

화자분할을 위한 지역적 특성 기반 밀도 클러스터링

Local Distribution Based Density Clustering for Speaker Diarization

노진상, 손수원*, 김성수**, 이재원**, 고한석*†

(Jinsang Rho, Suwon Shon,* Sung Soo Kim,** Jae-Won Lee,** and Hanseok Ko*†)

고려대학교 영상정보처리협동, *고려대학교 전기전자공학과, **삼성전자
(Received March 3, 2015; revised June 6, 2015; accepted July 6, 2015)

초 록: 화자 분할은 사전에 분류되지 않은 데이터를 각각의 화자로 분류하는 연구이며 DBSCAN(Density-Based Spatial Clustering of Applications with Noise)은 간결함과 계산의 효율성으로 인해 화자분할 분야에 널리 사용되어 왔다. 그러나 클러스터의 데이터들이 공간적이지 않으며 서로 다른 클러스터가 근접하여 경계를 공유할 때 오버클러스터링 문제가 발생하여 DBSCAN의 성능이 하락한다. 본 논문에서는 DBSCAN과 문제점을 설명하고, 개체의 지역적 특성에 기반한 밀도 기반 클러스터링 알고리즘을 제안한다. 제안하는 알고리즘은 개체의 지역적 밀도와 분산의 정도에 따라 가변적인 판단 기준을 탐색에 이용한다. DBSCAN과 제안 기법의 실험을 통해 성능을 비교하고 제안 기법의 효율을 보인다. 실험 결과 제안한 방법은 오버클러스터링이 발생하지 않으며 DBSCAN에 비해 보다 높은 정확도를 보여 지역적 특성을 이용한 접근 방법이 효과적임을 증명한다.

핵심용어: 밀도 기반 클러스터링, 화자 분할, DBSCAN, 지역적 밀도, 오버클러스터링

ABSTRACT: Speaker diarization is the task of determining the speakers for unlabeled data, and DBSCAN (Density-Based Spatial Clustering of Applications with Noise) has been widely used in the field of speaker diarization for its simplicity and computational efficiency. One challenging issue, however, is that if different clusters in non-spatial dataset are adjacent to each other, over-clustering may occur which subsequently degrades the performance of DBSCAN. In this paper, we identify the drawbacks of DBSCAN and propose a new density clustering algorithm based on local distribution property around object. Variable density criteria for local density and spreadness of object are used for effective data clustering. We compare the proposed algorithm to DBSCAN in terms of clustering accuracy. Experimental results confirm that the proposed algorithm exhibits higher accuracy than DBSCAN without over-clustering and confirm that the new approach based on local density and object spreadness is efficient.

Keywords: Density based clustering, Speaker diarization, DBSCAN, Local density, Over-clustering

PACS numbers: 43.71.Bp, 43.72.Ar

1. 서 론

최근 오디오 미디어의 수는 빠르게 증가하고 있으며, 사용자가 수동적으로 미디어에 대한 정보를 분석하는 것 대신 자동적으로 막대한 양의 데이터를 분석이 가능한 방법에 대한 수요가 증가하고 있다. 화자 분할에 대한 연구는 연속된 음향 정보로부터 음성에 관련된 정보는 화자에게, 음성외의 정보는

비음성으로 분류하는 있는 연구로서 정의할 수 있으며, 요약하자면 “who spoke when?”이라는 질문에 대답하는 것이라고 할 수 있다.^[1,2]

화자 분할은 일반적으로 동영상, 음성 미디어에 대해 신호처리를 적용하는 분야에서 주로 연구되고 있다. 전화통화, TV방송 프로그램, 교육 강좌, 라디오와 같은 미디어들이 사용 대상으로 고려될 수 있다. 미디어의 종류에 따라 음향의 품질, 잡음의 종류, 발화의 형식이 달라질 수 있으며 각각의 경우에 대한 연구들이 진행되어왔다.^[1,3]

†Corresponding author: Hanseok Ko (hsko@korea.ac.kr)
Engineering building room 419, Korea University, 145 Anam-Ro, Seongbuk-Gu, Seoul 136-713, Republic of Korea
(Tel: 82-2-3290-3239, Fax: 82-2-3291-2450)

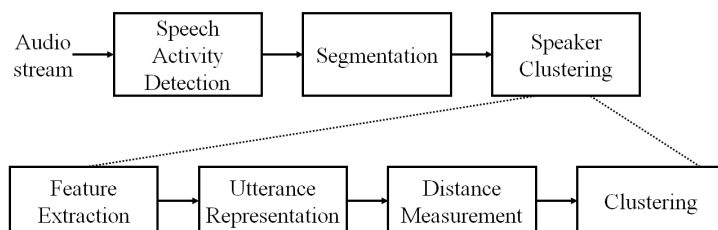


Fig. 1. The general process of speaker diarization.

