

한국에서의 음성 신호 처리 기술의 현황과 전망

안수길
서울대학교 전자공학과

Current Status and the Prospect of Speech Signal Processing Technology in Korea

Souguil ANN
Department of Electronics Engineering, Seoul National University

요 약

자 한다.

최근 우리나라에서는 음성신호처리 기술을 바탕으로한 여러 가지 시스템이 상용화되고, 또 그에 따라 관련분야의 연구도 더욱 활발해지고 있다. 본 고에서는 최근 몇년간 발표되었던 연구 결과들을 바탕으로 현재 국내에서 음성신호처리 관련 분야에서의 연구현황을 소개하고 향후의 연구 방향 및 미래의 연구 경향을 예측해보고자 한다. 이를 위하여 음성신호처리 분야를 음성 분석, 음성 합성, 음성 인식, 음성 부호화의 네 세부 분야로 나누고 각 분야별로 국내 현황 그리고 앞으로의 전망을 제시한다.

1. 서 론

음성신호는 인간의 정보 교류 수단으로서 가장 편리하고 빠른 매체로, 전기들 사용가능하게 된 이후 인간의 가장 주요한 통신 수단으로서 사용되어 왔다. 근래에 디지털 신호처리기술의 발전과 함께 음성신호의 디지털화 및 그에 따른 여러 응용분야에 대한 연구가 활발히 진행되어 왔다. 인간에 대한 음성신호의 정보전달 측면에 있어서의 신속성 및 사용의 용이함은, 음성용 중요한 정보 교환의 매체로서 자리잡게 하였다. 뿐만아니라 인간과 기계사이의 정보 교환을 위한 MMI (Man-Machine Interface)에서 음성은 더욱 중요한 의미를 가진다. 이러한 이유로 인해 음성은 아주 오랜동안 중요한 연구분야로서 매우 활발히 연구되어 왔다.

80년대까지의 디지털 음성신호처리 기술은 실용화를 위한 기반조성 및 기초연구가 그 목표였다고 할 수 있다. 여러가지 디지털 음성신호처리 기술을 기반으로 한 시스템이 연구, 개발되어, 제한적으로나마 일부 분야에서 실용화되기도 하였다. 그러나 최근 수년간의 국내 멀티미디어 및 이동통신 등의 새로운 응용분야에서의 수요 증대와 하드웨어 분야의 눈부신 발전은, 음성신호처리 기술의 실용화를 촉진시키는 촉매가 되었다. 그 결과, 90년대 들어 국내 음성신호처리 분야는 양적인 측면으로나 질적인 측면에서, 어느때보다 내실있고 빠른 발전을 이루고 있다. 현재 제한적으로나마 음성 인식, 또는 합성 기술을 바탕으로한 여러가지 상용 시스템이 개발되어, 가정에까지 보급되었으며, 보다 우수한 성능의 발전된 시스템들이 상용화를 눈앞에 두고 있다. 또한 음성 부호화 기술을 이용한 디지털 이동통신도 수년내에 서비스를 개시할 예정이다.

본 고에서는 현 시점에서 국내 연구 현황을 최근 몇년간의 연구 결과를 바탕으로 정리, 소개하고자 한다. 이를 위하여 음성신호처리의 연구 분야를 크게 음성 분석, 음성 합성, 음성 인식, 그리고 음성 부호화의 네가지 세부분야로 나누었다. 이러한 세부분야를 각각에 대하여 국내 연구 현황 및 그 전망을 기술하고

2. 음성 분석 (speech analysis) 분야

음성 분석은 이미 모든 음성신호처리 관련 응용 기술의 기초로서 그 동안 많은 연구와 연구를 거듭해온 분야이다. 음성분석은 음성의 발생 모델의 모델링을 통해 시간적 또는 주파수에 따른 파라미터화를 통해 보다 간접하게 음성을 모델링 하고자 하는데서 부터 비롯되었다. 이러한 연구 방향은 상도 모델과 성도 모델의 선형적 모델을 통해 가가의 파라미터의 추측에 많은 심혈을 기울였으나, 현재는 선형적 모델 보다는 비교적 복잡하고 어려운 비선형적 모델링을 통해 정확한 발생과정을 알아 내하고자 하는 방향으로 연구가 진행되어 가고 있다. 따라서 음성 분석에는 많은 연구 분야가 이루어졌으며, 음성 신호 처리에 있어서 기술적인 문제의 해결의 기본 역활은 담당하는 것뿐만 아니라 음성 신호 그 자체의 연구에도 많은 연구를 해 왔다. 음성 분석의 연구 결과는 음성 합성, 인식, 코딩등 여러 응용 분야에 과급적인 효과를 주어 왔고 더욱더 질 좋은 분석 기법은 이러한 응용 분야에 큰 기여를 해 왔다. 그런 본 정에서는 한국내에서 행해은 음성분석 연구 현황과 앞으로의 연구 방향에 대해 알아 보겠다.

2.1. 음성 분석의 연구 현황

2.1.1. 선형예측기법및 성도파라미터 추측

선형 예측 기법은 비교적 간단한 예측 기능임에도 불구하고 그 성능이나 실제적인 응용면에서도 좋은 성과를 제공하고 있는 매우 매력적인 기법이다. 이런 선형 예측 파라미터는 성도 모델 뿐만아니라 코딩, 인식, 합성에 중요한 정보를 제공하고 있다. 이러한 이유로 예전 부터 국내는 물론 국외에서도 많은 연구가 되어왔으며 이제 어느 정도의 훌륭한 결과를 보이고 있다. 흔히 사용되는 예측 방법은 최소 자승법 (least squares method) 이며 흔히 신호의 특성을 고려하여 자기 상관 방법과 공분산 방법을 사용하고 있다. 또한 파라미터의 계산에 있어서는 크게 직접 (direct) 추정 방법과 반복(iterative) 추정 방법으로 나눌수 있는데, 이에 대해 많은 연구가 수행되어 왔다 [7].

현재는 이러한 선형 예측 기법의 단점을 어느 정도 보정하고자 강인 (Robust) 기법을 사용한 연구가 발표 되었다 [3], [5], [6]. 즉 음성의 유성음의 여기 신호가 순수 백색 잡음이 아닌 임펄스 열이기 때문에 선형 예측 기법을 사용할 경우 이러한 오차는 피할 수 없게 된다. 따라서 이러한 비 가우시안 성질을 제거 해 주는 방법으로 강인 기법을 사용하여 보다 정확한 음성 파라

한국에서의 음성 신호 처리 기술의 현황과 전망

메디를 구해내는 방법이 제시되었다. 이러한 강한 기법은 그 음성 특성을 시분변으로 국한 하지 않고 시변의 성질을 가미하여 더더욱 복잡하고 난해한 음성신호의 분석을 도모하고 있다. 즉 음성의 정상성(stationary) 외에 모음에서 자음의 차이 구간이나 또는 자음에서 모음의 차이 구간 등에서 발생하는 시변 성질을 모델링하고 이를 추정함으로써 보다 정확하고 명확한 음성의 특성을 파악하는데 의의가 있다. 이러한 연구는 특히 합성음의 자연성과 명료도에 큰 기여를 할 수 있다.

