

GMM-UBM 기반 KL 거리를 활용한 화자변화 검증에 대한 연구

조준범¹, 이지은², 이경록^{3*}

¹남부대학교 간호학과, ²전남과학대학교 생활체육과, ³남부대학교 IT·디자인학과

The Study on the Verification of Speaker Change using GMM-UBM based KL distance

Joon-Beom Cho¹, Ji-eun Lee², Kyong-Rok Lee^{3*}

¹Dept. of Nursing, Nambu University

²Dept. of Living physical Training Special Study, Chunnam Techno University

³Dept. of IT & Design, Nambu University

요 약 본 논문에서는 기존의 BIC(Bayesian Information Criterion) 기반 화자변화의 성능 향상을 위하여 GMM-UBM(Gaussian Mixture Model-Universal Background Model) 기반 KL(Kullback Leibler) 거리를 활용한 화자변화 검증을 제안하였다. 정보량의 차이에 민감한 기존의 BIC 기반 화자변화검출 알고리즘을 상대적으로 정보량 차이에 견인한 KL 거리 알고리즘으로 검증하였고, 정보량의 비대칭을 보상하기 위해서 GMM-UBM을 활용하였다. 기존의 BIC 기반 화자변화 검출은 1단계로 비유사도 d 가 양수인 구간의 국소 최댓값인 지점을 화자변화 후보지점으로 검출하였고, 2단계로 검출된 화자변화 후보지점 중 ΔBIC 가 양수인 지점을 화자변화지점으로 결정하였다. 본 논문에서는 BIC 기반 화자변화 검출에 의해 결정된 화자변화지점에 대하여 GMM-UBM 기반 KL 거리 D 가 문턱치(threshold)보다 높은 지점을 최종 화자변화 지점으로 검증하였다. 실험결과, MDR(Missed Detection Rate)이 0인 조건에서 문턱치 0.028일 때 FAR(False Alarm Rate) 60.4%로 성능이 향상되었다.

키워드 : 화자변화검출, Bayesian Information Criterion, Gaussian Mixture Model-Universal Background Model, Kullback Leibler 거리

Abstract In this paper, we proposed a verification of speaker change utilizing the KL distance based on GMM-UBM to improve the performance of conventional BIC based Speaker Change Detection(SCD). We have verified Conventional BIC-based SCD using KL-distance based SCD which is robust against difference of information volume than BIC-based SCD. And we have applied GMM-UBM to compensate asymmetric information volume. Conventional BIC-based SCD was composed of two steps. Step 1, to detect the Speaker Change Candidate Point(SCCP). SCCP is positive local maximum point of dissimilarity d . Step 2, to determine the Speaker Change Point(SCP). If ΔBIC of SCCP is positive, it decides to SCP. We examined verification of SCP using GMM-UBM based KL distance D . If the value of D on each SCP is higher than threshold, we accepted that point to the final SCP. In the experimental condition MDR(Missed Detection Rate) is 0, FAR(False Alarm Rate) when the threshold value of 0.028 has been improved to 60.7%.

Key Words : Speaker Change Detection, Bayesian Information Criterion, Gaussian Mixture Model-Universal Background Model, Kullback Leibler distance

1. 서론

화자변화 검출은 입력 음성 데이터로부터 화자가 변화하는 지점을 검출하여 동일 화자에서 발생된 음성 클러스터를 생성하는 것이다. 요즘 큰 관심을 받고 있는 음향 빅데이터의 오디오 인덱싱, 화자인식과 인증, 화자 추적 등은 전처리부로 화자변화 검출을 통한 화자별 클러스터링 과정이 중요하다[1]. 뉴스 데이터 인덱싱을 중심으로 살펴보면 핵심화자인 앵커의 이야기를 분석하면 해당 기사의 주제를 추출할 수 있고, 이를 바탕으로 뉴스 데이터를 기사 단위로 분할/분석하는데 중요한 정보를 얻을 수 있다.

뉴스데이터 등 멀티미디어 데이터를 대상으로 한 화자변화검출은 일반적으로 화자에 대한 사전정보(발성화자의 수, 성별, 음향모델 등)가 없는 상황을 전제로 하기 때문에 화자변화 검출의 난이도가 높고, 적용할 수 있는 알고리즘도 제한적이다[2].

