1}, 1 \leq i \leq P \quad (11)$$

이러한 과정에 의해서 구해진 상태에서 특징 파라미터의 가중치 $w_s(p)$ 의 값을 위의 식 3에 대입하여 가중된 특징 파라미터의 출력 값의 합의 형태로써, 상태별 출력확률들을 재 추정하게 된다.

3.2.2 가변형 특징 파라미터 가중

앞서 설명한 방법은 특징 파라미터의 가중치를 상태별로 구한 후, 계산된 가중치를 인식 시에 상태에서의 특징 파라미터의 출력확률들을 곱의 형태로 적용하는 방법이다. 이러한 방법과는 달리, 인식에서 특징 파라미터의 출력값을 특징 파라미터의 가중치 행렬을 이용해서 가중된 특징 파라미터의 출력 값을 구하고, 기존의 방법과 동일한 방법으로 상태별 출력확률을 계산한다.

이를 위해서, 상태별로 수집된 특징 파라미터들의 열을 대상으로 매 시점에서 최적 특징 파라미터의 빈도와 함께 특징 파라미터 각각의 누적확률을 계산하게 된다. 이를 위해서, (특징 파라미터 수 \times 특징 파라미터 수)의 크기를 갖는 누적확률 행렬 T 를 사용한다. 행렬 T 에서의 셀 $T[i, j]$ 는 최적 특징 파라미터가 i 일 때, j 번째 특징 파라미터의 누적 출력확률을 나타낸다. 상태 s 에서 최적 특징 파라미터 p 의 빈도를 $N_s(p)$ 라고 하면, 누적빈도로 정규화된 행렬 \bar{T} 는 다음과 같다.

$$\bar{T}_{i,j} = T_{i,j} / N_s(i), 1 \leq i, j \leq P \quad (12)$$

이렇게 구해진 정규화된 \bar{T} 를 이용해서, 특징 파라미터의 색인별로 앞서 설명한 퍼지 목적 함수를 이용해서 최적 특징 파라미터와 다른 특징 파라미터간의 동시 기여도를 계산하게 된다. 이렇게 구해진 각 특징 파라미터별 퍼지 기여도 벡터를 이용해서, 상태별 특징 파라미터의 출력 확률 값을 아래와 같이 평활화한다.

$$\log \hat{P}(\vec{x}_p(t) | s) = \sum_{k=1}^P \log P(\vec{x}_p(t) | s) \cdot \bar{T}_{p,k}, 1 \leq p \leq P \quad (13)$$

임의의 시점에서 입력된 각 특징 파라미터별 출력확률을 위의 식 13과 같이 평활화한 후, 기존의 상태별 출력확률의 추정식과 동일한 방식으로 상태별 출력확률을 계산한다.

$$\log P_s(t) = \sum_{p=1}^P \log \hat{P}(\vec{x}_p(t) | s) \quad (14)$$

본 절에서는 상태별 특징 파라미터의 가중을 위해서 상태별 특징 파라미터의 최적 출력확률의 평균값에 기반한 출력확률의 추정식을 제안하였다. 학습과정에서 추정된 고정된 상태별 특징 파라미터 가중치를 적용하여 상태별 출력확률을 계산하는 방법과 인식 시 상태별 특징 파라미터의 출력확률 값을 특징 파라미터의 가중행렬을 이용해서 평활화하고, 평활화된 특징 파라미터의 출력확률을 기존의 방법과 동일한 방법을 적용하여 상태별 출력확률을 추정하는 방법에 대해서 설명하였다. 이러한 제안된 방법을 이용한 인식 실험결과는 다음의 4장에서 기술하기로 한다.

4. 실험 및 결과

4.1 실험환경 및 음성 인식 시스템의 구성

본 연구에서는 제안된 방법들의 성능평가를 위해서 이산형 은닉 마르코프 모델에 기반한 연속 음성인식 시스템을 구성하였으며, 그 개요도는 아래의 그림 3과 같다.

이산형 은닉 마르코프 모델에 기반하여 음성인식을 수행하기 위해서는 크게 세가지 과정이 필요하게 된다. 첫 번째가 특징 파라미터들에 대한 양자화된 코드워드의 사상을 위해서 사용하는 코드북 생성과정이 있으며, 두 번째로 음성인식에 사용된 인식모델을 학습하고, 학습된 인식모델들을 접속해서 단어모델로 구성하는 학습단계가 있다. 그리고 마지막으로 실험단계가 있는데, 실험단계에서는 생성된 코드북과 학습과정에서 얻어진 인식모델을 이용해서 실제적인 인식을 수행하게 된다.

