

## 잡음 환경 하에서의 입술 정보와 PSO-NCM 최적화를 통한 거절 기능 성능 향상

### Improvement of Rejection Performance using the Lip Image and the PSO-NCM Optimization in Noisy Environment

김 병 돈<sup>1)</sup> · 최 승 호<sup>2)</sup>

Kim, Byoungdon · Choi, Seunggho

#### ABSTRACT

Recently, audio-visual speech recognition (AVSR) has been studied to cope with noise problems in speech recognition. In this paper we propose a novel method of deciding weighting factors for audio-visual information fusion. We adopt the particle swarm optimization (PSO) to weighting factor determination. The AVSR experiments show that PSO-based normalized confidence measures (NCM) improve the rejection performance of mis-recognized words by 33%.

**Keywords:** audio-visual speech recognition, particle swarm optimization, normalized confidence measure, rejection performance

#### 1. 서론

인간의 목소리를 분석하고 이해하여 응답할 수 있는 인지 기능을 갖춘 지능형기가 최근 여러 분야에서 사용되고 있다. 인체공학적인 기기들의 성능이 지속적으로 발전하여도 인간의 목소리보다 더 편안한 환경을 제공할 수는 없기 때문에 HCI(Human-Computer Interface) 기술 중의 하나인 음성인식 기술은 현재 실험실 환경을 벗어나 일상생활의 영역까지 그 범위가 확대되고 있다. 그러나 실제 음성인식 기술은 잡음환경에서 성능이 매우 열악하다[1].

잡음의 종류에는 배경잡음, 채널왜곡, 배경잡음에 의한 음성 신호 자체에 왜곡이 오는 Lombard 효과[2] 등이 있다. 잡음에 강인한 음성인식은 인식부에서 잡음을 처리하는 방식보다 계산량이 적은 특징벡터 추출 후 보상 방법을 많이 활용하고 있다.

그러나 실생활에서의 잡음은 매우 다양하기 때문에 이를 모두 처리하기는 쉽지 않은 문제이다.

요즈음 이러한 인식성능 저하 문제를 해결하기 위해 인식과정에서 음성 이외의 정보를 추가하는 연구가 진행되고 있다. 일반적으로 의사소통은 음향정보 뿐만 아니라 시각정보도 복합적으로 사용된다. 시각정보는 음성잡음의 영향을 받지 않을 뿐 아니라 발성과정에서 다양하게 변화하기 때문에 음성잡음에 의한 인식을 저하를 보정할 수 있게 된다. 특히 시각정보 중 발음과정에서 변화하는 입술정보를 이용한 시청각 음성인식(audio-visual speech recognition) 분야가 활발히 연구되고 있다 [3]-[5].

본 논문에서는 시각정보 중 발화된 음성에 대한 입술정보를 이용하여 기존의 잡음환경에 적용된 음성인식과 입술정보를 통합한다. 이를 통해 잡음에 적용된 시청각 음성인식의 통합방법을 제안하고 그 타당성을 입증한다.

#### 2. 잡음환경에서 시청각 음성인식 통합방법

##### 2.1 시청각 음성인식 통합방법

시청각 정보의 통합에는 여러 방법이 존재하며, 두 입력 정보를 언제 통합하느냐에 따라 구분할 수 있다. 첫째, 음성과 영

1) 동신대학교 공과대학 컴퓨터학과, doni96@empal.com  
2) 동신대학교 공과대학 컴퓨터학과, shchoi@dsu.ac.kr, 교신저자  
이 논문은 2007년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(지원번호: KRF-2007-521-D00447).

상의 신호단계에서 통합하는 방법 둘째, 음성 파라미터와 영상 파라미터를 인식과정 이전에 통합하여 하나의 파라미터 벡터를 만들어 인식하는 방법 셋째, 모델학습에서 이루어지는 중간 통합하는 방법 넷째, 인식 후 통합하는 방법 등이 있다.

