Robust source counting and acoustic DOA estimation using density-based clustering

Sina Hafezi, Alastair H. Moore and Patrick A. Naylor - Imperial College London

Imperial College London

by Sina Hafezi

July 10, 2018

Abstract

In this presentation, we:

- address the problem of source counting and sources' DOA extraction from a set of local DOAs.
- start from a density-based clustering technique based on which an autonomous method in an evolutionary framework is proposed specifically for the addressed problem.

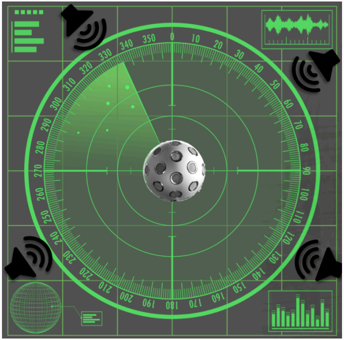
Outline

Introduction

Direction-of-Arrival (DOA) estimation

Wide range of applications:

- Source localization/tracking/separation
- Environment mapping
- Dereverberation
- Speech Enhancement
- Robot Audition



Challenges:

- Reverberation, Noise
- Number and separation of sources
- Sources' activity, motion and distance-to-mic

Typical procedure

Conventional system for scenarios with multiple wideband sources consists of two stages:

1. Temporal narrowband (local) DOA estimation in the Time-Frequency (TF) domain
2. Sources' DOA extraction from the set of estimated local DOAs

The last stage has direct impact on the source's DOA accuracy.

Therefore it is important to be robust to erroneous local DOAs (referred to as outlier DOAs).

This work focuses on the last stage only.

Interested Scenario

Our interested scenario:

- Multiple simultaneously active wideband sound sources
- No knowledge of the number of sources and their separation

The problem: Source counting and sources' DOA extraction from a given set of local DOAs.

Conventional methods

The conventional methods for the extraction of sources' DOA from a set of local DOA:

1. Peak Picking
2. Clustering

Peak picking

Directly or iteratively detects the peaks in the smoothed 2D histogram (azimuth $\times$ inclination) of the local DOAs as the sources DOA.

Limitations:

- number of sources is known *a priori*.
- requires some knowledge of sources separation to avoid merge of peaks due to oversmoothing.
- static setting for smoothing may fail for scenarios with varying peaks irregularity among multiple clusters of DOAs.

Clustering

Applies distance-based clustering techniques such as Kmeans or model-based such as Gaussian, Laplacian or Von Mises mixture models on the set of local DOAs.

Limitations:

- Mostly requires *a priori* knowledge of the source number
- Although various Information Criterion (IC) (e.g. Akaike or Baysian) can be used to estimate the number of clusters, it is highly prone to outlier DOAs.

Density- vs. Distance-based clustering

Distance-based: has the number of cluster as its constraint.

- Requires the number of clusters as *a priori*
- Clusters all data including the outliers

Density-based: has the clusters' minimum density as its constraint.

- Does not require the number of clusters as *a priori*
- Robust to outliers
- Requires the trade-off density between the outlier and cluster

Density-based clustering in DOA est.

Density-based clustering has received much less attention than distance- or model-based technique clustering techniques in the context of acoustic DOA estimation.

Suitable for sources' DOA extraction where:

- the set of local DOA could contain outliers
- the number of sources is unknown

Motivation: To study the utilisation of density-based clustering for source counting and sources' DOA extraction.
Baseline technique: Density-based Spatial Clustering of Application with Noise (DBSCAN)

DBSCAN

DBSCAN

Based on two parameters:

- **Local density metric:** Number of points including itself within the neighbourhood of distance ε.
- **Threshold density:** MinPts, trade-off density between the outlier and clusters.

DBSCAN in a nutshell: Points with density higher than threshold density (referred to as core points) are clustered using density connectivity.

Terms and Definitions in DBSCAN

Consider D as the points database.

1) Neighbourhood Points: the set of points

$$N_\varepsilon(p) = \{q \in D | dist(p, q) \leq \varepsilon\}, \tag{1}$$

where $dist(p, q)$ is a distance function for points p and q.

2) Directly density-reachable: a point p is directly density-reachable from a point q if

- $p \in N_\varepsilon(q)$ and
- $|N_\varepsilon(q)| \geq \mathrm{MinPts}$ (core point condition).