미디어 정보에서 영상 정보를 제외한 오디오 정보를 추출하고 이를 입력으로 사용하는 화자 분할의 수행은 주로 Fig. 1과 같은 순서에 따라 이루어진다.^[4]

첫 번째 단계인 음성 구간 검출(Speech Activity Detection, SAD)에서는 인공적인 효과음, 묵음, 사람의 비언어적인 음향들을 제거하고 발화들만을 남김으로써, 이후의 단계에서 불필요한 노이즈로 인해 알고리즘 결과에 무의미한 정보나 오류가 포함되지 않는 것을 목표로 한다.^[5]

이러한 이유로 음성 인식, 음성-문자 변환과 같이 음성 신호처리와 관련된 연구에서 전처리를 위해 일반적으로 SAD를 채용한다. 일반적으로 GMM(Gaussian Mixture Model), NMF(Non-negative Matrix Factorization)와 같은 모델에 기반한 접근 방법이 주로 사용되고 있으며^[6,7] 사전에 음성과 배경 소음, 종이 넘기는 소리, 두드리는 소리와 같은 비음성들의 모델을 미리 훈련하고 상황에 적용한다.

음성 분할(speech segmentation)은 연속된 오디오 정보로부터 음향적으로 성질이 균일한 구획을 분리하며, 이러한 구획들은 각각 한 명의 화자로부터 발생된 것으로 간주한다.^[8]

화자 분할을 거친 발화들 간의 유사성을 비교하기 위해 음성의 특징을 추출한다. 결과적으로 각 발화는 일련의 특징 벡터로 표현되며 이를 사용하여 개체 간의 상대적 거리를 측정할 수 있다. 화자별로 발화를 분류하는 과정은 일반적으로 레이블 되지 않은 데이터를 본래의 집단 별로 분리해내는 데이터 클러스터링과 유사한 과정을 따르게 된다. 화자 분할에서는 각 화자 별로 클러스터를 구성하며 각 클러스터는 오직 한 화자의 발화만이 포함되어야 한다.^[9-11]

데이터 클러스터링은 서로 다른 개체들로 이루어진 집단 내에서 유사한 특성을 갖는 개체들끼리 모

아 클러스터를 형성하는 것을 의미하며, 일반적으로 진행 과정에 따라 하향식(top-down)과 상향식(bottom-up) 접근방법으로 나누어진다. 두 방식의 구분은 최적화된 클러스터를 탐색하기 위해 사전에 결정된 초기 클러스터를 확장 혹은 새로운 클러스터를 추가하는 방법 하향식과 기존의 클러스터들 간의 관계를 비교, 병합을 반복하는 접근 방법 상향식이다.

하향식 접근 방법의 경우 레이블 되지 않은 데이터를 차례로 사전에 정해진 클러스터로 분류할 때, 초기 클러스터의 결정 방법과 클러스터의 추가 여부를 판단하기 위한 판단 기준에 대해 주로 연구되어 왔다. 반대로 상향식 접근 방법은 기존의 클러스터(혹은 개체)들간의 병합을 반복하며, 최적의 클러스터 수에 도달할 때까지 반복하기 위한 방법을 연구한다.

일반적으로 상향식 접근 방법의 클러스터링 알고리즘은 초기 단계에서 최적화된 클러스터의 수보다 많은 클러스터를 가정한다. 병합 후 클러스터의 수가 수렴할 때까지 반복했을 때의 클러스터가 최적화된 클러스터보다 적을 경우 오버 클러스터링 되었다고 표현하며 이는 상향식 방식이 일으킬 수 있는 문제로서 지적되었다.^[12-13]

본 논문에서는 실제 상황에서 오버 클러스터링 되기 쉬운 DBSCAN(Density-Based Spatial Clustering of Applications with Noise)의 문제점을 설명하고 이를 개선한 방법을 제안하고자 한다. 인접한 클러스터들이 존재하는 실제 데이터 집단에서는 오버 클러스터링 문제가 발생할 수 있다. 지역적인 특성을 이용한 알고리즘은 인접한 클러스터간을 구별할 수 있는 기준을 사용하여 문제를 해결한다.