강인 기법 이외에도 자음과 모음의 여기 신호를 차별화 하여 모델링 한 방법도 연구 되었다 [1]. 즉 모음의 여기 신호인 임펄스 역을 따로 파라미터화 하여 이 파라미터 값도 추정해 내는 방법으로 전체 과정이 어렵고 복잡하여 반복적(iterative) 방식으로 파라미터를 추정하고 있다.

같은 환경에서의 음성 분석 기법도 많은 연구가 되어 왔다 [2], [8]. 이러한 연구의 응용 분야로는 음성 감음 제거, 음성 향상(enhancement), 음성 감출 등에 사용되며 이러한 상황은 보다 현실적인 환경으로 실제 응용 가능한 분야이다

2.1.2. 음원 추출 및 피치 검출

음원이란 음성 모델링 및 신호 파라미터의 추정과 밀접하게 연관이 되어 있는 것으로, 음성 신호의 발생 메카니즘을 시스템 측면에서 고찰할 때 입력에 해당하는 부분이다. 음원 관련 연구로는 피치 검출과 음부성을 판정, 음원의 모델링 및 음원 파라미터 추정 등이 있다.

현재 한국내에서의 피치 검출은 활발한 연구가 계속되고 있는데 이는 같은 환경도 고려한 피치 검출 방식도 제시되었다 [4], [11]. 이러한 피치 검출은 주로 시간 영역, 주파수 영역, 그리고 시간-주파수 영역에서 많이 이루어지고 있으며, 피치 검출은 또한 음 부성을 판별 또는 분석을 동시에 행하는 경우도 있다 [9].

음원 모델링은 외국의 연구 개발에 비해 많은 연구가 이루어지지 않은 편이다. 음원 신호는 해부학적으로 성대에서 발생하는 진동이나, 성분이 좁아져 생기는 난류에 의하여 유성음 및 무성음이 발생한다고 알려져 있다. 지금까지는 LF(Liljencrants-Fant) 모델은 사용하거나, 여기에 다른 유성음원의 성질을 추가하는 형태로 음원을 모델링하였다. 또한, SVD(singular value decomposition) 와 펄스 망원류 이용한 기저신호의 형태로 음원을 발생한 방법도 제시 되었다 [12]. 즉 음원 신호를 SVD를 통해 적은 차수의 행렬을 갖는 유효 신호로 분류하고 이 신호들의 기저 형태의 함수로서 각각의 펄스 망원의 계수들을 계산하여 음원을 발생하도록 하였다. 기저 신호를 이용한 음원의 발생은 외국에서도 많은 연구가 되어 있으며 이미 데이터로도 많이 제공되어 있으나 한국어의 경우 다 언어의 경우와 달리 독자적인 연구가 이루어 지야 겠다.

2.1.3. 음성신호 모델링

음성 신호 모델링은 보다 정확한 모델링을 통해 음성의 성질을 규명하여 음성 신호를 전기적으로 표현하고자 하는데 주력하고 있다. 이러한 연구는 거의 40여년간 지속되어 왔으며 현재도 미쳐의 과제로 남아 있다. 음성 모델링에 대한 연구 과정: 음성 신호의 성질을 유행적, 선형적, 그리고 신호 차라적으로 해석하여 연구하는 것이 대부분이다. 초기에는 주로 음성의 발생과정을 전기적으로 이해하여 전기적 능가 회로로 표현하고자 했으나, 그 후 성도를 유행 튜브로 모델링하여 그에 따른 적합한 필터 구조를 가정하고 스펙트럼을 계산하였다. 음성 신호는 대개 극점(pole-zero) 모델로 표현할 수 있다. 이 모델에서는 적절한 여기 신호에 의해 구동된 필터의 출력이 음성 신호라고 가정하고, 또한 필터의 전달함수 형태에 따라 ARMA(auto regressive

moving average)와 AR도 모델링 하고 있다.

2.2. 음성 분석 분야의 전망

음성 분석은 앞서 언급한 바와 같이, 음성의 모든 응용 분야의 기초로서, 좀더 깊 좋은 제품을 위해서는 꼭 필요한 분야이다. 즉 음성 분석은 음성 신호처리의 필수적인 요소이기 때문에 많은 연구가 앞으로 계속 되어야 하며 보다 정확한 음성의 분석을 위해서는 선형성과 정상성(stationary)의 성질을 벗어난 비선형성과 시변의 형태로 음성을 모델링하고 분석해야만 할 것이다. 이러한 시도는 음성 합성뿐 아니라 코딩과 인식에 과급적인 효과를 줄 뿐만 아니라 더 나아가 음성 전체의 신호처리 분야의 수준을 한층 더 끌어 올릴 수 있는 계기가 될 것이다. 또한 이제까지의 음성 분석이 음성 신호만 가지고 하는 신호 처리적인 기술이었다면 앞으로는 음운학과 음운의 통계적 성질까지도 고려하여 한국어에 보다 적합한 음성 분석 기법을 제공하자는 것이 바람직하다고 생각된다.

3. 음성 합성 (speech synthesis) 분야

음성 합성이라고 하면 일반적으로, 문자로 입력된 정보를 여러가지 합성 규칙과 이미 저장되어 있는 무호화되어 있는 음성 분석 정보를 이용하여 대응되는 음성중역으로 바꿔주는 문자-음성 변환 시스템, 즉 TTS(text-to-speech) 시스템을 말한다. 문자-음성 변환은 인간과 기계사이에 음성으로 정보를 주고 받기 위하여 꼭 필요한 기술로서 전자 전화번호부 서비스, 음성인식, 주가조회, 오디오텍스트(audiotext) 서비스 등의 응용분야에서 사용되어 갈 수 있다.

음성 합성에 관한 연구는 외국의 경우 상당히 오랜 역사를 갖고 있으며 많은 연구결과와 기술이 축적되어 있다. 이러한 것은 기원으로 1980년대 초반에 몇몇의 TTS 시스템이 개발되었다. 그 중 대표적인 것으로 미국 MIT 대학의 MITalk, 스웨덴 STL(speech transmission laboratory)의 워싱턴 시스템, 일본의 몇몇 시스템 등이 있으며, 국내에서도 1980년대 후반부터 대학과 연구소에서 여러 시스템들이 개발되었으나 아직도 음질향상의 여지가 많다고 할 수 있다.