본 논문에서는 음성과 영상정보에서 추출된 파라미터에 대한 인식과정을 수행한 다음 인식결과에 가중치를 이용하여 통합하는 방법 즉, 네 번째 방법을 사용하였다.

신뢰도 정규화는 각기 다른 두개의 신뢰도의 분포를 동일한 범위를 갖도록 조정하며, 통합은 정규화 신뢰도들의 선형결합에 의하여 최종 신뢰도를 얻는다. 이는 식 (1)과 같다.

$$C = \alpha C_A + (1 - \alpha) C_V \quad (1)$$

식 (1)에서  $C_A$ 는 음소 신뢰도,  $C_V$ 는 입술정보 신뢰도이다. 잡음환경에 적용된 파라미터를 선택하기 위하여 PSO 최적화 방안을 적용한다. 이 방법은 [6]에서 음소 신뢰도에 적용하여 잡음 환경하의 음성 인식 영역에서 그 성능이 충분히 검증되었다.

## 2.2 PSO를 사용한 시청각 음성인식 통합

PSO(Particle Swarm Optimization)는 수학적으로 계산하기 힘든 비선형 문제에서 널리 응용되며 최적화 함수  $f(\cdot)$ 가 비선형일 경우 PSO 방법은 해가 국부적인 최적값을 피하면서 전체 최적의 해를 구하는데 사용한다[7]-[8]. PSO 방법은 아래와 같다.

- 1) Random한 잠재적인 해  $\{P_{i0}\}$ 를 결정한다.
- 2) 각 iteration  $j$  에 대하여 아래를 반복한다.
  - 2-1) 각  $P_{ij}$ 에 대하여  $f(P_{ij})$ 를 계산한다.
  - 2-2) 최적  $f$ 값의 변화를 계산하고 수렴한 경우 루프를 빠져나간다.
  - 2-3) 각  $i$ 에 대하여  $\{0, \dots, j-1\}$ 에 대하여 가장 최적인(본 논문에서는 가장 작은)해를 저장한다. 이 값을  $pbest_{ij}$ 라고 한다.
  - 2-4) 모든  $pbest_{ij}$ 를 대상으로 가장 최적인 해를 저장하여 이를  $gbest_j$ 라고 한다.
  - 2-5) 각 particle의 속도는 다음과 같이 계산한다.

$$v_{ij} = v_{ij-1} + c_1 r_1 (pbest_{ij} - P_{ij-1}) + c_2 r_2 (gbest_j - P_{ij}) \quad (2)$$

위 식에서  $c_1$ 과  $c_2$ 는 상수이며,  $r_1$ 과  $r_2$ 는 임의의 수이다.

- 2-6) 각 particle 값을 갱신한다.

$$P_{ij} = P_{ij-1} + v_{ij} \quad (3)$$

- 3)  $gbest_j$ 를 최적의 해로 결정한다.

단어단위 신뢰도는 음소 신뢰도의 가중 합으로 계산된다. 그러나 잡음이 포함된 음성의 음소 신뢰도는 매우 낮은 확률값을 갖게 되며 일정하지 않다. 따라서 거절성능 향상을 위해 잡음에 따른 음소 신뢰도에 대한 보상이 필요하다. PSO 기반 최적화 함수의 파라미터는 음소에 따라 모두 다른 값을 갖게 되며 음소 신뢰도의 평균분포에 따라 재구성된다. [6]에서 알 수 있듯이 각 잡음에 따른 음소별 파라미터의 최적화를 위해 적용된 음소 신뢰도의 평균은 SNR의 함수로 표현하였으며, 분석 결과 평균은 <그림 1>의 음소 신뢰도 평균에서 알 수 있듯이 잡음의 양에 따라 일정하게 증가하는 형태를 가지게 된다. 따라서 최적화에 사용되는 매핑 함수는 시그모이드(sigmoid) 함수를 이용하였다. 각 음소별 신뢰도 매핑 함수는 식 (4)와 같이 정의된다.