첫 번째 과정인 "코드북 생성"은 이산형 HMM에 기반한 음성 모델링을 위해서 반드시 필요한 단계로, 이산형 HMM을 구성하는 각 상태에서의 출력확률이 코드북을 구성하는 코드워드들의 이산형 확률 밀도 함수 형태로 표시 된다. 생성된 코드북으로부터 주어진 특징 벡터에 대응되는 최적 코드워드의 색인(index)을 얻기 위해서 주어진 특징 벡터와 코드워드 벡터간의 거리 계산을 위해 Euclidean 거리 척도가 사용된다[14]. 본 연구에서는 특징 파라미터에 대한 코드북 생성을 위해서 일반적으로 사용되는 방법인 LBG방법[15]을 사용하였으며, 코드북의 생성을 위한 최대 반복횟수(epoch)는 100, 종료조건은 이전 단계에서 얻어진 코드북과 총 특징 벡터들간의 평균 거리값과 현재 시점에서 얻는 평균 거리값과의 차이가 0.00001이하일 경우에 종료하였다. 음성인식을 위한 특징 파라미터로써 14차 멜켵스트럼, 14차 차분 멜 켈스트럼, 에너지와 차분 에너지의 2차 특징 파라미터등 총 세가지의 특징 파라미터들을 사용하였으며, 벡터 양자화를 위한 코드북의 크기는 멜켵스트럼과 차분 멜켵스트럼의 경우는 256씩을 사용하였으며, 에너지와 차분 에너지를 병

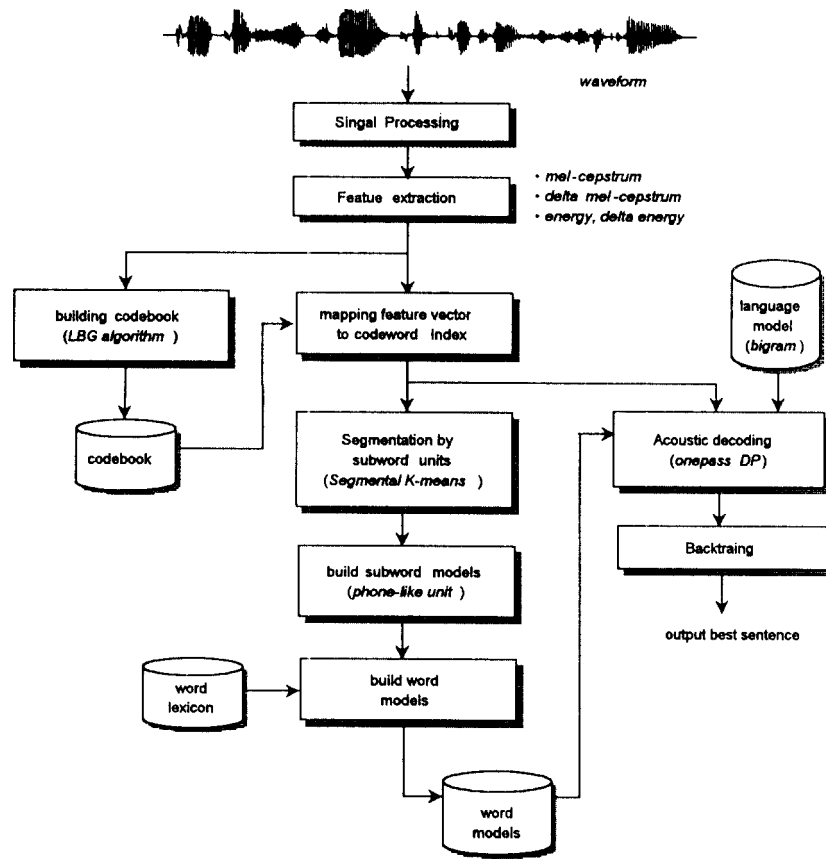


그림 3. Baseline 음성인식 시스템의 구성도

합한 특징 파라미터의 경우에는 128개를 각기 사용하였다.