음성 합성 시스템은 다음의 순서에 따라서 구성된다. 문자열이 입력되면 전처리부에서 영어, 숫자, 아어, 난류등을 나타내는 기호들은 변환규칙과 사진은 이용하여 발음되는 한글로 변환시키고 구분문자를 통해 한국어 문장안에서 이전 간의 문법적 연관관계를 밝혀내어 이로부터 음운처리에 필요한 정보가 제공되며 이어서 표기된 발음을 발음되는 음소로 변환시키며 음운처리부에서는 음운제어에 필요한 피치, 지속시간, 휴지기등의 파라미터를 생성하고 합성기에서는 앞에서 구한 파라미터들을 이용하여 실제 음성을 출력하게 된다. 이러한 합성음의 음질과 자연성은 음운제어 규칙, 음원생성 모델, 음성합성단위의 연결방법, 합성방식 등에 의해서 결정된다.

3.1. 음성 합성의 연구 현황

합성 방식에 있어서는 초기에는 비교적 시스템이 간단하고 구현이 쉬운 LPC 계열의 방법들이 주로 연구되었으나 최근에는 포먼트(formant) 합성방식은 미묘해서 [13], [14], 일종의 피형 편지법적인 PSOLA(pitch synchronous overlap and add) 방식 등에 의한 합성 시스템들이 연구 개발되고 있다 [15]-[20]. 포먼트 방식은 MITalk에서 사용된 방식으로 명대안 양의 데이터 처리를 통한 파라미터 추출 및 규칙도출이 필요하므로 구현이 쉽지가 않으나 음질과 자연성 면에서는 데이터 베이스의 크기, 음색의 민감도와 관련된 문제점이 없다는 장점을 가지고 있다. PSOLA 방식은 시간축에서의 피형을 그대로 이용하기 때문에

깨끗한 음성을 구현할 수 있지만 파형을 그대로 담고 있으므로 많은 기억 용량을 필요로 하며 연결부위에서의 불연속을 제거하기 위한 세심한 고려가 필요하다.

음성 합성을 위한 합성 단위로는 음소, 반음절, 다이폰(diphone) 등과 이들의 변형된 형태들이 사용된다. 일반적으로 합성 단위가 작으면 필요한 데이터의 양은 작아지게 되나 합성 단위를 연결시 연결하는 음사이의 전이를 실현하기 어렵게 된다. 한국어의 합성 단위로는 안정된 모음구간에서 연결을 하는 반음절 단위가 (변이음을 포함하) 주로 사용되었으며 근래에는 다이폰 단위가 많이 사용되고 있으며 [21]-[23], 음소물 정상상태 시점을 기준으로 양분하여 얻은 반음소 단위도 제안되었다 [17], [25]. 앞으로는 여러 형태의 합성단위가 경우에 따라서 선택되는 방법과 COC(Context Oriented Clustering) 방법등의 사용이 기대가 된다.

초기의 합성 시스템은 우선 명료성의 확보에 중점을 두었으나 자연성도 합성음의 평가에 있어서 중요한 요소이다. 자연성은 억양, 강세, 리듬등의 유포와 관계가 있는데 유포를 피치, 지속시간, 휴지기, 크기등에 의해서 결정된다. 한편 이 유포 요소들은 문장의 의미론적 요인과 구문론적 요인에 의해서 영향을 받는다. 따라서 올바른 유포 생성을 위해서는 한국어 문장 안에서 어절간의 문법적인 연결 관계를 자동적으로 추출하는 과정인 구분분석이 필요하다. 영미의 경우, 유포와 구분분석에 관한 연구가 일찍이 시작되어 왔으나 한국어의 경우에는 최근에 와서 연구가 진행되고 있다 [26]-[30]. 한편 한국어 합성음성의 품질 평가 방법에 관한 연구 역시 중요한 과제인데 최근에 와서 이에 관한 연구가 활발히 이루어지고 있다 [31]-[33].

그 밖에 연속 음성에서의 음운 변형 규칙과 조음 결합 규칙의 확립, 음소 전이의 동적특성 해명과 합성단위 연결에의 적용, 합성을 위한 한국어의 음성, 음운론적 특성 연구 등의 한국어 음성합성을 위한 기초 연구가 진행되고 있으며 [34]-[37], 앞으로 더욱 더 그 필요성이 요구된다고 할 수 있다.

3.2 음성 합성 분야의 전망

음성합성이 실용화되기 위해서는 현재의 명료성과 자연성이 훨씬 개선되어야 할 것이다.

최근의 연구방향은 명료성의 향상을 위해 PSOLA 방식과 같은 시간축상의 코딩방법을 많이 사용하고 있으며 합성 단위에는 음소의 환경을 충분히 흡수하고자 많은 수의 합성단위를 이용하는 추세이다. 앞으로 한국어 음성합성에 적합한 합성단위에 대한 연구가 더욱 필요하다고 할 수 있다.

합성음의 자연성을 향상시키기 위한 한국어 처리 기술에 대한 연구가 아직은 미흡한 상태이며 이를 개선하기 위해서는 언어학 및 실험음성학적인 연구와의 연계가 필수적이다.

한편, 음색을 변화함으로써 다양한 목소리의 합성음으로 사용자들의 요구를 만족시키기 위한 연구가 필요하며 이를 위해서는 음색을 결정하는 요인에 관한 연구와 음원생성모델에 관한 연구가 진행되어야 할 것이다.

4. 음성 인식 (speech recognition) 분야

인간이 기계와 통신을 하는 방법으로서 가장 바람직한 것은 인간의 음성을 통신수단으로 이용하는 것이다. 이러한 목적을 달성하기 위해서 기계가 음성을 인식하게 하는 음성인식 기술이 지난 몇십년간 빠르게 발전해 왔다. 이러한 음성인식기술의 최종목표는 임의의 화자가 발생한 연속적인 음성용 실시간에 높은 인식률로서 인식하는 인식기를 만드는 것이다. 그러나 이러한 화자에 독립적인 연속음을 인식하는 것은 대단히 어려운 일이므로 기존의 인식기는 좀 더 제한적인 기능을 갖는 것이 연구되어 왔다. 이러한 인식기는 대상화자의 수에 따라 화자중속, 화자독

립으로 나뉘고, 음성의 형태에 따라 고립단어인식, 연결음인식, 연속음인식으로 나뉠 수 있다. 현재까지는 화자독립 고립단어인식, 화자독립 연결음인식에 대한 많은 연구가 수행되었고, 고립단어 인식시스템은 이미 하드웨어로 구현할 수 있는 단계에 이르렀으나 [38], 연속음인식에 대해서는 아직은 시작단계에 있다고 할 수 있다 [39]. 음성인식기의 구성은 크게 특징벡터를 추출하는 부분과 패턴을 학습시키는 부분으로 크게 나눌 수 있다. 본 고에서는 음성인식기의 구성에 따라 현재까지 수행되어온 음성인식기술의 동향을 알아보고 음성인식 기술의 앞으로의 전망과 문제점 등을 기술하고자 한다.