화자변화검출을 위한 알고리즘으로 거리(매트릭스) 기반 분할방식, 모델기반 분할 방식을 주로 사용한다[3]. 거리 기반 분할 방식은 화자에 대한 사전정보가 필요하지 않지만, 분석 윈도우의 길이(정보량)에 영향을 많이 받는다. 화자모델 기반 분할 방식은 화자에 대한 사전정보를 활용하여 미리 구축된 화자별 음향모델을 사용하기 때문에 분석윈도우의 길이에 상대적으로 덜 민감한 장점이 있지만 화자정보가 없는 환경에서는 활용하는데 제한 사항이 있다.

본 논문에서는 다양한 배경환경잡음이 혼재하고, 다양한 화자의 짧고 빈번한 발화경계가 나타나는 방송 뉴스 환경에서의 화자변화검출에 대해 실험했기 때문에 화자 수 및 화자에 대한 사전정보를 획득하기 어렵다. 이에 화자에 대한 사전정보가 필요하지 않은 거리 기반 분할방식을 선택하였다.

거리 기반 분할방식은 음성 데이터를 인접한 두 개의 분석 윈도우를 이용하여 분석한다. 두 분석 윈도우는 일정 시간 간격으로 이동하면서 두 분석 윈도우 간의 유사도를 측정하여 화자변화 지점을 검출하게 된다. 두 분석 윈도우 간의 유사도를 측정하는데 BIC(Bayesian Information Criterion), KL(Kullback Leibler), GLR(Generalized Likelihood Ratio)와 같은 유사도 함수를 사용한다. 일반적으로 거리 기반 분할 방식은 정보량에 따라 성능이 영향을 받는 것으로 알려져 있다[3-7].

본 논문에서는 [6]에서 MDR(Missed Detection Rate)가 대개의 조건에서 고루 좋은 결과를 보이고, 정보량의 변화에 상대적으로 민감한 BIC를 기본 화자변화 검출 알고리즘으로 선택하였다. 입력 데이터의 양에 따라 성능의 차이를 보이는 BIC 기반 화자변화 검출의 성능을 향상하기 위해서 상대적으로 정보량의 변화에 강한 KL 거리를 사용하였고, GMM(Gaussian Mixture Model) 훈련 데이터의 부족으로 인한 왜곡을 보정하기 위해서 GMM-UBM(Gaussian Mixture Model-Universal Background Model)을 적용하여 정보량 비대칭을 보정하였다.

2. BIC 기반 화자변화 검출

2.1 BIC 기반 화자변화 검출 1단계

BIC 기반 화자변화 검출은 시간 T_i 를 일정시간만큼 이동시켜가면서 T_i 에서 인접한 두 분석 윈도우 X, Y를 분석하여 화자변화의 발생 여부를 결정한다. 본 논문에서는 10초 길이의 분석윈도우를 0.5초씩 이동하면서 계산하였다. 분석 윈도우 $X = \{x_1, x_2, \dots, x_{N_x}\}$, $Y = \{y_1, y_2, \dots, y_{N_y}\}$ 이며 $Z = XU Y$ 이다.

T_i 에서 인접한 두 분석 윈도우의 발생화자가 동일하다는 가정, 즉 화자변화가 일어나지 않았다는 가정 H_0 의 로그 유사도 L_0 를 정의하면 다음과 같다.

$$L_0 = \sum_{i=1}^{N_x} \log p(x_i | \theta_z) + \sum_{i=1}^{N_y} \log p(y_i | \theta_z) \quad (1)$$

θ_z 는 분석 윈도우 Z의 데이터를 대상으로 EM(Expectation Maximization) 알고리즘을 적용하여 모델링한 GMM(Gaussian Mixture Model) 파라미터들이다.

T_i 에서 두 분석윈도우의 발생화자가 동일하지 않다는 가정, 즉, 화자변화가 일어났다는 가정 H_1 의 로그 유사도 L_1 을 정의하면 다음과 같다.

$$L_1 = \sum_{i=1}^{N_x} \log p(x_i | \theta_x) + \sum_{i=1}^{N_y} \log p(y_i | \theta_y) \quad (2)$$

θ_x, θ_y 는 각 분석윈도우 X, Y에서 EM 알고리즘을 적

용하여 모델링된 GMM 파라미터들이다.