$$M_P(x_{SNR}) = a \left[ \frac{1}{1 + e^{-b(x_{SNR} + c)}} \right] + d \quad (4)$$

위 식에서  $a$ 는 스케일 파라미터,  $b$ 는 기울기 파라미터,  $c$ ,  $d$ 는  $x$ ,  $y$ 축의 이동 파라미터이다. SNR에 따른 음소 신뢰도의 PSO 최적화 함수  $f_A(\cdot)$ 를 사용하며, 식 (5)와 같다. PSO 최적화를 위해 적용된 각 파라미터의 초기값은  $0 \leq a \leq 0.5$ ,  $0 \leq b \leq 0.5$ ,  $-40 \leq c \leq 0$ ,  $-0.5 \leq d \leq 0$ 에서 결정하였으며, 최적화를 위한 particle 수는 500, 각 음소에 따라 200번의 최적화를 진행한다. 위의 값은 실험에 의해 얻어진 결과이다.

$$f_A(a, b, c, d) = \frac{1}{m} \sqrt{\sum_{k=1}^m N_{x_k} [M_P(x_k) - M(x_k)]^2} \quad (5)$$

위 식에서  $x_k$ 는 실험에 사용된 51개의 음소,  $m$ 은 최적화 함수에 사용된 음소의 개수,  $N_{x_k}$ 는 음소 신뢰도의 수,  $M(x_k)$ 는 실제 음소 신뢰도의 평균값이다.

<그림 1>은 식 (4)와 식 (5)에 의해 PSO 방법이 적용된 음소 신뢰도의 평균을 구하고 이를 비교한 것이다. PSO 적용 후의 최적화 함수는 일정한 기울기를 가지며 수렴하고 큰 변화를 일으키는 영역에 대해서는 보상이 이루어짐을 알 수 있다. 본 논문에서는 PSO 적용된 음소 신뢰도를 사용한다.

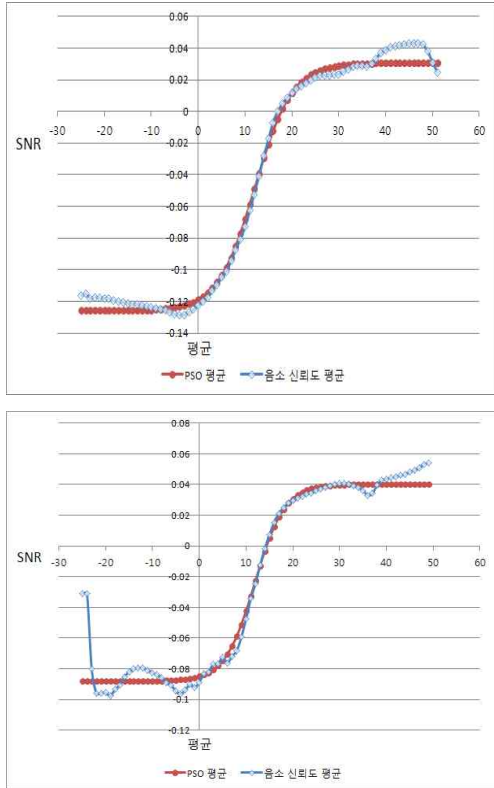


그림 1. 음소 신뢰도 평균과 PSO 최적화 평균 비교 (음소 /ii/, /wa/)

Figure 1. Comparison between the means of phoneme confidence and PSO optimization

입술 영상에 사용된 HMM은 연속적인 확률통계로써 음성인식과 동일한 방식을 적용하였으며, 발음한 영상 관찰  $O$ 는 입술 파라미터 추적으로 구해진 영상 특징벡터의 연속적인 열로 구할 수 있으며 그 식은 아래와 같다.

$$O = o(1), o(2), \dots, o(t) \tag{6}$$

본 논문에서 사용된 시각정보는 실험실에서 녹화된 입술 영상으로 잡음의 영향이 없어 잡음이 섞인 음성이 입력되었을 때 입술 파라미터 값은 변하지 않는다. 따라서 음성인식과정에서 SNR이 클수록  $C_V$ 는 감소하고  $C_A$ 는 증가한다.