4.1. 음성 인식의 연구 현황

4.1.1. 특징벡터추출

특징벡터라는 것은 음성인식기를 만드는 데 있어서 처리의 기본이 되는 단위를 말한다. 즉, 음성의 파형을 그 자체로서 다루지 않고 일정한 구간마다 LPC 계수나 켈스트럼 계수 등을 추출하여 음성인식에 사용한다. 특징벡터의 추출과 그들 사이의 거리를 나타내는 거리측정도 (distance measure)는 동시에 고려되어야 할 문제로 최근에는 기존의 LPC 계수를 이용하는 것보다 유포리드 거리를 이용할 수 있는 LPC 켈스트럼 (cepstrum)을 이용하는 방법에 대한 연구가 많이 되고 있다 [40]. 또한 양자화 성능이 뛰어난 LSP 계수와 켈스트럼의 장점을 이용한 의사켈스트럼 계수를 이용한 인식방법에 대해서도 연구되고 있다 [41]. 한편 기존의 LPC 계수를 구하는 방법은 자기상관함수나 자기공분산함수를 이용한 2차 모델트를 이용하는 것이었으나 유적잡음 환경에서의 성능개선을 위해 3차 이상의 고차 누적관료 (cumulant)를 이용하여 LPC 계수를 구해서 인식기에 사용하는 방법에 대해서도 연구되고 있다 [42].

4.1.2. 패턴학습기술

패턴학습이라는 것은 기존 음성모델을 학습시키는 기술로서 음성모델의 유사한 정도를 측정하는 방법이라고 할 수 있다. 즉 어떠한 발성 한 음소 또는 단어에 대한 대표패턴을 찾아내는 방법을 말하며, 이렇게 구해진 기준 음성 패턴을 이용하여 실제로 음성인식을 할 때 사용하는 것으로 음성인식기의 핵심부이다 할 수 있다. 여기에 쓰이는 기술로는 대표적인 것으로 VQ (vector quantization), HMM (hidden Markov model), DTW (dynamic time warping), NN (neural network) 등이 있다.

VQ는 입력신호를 개별적인 값이 아닌 상관성이 있는 신호의 열을 하나의 벡터로 정하고 이를 하나의 양자화 기본단위로 삼은 일종의 데이터 압축방법으로 음성인식에 있어서 대이타압축을 통한 표준패턴을 생성하는 데 사용된다. 즉 입력특징벡터를 VQ 코드북과 비교한 결과인 코드북의 인덱스열을 이용하여 음성을 인식한다. VQ는 HMM이나 DTW, NN등과 연결되기도 하여 패턴학습기를 구성하게 된다.

DTW는 패턴간의 비선형적인 사상을 이용해 패턴간의 유사도를 측정하는 방법이고, HMM은 음성이 불어 정상성을 이용하여 음성을 정상 구간으로 확률적으로 나누어서 각 불리에 대한 확률분포를 구해서 패턴간의 유사도를 구하는 방법이고, NN는 인간의 뇌의 원리를 이용하여, 각 뉴런의 가중치들 구한 뒤 이것을 이용해서 다른 음성과의 유사도를 구하는 방법이다.

과거에는 VQ와 DTW 또는 VQ와 HMM, VQ와 NN를 이용한 패턴학습을 많이 수행하였다. 그러나 최근에는 DTW와 HMM, NN, VQ등의 특성을 개선하거나 이들의 장점을 서로 결합해서 새로운 알고리즘을 얻는 방법이 많이 등장하고 있다. 대표적인 것으로는 VQ의 학습알고리즘을 개선하여 VQ 코드북을 학습할 때, 오차제약 뿐만 아니라 오인식이 줄어들게 하기 위한

방법으로 LVQ (learning VQ)를 이용한 방법이 있으며 [43], 유사한 VQ 코드를 하나로 묶어서 인식에 이용하는 코트워드 연결 알고리즘 [44], HMM을 학습시키는 데 있어서 오인식된 결과들 이용하여 더 나은 인식결과를 얻고자하는 적응오차 학습 알고리즘 [45], HMM을 학습시키는 대표적인 알고리즘인 Viterbi 알고리즘에서 전체경로제한 상태에서 학습시키는 알고리즘 [46] 등이 있다. 또한 각각의 알고리즘을 개선시키는 데서 한 실용적 나아가서 위에서 언급한 여러 패턴 학습기의 장점들을 이용하는 연구로 VQ코드의 최이행법과 HMM을 결합해서 인식성능을 개선한 방법 [47], NN과 DTW를 결합하는 연구 [48] 등이 있다. 또한 최근에는 인식기의 변별능력을 향상시키기 위한 변별학습에 관한 연구가 활발하다. 대표적인 방법으로 최근에 개발된 GPD(Generalized Probabilistic Descent) 방법을 들 수 있다. GPD 방법은 최소오차를 분류를 인식화하기 위한 수식화해 Amari가 제안한 PDP(Probabilistic Descent) 알고리즘을 적용하여 일반화적인 방법으로 국내에서도 예측신경회로망에 적용되고 있다 [49]. 또 다른 변별학습으로 상태가중함수(state weighting function)를 도입하는 연구도 진행되고 있다. 이는 한 모델이 여러개의 유한 상태로 구성된 경우 각 상태가 담당하는 중요도에 따라 가중치를 가해 변별력을 향상시키는 것이다. 최근에 GPD 방법을 사용하여 예측신경회로망과 HMM에 이러한 방법을 적용시키고 있다 [50]. 이와같은 방법학습은 특히 유사한 이득가 많은 인식부문에 유리하고 전체적인 인식기의 성능에 큰 향상을 줄 수 있는 것으로 기대된다 따라서 앞으로 개발된 연속음 혹은 연결된 인식기의 원천기술로서 중요한 역할을 할 것이다. 그리고 이것과는 좀 다른 배경이지만 위에서 언급한 알고리즘을 약간 개선해서 화자가 읽을 수 존재하는 환경하에서 말성학 예를 대비한 읽음에 강한 인식기에 대한 연구도 많이 수행되고 있다 [41, 42, 51, 52].