두 가정을 활용한 비유사도 d 는 다음과 같이 정의할 수 있다.

$$d = L_1 - L_0 - \frac{\lambda_d}{2} \cdot \Delta K \cdot \log N_z \quad (3)$$

N_z 은 분석 윈도우 Z 의 총 샘플 수이고, λ_d 는 실험을 통해 조절되는 페널티 팩터, ΔK 는 모델 파라미터의 수이다. 비유사도 $d > 0$ 인 구간 중에서 국소 최대가 되는 지점 T_i 에서 화자변화가 발생했다고 정의한다.[1, 3, 8-10].

2.2 BIC 기반 화자변화 검출 2단계

[1]에서는 비유사도 d 를 사용하여 검출된 화자변화 지점에 대하여 ΔBIC 를 적용하여 FA(False Alarm)을 줄이는 방법을 사용하였다. 화자변화 후보지점 T_i 에서 분석윈도우 X 는 $T_{i-1} \sim T_i$, 분석윈도우 Y 는 $T_i \sim T_{i+1}$, 분석윈도우 Z 는 $T_{i-1} \sim T_{i+1}$ 이라고 정의하고, 분석윈도우 X, Y, Z 로부터 훈련된 가우시안 모델을 $N(\mu_x, \Sigma_x)$, $N(\mu_y, \Sigma_y)$, $N(\mu_z, \Sigma_z)$ 라 정의하면 ΔBIC 는 다음과 같다.

$$\Delta BIC = \frac{1}{2} (N_z \log |\Sigma_z| - N_x \log |\Sigma_x| - N_y \log |\Sigma_y|) \quad (4)$$

$$- \frac{1}{2} \cdot \lambda \left(\Delta K + \frac{1}{2} \Delta K (\Delta K + 1) \right) \log (N_x + N_y)$$

λ 는 모델 복잡도에 대한 페널티 팩터이고, Σ_x 는 분석윈도우 X 에서 모델링된 GMM의 공분산, Σ_y 는 분석윈도우 Y 에서 모델링된 GMM의 공분산, Σ_z 는 분석윈도우 Z 에서 모델링된 GMM의 공분산이다. $\Delta BIC > 0$ 이면 화자변화 후보지점 T_i 에서 화자변화가 발생한 것이다.

3. GMM-UBM 기반 KL 거리를 활용한 검증

3.1 GMM-UBM

일반적으로 GMM은 음성인식, 화자인식 등 다양한 분야에서 활용된다[11, 12]. GMM의 정밀도를 높이기 위해서는 대량의 데이터가 필요하다. 실험환경의 특성상, 짧

은 길이의 분석윈도우의 데이터를 가지고 모델링해야 하므로 데이터 부족으로 인한 GMM의 왜곡이 문제가 된다.

본 논문에서는 이러한 데이터 부족을 해소하기 위해서 UBM(Universal Background Model)을 적용하였다. 개념적으로 모든 화자의 음성을 대표하며, MAP(Maximum A Posteriori)를 사용하여 각 화자 모델에 적용된다. 화자인증시스템에서는 승인 혹은 거절의 결정이 필요한 때, 특정 화자 정보에 의한 모델과 비교하기 위한 화자독립 모델로 사용된다[13-17].

3.2 KL 거리 D

입력정보량에 따라 성능의 차이를 보이는 BIC 기반 화자변화검출 방법을 보완하기 위해서 상대적으로 강한 KL 거리 방법을 적용하였다.