논문은 다음과 같이 구성되어 있다. II장에서는 종래의 방법과 이슈를 소개하고 이를 개선한 제안 기

법을 설명한다. III장에서는 기존 방법과 제안한 방법의 비교 실험을 통하여 성능을 검증할 것이다.

II. 기존 연구 및 제안 기법

2.1 DBSCAN

밀도 기반 클러스터링은 다른 개체와의 상대적인 관계를 밀도로 정의하여 이를 클러스터링의 지표로써 활용한다.^[14] 서로 다른 클러스터들을 적절히 분리할 수 있으면서 밀도가 낮은 개체가 노이즈로 분류되는 것을 방지할 수 있는 평가 기준을 사용하고 클러스터를 탐색하는 것이 목표이다.

DBSCAN은 N 개의 개체로 이루어진 집단의 클러스터를 분류하기 위해 각 개체별로 정해진 최소 기준거리 내의 근접 개체 수가 기준 수 이상인지에 따라 클러스터를 결정한다.^[15] DBSCAN에서는 클러스터의 확장과정에서 코어포인트를 기준치 이상의 밀도를 갖는 개체로 정의하며, 보더포인트는 상대적으로 밀도가 떨어지는 개체로 정의한다. 따라서 클러스터의 확장과정에서 낮은 밀도의 보더포인트는 고려되지 않으며 근접한 코어포인트들 간의 관계만을 고려한다.

DBSCAN을 사용하여 클러스터를 수행하기 위해서는 크게 두 단계의 과정을 거치게 된다. 첫째로 임의의 코어포인트 p 를 결정하고 p 의 모든 밀도 근접 포인트를 탐색한다. 두 번째로는 p 로부터 밀도 근접 포인트들을 재귀적으로 탐색하여 연결고리를 구성한다. 이러한 과정은 Fig. 2에서 코어포인트 p, q 사이의 연결고리를 탐색하여 p, q, m, n 이 하나의 클러스터로 묶이는 것을 볼 수 있다.

2.2 DBSCAN의 한계점

DBSCAN은 공간적 데이터베이스에서 클러스터와 노이즈를 구분하기 위해 설계되었으며 알고리즘

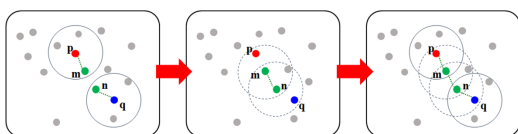


Fig. 2. The general process of DBSCAN.

이 복잡하지 않으며 다른 기법에 비해 상대적으로 빠른 연산 속도를 갖는다는 장점이 있다.

Fig. 3(a)과 같이 각 클러스터에 속하는 개체와 노이즈가 구분될 수 있는 데이터베이스에 대해 DBSCAN은 효과적인 방법이 될 수 있다. 그러나 Fig. 3(b)같이 공간적 데이터가 아닌 서로 다른 클러스터가 상대적으로 근접한 위치에 있어 둘의 보더포인트가 구분이 어려워질 경우, 클러스터의 경계에서 다른 클러스터의 점과 연결될 수 있는 가능성이 증가한다.^[16] 결과적으로 DBSCAN 알고리즘 내 클러스터의 확장 단계에서 클러스터의 지나친 확장, 혹은 외부 클러스터로부터 침범당하는 결과를 관측할 수 있다.^[17] 이러한 현상은 DBSCAN의 단점으로 지적되었으며 이를 해결하기 위한 접근 방법이 필요하다.

2.3 DBSCAN을 개선한 제안 방법

2.2에서 언급하였던 것과 같이 인접한 클러스터의 경계에서 발생할 수 있는 DBSCAN의 문제는 지역에 따른 개체의 차이를 반영하지 않고 동일한 코어포인트로서 판단하는 것에서 기인한다고 할 수 있다. 따라서 본 논문에서는 서로 다른 코어포인트들의 밀도와 지역적인 특성을 반영함으로써 이를 해결하고자 한다.