4.3. 음성 인식 분야의 전망

향후 한국에서의 음성인식에 관한 연구는 크게 세가지 방향으로 이루어질 것으로 보인다. 첫째는 현재의 제한된 범위의 고립단어와 연결단어 인식에서 연속음성인식으로의 분야의 확대이다. 특히, 음성인식기술의 당국적인 복표인 화자독립 연속음성 인식 시스템의 개발에 많은 노력이 기대된다. 그러한 기술은 음성합성기술과 함께 자동번역기의 구현에 필수적이다. 또한, 그러한 시스템은 실시간에 가까운 구현하기 위한 연구도 활발한 것이다.

둘째로는 새로운 인식모델의 개발에 관한 연구로서 현재 많은 관심을 모으고 있는 신경회로망과 기존의 HMM, DTW 등과의 혼합형태가 일차적인 대상이 될 것이다. 특히, 신경회로망은 음성의 분별정보를 쉽게 사용할 수 있고 그 구현이 간단하다는 장점이 있어서 앞으로의 인식연구에 많이 사용될 것으로 보인다.

마지막으로는 기존의 모델에 대한 효과적인 학습알고리즘의 개발을 할 수 있다. 특히, 최근에 많은 관심을 모으고 있는 변별 학습 (discriminative training)에 대한 다양한 연구가 기대된다. 유사한 이득간의 최대 격리도를 최대화하는 이러한 학습방법은 전체 인식기의 하부에서의 성능을 향상시키는 중요한 역할을 할 것이다.

5. 음성 부호화 (speech coding) 분야

디지털 음성 부호화 방식은 전송률에 따라서 그 특성이 크게 달라진다. 초기에 매우 높은 전송율을 가진 PCM (pulse coded modulation), DPCM (delta PCM) 등의 방식들로부터 출발하여, ADPCM (adaptive DPCM) 까지 발전된 시간축 상에서의 신호 부호화 방식과, SBC (sub-band coding), APC (adaptive predictive coding) 등의 주파수 축 상에서의 신호 부호화 방식은 16~64kbps 대의 전송율을 요구하는 시스템에의 응용을 목표로

로 개발되었다. 이러한 방식은 지연시간이 매우 짧고, 원신호에 충실한 음을 얻을 수 있다는 장점으로 인해, ISDN (integrated services digital network)과 같이 지연시간이 큰 의미를 가지는 분야에서 활용될 수 있다. 그러나, 디지털 이동통신 분야와 같이 8kbps~16kbps 대의 낮은 전송율을 요구하는 경우에는 이러한 방식의 부호화 방식은 사용하기 어렵다. 이러한 낮은 전송률에 적합한 방식들로는 LPC (linear predictive coding) 기법을 바탕으로 한 혼성 부호화 (hybrid coding) 기법이 대표적인데, CELP (code excited linear predictive coding), VSELP (vector sum excited linear predictive coding) 등이 가장 대표적이다. 1) 밤에 4.8kbps이하의 전송율을 가지는 부호화 기법의 구현을 위하여 LPC10, CELP, MBE (multi-band excitation) 방식등의 음원 부호화 (source coding) 또는 혼성 부호화 방식들이 제안되었다. 이러한 음원 부호화 및 혼성 부호화 방식은 낮은 전송률에서 우수한 음질을 얻을 수 있다는 장점이 있으나, 지연시간이 매우 길고, 구조가 매우 복잡하다는 단점을 갖는다.

국내에서는 디지털 이동통신 서비스의 개시를 눈 앞에 두고 있는 시점에서, 실용화를 목표로한 저전송율을 갖는 혼성 부호화 기법에 대한 연구와 개발 노력이 가장 활발히 이루어지고 있다. 이 밖에, 차세대 디지털 이동통신 및 ISDN에서의 응용을 위한 여러가지 기초 연구도 활발히 이루어지고 있다.

5.1. 음성 부호화의 연구 현황

국내에서의 음성 부호화 분야 대한 연구는 90년대 들어, 멀티미디어 및 디지털 이동통신 분야에서의 수요에 의해 매우 활발해졌다. 특히 이동통신 분야의 수요는 저전송율을 가지는 혼성부호화 시스템에 대한 연구 노력을 촉진하여, 음성 부호화 분야에서는 현재까지도 가장 활발히 연구되고 있는 과제이다. 디지털 이동통신 시스템의 개발을 눈앞에 두고 있는 현실에 비추어 볼 때, 지금까지의 음성 부호화에 대한 80년대까지의 국내 연구는 그리 미진하였다 할 수 있다. 1980년대 후반에 CELP, VSELP 방식의 상용화 가능성 확인에 따라, 미국, 일본등에서는 관련 분야의 표준제정 및 상용화 시스템 개발에 박차를 가하고 있는 상황에서, 국내 연구 노력은 상대적으로 크게 부족하였다. 그 결과 현재 상용 서비스 된 디지털 이동통신 시스템에서의 음성 부호화기술은 미국 Qualcomm 사의 QCELP (Qualcom CELP) 부호화기술 미방으로 개발되었다. 이 시스템의 바탕으로 한국통신, 전자통신연구소 및 휴대전화 단말기를 생산하는 삼성, 삼성등의 업체에서 상용화 시스템의 개발을 발표하기도 하였다.

국내에서, 이러한 상이계의 상용화 연구에 발맞추어, CELP, VSELP 등의 저전송율 부호화기술의 구현은 용이하게 하기 위한 노력의 일환으로 여러가지 계산량 감소 기법에 대해 활발한 연구를 수행하였다. 혼성 부호화기술의 대부분의 계산량은 음성의 음원 신호(즉, 모델링하는 과정에서 요구된다. 혼성 부호화기술은 일반적으로 음성의 장기상관관계 (long-term correlation), 즉 과거를 모델링하는 적응형 코드북 (adaptive codebook)과 나머지 잔여 신호를 모델링하기 위한 고정형 코드북 (fixed codebook)을 사용한다. 이 두 과정에서 요구되는 계산량은 크게 감소시키면서 그에 따른 음질 저하를 최소화 하려는 연구가 매우 활발히 이루어지고 있는데, 특히 CELP 부호화기술 바탕으로, 폐지정보를 추출하는 적응형 코드북 구조에 대한 계산량 감소 연구가 가장 활발히 이루어지고 있다. 폐지정보 추출 과정에서의 계산량 감소 기법은 음성의 기본 특성을 이용하여 검색 범위를 제한하거나, 개방형회로 구조 (open-loop configuration)를 통해 얻은 추정값을 사용하여 검색 범위를 제한하는 방식등이 주로 연구되었다 [53]~[56].

