DISTRIBUTED LABELLING OF AUDIO SOURCES IN WIRELESS ACOUSTIC SENSOR NETWORKS USING CONSENSUS AND MATCHING

Mohamad Hasan Bahari *Jorge Plata-Chaves* *Alexander Bertrand* *Marc Moonen*

KU Leuven, Dept. Electrical Engineering ESAT
STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics
{Mohamadhasan.Bahari,Jorge.Plata-Chaves,Alexander.Bertrand,Marc.Moonen}@esat.kuleuven.be

ABSTRACT

In this paper, we propose a new method for distributed labelling of audio sources in wireless acoustic sensor networks (WASNs). We consider WASNs comprising of nodes equipped with multiple microphones observing signals transmitted by multiple sources. An important step toward a cooperation between the nodes, e.g. for a voice-activity-detection, is a network-wide consensus on the source labelling such that all nodes assign the same unique label to each source. In this paper, a hierarchical approach is applied such that first a network clustering algorithm is performed and then in each sub-network, the energy signatures of the sources are estimated using a non-negative independent component analysis over the energy patterns observed by the different nodes. Finally the source labels are obtained by an iterative consensus and matching algorithm, which compares and matches the energy signatures estimated in different sub-networks. The experimental results show the effectiveness of the proposed method.

Index Terms— Distributed labelling, consensus and matching, wireless acoustic sensor networks, energy signatures, non-negative independent component analysis

1. INTRODUCTION

A wireless acoustic sensor network (WASN) typically consists of spatially distributed wireless nodes equipped with one or more microphones observing signals transmitted by multiple sources [1, 2, 3]. In 'multiple devices for multiple tasks' (MDMT) paradigm [4, 5, 6], different nodes cooperate with each other to carry out different node-specific tasks. In such an MDMT-based WASN, a critical step to empower the cooperation between the nodes, e.g., for a better node-specific signal enhancement or a distributed voice-activity-detection (VAD), is a network-wide consensus on the source labelling such that all nodes assign the same unique label to each source [4]. In this setting, each node observes mixtures of interfering signals transmitted by different sources, while labelling the sources requires source-specific information in each node. In this work, we use energy envelopes as a signature to label the sources. Unmixing the observed microphone signals to extract source-specific features (signatures) is a challenging and computationally expensive task [7]. However, energy envelopes are non-negative and have a low sampling rate, which allows us to rely on cheap non-negative source separation methods operating on low-rate microphone signal energy envelopes[8, 9].

In [8] and [10], a non-negative principal component analysis (NPCA) and a multiplicative non-negative independent component analysis (MNICA) have been proposed to unmix the non-negative signals respectively. Although these methods are attractive from several aspects, they use the observations of all nodes in a fusion centre, which requires a large communication bandwidth, and hence the are very energy-inefficient. On the other hand, performing a NPCA or a MNICA on the microphone signals within a single node results in poor estimation since these methods typically require sufficient spatial diversity in the observed signals to yield satisfactory results [9, 4].

Chouvardas et al. [4] tackled this problem by introducing a hierarchical approach such that first a network clustering algorithm is performed and then an MNICA is applied for each sub-network to estimate the energy signatures of the sources. The required communication bandwidth and the estimation accuracy of this method is less compared to the centralized estimation and more compared to the node-level estimation. Since the estimated energy signatures corresponding to a specific source are expected to be similar in different sub-networks, the sources can be labelled by comparing and

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of KU Leuven Research Council CoE PFV/10/002 (OPTEC), BOF/STG-14-005, the Interuniversity Attractive Poles Programme initiated by the Belgian Science Policy Office: IUAP P7/23 'Belgian network on stochastic modeling analysis design and optimization of communication systems' (BESTCOM) 2012-2017, Research Project FWO nr. G.0763.12 'Wireless Acoustic Sensor Networks for Extended Auditory Communication', Research Project FWO nr. G.0931.14 'Design of distributed signal processing algorithms and scalable hardware platforms for energy-vs-performance adaptive wireless acoustic sensor networks', and the FP7-ICT FET-Open Project 'Heterogeneous Ad-hoc Networks for Distributed, Cooperative and Adaptive Multimedia Signal Processing (HANDiCAMS)', funded by the European Commission under Grant Agreement no. 323944, The scientific responsibility is assumed by its authors.