정의 1: 개체의 지역밀도

$$\rho_p = \frac{1}{N} \sum_{j=1}^N distance(p, NN_{p,j}), \quad (1)$$

$$distance(p, q) = \sqrt{(p - q)^T (p - q)}. \quad (2)$$

점 p 의 지역밀도 ρ_p 는 점 p 를 기준으로 N 개의 최

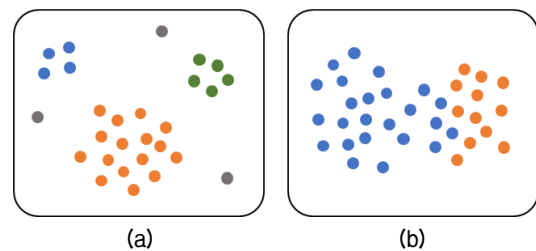


Fig. 3. Spatial sample data for DBSCAN (a) and issued sample data (b).

근린 점(nearest neighbor)과의 거리 평균을 사용하며, $NN_{p,j}$ 는 점 p 주변에서 j 번째의 최근린 점을 의미한다. 하나의 클러스터 내에 속하는 개체들의 밀도는 개체 주변의 분포에 의해 결정되며 서로 근접한 개체들은 유사한 지역밀도를 보이게 된다.

정의 2: 개체의 거리 가중치

$$\omega_p = \text{diag} \left[\frac{1}{N-1} \sum_{j=1}^N (p - NN_{p,j})(p - NN_{p,j})^T \right]. \quad (3)$$

비슷한 지역밀도를 갖는 개체들 일지라도 클러스터 내 위치에 따라 서로 다른 분포를 갖는다. 클러스터의 확장 단계에서 이러한 점 p 의 주변 분포 특성 분산도를 확장 가능한 점의 탐색을 위해 가중치로 사용한다.

정의 3: 거리 내 확장성 판단

$$1) q \in N_{Eps}(p) \text{ and} \quad (4)$$

$$2) \sqrt{(p-q)^T W_p(p-q)} < \omega(p) * Eps. \quad (5)$$

점 p 의 최소 기준거리 $N_{Eps}(p)$ 내에 존재하는 점 q 으로부터 점 p 까지의 거리가 지역밀도를 가중치로 적용한 최소 기준거리 내라면 q 는 p 로부터 클러스터의 확장이 가능한 후보 근린점이 된다.

정의 4: 클러스터의 확장

점 q 가 점 p 로부터 근접한 밀도를 갖는다면 q 는 p 의 클러스터에 속하는 것으로 정의한다. 이 때 p 는 q 보다 지역밀도가 높은 개체이다.

제안하는 방법에서 지역밀도와 분산도는 다음과 같이 이용된다. 첫 번째로는 지역밀도를 사용해 지역에 따라 다른 밀도를 갖는 개체간의 차이를 나타낼 수 있다. DBSCAN에서는 하나의 코어포인트부터 시작하여 재귀적 탐색을 통해 더 이상 새로운 밀도 근접 포인트를 찾을 수 없을 때까지 클러스터의 확장을 계속한다. 그러나 이는 근접한 클러스터들의 경계를 넘어 과도한 확장으로 이어질 수 있다. 따라서 수정된 클러스터의 확장 과정은 상대적으로 지역

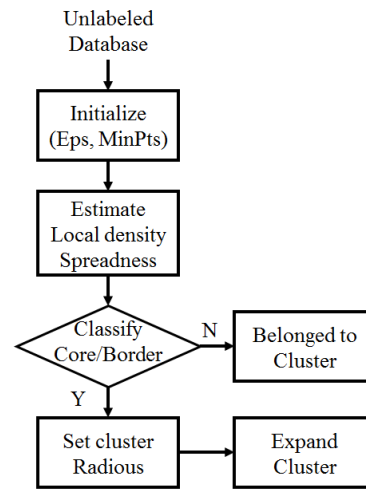


Fig. 4. Flow chart for local distribution based density clustering.

밀도가 높은 개체부터 시작하여, 한 번의 과정 동안 하나의 클러스터를 수렴할 때까지 확장하는 것이 아니라 복수의 클러스터가 동시에 점진적인 확장을 진행하게 된다.