각 분석 윈도우의 데이터를 사용하여 GMM을 계산하였다. 구축된 각 분석윈도우의 GMM을 비교하기 위해서 KL 거리를 적용하여 모델간 거리를 측정하였다. KL 거리에서 모델의 평균은 환경조건의 변화(배경잡음 등)에 영향을 많이 받기 때문에 이를 제외한 공분산만을 사용하였다. 다음은 두 분석 윈도우 X, Y 에서 훈련된 GMM의 공분산을 활용한 각 모델간 거리를 계산하는 공식이다[15, 18, 19].

$$D = \frac{1}{2} \text{tr} \left[(\Sigma_x - \Sigma_y) (\Sigma_y^{-1} - \Sigma_x^{-1}) \right] \quad (5)$$

Σ_x, Σ_y 는 분석 윈도우 X, Y 의 GMM의 공분산이다. D 가 작으면 두 GMM간의 거리가 가깝다는 뜻이므로 두 분석 윈도우가 동일 화자에게서 발생되었다는 것이다. 실험에서는 문턱치를 조절하면서 문턱치보다 낮은 D 값을 가진 화자변화 지점은 화자변화가 일어나지 않은 것으로 처리하였다.

GMM을 활용한 거리 기반 분할방식 알고리즘들의 일반적인 문제점인 GMM의 훈련 데이터 부족으로 인한 왜곡을 보상하기 위해서 기존의 KL 거리에 GMM-UBM을 적용하였다.

$$D = \frac{1}{2} \text{tr} \left[(\Sigma_{ux} - \Sigma_{uy}) (\Sigma_{uy}^{-1} - \Sigma_{ux}^{-1}) \right] \quad (6)$$

Σ_{ux}, Σ_{uy} 는 분석 윈도우 X, Y 의 GMM-UBM의 공분산이다. 실험에서는 KL 거리 방법과 동일하게 문턱치를

조절하면서 문턱치보다 낮은 D값을 가진 화자변화 지점을 거절하였다.

Fig. 1은 기존의 BIC 기반 화자변화검출 알고리즘에 의해 검출된 화자변화 후보지점을 본 논문에서 제안한 KL 거리 D(GMM-UBM 미적용, GMM-UBM 적용)를 활용하여 검증하는 과정을 간략하게 나타낸 것이다.

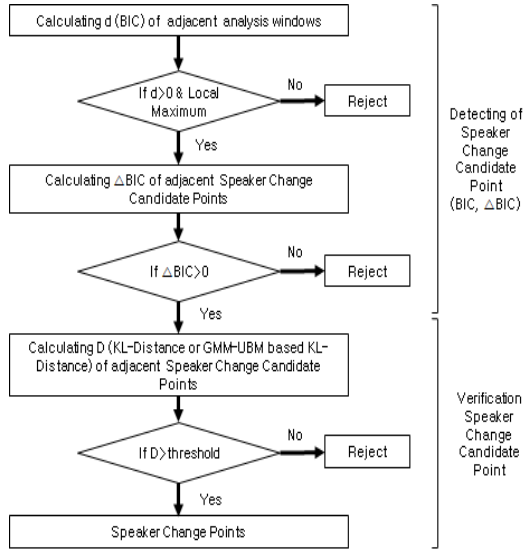


Fig. 1. A brief block diagram of the verification of Speaker Change Candidate Points using KL-Distance

4. 화자변화검출 실험

4.1 실험환경

화자변화 검출을 위하여 실제 뉴스 데이터 1회분을 분석하여 잡음환경, 화자정보(성별) 등의 특징을 고려하여 테스트 데이터를 구축하였다. 특히, 실제 뉴스 데이터에서 잡음환경의 비중이 더 큰 것을 반영하여 잡음 환경에서의 화자변화지점을 더 많이 고려하였다. 실험을 위한 데이터베이스 구성은 <Table 1>과 같다. 데이터베이스는 남성화자 5명, 여성화자 4명, 총 9명의 발성 데이터를 중심으로, 배경잡음의 존재여부, 이성간 혹은 동성 간 화자변화 패턴을 고려하여 구성하였다. UBM은 테스트 데이터를 추출한 뉴스 1회분 전체를 사용하여 훈련하였다.

선택된 뉴스 1회분에는 배경잡음이 없는 환경에서의 화자변화, 다양한 강도의 배경잡음환경에서의 화자변화

등, 뉴스 데이터에서 발생하는 화자변화의 형태가 대부분 포함되어 있다.