<그림 2>는 잡음이 섞인 음성과 입술 영상이 동시에 입력되었을 때 통합 신뢰도를 계산하는 과정을 나타낸 것이다. 음소 구간 내의 정확한 SNR 정보를 얻기 위하여 clean speech와 noisy speech를 비교, 음소 구간에 삽입된 잡음의 양을 계산하고 이 결과를 통해 음소 구간별 SNR을 얻을 수 있다. 이때 계산된 SNR에 따라 가중치  $\alpha$ 를 결정하여 식 (1)에 적용된다.

식 (1)에서 두 파라미터  $C_A$ 와  $C_V$ 는 PSO를 적용하여 SNR에 따른 최적화된 값을 선택하고 잡음의 정규화를 진행한다. 정규화 과정에서  $\alpha$ 는 잡음에 따라 비례적으로 증가하며 잡음이 적

은(SNR 30dB 이상) 영역에서는 1에 근접하는 값을 갖는다. 따라서 잡음에 따른 파라미터  $\alpha$ 를 구성하는 최적화 함수는 식 (7)과 같다. 식에서  $e$ 는 스케일 파라미터,  $g$ 는 이동 파라미터이다.  $e$ 와  $g$ 는 식 (1)에서 가장 높은 인식 성능을 갖는  $C$ 를 데이터베이스에 따라 계산하였고, PSO 이론이 적용된 값이다.  $e = -0.61$ ,  $g = 13$ 을 실험에 적용하였다.

$$\alpha(SNR) = \frac{1}{1 + e^{(SNR-g)}} \tag{7}$$

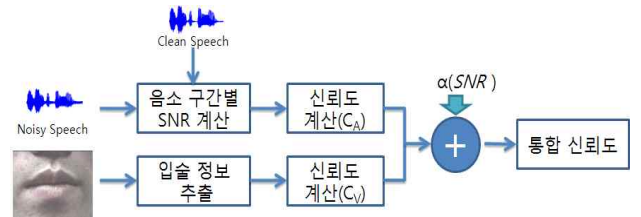


그림 2. SNR에 따른 통합 신뢰도 계산 과정  
Figure 2. Calculation process of integration confidence following SNR

### 3. 실험 환경

#### 3.1 데이터베이스

잡음에 적용된 시청각 인식기를 구현하기 위해 사용된 음성 인식은 HMM 기반 고립단어인식기이다. DB는 뉴스에서 발생되는 도로명과 증권 관련 100개의 단어이며 남성화자 50명을 대상으로 조용한 환경에서 녹음하였다. 녹음된 음성 DB는 SNR에 따른 white gaussian noise를 첨가하였다. 인식모델의 훈련용 DB는 30명, 평가용 DB는 20명이며 HMM 학습은 Forward 및 Backward 알고리즘 그리고 음성인식은 Mel-cepstrum 파라미터를 적용하였다.