두 번째로는 목표 개체의 주변 밀도 근접 포인트를 클러스터로 병합할 때 개체의 특성별로 지역밀도, 분산도를 최소 기준거리와 거리값의 가중치로서 적용한다. 결과적으로 밀도의 높낮음, 개체의 특성별 분산에 따라 가변적인 판단 기준을 사용할 수 있게 된다. 이와 같은 제안 방법으로 실제 데이터를 사용한 실험에서는 기존 방법과 비교한 성능을 확인해 볼 수 있다.

Fig. 4에서는 이러한 과정을 도식화하여 설명하고 있다. 초기 입력으로 최소 기준거리와 기준 수를 사용하여 임의의 개체의 밀도를 측정하고 코어포인트와 보더포인트로 분류한 후, 지역밀도와 분산도로써 계산된 가중치를 적용하여 가변적 클러스터 반경을 이용한 클러스터의 확장을 수행한다.

III. 실험 결과 및 분석

3.1 실험 데이터베이스 설명 및 성능 지표

실험은 2개의 데이터 집단에 대해 성능을 평가할 것이다. 첫 번째 데이터 집단은 2차원 상에 존재하는 N 개의 클러스터로 구성되며 일부 클러스터는 코어

```

Proposed method (data, Eps, MinPts)
  ClusterNum = 0

  For all data
    FindNearestNeighbor(data)
    EstimateParameters(data, local density, spreadness)
  End

  For target  $x_i$  in descending density
    If  $NN_i$  in Eps > MinPts
       $x_i$  = core
    Else
       $x_i$  = border
    End

    If  $x_i$  is core
      SetClusterRadius( $x_i$ , Eps, local density, spreadness)

      If  $x_i$  is new cluster
        ClusterNum = ClusterNum + 1
         $x_i$  = ClusterNum
      End

       $NN_i$  = ClusterNum( $x_i$ )
    End
  End
End
    
```

Fig. 5. Pseudocode for local distribution based density clustering.

포인트의 최소 기준거리 범위 내에 존재하는 보더포인트를 공유한다. 보더포인트들은 서로 다른 클러스터간의 밀도 연결 고리를 잇게 되며 이를 방지하지 못할 경우 분류성능은 하락하게 된다. 다른 데이터 집단은 ETRI(Electronics and Telecommunications Research Institute) 화자인식용 데이터베이스에서 200명의 음성을 사용한다. 한 명의 화자 당 총 50문장의 발화를 보유하고 있으며 모든 발화는 임의의 문장으로 구성되어있다. 우리는 발화에서 MFCC(Mel-Frequency Cepstral Coefficient) 39차의 특징을 추출하였으며, 하나의 발화내에 포함된 특징 벡터들은 150차의 i-vector로 변환되었다.^[18]

분류성능은 사전에 분류되지 않은 데이터들로 구성된 클러스터 내에서 우세한 데이터의 비율을 사용한다. 이를 정확도로 지칭하며, 이 수치는 결과 클러스터가 본래의 클러스터를 전부하지 않더라도 일부의 특성을 포함하고 있음을 말한다.

3.2 2차원 샘플 클러스터링 성능 평가

Fig. 6은 2차원 데이터의 클러스터링 결과를 보여주고 있다. Fig. 6(a)는 총 6개의 클러스터로 구성된

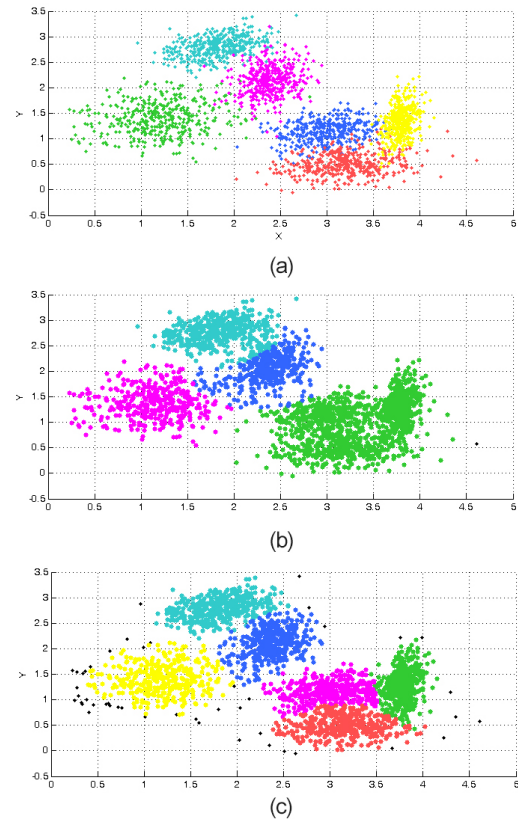


Fig. 6. Visualization of clustering result for 2-dimensional sample data. (a) Distribution of actual data. (b) Clustering result of DBSCAN. (c) Clustering result of proposed method.