이러한 적응형 코드북 검색 과정에서 요구되는 계산량을 감소시키고자 하는 연구이외에, 혼성 부호화 방식에서 가장 많은 계산량을 요구하는 고정형 코드북 구조의 계산량 감소를 위한 연구도 진행되고 있다. 이를 위해서 예이브트 변환과 대역선택

구조 방식을 이용한 개발형태로 구조의 고정형 코덱 구조가 제안되었다 [57]. 필리핀을 요구하는 폐쇄형회로 구조 (closed-loop configuration) 구조와는 달리 개방형회로 구조를 사용함으로써 개선량을 감소시켰다.

이외에 음원 모델이나 코덱 구조를 변화시킴으로써 성능을 개선시키고자 하는 여러가지 방식도 연구되었다. Trellis VQ 구조를 CELP 부호화기에 도입하여 음질 향상을 개선하고자 하는 방식과, 분석 구간에 변화를 주어 적응형 코덱 구조의 모델링 효율을 높인 방식 등이 최근에 제안되었다 [58], [59].

이처럼 국내의 음성 부호화 관련 연구는 실용화를 위한 CELP 음성 부호화기에 대해 집중되었지만, 그 밖에 초저전송률의 음성 부호화기에 관한 연구와 [60], LPC 방식을 채택하고 있는 부호화기에 필수적인 LPC 정보의 효율적인 양자화 기법등도 제안되었다 [61], [62]. 이러한 연구 결과는 차세대의 음성 부호화기의 표준 제정에 있어서 매우 중요한 기반이 될 것으로 기대된다. 또한 최근 연구 결과의 발표는 활발하지 않으나, BISDN에의 응용을 위한 16kbps대의 ADPCM 부호화기나 LD-CELP (low delay-CELP)등도 기존의 확보된 연구 결과들을 바탕으로 실용화를 위한 연구를 가일층 강화하고 있다.

5.2 음성 부호화 분야의 연구 전망

현재 마주지역의 8kbps대의 음성 부호화기 표준은 VSELP 기법을 바탕으로한 음성 부호화기를 채택하고 있다. 한편 Qualcomm 사는 CDMA (code division multiple access)를 바탕으로 자사에서 개발한 QCELP라는 CELP 방식에 기초한 부호화 방식을 제안하여, 국내에서 표준으로 채택될 것으로 보인다. 그러나, 향후에는 VSELP나 CELP 방식의 음원 모델의 미비점을 보완한 새로운 모델에 대한 수요가 발생할 것이며, 그에 따라 관련 연구가 더욱 활발해질 것이다. 이는 계산량 감소와 성능 개선이라는 두 가지 방향으로 지속될 것이며, MBE나 정형파형 (prototype waveform) 등을 이용하는 새로운 음원 모델에 대한 연구도 계속될 것이다.

한편 제한된 전송률하에서 음질을 개선하기 위해서는 음성의 특성을 이용한 구조를 도입하는 방법이 고려해야 하는데, 최근에 활발히 연구되고 있는 음운적 구분 (phonetic classification)에 따라 부호화 방식 및 비트 할당 방식을 변경하는 기법이 가장 높은 가능성을 보이고 있다. 이러한 방식의 연구는 국내에 아직 뚜렷한 결과를 나타내고 있지는 않으나 음성 부호화에 대한 연구가 지속됨에 따라 점차로 활성화될 전망이다.

이러한 기본적인 음성 부호화기에 대한 연구외에, 좁은에 강한 음성 부호화 기법에 대한 연구도 신호화 단계에서 매우 중요한 의미를 갖는 관계로, 연구에 대한 수요가 급증할 것으로 보인다. 특히 CELP 등의 음성 부호화 방식은 잡음에 의해 스펙트럼 정보인 LPC 계수에 변화가 생길 경우, 그 음질이 크게 저하되는 단점이 있으므로, 잡음에 강한 시스템의 개발이 다른 어느 기법에 비해 중요하다고 할 수 있다. 특히 실제 환경에서 발생하는 잡음에 대한 강인성 확보에 관한 연구는 매우 어려운 작업이므로, 장기적인 관점에서 꾸준히 진행될 것이다.

6. 결 론

최근의 음성신호처리 분야의 연구 동향은 80년대까지의 다양한 기초 연구들을 바탕으로 실제 실용화가 가능한 시스템 개발 및 개선에 관한 연구가 주류를 이루고 있다. 이와 병행하여 기존의 기술로 해결이 어려운 과제에 대한 장기적인 안목의 연구도 꾸준히 수행되고 있다.

음성 분석 분야의 경우 기존의 선형예측기법과 성도파라미터 추정등의 연구과제에 강한 기법이 결합되어 가는 추세로 발전

중에 있으며, 음원 추출 및 음성 신호의 모델링도 선형성 및 정상상태등의 가정에서 벗어나 음성 신호의 파라미터들 보다 정확히 추정할 수 있는 비선형, 선형 개념의 기법들이 연구중에 있다. 이러한 경향은 한국어에 보다 적합한 방식을 찾아내고자 하는 연구와 결합되어 보다 정확한 음성 파라미터 추출을 위한 기법 개발될 때 까지 국내 연구계에 지속적으로 존재할 것으로 전망된다.

음성 합성 분야의 경우, 그간 외국의 활발한 실용화 연구에 비해 국내의 연구가 미약하기는 했으나, 최근의 잇따른 시스템 개발로 인하여 이 연구 분위기가 매우 고조되어 있는 상황이다. 그러나 시스템의 개발과 병행하여, 한국어 합성기의 음질을 높이기 위해서는 한국어에 대한 구분 분석 및 음운학적인 연구가 필수적이다. 최근 이러한 연구가 여러 기관에서 시도되고 있는 만큼 머지않아 한국어 음성 합성기의 음질 수준도 크게 향상될 것으로 보인다.

음성 인식 분야에서는 화자 중심의 고립 단어 인식을 목표로 한 여러 시스템이 현재 실용화 되었다. 인식 분야에서 가장 큰 연구 과제라 할 수 있는 특징패턴 추출 및 패턴확률기법에 관한 연구는 중국의 목표인 무제한 단어 인식을 위한 시스템 개발을 위하여 계속적으로 연구되고 있다. 기존의 DTW나 HMM등의 모델을 개선하려는 노력과 병행하여 최근에는 신경망 회로와 결합된 음성 인식 시스템 개발에 관한 연구가 매우 활발하게 이루어지고 있다. 이러한 연구는 화자특원 연속음 인식 시스템 구현을 위하여 더욱 활발해 질 것이다.