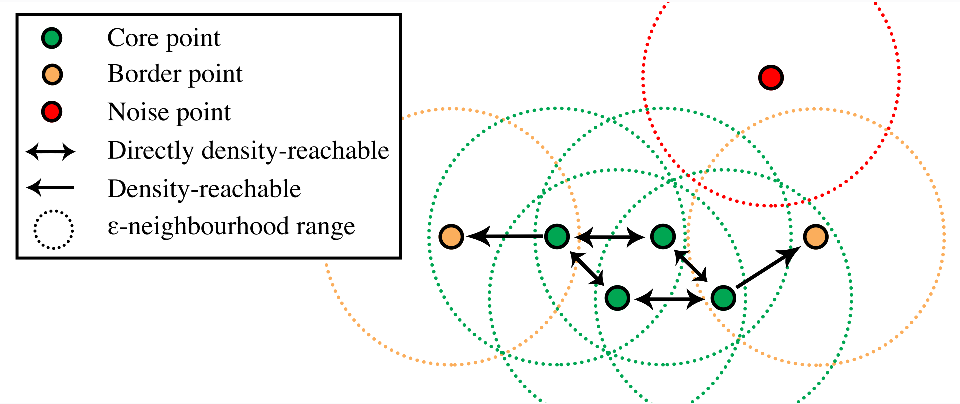
Terms and Definitions in DBSCAN

3) Density-reachable: a point p is density-reachable from a point q if there is a chain of points $\{p_i\}_{i=1}^{J}$, where $p_1 = q$ and $p_J = p$, such that p_{i+1} is directly density-reachable from p_i.
4) Density-connected: a point p is density-connected to a point q if there is a point m such that both p and q are density-reachable from it.
5) Cluster: a cluster C is a subset of D satisfying:

- $\forall p, q :$ if $p \in C$ and q is density-reachable from p, then $q \in C$ and
- $\forall p, q \in C : p$ is density-connected to q.

6) Noise: a subset of points in D not belonging to any cluster.

Example of DBSCAN points labelling



Points labelling by DBSCAN with MinPts=3.

DBSCAN algorithm

Given the user-defined parameters ε and MinPts:

1. Detects all the core points.
2. Chooses an arbitrary unvisited core point
3. Clusters all points which are density-reachable from it.
4. Repeats 2-4 until all core points are clustered.

Right choice of setting

DBSCAN performance highly depends on (ε,MinPts).
Knee in k-dist graph: A simple heuristic approach to choose (ε,MinPts)

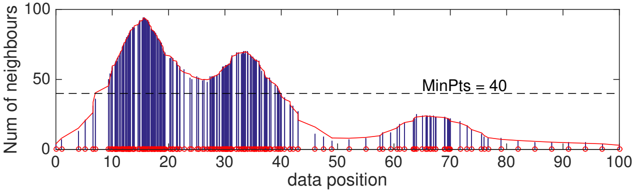
- defines k-dist as the distance to the kth nearest neighbour. Choose a k and set MinPts$=k$.
- ε is chosen as the knee in the graph of sorted k-dist values for all points.

Although the tuning is simplified, it still suffers from dependency on the choice of k and is not fully autonomous.

Problematic Case

Imperial College London

Cases where there is no set of settings which gives the right clustering.



Common scenario in DOA estimation: Mixtures of distributions with widely varying density due to difference in sources activity or distance-to-mic.

Solution for problematic cases?

Imperial College London

- Variations of DBSCAN [**Ram2009**, **Liu2007**, **Xiaoyun2008**, **Xiong2012**, **Uncu2006**] are proposed
- All somehow require user engagement for setting parameters.
- All aim to come up with one final optimum clustering.
- In the context of DOA estimation: Only final sources' direction are needed and not the clusters
- Therefore a new variation of DBSCAN is proposed specifically for DOA estimation.

Proposed: Evolutive DBSCAN

Evolutive DBSCAN

Evoutive DBSCAN in a nutshell:

1. **Denoising of distribution**: Iteratively run DBSCAN with descending threshold density and store the reliable centroids and their associated weights. The distribution of the weighted centroids is shown to be significantly more sparse and less noisy than the distribution of the DOAs.
2. **Source counting & sources DOA extraction:** Perform a final DBSCAN on the set of resulting weighted centroids, where the threshold density of DBSCAN is autonomously estimated.