Table 1. Description of database

Background Noise(Y/N)	Number of speaker change	Pattern of speaker change
Yes	15	Male→Male
	11	Male→Female
No	8	Male→Male
	6	Male→Female

4.2 BIC 기반 화자변화 검출

BIC 기반 화자변화 검출은 2단계로 구축되었다.

1단계는 화자변화 후보 지점을 검출하기 위하여 입력 데이터의 비유사도 d 를 계산하고 $d > 0$ 이고 국소 최대가 되는 T_i 지점을 화자변화 후보 지점으로 검출하였다. 다음 단계를 위해서 페널티 벡터 λ_d 를 변경해가면서 MDR이 0일 때 FAR(False Alarm Rate)이 가장 낮은 조건을 찾았다.

2단계, 화자변화 후보지점 검증을 위해서 각 화자변화 후보지점의 ΔBIC 를 계산하여 $\Delta BIC > 0$ 인 경우 화자변화 지점으로 결정하였다. MDR을 0으로 고정한 상태에서 λ 를 변경해가면서 실험하였다.

실험결과 분석을 위하여 MDR, FAR을 다음과 같이 정의하였다.

$$MDR = \frac{N_{scd} - N_{cd}}{N_{scd}} \times 100 \quad (7)$$

$$FAR = \frac{N_{all} - N_{cd}}{N_{all}} \times 100$$

N_{scd} 는 실제 화자변화 지점의 수, N_{cd} 는 실제 화자변화 지점 중 검출된 수, N_{all} 는 검출된 전체 화자변화 지점의 수이다. N_{cd} 를 판별하기 위해서 검출된 화자변화 후보지점과 실제 화자변화 지점 간에 0.5 sec이내의 오차를 허용하였다.

BIC 기반 화자변화 검출을 실험한 결과, BIC 2단계에서 1단계에 비해 FAR이 15.7% 향상되었다. 화자변화 검출 2단계에서 인접한 분석 윈도우 길이(정보량)의 비대칭이 큰 화자변화 후보지점, 분석 윈도우 길이가 짧은 화자변화 후보지점에서 MDR, FAR이 다수 발생하였다.

Table 2. Experiment result of the BIC based SCD

Categories	Penalty factor	MDR	FAR
BIC Step 1.	$\lambda_d = 0.14$	0	85.4
BIC Step 2.	$\lambda = 0.16$	0	69.7

4.3 GMM-UBM 기반 KL 거리 D를 활용한 화자변화 검증

4.2의 실험결과 분석에서 나타난 취약점을 보완하기 위해서 정보량의 변화에 상대적으로 견인한 KL 거리를 적용하여 검증을 실시하였고 정보량의 비대칭성을 보상하기 위해서 GMM-UBM을 활용하였다.

BIC 2단계에서 결정된 화자변화지점에 대하여 문턱치를 변경해가면서, D가 문턱치(threshold)보다 클 경우에만 최종 화자변화 지점으로 인정하였다.

실험결과, KL 거리를 이용한 화자변화 후보지점 검증은 기존 BIC 기반 화자변화 검증에 비해서 성능이 향상되었다. GMM-UBM을 적용하지 않은 경우 MDR이 0인 조건에서 문턱치 0.029에서 FAR이 6.4% 향상되었고, GMM-UBM을 적용한 경우 MDR이 0인 조건에서 문턱치 0.028에서 FAR이 9.3% 향상되었다. 이를 통해서 GMM-UBM을 통한 분석원도우 길이 비대칭 보상이 효과가 있음을 알 수 있었다.

Fig. 2는 문턱치에 따라 변화하는 GMM-UBM 기반 KL 거리를 활용한 화자변화 검증실험의 MDR, FAR를 나타낸 것이다. 최종화자변화 지점으로 검증받지 못한 화자변화지점은 분석원도우 확장에 의해 분석원도우 X로 통합되기 때문에 MDR과 FAR이 변화하는 것을 볼 수 있다.

화자변화 검증 결과를 분석한 결과 배경잡음환경에서 화자변화가 일어나는 경우, 짧은 발성구간을 가진 이상 화자간의 화자변화가 일어나는 경우에 화자변화 지점의 검증 변별력이 감소하는 것을 확인할 수 있었다. 이것은 실제 뉴스 데이터에서 빈발하는 화자변화 유형으로서 향후 연구를 통해 개선할 필요성이 있다.