음성 부호화 분야에서는 디지털 이동통신에의 응용을 위한 실용화 단계의 시스템들이 QCELP 부호화기를 기반으로 하여 개발을 완료한 단계에 있다. 따라서 최근에는 이러한 CELP 방식의 부호화기의 계산량 감소 및 성능 개선에 대한 연구가 매우 활발히 이루어지고 있다. 그러나 장기적인 안목에서는 차세대의 국내 자체의 표준화를 목표로 여러 새로운 방식의 부호화 기법들에 대한 연구가 시도되고 있으며, 초저전송률의 부호화기에 대한 관심이 매우 고조되고 있다. 향후에는 이동통신 분야에 응용 가능한 저전송률 부호화기 뿐만 아니라, ISDN 또는 멀티미디어 분야등에서 사용할 수 있는 여러가지 음성 부호화기에 대한 연구가 보다 폭넓게 수행될 것이다.

참고 문헌

- [1] K.Y. Lee, Y.T. Kang, and B.G. Lee, "Speech Parameter Estimation and Enhancement Using the EM Algorithm," 한국음향학회 논문집, vol. 13, no. 2E, pp. 68-75, 1994.
- [2] 구분용, "광반 필터를 이용한 음성 신호에서의 중첩된 유색 잡음의 감쇠," 한국음향학회지, vol. 13, no. 2, pp. 76-83, 1994.
- [3] K. Y. Lee, P. Ha, J. Rheem, S. Ann, and I. Song, "Robust Estimation of Time-Varying LP Parameters on Speech," Proc. ICSP-92, vol. 2, pp. 1589-1592, 1992.
- [4] J. Rheem, M. J. Bae, and S. Ann, "A Spectral AMDF Method for Pitch Extraction of Noise-Corrupted Speech," Proc. Eurospeech-93, vol. 3, pp. 2021-2024, 1993.
- [5] P. Ha, and S. Ann, "Robust Time-Varying Parametric Modeling of Voiced Speech," Signal Processing, vol. 42, no. 3, Mar. 1995.
- [6] H.S. Choi, S.C. Bang, and S. Ann, "A Robust Sequential Parameter Estimation for Time-Varying Speech Signal Analysis," Proc. ICASSP-94, vol. 1, pp. 457-460, 1994.
- [7] K.Y. Yoo, and N. Hubing, "The WRLS Algorithm for Speech Processing," Proc. ICASSP-94, vol. 3, pp. 377-380, 1994.
- [8] 강해동, 배건성, "음성신호의 난인 입력 적응 잡음 제거," 한국음향학회지, vol. 13, no. 3, pp. 16-24, 1994.

- [19] 최현영, 김형순, 김재호, 손강식, "포맷된 자원을 이용한 다음절 음성의 사. 모음 분리," 한국음향학회지, vol. 13, no. 3, pp. 51-62, 1994.
- [10] 김종기, "자음의 단이내 음운 환경별로 분 음가 변화," 한국음향학회지, vol. 13, no. 5, pp. 69-76, 1994.
- [11] 김 흥, 백규환, 배병진, 장호성, "가변 대역폭 LPF를 이용한 피치 검출," 한국음향학회지, vol. 13, no. 5, pp. 77-82, 1994.
- [12] S. H. Hong, S. K. Kang, and S. Ann, "Voice Parameter Estimation Using Sequential SVD and Wave Shaping Filter Bank," *Proc. ICSLP-94*, 198, 1994.
- [13] 안승권, 성영호, "포맷된 제2 중점방법에 의한 한국어 문자-음성 변환," 제9회 음성통신 및 신호처리 워크샵 논문집, pp. 197-200, 1992.
- [14] 조철우 외, "A Study on the implementation of Korean Synthesis-By-Rule System Using Formant Synthesis Method," 한국음향학회지, vol. 9, no. 6, 1990.
- [15] 최운권 외, "고품질의 한국어 문장 음성변환 시스템: 급조리 II," 제9회 음성통신 및 신호처리 워크샵 논문집, pp. 193-196, 1992.
- [16] 김상훈 외, "한국어 분절/음성 변환에서의 TD-PSOLA 적용," 제10회 음성통신 및 신호처리 워크샵 논문집, pp. 291-294, 1993.
- [17] 강용인, 김재인, "한소리: 무제한 음성합성 시스템," 제11회 음성통신 및 신호처리 워크샵 논문집, pp. 342-345, 1994.
- [18] 김성일 외, "한국어 음성 파형의 편집에 의한 한국어 음운 변화에 관한 연구," 한국음향학회지, vol. 13, no. 6, pp. 60-65, 1994.
- [19] 배병진, "위상 보상된 고조파 스케인링에 의한 음성합성용 피치변경법," 한국음향학회지, vol.13, no. 6, pp. 91-97, 1994.
- [20] 장천희 외, "사간영역에서의 피형분석에 의한 무제한 이휘 합성 및 음운 유형별 규칙합성은 음운 평가," 한국음향학회지, vol. 13, no. 1, pp. 71-83, 1994.
- [21] 정원경 외, "다이톤 이용한 한국어 음성의 규칙합성 시스템에 관한 연구," 제10회 음성통신 및 신호처리 워크샵 논문집, pp. 295-300, 1993.
- [22] 정준구 외, "다이톤을 이용한 한국어 문자-음성 변환 시스템의 설계 및 구현," 제11회 음성통신 및 신호처리 워크샵 논문집, pp. 91-94, 1994.
- [23] 이승준 외, "다이톤 단위의 합성방법을 이용한 오디오 텍스트 시스템의 구현에 관한 연구," 제11회 음성통신 및 신호처리 워크샵 논문집, pp. 99-102, 1994.
- [24] 정인종 외, "An Automatic Diphone Segmentation for Korean Speech Synthesis-by-Rule," 한국음향학회지, vol. 12, no. 2E, 1993.
- [25] 이종락, "만음소: 새로운 음성합성 및 인식단위," 제10회 음성통신 및 신호처리 워크샵 논문집, pp. 208-212, 1993.
- [26] 김성용 외, "형태소 해소를 이용한 합성음성의 음운 및 운율치리," 대한전자공학학회지, vol. 20, no. 5, 1993.
- [27] 김형운 외, "한국어 분절-음성 변환기의 음운재이용 구분분석기," 제9회 음성통신 및 신호처리 워크샵 논문집, pp. 218-223, 1992.
- [28] 김연희 외, "한국어 분시-음성 변환 시스템에서의 구분분석에 의한 음운추출에 관한 연구," 제10회 음성통신 및 신호처리 워크샵 논문집, pp. 285-290, 1993.
- [29] 허경희 외, "합성음 구현을 위한 음의 영향과 장단변환 연구," 제11회 음성통신 및 신호처리 워크샵 논문집, pp. 328-333, 1994.
- [30] 이현복 외, "An Experimental Phonetic Study of Speech Rhythm in Standard Korean," *Proc. ICSLP-94*, 204, 1994.
- [31] 조철우 외, "Generation of Multi-Syllable Nonsense Words for the Assessment of Korean Text-To-Speech System," *Proc. ICSLP-94*, 224, 1994.
- [32] 김정현 외, "명료도 평가용 다음절 목록의 개발," 음향학회지, vol. 13, no. 4, pp. 69-76, 1994.
- [33] 조철우 외, "합성음성 평가를 위한 다음절 부의미 단어 생성과 이용에 관한 연구," 한국음향학회지, vol. 13, no. 5, pp. 51-58, 1994.
- [34] 장국 외, "음성 인식/원상을 위한 국어의 음성-음운론적 특성 연구," 한국음향학회지, vol. 13, no. 6, pp. 31-44, 1994.
- [35] Yun-Keun Lee, Seung-Kwon Ahn, "Effects of the Phase Jitters on Naturalness of Synthesized Speech," *Proc. Eurospeech-93*, vol. 1, pp. 933-936, 1993.
- [36] 공영구 외, "이점음 검출에 의한 음절지하 및 남북대역 연구," 제10회 음성통신 및 신호처리 워크샵 논문집, pp. 279-284, 1993.
- [37] 김대원, "한국어 문법에 관한 음향음성학적 연구," 제11회 음성통신 및 신호처리 워크샵 논문집, pp. 323-327, 1994.
- [38] Y. Cho, J.Y. Kim, K.S. Oh, H.S. Lee, "A Real Time Implementation of Isolated Word Recognition System Based on Hardware Efficient Viterbi Scorer," 한국음향학회지, vol. 13, no. 2E, pp. 58-67, 1994.
- [39] H.R. Kim, K.W. Hwang, Y.M. Ahn, J.H. Ryu, "A Continuous Speech Recognition System Using Finite State Network and Viterbi Beam Search for the Automatic Interpretation," *Proc. ICASSP-95*, vol. 1, pp. 117-120, 1995.
- [40] 최승호, 김형식, "신형 예측 세수에 기초한 피치추론 단어 인식," 한국음향학회지, pp. 5-13, vol. 13, no. 1, 1994.
- [41] H.K. Kim, H.S. Lee, "Pseudo-Cepstral Representation of Speech Signal and Its Application to Speech Recognition," 한국음향학회지, vol. 13, no. 1E, pp. 71-81, 1994.
- [42] 이창식, 양원영, 조홍수, "유색잡음 환경하에서 Cumulant를 이용한 한국어 모음인식," 한국음향학회지:문지, vol. 13, no. 2, pp. 50-59, 1994.
- [43] M.W. Koo, C.K. Un, "A Speaker Adaptation Based on Corrective Training and Learning Vector Quantization," *Proc. ICSLP-92*, pp. 1475-1478, vol. 2, 1992.
- [44] 김도영, 김남수, 은종권, "이산분포 HMM을 이용한 음성 인식에서의 코트워드 Tying 알고리즘," 한국음향학회지, vol. 13, no. 3, pp. 63-70, 1994.
- [45] H.R. Kim, H.S. Lee, "Segmental Corrective Training for HMM Parameter Estimation in Speech Recognition," 한국음향학회지, vol. 12, no. 2E, pp. 5-11, 1993.
- [46] W.K. Kim, J.Y. Choi, D.H. Youn, "HMM with Global Path Constraint in Viterbi Decoding for Isolated Word Recognition," *Proc. ICASSP-94*, vol. 1, pp. 605-608, 1994.
- [47] 김광우, 홍광식, 마병천, "VQ SLD의 원리행렬과 이산 HMM을 이용한 한국어 단어 인식," 한국음향학회지, vol. 13, no. 4, pp. 40-49, 1994.
- [48] 황영수, "TDNN과 DTW를 이용한 소리 단어 인식," 한국음향학회지, vol. 12, no. 2, pp. 45-50, 1993.
- [49] K.M. Na, J.Y. Rheem and S. Ann, "Discriminative Training of Predictive Neural Network models," 한국음향학회지, vol. 13, no. 1E, pp. 64-70, 1994.
- [50] K.M. Na, J.Y. Rheem and S. Ann, "Minimum-error-rate Training of Predictive Neural Network Models," *Proc. ICSLP-94*, pp. 1479-1482, 1994.
- [51] 이기영, "스펙트럼 사상 학습을 이용한 잡음환경에서의 한국어 숫자음 인식," 한국음향학회지, vol. 13, no. 3, pp. 25-32, 1994.
- [52] K. Lee, "Noisy Speech Recognition Based on Spectral