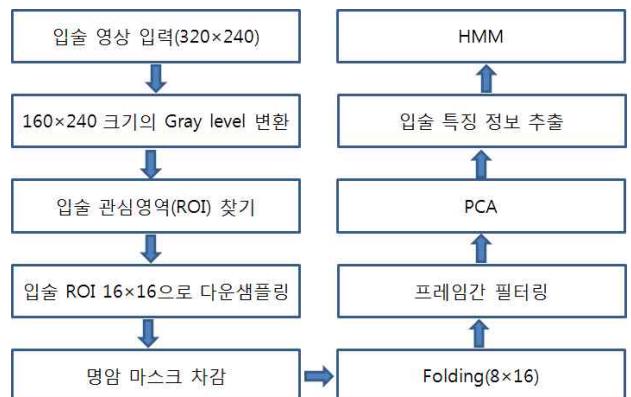


그림 3. 입술 영상 특징정보 추출 알고리즘  
Figure 3. Extraction algorithm of lip-image feature

입술정보를 이용한 인식 DB는 음성인식에 사용된 것과 동일한 고립단어로서 남성 50명의 360×240 pixel DB로 30frame/sec로 샘플링 된 영상이다. 입술정보는 <그림 3>과 같은 과정을 통해 특징 파라미터를 추출한다. 입술정보는 DCT기반 특징 파라미터 또는 입술의 기하학적 특징(폭, 높이, 넓이) 등으로 분석된다. 일반적으로 입술 파라미터는 DCT를 사용하지만 최근에는 DCT의 변형으로서 DCT-mod2라는 파라미터가 제안되었으며 이것은 조명의 변환에 강하도록 delta-parameter를 첨가한 방법이다. 입술영역의 전처리 과정은 컬러이미지를 그레이 이미지(160×240)로 변환하고 그레이 이미지 4분면 히스토그램을 만들어 입술 관심영역을 찾고 입술 위치를 16×16으로 down-sampling한 뒤 DCT 변환한다. 변환 후 PCA를 통해 입술 특징 정보를 추출하고 영상에 대한 각 프레임 PCA 결과는 HMM 인식 파라미터로서 이를 바탕으로 가장 유사도가 높은 5개의 단어를 출력한다. 입술정보를 이용한 인식은 Visual C++ 6.0로 구현되며 <그림 4>와 같다. 그림은 위에서 설명한 과정을 순서대로 나타낸 것이다.

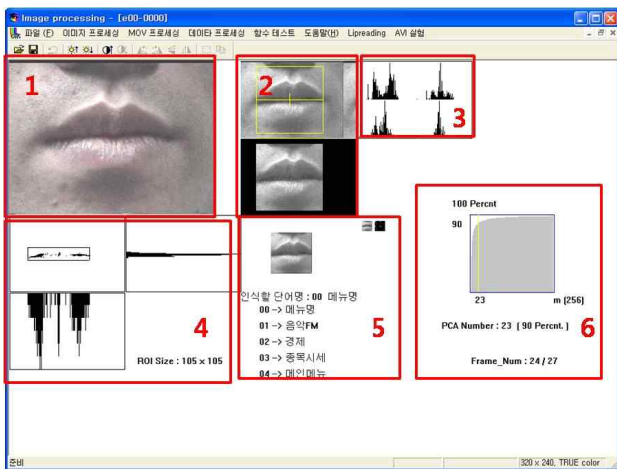


그림 4. 입술 영상인식 시스템  
Fig. 4. Lip image recognition system

3.2 거절성능 평가

시청각 인식 거절성능을 평가하기 위하여 다음과 같은 평가 기준을 선정하였다. 신뢰도의 거절성능 평가에서는 MDR(miss detection ratio)과 FAR(false alarm ratio)을 사용하며 MDR은 인식과정에서 인식단어가 정확하게 인식되지 못하는 경우이고 FAR은 사전 외 단어가 인식된 경우이다.

$$MDR = \frac{\text{Number of missed detection}}{\text{Total number of recognition word}} \quad (8)$$

$$FAR = \frac{\text{Number of false alarm}}{\text{Total number of recognition word}} \quad (9)$$

MDR은 인식과정에서 모든 단어에 대하여 입력단어와 인식 결과가 불일치한 경우의 비율을 나타내고, FAR은 모든 단어에 대하여 인식결과가 거절된 비율을 나타낸 것이다. MDR은 전체 인식성능을 나타낼 수 있으며, 임계값의 변화를 통하여 거절비율의 조정이 가능하다. 따라서 우수한 거절성능은 MDR과 FAR 모두 최소값을 갖도록 하는 임계값을 찾는 데 있다.