두 번째 과정인 학습 과정에서는 인식에 사용될 인식 단위모델을 학습하고, 학습이 완료된 인식모델을 접속하여 인식 어휘모델을 구성하게 된다. 인식단위로는 triphone 모델을 기본 인식단위로 사용하였다[9, 13]. 인식 단위모델을 학습하기 위해서 segmental K-mean방법[16]을 사용하여 학습자료를 인식단위로 분할하였다. 인식단위별로 학습이 완료되면, 음성인식을 위해 어휘사전(lexicon)을 사용하여, 어휘별 인식모델을 생성하게 된다. 본 기본 시스템에서는 어휘사전에 들어있는 각 어휘의 발음표기를 사용하여 기본 인식단위인 triphone모델을 접속해서 어휘모델을 구성하였다.

세 번째 과정인 실험단계는 학습된 인식모델을 이용해서 실제적인 인식을 수행하는 단계로 특징 파라미터별 코드북과 실제적으로 인식할 단어형태의 어휘모델들, 그리고 인식대상 어휘간의 접속관계에 대해 제약을 위한 언어모델을 사용하여 화자에 의해서 발생된 음성에 대해서 단어열 형태의 인식결과를 얻게 된다. 기본 인식 시스템을 위해서 언어모델로는 bigram이 사용되며, 화자의 발성음에 대한 최적 단어열을 구하기 위해서 onepass DP (dynamic programming) 탐색방법[17]을 사용하여 가장 큰 출력확률을 가지는 경로를 구하고, 이로부터 최적 단

어열을 얻으므로써 최종적인 문장인식이 완료되게 된다. 인식 시스템의 성능평가를 위해 인식된 단어열과 인식할 단어열과의 정합을 통해서 삽입(insertion), 삭제(deletion), 대치(substitution) 및 정인식된 단어의 빈도를 구하고, 그러한 값들에 기반하여 단어 정확도(word accuracy)나 문장 인식율(sentence accuracy)등을 계산하였다[18].

본 논문에서는 제안된 방법들의 평가를 위해서 연속 발생된 문장 단위의 데이터를 사용하였으며, 90명의 화자에 의해서 발생된 1074단어로 구성되어 있다. 무역상단에 관련된 회화체의 문장들을 대상으로 수집된 초기 문장집합에 대해서 단어 조정, 문장의 발생 및 화자별 문장 배분등의 절차를 통해서 재구성된 문장들을 남성화자 100명이 평균적으로 98.3문장씩 발생하였으며, 녹음은 조용한 사무실 환경에서 Sennheiser HMD224X headset 마이크로폰을 사용하여 Ariel ProPort 656로 16KHz, 16bit 선형 PCM으로 A/D변환하였다. 음성자료의 구성은 표 1과 같다.

본 연구에서는 이러한 데이터베이스 중에서 빈도수가 높은 1074단어를 추출하고, 이러한 단어들로 구성된 학습문장과 실험문장을 대상으로 인식실험을 수행하였다. 음성인식을 위한 기본적인 인식단위로는 음소모델을 사용하며, 음소의 좌/우 문맥을 반영하기 위해서 triphone

표 1. 음성자료의 구성

항목	항목치	항목	항목치
샘플링율	16 KHz	분석창 이동길이	10 msec
샘플당 할당비트	16 bits	인식 어휘수	1074
pre-emphasis	$1.0 - 0.972^{-1}$	단어당 평균 음소수	6.04
프레임 길이	20 msec	음소모델수	224 triphone

표 2. 학습 및 실험자료의 구성

자료집합	문장 수	화자 수	복잡도(perplexity)
학습자료	2981	55	5.6
실험자료	1322	35	4.0
총합	4203	90	6.2

을 사용하였다. 부적절한 triphone 모델링에 의한 성능저하를 방지하기 위해서, triphone 축소 규칙(reduction rule) [19]을 이용하여 triphone 수를 224개로 줄여서 사용하였다. 본 논문에서 사용된 음성 데이터베이스의 구성은 표 2와 같다.

4.2 특징 파라미터 가중치의 특성

먼저, 상태에서 특정 파라미터들간의 특성을 고찰해보기로 하자. 아래의 그림 4는 특정 파라미터별 평균 중요도를 표시한 것이다.