- Mapping Techniques," 한국음향학회지, vol. 14, no. 1E, pp. 39-45, 1995.
- [53] J.Y. Rheem, M.J. Bae, S. Ann, "Reduction of Pitch Searching Range in CELP Vocoder," 한국음향학회지, vol. 12, no. 1E, pp. 33-37, 1993.
- [54] J.H. Lee, M.J. Bae, S. Ann, "Fast Pitch Searching Algorithm Using Correlation Characteristics in CELP Vocoder," 한국음향학회지, vol. 13, no. 2E, pp. 20-25, 1994.
- [55] D. Kim et al., "On a Reduction of Pitch Searching Time by Preliminary Pitch in CELP Vocoder," 한국음향학회지, vol. 13, no. 2E, pp. 51-57, 1994.
- [56] J.I. Hyun et al., "On a Reduction of Pitch Searching Time by Separating the Speech Components in the CELP Vocoder," 한국음향학회지, vol. 14, no. 1E, pp. 22-29, 1995.
- [57] D. Chang, Y.K. Cho, S. Ann, "Enhanced Wavelet Transform-based CELP Coder with Band Selection and Selective VQ," 한국음향학회지, vol. 14, no. 1E, pp. 46-55, 1995.
- [58] 강상원, 조성호, 이인성, "효율적인 Trellis Excitation 음성 부호화 방식," 제 11회 음성통신 및 신호처리 워크샵 논문집, pp. 210-223, 1994.
- [59] C. Kwon and C.K. Un, "Improving the Adaptive Source Model for CELP Coding with Long Analysis Frame Size," *Speech Communication*, vol. 16, no. 4, 1995.
- [60] 박승중, 김승주, 오영환, 이양희, "2.9 kbps LP-SMBE 음성 부호기 개발," 제 11회 음성통신 및 신호처리 워크샵 논문집, pp. 175-178, 1994.
- [61] 강정원 외, "음성 신호의 Line Spectral Pair (LSP) 주파수의 2단 벡터 양자화," 제 7회 신호처리 합동 학술대회, pp. 288-291, 1994.
- [62] D. Chang, Y.K. Cho, S. Ann, "Efficient Quantization of LSP Parameters Using Classified SVQ Combined with Conditional Splitting," *Proc. ICASSP-95*, vol. 1, pp. 736-739, 1995.