Denoising of distribution

Same density is used as in DBSCAN with $\varepsilon = 10^\circ$ defined as a base unit. (Note that it is not a user parameter.)

MinPts Range: $[\mathrm{min}(|N_\varepsilon(.)|) + 1, \mathrm{max}(|N_\varepsilon(.)|) - 1]$

Step of iteration: Calculated depending on the number of iterations set by the user. (MinPts Range)/NumIt

Denoising of distribution

Clusters at each iteration, compared with the previous clustering, are labelled as 'dead' or 'alive' each defined as:

1) Dead: A cluster is dead if:

- it has a shared member with more than one alive cluster in the last iteration (merge condition) or
- it has a shared member with any previously dead cluster (re-occurrence of a previously merged cluster).

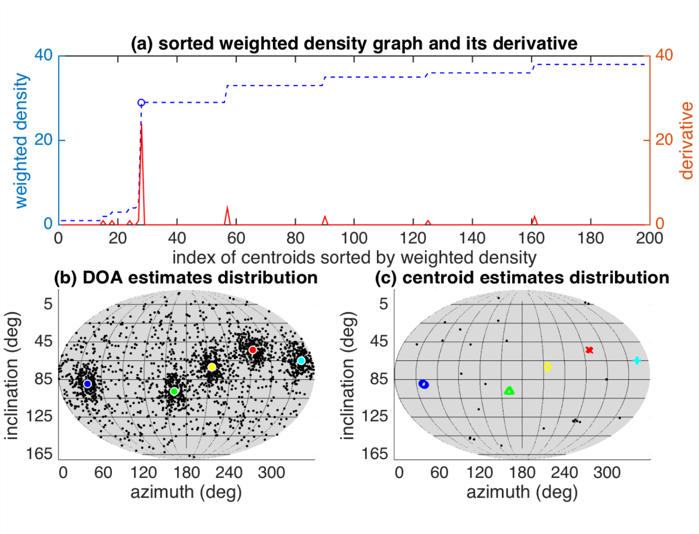
2) Alive: A cluster is alive if it is not dead.

Cluster's weight: mean density of the cluster's members. The weight and centroid of the alive clusters are iteratively stored. The pseudocode can be found in the paper.

Source counting and DOA extraction

- Having obtained M centroids $\{c_i\}_{i=1}^{M}$ and their associated weights $\{w_i\}_{i=1}^{M}$
- One final DBSCAN is performed on the centroids
- Results in L clusters with centroids $\{d_i\}_{i=1}^{L}$ as source directions
- The MinPts for the final DBSCAN is the position of the first peak in the derivative of the sorted weighted centroid graph.

Illustration of an example



Evaluations

Evaluations: metrics and methods

Imperial College London

Two metrics for performance evaluations:

- **Successful Localization Rate (SLR)**: % of the trials where the correct number of sources was detected AND all the best case data associated pairs of estimated-true DOAs have less than 5° error in azimuth and inclination.
- **Mean Error:** average DOA estimation error across the successfully localized cases only.

Evaluated methods: Peak picking, original DBSCAN, Kmeans (K=N_s), AIC+Kmeans, BIC+Kmeans

Experiment 1: using generated DOAs

DOAs generated for different cases of distribution controlled by:

- N_s: number of sources
- N_p: ratio of the number of outlier (noise) DOAs to total DOAs
- S_p: ratio of the DOAs from a single distribution associated with a single source to all DOAs (excluding the noise DOAs)
- Sep: angular separation between the distributions' centroid

Example: 500 total DOAs and $\{N_s, N_p, S_p\} = \{2, 0.8, 0.3\}$, we have 400 noise DOAs with 30 and 70 DOAs generated for source 1 and 2 respectively.