4. 실험 결과

4.1 인식성능

본 논문에서 적용된 신뢰도는 [9]에서 제안된 정규화 신뢰도를 사용한다. 제안된 신뢰도는 음소 기반 신뢰도를 단어단위로 확장하여 신뢰도를 계산하는 RLJC-CM [10]의 평균과 표준편차로 정규화한 신뢰도이다. 음성만을 이용한 인식성능을 측정하기 위하여 고립단어 100개를 HMM 학습 사전에 재 트레이닝하여 구성하였다. SNR은 <그림 2>에서와 같이 입력 음성에 의해 계산될 수 있으며, 계산된 SNR에 따라 입술영상과 음성의 가중치를 결정한다.

<그림 5>는 SNR 25dB에서의 거절성능의 결과로서 PSO-NCM은 NCM의 성능보다 우수하고 잡음이 많을수록 PSO-NCM의 성능이 우수함을 [6]에서 확인하였다. <그림 5>의 MDR 13.4% 부근에서 NCM의 FAR은 0.31이며, PSO-NCM은 0.27로 PSO-NCM의 성능이 13% 향상됨을 의미한다.

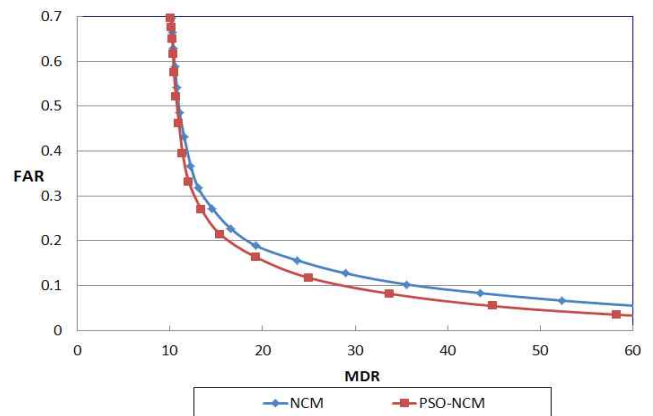





그림 5. 25dB에서의 거절성능  
Figure 5. Rejection performance in 25dB

<표 1>은 입술영상에 대한 인식성공 여부와 입술영상을 나타낸 것이다. 실험은 평가용 DB인 20명의 화자가 발성한 2,000 단어를 사용하였으며 83.1%의 인식 성공을 보인다. 인식결과는 평가 단어 중 가장 높은 신뢰도를 갖는 5개의 단어와 입력단어를 비교한 후 인식 유무를 판단한다. 화자에 따른 입력단어 “중합지수”에 대한 인식이 실패한 경우와 성공한 경우로 구분하여 화자를 분석하였다.

<표 1>에서 화자 3의 경우 확률이 가장 높은 단어는 “메뉴명”으로 이를 분석한 결과 단어 발생 시 입술의 변화가 거의 없이 발생하였고, 입술 ROI를 찾는 과정에서 입술 이외의 영역(코)이 포함되었음을 확인하였다. 이것은 인식이 실패한 주요인이라고 사료되며 다른 요인으로는 화자의 발음습관에 따라 입술의 움직임은 일정하지 못하다는 사실을 알 수 있었다. 그리고 화자 10, 17은 인식이 성공하였음을 보여준다.