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

matching the signatures obtained in different sub-networks.

In [4], a distributed k-means approach is applied to measure the similarity of the energy signatures obtained in different sub-networks and achieve a network-wide consensus on the labels of the sources. While effective, the accuracy of this method drops down in the presence of uncorrelated noise, as shown in our simulation.

In this paper, we adopt the method of Chouvardas et al. [4] to estimate the energy signatures in each sub-network and replace the distributed k-means approach by proposing an accurate signature matching method. In each iteration of the proposed matching method, referred to as the consensus and matching (CM) algorithm, we first compute a network-wide consensus between the sub-networks on the energy signatures of all the sources and then we match the signatures of each sub-network with respect to the obtained network-wide consented signatures. This method has two distinct advantages: (1) it yields more accurate results in presence of an uncorrelated noise compared to the distributed k-means [4], and (2) it results in network-wide consented energy signatures, rather than per-sub-network estimates. The latter will result in a better estimation performance, which is important if the energy signatures are also used in the further processing pipeline, e.g., for VAD. Experimental results show that the consented signatures are more accurate than the estimated signatures in each sub-network independently and the CM algorithm yields more accurate labelling compared to the benchmark method of [4].

2. DISTRIBUTED AUDIO SOURCE LABELLING

2.1. Problem Formulation

Consider a WASN with N sources and D nodes, where node d is equipped with J_d microphones, d being the node index. The total number of microphones in the network is $J = \sum_d J_d$. The nodes need to label N sources such that a unique label is assigned to each source throughout the network to facilitate the collaboration between the nodes.

2.2. Energy Signatures

We denote the i^{th} sample of the signal of the n^{th} source as $\tilde{s}_n[i]$, $n = 1, \cdots, N$. Given a block of length L the instantaneous energy of this signal at sample iL is computed as

$$s_n[i] = \sum_{l=0}^{L-1} \tilde{s}_n^2[iL + l]. \tag{1}$$

Similarly we denote the i^{th} sample of the j^{th} microphone signal as $\tilde{y}_j[i]$, $j = 1, \cdots, J$ and the instantaneous energy of this signal at sample iL is computed as

$$y_j[i] = \sum_{l=0}^{L-1} \tilde{y}_j^2[iL + l]. \tag{2}$$

As discussed in [9], assuming that the source signals are mutually independent and that the reverberation has a negligible effect across the block edges, we can model $\mathbf{y}[i]$ as

$$\mathbf{y}[i] \approx \mathbf{A}\mathbf{s}[i], \tag{3}$$

where $\mathbf{A}$ is a mixing matrix of size $J \times N$ describing the power attenuation between the speech sources and the microphones and

$$\mathbf{s}[i] = [s_1[i], \cdots, s_N[i]]' \tag{4}$$

$$\mathbf{y}[i] = [y_1[i], \cdots, y_J[i]]', \tag{5}$$

where $'$ denotes the transpose operation.

In practice, $\mathbf{s}[i]$ and $\mathbf{A}$ are not available and we have to estimate them. Given $\mathbf{y}[i]$, NPCA [8] and MNICA [10] estimate the mixing matrix and the energy signature of the sources, where the estimates are denoted as $\hat{\mathbf{A}}$ and $\hat{\mathbf{s}}[i]$ respectively.

$$\mathbf{Y} \approx \hat{\mathbf{A}}\hat{\mathbf{S}}^{cent}, \tag{6}$$

where $\mathbf{Y} = [\mathbf{y}[1], \mathbf{y}[2], \cdots, \mathbf{y}[\Gamma]]$ is the entire observation energy matrix with Γ being the number of observed blocks of length L, and $\hat{\mathbf{S}}^{cent} = [\hat{\mathbf{s}}[1], \hat{\mathbf{s}}[2], \cdots, \hat{\mathbf{s}}[\Gamma]]$ and $\hat{\mathbf{A}}$ are the source energy signatures and their corresponding mixing matrix estimated by NPCA respectively.