데이터 집단을 보여주며 이 데이터에 DBSCAN을 적용한 실험 결과는 Fig. 6(b)와 같다. 인접한 클러스터간의 보더포인트를 통해 서로 다른 클러스터가 하나의 클러스터로 병합되거나, 경계를 넘어 다른 클러스터까지 침범하는 것을 볼 수 있다. Fig. 6(c)의 경우는 제안하는 방법을 통해 실험한 것으로 경계를 공유하는 클러스터의 경우에도 서로 침범하지 않고 각 클러스터가 최대화되는 경계에서 확장이 멈추는 것을 확인할 수 있다. Fig. 6(b), (c)에서 검은 색의 점은 노이즈로 간주된 경우인데 Fig. 6(c)에서는 노이즈로 판단된 점의 수가 Fig. 6(b)에 비해 많으나 단일 클러스터 내의 정확도가 보다 높은 것을 확인할 수 있다.

3.3 ETRI DB 클러스터링 성능 평가

Table 1에서는 총 10, 20, 50, 100명의 화자의 음성 데이터를 각각 하나의 클러스터로서 가정하고 ETRI

Table 1. Clustering performances of DBSCAN and proposed method.

Method	DBSCAN		Proposed method	
	# cluster	Average accuracy (%)	# cluster	Average accuracy (%)
10	8.4	72.5	12.3	85.4
20	17.6	69.6	28.5	84.2
50	42.1	68.4	67.9	81.1
100	160.9	60.1	262.1	79.5

화자 인식용 데이터베이스 중에서 교차검증을 통해 10개의 조합을 구성하였다. Table 1은 실험에 대한 성능을 표현하고 있는데 실제 클러스터의 수에 대비해 분류 결과 결과로 분류된 클러스터 수의 평균과 각 클러스터내에서 우세한 비율을 차지하는 클러스터를 기준으로 정확도를 계산하였다.

DBSCAN을 이용한 실험결과에서는 클러스터 간의 병합이 발생하여 실제 클러스터의 수보다 낮은 개수로 추출되었으며 각 클러스터의 정확도 또한 60 ~ 75 % 사이로 낮은 성능을 보였다. 이는 하나의 클러스터내에 여러 클러스터가 섞여 정확도가 떨어지는 것임을 말한다. 그러나 제안하는 방법의 실험 성능은 비록 실제 클러스터의 수보다 훨씬 많은 클러스터로 나뉘었으나 평균적인 정확도가 높게 나왔으므로 하나의 클러스터가 여러 개의 클러스터로 나뉘었지만 보다 정확하게 분류가 되었다고 이해할 수 있다.

IV. 결 론

본 논문은 기존의 방법인 DBSCAN에서 일어날 수 있는 오버 클러스터링 문제를 지적하고 이를 해결하기 위해, 지역적 특성에 기반한 화자 분할 알고리즘을 제안하였다. 랜덤한 샘플과 ETRI 데이터베이스를 사용하여 성능을 비교한 결과, 제안한 방법은 과군 집화가 발생하지 않았으며 DBSCAN에 비해 클러스터 내에서 보다 높은 정확도를 확보하였다. 차후에는 보다 정확한 클러스터 수의 추정을 위한 연구를 진행할 예정이다.

감사의 글

본 논문은 삼성전자의 지원을 받아 수행한 연구 과제 결과 중 일부이다.