Table 3. Experiment result of the verification of SCP using KL distance D

GMM-UBM (Y/N)	Threshold	MDR	FAR
No	0.029	0	63.3
Yes	0.028	0	60.4

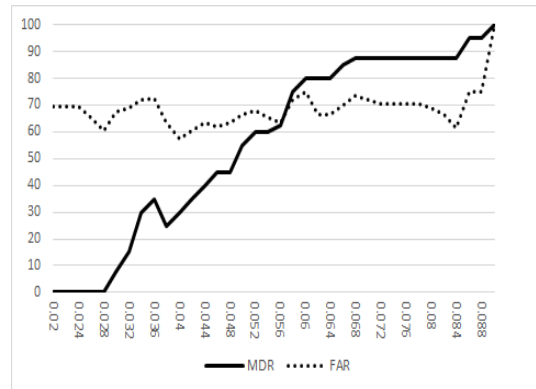


Fig. 2. Experiment result of the verification of SCP using GMM-UBM based KL distance D

5. 결론

본 논문에서는 기존의 BIC 기반 화자변화 검출의 성능 향상을 위하여 GMM-UBM 기반 KL 거리를 활용한 화자변화 검증을 제안하였다.

기존의 BIC 기반 화자변화 검출은 1단계로 비유사도 d가 양수인 구간의 국소 최댓값인 지점을 화자변화 후보 지점을 검출하였다. 2단계로 검출된 화자변화 후보지점 중 ΔBIC 가 양수인 지점을 화자변화지점으로 결정하였다.

2단계에서 인접한 분석 윈도우의 비대칭으로 발생한 정보량의 차이가 큰 지점에서 인식오류가 발생하는 것을 보완하기 위해서 상대적으로 정보량의 변화에 견인한 KL 거리를 이용하여 검증을 실시하였다. 또한, 정보량 비대칭을 보상하기 위해서 KL 거리에 GMM-UBM을 적용하였다. BIC 기반 화자변화 검출에 의해 결정된 화자변화지점에 대하여 GMM-UBM 기반 KL 거리 D가 문턱치보다 높은 지점을 최종 화자변화 지점으로 검증하였다.

실험결과, MDR이 0인 조건에서, 문턱치 0.028일 때 FAR 60.4%로 성능이 향상되었다.

향후 정보량의 변화에 견인한 복수개의 알고리즘을 이용한 하이브리드 검증에 대해 연구할 계획이다.

REFERENCES

[1] M. Kotti, E. Benetos, C. Kotropoulos, L. Gustavo and P. M. Martins, "Speaker Change Detection using BIC :