표 1. 화자에 따른 인식 단어 영상(단어명: 종합지수)  
Table 1. Recognition word image in speaker

| 화자    | 인식 단어   | 인식 결과 | 입력영상  |
|-------|---------|-------|---|
| 화자 3  | 1. 메뉴명  | 실패    |    |
|       | 2. 교통정보 |       |   |
|       | 3. 종합지수 |       |   |
|       | 4. 증권정보 |       |   |
|       | 5. 종합지수 |       |   |
| 화자 10 | 1. 종합지수 | 성공    |   |
|       | 2. 표준FM |       |   |
|       | 3. 음악FM |       |   |
|       | 4. 문화정보 |       |   |
|       | 5. 투자정보 |       |   |
| 화자 17 | 1. 종합지수 | 성공    |  |
|       | 2. 문화정보 |       |   |
|       | 3. 투자정보 |       |   |
|       | 4. 방송정보 |       |   |
|       | 5. 음악FM |       |   |

4.2 통합 후의 인식성능

잡음이 포함된 음성에 대한 인식성능 개선을 위해 입술과 음성이 통합된 검증시스템을 바탕으로 실험을 실시하였다. 기존의 음성만을 이용한 인식성능과 통합된 인식성능을 비교실험하였다. 통합 신뢰도 계산은 SNR에 적응된 파라미터를 사용하여 잡음이 증가할수록 입술정보의 가중치가 증가하게 된다.

<표 2>는 통합 전 후의 인식성능을 비교 실험한 결과이다. 실험결과 잡음이 많은 구간인 10dB에서는 음성 신뢰도보다는 입술정보 신뢰도의 가중치가 높게 적용되어 인식성능이 향상됨을 알 수 있다. 잡음이 거의 존재하지 않는 SNR 30dB이상에서는 기존의 인식성능과 유사하다.

<그림 6>은 SNR 20dB에서 입술과 음성의 통합 전 후의 거절성능을 나타낸 것이다. 실험결과 통합 후의 NCM과 PSO-NCM의 성능이 우수함을 알 수 있다. 그림에서 MDR 20% 부근에서 통합 PSO-NCM의 FAR은 0.3이며, 통합 NCM은 0.45이다. 이는 통합 후의 PSO-NCM의 성능이 33% 향상됨을 알 수 있다. 따라서 PSO 최적화를 통한 음성과 입술정보의 통합은 잡음에 적합함을 알 수 있다.

음에 적합함을 알 수 있다.

표 2. 통합 전후의 인식 성능  
Table 2. Recognition performance before integration and after integration

| SNR   | 음성인식 성능(%) | 통합인식 성능(%) |
|-------|------------|------------|
| 10dB  | 45.16      | 83.12      |
| 15dB  | 78.11      | 82.46      |
| 20dB  | 85.49      | 85.41      |
| 25dB  | 87.69      | 87.62      |
| 30dB  | 89.40      | 89.10      |
| 40dB  | 90.18      | 90.02      |
| clean | 91.10      | 91.10      |

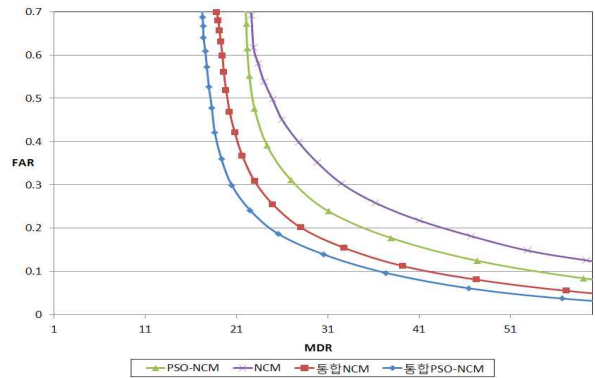


그림 6. 통합 전 후의 거절성능(SNR 20dB)  
Figure 6. Rejection performance before integration and after integration(SNR 20dB)

5. 결론

기존의 잡음이 포함된 음성인식은 음성정보만을 활용하기 때문에 인식성능이 매우 낮다. 이를 해결하기 위해 본 논문에서는 음성인식 신뢰도와 입술정보 신뢰도를 이용해 잡음에 종속된 통합방법을 제안하였고 잡음에 따른 신뢰도의 가중치를 이용하여 PSO 방법을 최적화하였다. 그 결과 SNR 20dB에서 통합 후의 PSO-NCM의 성능이 33% 향상됨을 알 수 있다. 향후 제한된 입술영역을 확장하여 시청각 음성인식에 대한 연구를 진행하고자 한다.