References

1. Y. Esteve, T. Bazillon, J.Y. Antoine, F. Béchet, and J. Farinas, "The EPAC corpus: manual and automatic annotations of conversational speech in french broadcast news," in LREC., 1686-1689 (2010).
2. E. El-Khoury, C. Senac, and J. Pinquier. "Improved speaker diarization system for meetings," in IEEE ICASSP., 4097-4100 (2009).
3. A. Vinciarelli, Alessandro, F. Fernandez, and S. Favre, "Semantic segmentation of radio programs using social network analysis and duration distribution modeling," in ICME., 779-782 (2007).
4. H. Tang, S. M. Chu, M. Hasegawa-Johnson, and T.S. Huang, "Partially supervised speaker clustering," IEEE Transactions on Pattern Analysis and Machine Intelligence **34**, 959-971 (2012).
5. T. Pfau, Daniel P.W. Ellis, and A. Stolcke, "Multispeaker speech activity detection for the icsi meeting recorder," in IEEE Workshop on Automatic Speech Recognition and Understanding, 107-110 (2001).
6. D. Wing, Y. Yan, J. Dang, and F. K. Soong, "Voice activity detection based on an unsupervised learning framework," IEEE Transactions on Audio, Speech, and Language Processing **19**, 2624-2633 (2011).
7. F. G. Germain, D. L. Sun, and G. J. Mysore. "Speaker and noise independent voice activity detection," in Interspeech **2013**, 732-736 (2013).
8. R. Sinha, S. Tranter, M. Gales, and P. Woodland, "The cambridge university march 2005 speaker diarisation system," in Interspeech, 2437-2440 (2005).
9. P. Jain and H. Hermansky, "Improved mean and variance normalization for robust speech recognition," in IEEE ICASSP., 4015-4015 (2001).
10. D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker verification using adapted Gaussian mixture models," in DSP., 19-41 (2000).
11. S. Meignier, D. Moraru, C. Fredouille, J. F. Bonastre, and L. besacier, "Step-by-step and integrated approaches in broadcast news speaker diarization," Computer Speech & Language **20**, 303-330 (2006).
12. X. Zhu, C. Barras, L. Lamel, and J-L. Gauvain, "Multi-stage speaker diarization for conference and lecture meetings," in *Multimodal Technologies for Perception of Humans* (Springer Berlin Heidelberg, Germany, 2008), pp. 533-542.

13. S. Salvador and P. Chan, "Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms," in 16th IEEE International Conference on Tools with Artificial Intelligence, 576-584 (2004).
14. H-P. Kriegelet, P. Kröger, J. Sander, and A. Zimek, "Density based clustering," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 231-240 (2011).
15. M. Ester, H-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in Knowledge Discovery and Data Mining **96**, 226-231 (1996).
16. C. Braune, S. Besecke, and R. Kruse, "Density Based Clustering: Alternatives to DBSCAN," Partitional Clustering Algorithms(Springer International Publishing, Switzerland, 2015), pp. 193-213.
17. Z. Aoying, Z. Shuigeng, C. Jing, F. Ye, and H. Yunfa, "Approaches for scaling DBSCAN algorithm to large spatial databases," JCST. **15**, 509-526 (2000).
18. N. Dehak, P. J. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Front-end factor analysis for speaker verification," IEEE Transactions on Audio, Speech, and Lang. Process. **19**, 788-798 (2011).

▶ 이 재 원 (Jae-Won Lee)



1993년: KAIST 컴퓨터과학부 석사
 1999년: KAIST 컴퓨터과학부 박사
 1999년 ~ 현재: 삼성전자 수석연구원

▶ 고 한 석 (Hansek Ko)



1982년: 미국 카네기 멜론 대학교 전기공학
 학사
 1986년: 미국 메릴랜드 대학교 시스템공학
 석사
 1988년: 미국 존스 홉킨스 대학교 전기공학
 석사
 1992년: 미국 카톨릭 대학교 전기공학 박사
 1995년 ~ 현재: 고려대학교 전기전자공학과
 교수

저자 약력

▶ 노 진 상 (Jinsang Rho)



2011년: 고려대학교 전기전자공학부 학사
 2011년 ~ 현재: 고려대학교 영상정보처리
 협동 석 박사통합과정 재학중

▶ 손 수 원 (Suwon Shon)



2010년: 고려대학교 전기전자전파공학부
 학사
 2012년: 고려대학교 전기전자공학과 석사
 2012년 ~ 현재: 고려대학교 전기전자공학과
 박사과정 재학중

▶ 김 성 수 (Sung Soo Kim)



2008년: 고려대학교 전기전자전파공학부
 학사
 2010년: 서울대학교 전기컴퓨터공학부
 석사
 2010년 ~ 현재: 삼성전자 선임연구원