- A Comparison on two dataset," *Proceedings of the 2006 Second International Symposium on Communications, Control and Signal Processing(ISCCSP 2006)*, pp. 1-4, 2006.
- [2] Rosenberg Aaron, M.C. Ivan and S. Parthasarathy, "Speaker detection in broadcast speech databases," *Proceedings of the 5th International Conference on Spoken Language Processing(ICSLP '98)*, pp. 1-4, 1998.
- [3] T. Theodorou, I. Mporas and N. Fakotakis, "An overview of automatic audio segmentation," *Proceedings of the I.J. Information Technology and Computer Science*, Vol. 11, pp. 1-9, 2014.
- [4] S. E. Tranter and D. A. Reynolds, "An overview of automatic speaker diarization systems," *Journal of the IEEE Trans. Audio, Speech and Language Processing*, Vol. 14, no. 5, pp. 1557-1565, 2006.
- [5] M. Y. Choi and H. S. Kim, "Speaker Segmentation System Using Eigenvoice-based Speaker Weight Distance Method," *Journal of the Acoustical Society of Korea*, Vol. 31, No. 4, pp. 266-272, 2012.
- [6] K. R. Lee, "The Study on Speaker Change Detection based on Matrix distance," *The Journal of Korean Institute of Information Technology*, Vol. 11, No. 1, pp. 71-77, 2013.
- [7] H. S. Kim, H. Y. Park and S. Y. Park, "Speaker Change Detection by Normalization of Phonetic Characteristics", *Journal of the Malsori*, Vol. 47, pp. 98-107, 2003.
- [8] S. Chen and Gopalakrishnan, "Speaker, Environment and Channel Change Detection and Clustering via the Bayesian Information Criterion," *Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop*, pp.127-132, 1998.
- [9] P. Delacourt and C. J. Wellekens, "DISTBIC: A speaker-based segmentation for audio data indexing," *Journal of the Speech Communication*, Vol. 32, Issue 1-2, pp.111-126, Sep. 2000.
- [10] I. Ajmera, I. McCowan, and H. Bourlard, "Robust speaker change detection," *IEEE Signal Processing Letters*, vol. 11, no. 8, pp.649-651, 2004.
- [11] J. S. Lee, "Vocabulary Recognition Performance Improvement using k-means Algorithm for GMM Support," *Journal of the Digital Convergence*, Vol. 13, No. 2, pp.135-140, 2015.
- [12] S. Y. Oh, "Voice Recognition Performance Improvement using a convergence of Voice Energy Distribution Process and Parameter," *Journal of the Digital Convergence*, Vol. 13, No. 10, pp.313-318, 2015.
- [13] D. Povey, S. M. Chu and B. Varadarajan, "Universal background model based speech recognition," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2008(ICASSP'08)*, pp. 4561-4564, 2008.
- [14] D. A. Reynolds, T. F. Quatieri, and R. Dunn, "Speaker verification using adapted Gaussian mixture models," *Journal of the Digital Signal Processing*, Vol. 10, No. 1-3, pp. 19-41, Jan. 2000.
- [15] T. Wu, L. Lu, K. Chen, and H. Zhang, "Universal Background Models for Real-time Speaker Change Detection," *Proceedings of the IEEE Int'l. Conf. on Multimedia Modeling*, pp.35-149, 2003.
- [16] R. Vogt, J. Pelecanos and S. Sridharan, "Dependence of GMM adaptation on feature post-processing for speaker recognition," *Proceedings of the 8th European Conference on Speech Communication and Technology(EUROSPPEECH'03)*, pp. 3013-3016, September 2003.
- [17] S. M. Chu, H.-K. Kuo, Y. Y. Liu, Y. Qin, Q. Shi, and G. Zweig, "The IBM Mandarin broadcast speech transcription system," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2007(ICASSP 2007)*, Vol. 2, pp. 345-348, May 2007.
- [18] J. P. Campbell. "Speaker Recognition: A Tutorial," *Proceedings of the IEEE*, Vol. 85, No. 9, pp. 1437-1462, 1997.
- [19] L. Lu, H. J. Zhang, "Speaker Change Detection and tracking in Real-time News Broadcasting Analysis," *Proceedings of the Proceedings of the tenth ACM international conference on Multimedia(MULTIMEDIA '02)*, pp. 602-610, 2002.

저 자 소 개

조 준 범(Joon-Beom Cho) [정회원]



- 1989년 2월: 원광대학교 무역학과 학사
- 1995년 2월: 원광대학교 영어영문학과 석사
- 2005년 8월: 원광대학교 영어영문학과 박사

▪ 1999년 3월 ~ 현재 : 남부대학교 교수

<관심분야> : 멀티미디어 콘텐츠 인텍싱, 언어학, 통사론

이 지 은(Ji-Eun Lee) [정회원]



- 2001년 2월: 조선대학교 무용과 학사
- 2003년 2월: 조선대학교 체육학 석사
- 2006년 8월: 전남대학교 체육학 박사

▪ 2015년 9월 ~ 현재: 전남과학대학교 부교수

<관심분야> : 스포츠 콘텐츠 인텍싱, 스포츠융합, 통계

이 경 록(Kynog-Rok Lee) [정회원]



- 1997년 2월 : 호남대학교 전자공학과 공학사
- 2001년 8월 : 전남대학교 정보통신협동과정 공학석사
- 2006년 2월 : 전남대학교 컴퓨터전자공학부 공학박사

▪ 2008년 4월 ~ 현재 : 남부대학교 IT·디자인학과 부교수

<관심분야> : 멀티미디어 콘텐츠 인텍싱, 화자인식, 음성인식