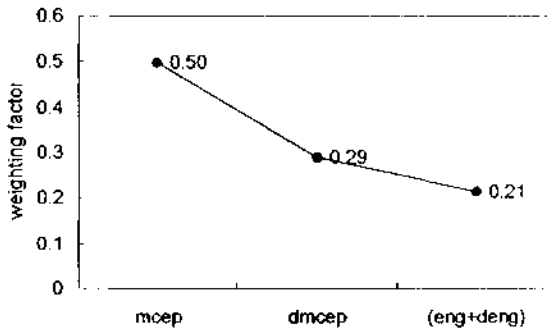


그림 4. 특정 파라미터별 평균 가중비율

한 상태에서 특정 파라미터의 가중치의 합을 1이라고 할 때, 음성 인식을 위해서 사용되는 인식모델의 모든 상태에서의 평균적인 특정 파라미터의 가중치는 아래의 그림에서 알 수 있듯이, 멜켵스트림이 0.5, 차분 멜켵스트림이 0.3, 그리고 (에너지+차분 에너지)가 0.2의 가중비율을 갖는다. 이러한 결과는, 앞의 4.2절에서 예비 실험을 통해서 얻은 특정 파라미터의 가중 비율인 5:3.5:1.5의 비율과 거의 유사한 특성을 갖는 것으로, 차분 멜켵스트림의 가중치가 다소 감소하고, (에너지+차분 에너지)의 가중치가 다소 증가한 점만이 차이가 있다.

다음으로, 상태별 특정 파라미터의 가중치의 특성을 살펴보도록 하자. 아래의 그림 5은 상태별 특정 파라미터 가중치의 추이선을 보이고 있다. (에너지+차분 에너지)에 해당되는 특정 파라미터의 경우, 상태에 따른 가중치의 변화가 없다. 즉, 모든 상태에서 0.2의 가중치를 갖는다. 반면, 멜켵스트림과 차분 멜켵스트림의 경우에는 다소 상태별 차이가 있다. 멜켵스트림의 경우, 가운데 상태에서 가장 큰 가중치를 가지며, 처음 상태보다는 마지막 상태가 보다 큰 가중치를 갖는다. 이러한 특성은 음소모델에서 멜켵스트림의 경우는 가운데 상태에 머무르는 신호특성을 다른 음소와의 접합부에 해당되는 상태보다 강조하며, 다른 음소와의 빈이 특성을 가지는 처음 상태와 마지막 상태의 경우에는 마지막 상태에서의 멜켵스트림이 처음 상태에서보다 강조되어야 함을 알 수 있다. 반면, 차분 멜켵스트림의 경우는 멜켵스트림과 반대의 현상을 보임을 실험결과로부터 알 수 있었다. 차분 멜켵스트림의 경우는 중간 상태에서 가장 낮은 가중치를 보이고, 상태의 처음과 마지막 부분에서 높은 가중치를 보였으며, 초기상태에서 가장 높은 가중치를 나타내었다. 이러한 결과로부터 차분 특징 파라미터는 천이특성을 갖는 음소 모델의 상태에서 유효하며, 음소 모델의 초기 상태의 가중에 있어서는 중요한 역할을 함을 알 수 있다.

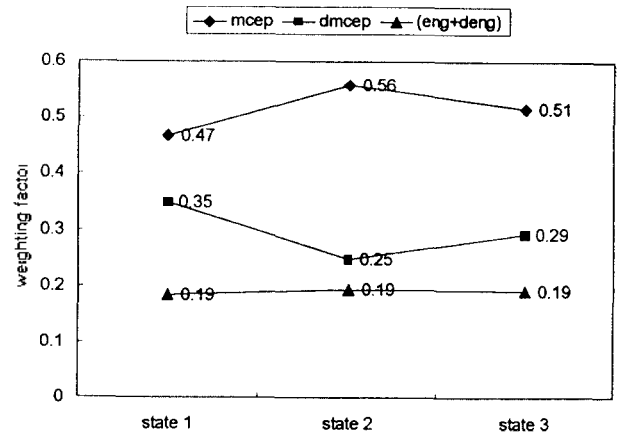


그림 5. 상태별 가중치의 변이

4.3 퍼지목적함수를 위한 퍼지 인수의 선정

본 절에서는 상태에서 특정 파라미터 가중을 위해서, 가중 인수를 유도하기 위해서 퍼지 목적 함수를 사용한다. 특정 파라미터를 위한 가중치의 유도에 있어서 사용되는 퍼지 목적함수는 퍼지정도(degree of fuzzy)라는 F 파라미터에 의해서 다른 형태의 퍼지 목적함수 결과를 가지게 된다. 따라서, 적절한 퍼지 파라미터를 정함으로써, 인식기의 성능을 향상시킬 수 있어야 한다. 이러한 퍼지 파라미터의 일반적인 특성은 $F > 1$ 이며, F의 값이 클

수록 완만한 형태를 가지게 된다.