## 참고문헌

- [1] Gong, Y. F. (1995). "Speech recognition in noisy environments: a survey", *Speech Communication*, Vol. 16, 261-291.
- [2] Skovronski, M. D., Harris, J. G. (2006). "Applied principles of clear and lombard speech for automated intelligibility enhancement in noisy environments", *Speech Communication*, Vol. 48, 549-558.
- [3] Gowdy, J. N., Subramanya, A., Bartels, C., Bilmes, J. (2004). "DBN-based multi-stream models for audio-visual speech recognition", *proc. IEEE Int. conf. Acoustics, Speech, and Signal Processing*, 993-996.
- [4] Bilmes, J. A., Bartels, C. (2005). "Graphical model architectures for speech recognition", *IEEE Signal Processing Magazine*, Vol. 22, 89-100.
- [5] Schwartz, J. L., Berthommier, F., Savariaux, C. (2004). "Seeing to hear Better: evidence for early audio-visual interactions in speech identification", *ERIC Journal Articles: Reports-Research Cognition*, Vol. 93 No. 2, B69-B78.
- [6] Kim, B. D., Song, M. G., Choi, S. H., Kim, J. Y. (2008). "Enhancement of rejection performance using the PSO-NCM in noisy environment", *Speech Sciences*, Vol. 15, No. 4, 85-96.  
(김병돈, 송민규, 최승호, 김진영 (2008). "잡음 환경하에서의 PSO-NCM을 이용한 거절기능 성능 향상", *음성과학* 제 15권, 제4호, 85-96)
- [7] Eberhart, R., Kennedy, J. (1995). "A new optimizer using particle swarm theory", *Proc. of Sixth International Symposium on Micro Machine and Human Science*, 39-43.
- [8] Min, S. H., Kim, J. Y., Song, M. G., Na, S. Y. (2007). "Performance enhancement of speaker identification in noisy environments by optimization membership function based on particle swarm", *Speech Sciences*, Vol. 14, No. 2, 105-114.  
(민소희, 김진영, 송민규, 나승유 (2007). "Particle Swarm 기반 최적화 멤버십 함수에 의한 잡음 환경에서의 화자인식 성능향상", *음성과학* 제14권, 제2호, 105-114)
- [9] Kim, C., Lee, K. R., Kim, J. Y., Choi, S. H., Choi, S. H. (2002). "Improvement of keyword spotting performance using normalized confidence measure", *The Journal of the Acoustical Society of Korea*, Vol. 21, No. 4, 380-386.  
(김철, 이경록, 김진영, 최승호, 최승호 (2002). "정규화 신뢰도를 이용한 핵심어 검출 성능 향상", *한국음향학회지* 제21권 제4호, 380-386)
- [10] Rahim, M. G., Lee, C. H., Juang, B. H., Chou, W. (1996). "Discriminative utterance verification using minimum string verification error (MSVE) training", *Proc of ICASSP 96*, 3585-3589.
- **김병돈 (Kim, Byoungdon)**  
동신대학교 공과대학 컴퓨터학과  
전라남도 나주시 대호동 252번지  
Tel: 061-330-3194 Fax: 061-330-3369  
Email: doni96@empal.com  
관심분야: 신호처리, 음성인식  
현재 동신대학교 컴퓨터학과 강사
  - **최승호 (Choi, Seungho)**, 교신저자  
동신대학교 공과대학 컴퓨터학과  
전라남도 나주시 대호동 252번지  
Tel: 061-330-3194 Fax: 061-330-3369  
Email: shchoi@dnu.ac.kr  
관심분야: 신호처리, 음성인식  
현재 동신대학교 컴퓨터학과 교수