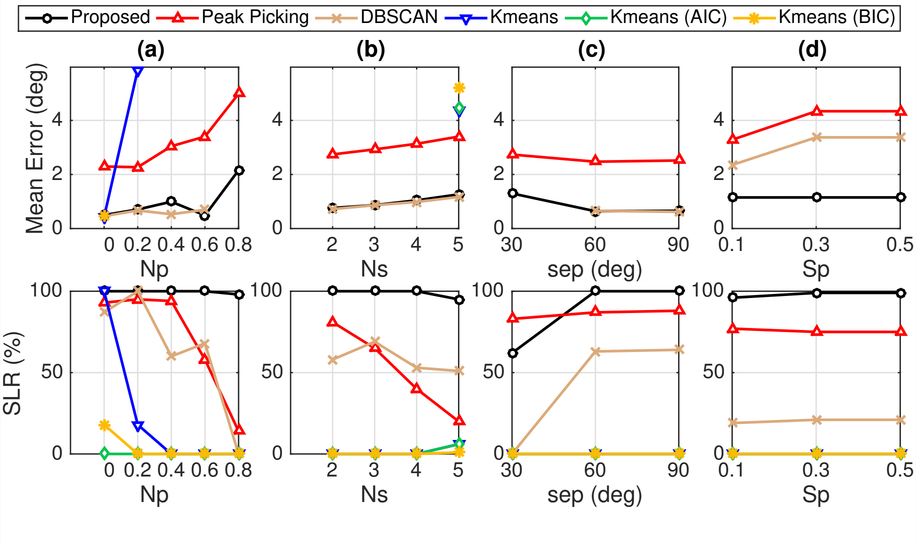
Experiment 1: setting

Distribution: Von Mises-Fisher with centroid set on true-DOA and $\kappa = 30$ (concentration parameter).

2000 DOAs with $\{N_s, N_p, S_p, Sep\} = \{2, 0.5, 0.5, 60°\}$ per trial unless otherwise stated.

100 trials per case with random first DOA for each trial.

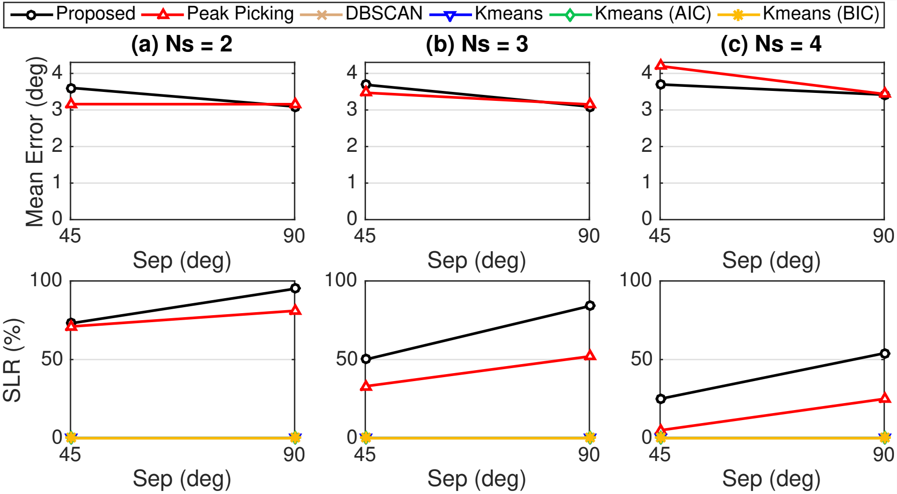
Experiment 1: Results



Experiment 2: using estimated DOAs

- DOA estimator: Direct Path Dominance Pseudointensity Vectors (DPD-PIV)
- Microphone Array: Rigid spherical (em32 Eigenmike®)
- Room Impulse Response: Simulation based on image method
- Room: $5 \times 6 \times 4$ m with $T_{60} = 0.4$ s
- Sensor noise: Spatially white with SNR$=25$ dB
- Sources content: anechoic speech (APLAWD database)
- Sources location: 1 m distance-to-array with random first DOA

Experiment 2: Results



Conclusions

- Investigated the use of density-based clustering for source counting and DOA estimation.
- Proposed an autonomous DBSCAN-based method in an evolutionary framework specifically designed for source counting and sources DOA extraction from a set of local DOAs.
- The evaluations using generated and estimated DOAs validate $\leq 4^{\circ}$ mean DOA estimation accuracy and the superior source counting accuracy compared to the comparative methods.

End of the presentation

Imperial College London

Thank you for listening! :)

Questions?