본 장에서의 실험을 위해서 사용될 퍼지 파라미터 F 를 위한 적절한 값을 선정하기 위해서 F 의 값을 변경하면서 인식 실험을 수행하였다. F 의 값을 1.15, 1.25, 1.5의 세가지 값으로 변경하면서, Baseline시스템에 상태별 특징 파라미터의 가중 실험을 수행하였다. 퍼지 파라미터 F 에 따른 인식기의 성능은 그림 6과 같다.

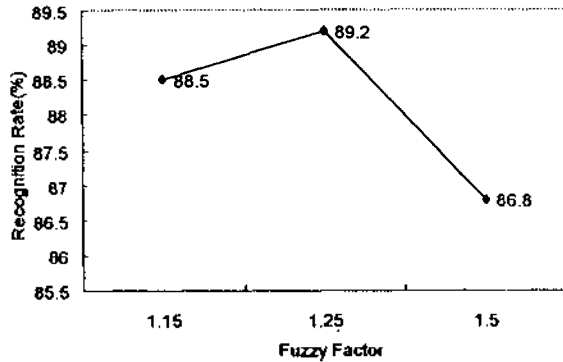


그림 6. 퍼지인수에 따른 성능평가

실험결과에서 알 수 있듯이, F 의 값으로 1.25를 사용하는 경우가 1.25보다 작은 1.15를 사용하는 경우에 비해서 좋은 결과를 나타내었으며, 1.25보다 큰 1.5의 경우는 약 1.7% 정도의 성능저하를 보였다. 이러한 결과를 기반으로 해서, 본 연구에서는 상태별 특징 파라미터의 가중을 위한 퍼지 목적함수의 퍼지 파라미터 값인 F 값을 1.25로 정하여 실험하였다. 다음으로, 상태별 특징 파라미터의 가중에 따른 Baseline 시스템과 제안된 다양한 가중방법들에 따른 실험결과에 대해서 기술하고자 한다.

4.4 실험 및 결과

제안된 방법의 성능 평가를 위해서 본 장에서 제안된 두가지 상태 가중방법들과 Baseline 인식 시스템과의 인식 실험을 수행하였다. 제안된 첫 번째 방법은 “고정 특징 파라미터 가중(FFPW: fixed feature parameter weighting)”으로, 학습과정에서 상태별 특징 파라미터의 가중치를 구하고, 이 값들을 실제 인식 시에 단순히 상태별 특징 파라미터의 로그 함수 값들과 선형 결합하여 최종적으로 상태에서의 가중된 출력확률을 유도하는 방법이다. 두 번째 방법은 “가변 특징 파라미터 가중(VFPW: variable feature parameter weighting)”으로 기존의 출력확률 추정 방법과 동일하나, 차이점은 상태에서 특징 파라미터별 출력확률을 평활화한다는 점이다. 최대의 출력확률을 갖는 특징 파라미터와 함께 나올 수 있는 다른 특징 파라미터의 출력 확률값을 추적하여 얻어진 가중행렬을 이용해서, 상태별 특징 파라미터의 출력확률을 평활화하고, 이러한 값들을 기존의 상태별 출력확률을 추정하는 방법과

동일한 방법으로 상태별 출력확률을 계산한다.

상태별 가중의 인식 성능에 미치는 결과를 알아보기 위해서, 기존의 Baseline 인식 시스템에 제안된 상태별 특징 파라미터 가중방법을 결합하여 인식실험을 수행하였다. 상태가중을 위해서 사용한 퍼지 목적함수를 위한 퍼지 파라미터 값이 F 은 앞의 소절에서 얻은 결과를 바탕으로, 1.25의 값을 사용하였다. 특징 파라미터의 상태별 가중에 따른 인식 실험결과는 그림 7과 같다.

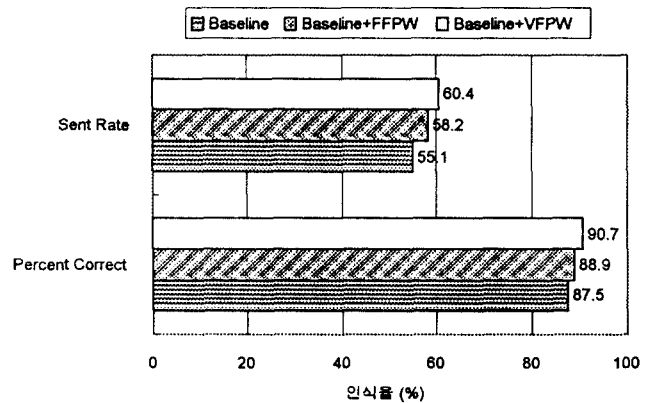


그림 7. 특징 파라미터 가중에 따른 인식기의 성능

그림에서 알 수 있듯이, Baseline 인식 시스템에 비해서, 제안된 상태가중 방법들이 보다 향상된 인식율을 보였으며, FFPW방법보다는 VFPW방법이 보다 유효함을 알 수 있었다. “Baseline + VFPW”방법은 기존의 Baseline 방법에 비해서, 단어 인식율 면에서 3.3%의 향상을 보였으며, 문장 인식면에서는 5.32%의 인식율을 보였다. FFPW를 적용한 방법에 비해서는 단어 인식율 면에서 1.6%, 문장 인식율 면에서 2.2%의 향상을 보였다. 이러한 결과로부터, 상태별 특징 파라미터의 가중이 인식율의 향상에 유효하고 상태에서 출력확률의 가중 시, 단순히 특징 파라미터별 가중치만을 적용하기보다는 가중에 있어서 다른 특징 파라미터도 함께 고려하는 방법이 보다 인식율 향상에 유효함을 알 수 있었다.

V. 결 론

본 논문에서는 음성인식에 있어서 특징 파라미터의 인식 성능에 미치는 영향의 차이를 인식 시스템에 반영하여 인식성능을 향상시키기 위한 방법으로, 상태별로 특징 파라미터의 가중치를 계산하고 그러한 가중치 값을 상태에서 출력확률의 추정 시에 특징 파라미터 각각의 상태에서의 출력확률에 가중하여 상태별 출력확률을 재추정하는 방법을 제안하였다.

상태에서의 특징 파라미터 가중을 위해서 본 장에서는 두가지 방법을 제안하였는데, 학습 시에 상태에서의 특

정 파라미터 가중치를 계산하고, 그 값을 실제적인 인식 시에 상태별 특징 파라미터의 출력확률에 가중하여 계산하는 방법인 "고정가중방법"과 상태별로 특징 파라미터의 출력확률을 가중행렬을 이용하여 재계산하고, 그러한 재 계산된 특징 파라미터의 출력 확률값을 사용하여 기존의 방법과 같은 형식으로 최종적인 상태별 출력확률을 추정하는 "가변가중방법"을 제안하였다. 실험결과, "가변가중방법"이 "고정가중방법"에 비해서 단어 인식의 경우 3.3%, 그리고 문장 인식율의 경우 5.3%의 성능향상을 보임으로써 상태별 특징 파라미터의 가중이 인식 성능 향상에 유효함을 알 수 있었다.

향후 연구방향으로는 상태 가중을 위한 연구가 수행되어야 할 것으로 생각된다. 인식모델에서 상태의 유효성은 음성신호의 특성에 따라 달라지게 된다. 은닉 마르코프 모델에 기반한 음성인식의 경우, 모델은 상태들로 구성되며 그러한 모델을 구성하는 상태들이 인식에 있어서 동일한 기여를 하는 것은 아니다. 모델링 대상의 음성신호의 특성에 따라, 인식에 유효한 상태들이 있으며, 그러한 상태들이 그렇지 않은 상태들에 비해서 인식에 보다 기여하므로 그들의 출력확률이 보다 가중되어야 한다. 이러한 상태가중은 혼동 가능한 인식모델들간의 변별력을 향상시켜, 인식기의 인식율을 향상시킬 것으로 기대됨으로써, 인식 모델에서의 상태별 가중을 위한 연구가 필요할 것으로 판단된다.

참 고 문 헌

1. L. R. Rabiner, B. H. Juang, Fundamentals of Speech Recognition, Prentice-Hall, 1993.
2. P. Brown, The acoustic-modeling problem in automatic speech recognition, PhD Thesis, Computer Science Department, Carnegie Mellon University, May, 1987.
3. K. F. Lee, et al., "An overview of the SPHINX Speech Recognition System," IEEE trans. on ASSP, Vol. 38, No. 1, pp. 35-45, 1990.
4. S. Seneff, "A joint synchrony/mean-rate model of auditory speech processing," Journal of Phonetics, Vol. 16, pp. 55-76, 1988.
5. G. M. White, R. B. Neely, "Speech Recognition Experiments with Linear Prediction, Bandpass Filtering, and Dynamic Programming," IEEE trans. on Acoust. Speech and Signal Processing(ASSP), Vol. 24, No. 2, pp. 183-188, 1976.
6. K. Shikano, "Evaluation of LPC Spectral Matching Measures for Phonetic Unit Recognition," Technical Report, Computer Science Department, Carnegie Mellon University, 1986.
7. S. Furui, "Speaker-independent isolated word recognition using dynamic features of speech spectrum," IEEE trans. on ASSP, Vol. 34, pp. 52-69, Feb., 1986.
8. L. R. Rabiner, K. C. Pan, et al., "On the Performance of Isolated Word Speech Recognizers Using Vector Quantization and Temporal Energy Contours," AT & T Bell Lab. Technical Journal, Vol. 63, No. 7, pp. 1245-1260, 1984.
9. C. H. Lee, "On the use of some robust modeling techniques for speech recognition," Computer Speech and Language, Vol. 3, No. 1, pp. 35-52, 1989.
10. H. J. Choi, S. J. Yun, Y. H. Oh, "Robust Estimation of Discrete HMM parameters using the Entropy based Feature-Parameter Weighting and the Source-Quantization Modeling," Artificial Intelligence in Engineering (to be published).
11. V. N. Gupta, M. Lenning, et al., "Integration of Acoustic Information in a Large Vocabulary Word Recognizer," Proc. of ICASSP, pp. 697-700, 1987.
12. R. Schwartz, Y. Chow, et al., "Improved hidden Markov modeling of phonemes for continuous speech recognition," Proc. of ICASSP, pp. 35.6.1-35.6.4, 1984.
13. L. R. Rabiner, J. G. Wilpon, et al., "High Performance connected digit recognition using Hidden Markov Models," IEEE trans. on ASSP, Vol. 37, No. 8, pp. 1214-1225, 1989.
14. R. O. Duda, P. E. Hart, Pattern Classification and Scene Analysis, John Wiley & Sons, 1973.
15. Y. Linde, A. Buzo and R. M. Gray, "An algorithm for vector quantizer design," IEEE trans. on Comm., Vol. 28, pp. 84-95, 1980.
16. L. R. Rabiner, J. G. Wilpon, et al., "A segmental k-means training procedure for hidden Markov models with continuous mixture densities," AT & T Technical Journal, Vol. 65, pp. 21-31, 1986.
17. H. Ney, "The use of a one-stage Dynamic Programming algorithm for connected word recognition," IEEE trans. on ASSP, Vol. 32, No. 2, pp. 236-271, 1984.
18. F. Jelinek, L. R. Bahl, et al., "Design of a linguistic statistical decoder for the recognition of continuous speech," IEEE trans. on Information Theory, IT-21, pp. 404-411, 1975.
19. E. P. Gialchin, et al., "On the use of inter-word context-dependent units for word juncture modeling," Computer Speech and Language, Vol. 6, pp. 197-213, 1992.

▲최 환 진(Hwan Jin Choi)

1990년 2월: 고려대학교 전산학과(학사)

1992년 2월: 한국과학기술원 전산학과(석사)

1997년 8월: 한국과학기술원 전산학과(박사)

1997년 9월~현재: 시스템공학연구소 자연어정보처리연구
구부(선임연구원)

※주관심분야: 음성인식, 화자인식, 신경회로망, 퍼지이론, 사용자 모델링