To avoid an energy inefficient centralized estimation, a hierarchical approach is applied such that first a network-clustering algorithm is performed to divide the network into K sub-networks. In this paper we use a distributed Fiedler vector algorithm [11], which identifies densely connected node clusters in a distributed fashion. Then NPCA[1] is applied on the sub-network level as follows:

$$\mathbf{Y}^k \approx \hat{\mathbf{A}}^k\hat{\mathbf{S}}^k, \tag{7}$$

with $k \in 1, \cdots, K$ denoting the sub-network index. Note that $\mathbf{Y}^k$ and $\hat{\mathbf{A}}^k$ denote a subset of the rows of $\mathbf{Y}$ and $\hat{\mathbf{A}}$ respectively, whereas $\hat{\mathbf{S}}^k$ is a sub-network estimate of the full matrix $\hat{\mathbf{S}}^{cent}$, i.e., it has the same dimensions as $\hat{\mathbf{S}}^{cent}$. To avoid scaling ambiguity, we apply a length normalization over the obtained energy signatures.

Remark 1: Note that NPCA or MNICA require sub-networks with sufficient spatial diversity to yield reasonable results [9, 4].

2.3. Labelling using Distributed k-means

Chouvardas et. al. [4] use a distributed k-means algorithm to label the sources given their energy signatures. In this method, first N centroids of dimension Γ are considered for each sub-network[2]. The centroids should be initialized such that they are the same in all sub-networks.Then each sub-network performs a local labelling scheme by employing a

[1]Both NPCA and MNICA can be used to find energy signatures. However, since our simulation results show that NPCA yields more accurate energy signatures, we applied this method in the sequel.

[2]The number of sources N is assumed to be known in [4] and also in this work. Note that many methods are suggested to estimate the number of sources such as [12].

k-means algorithm using the computed energy signatures and the previously computed centroids such that each energy signature is assigned to the cluster in which the correlation between the energy signatures is maximized. Finally clusters update their centroids in cooperation with the neighbouring sub-networks. After convergence of the k-means labelling procedure, the label of each signal is set to the number of the class, in which the respective signature is assigned. Although this method is effective, it does not yield accurate results in the presence of uncorrelated noise, as will be demonstrated in our simulations.

Remark 2: The distributed k-means algorithm is originally developed for an unsupervised clustering [13], while Chouvardas et. al. [4] modify it for a distributed labelling.

3. LABELLING USING THE CM ALGORITHM

To improve the labelling accuracy, we introduce a robust labelling method based on an iterative consensus on the energy signatures and matching the local energy signatures in each sub-network to the obtained consented signatures. The proposed CM algorithm relies on the following relation between the true signatures of the sources and the estimated signatures in each sub-network locally:

$$\hat{\mathbf{S}}^k = \mathbf{P}^k\mathbf{S} + \mathbf{E}^k, k = 1, \cdots, K, \tag{8}$$

where $\mathbf{S}$ represents the true energy signatures, $\mathbf{P}^k$ is a permutation matrix for sub-network k and $\mathbf{E}^k$ is the corresponding error matrix. Eq. (8) implies that the estimated energy signatures in each sub-network are equal to a permutation of the true signatures up to an estimation error. Therefore, assuming that $\mathbf{S}$ is available, finding the permutation matrix $\mathbf{P}^k$ in each sub-network is trivial. In practice, however, neither the permutation matrix $\mathbf{P}^k$ nor the true energy signatures $\mathbf{S}$ are available and we estimate them given the locally estimated energy signatures $\hat{\mathbf{S}^k}$ through minimizing the error Frobenius-norm $||\mathbf{E}^k||_F = ||\hat{\mathbf{S}^k} - \hat{\mathbf{P}}^k\hat{\mathbf{S}}||_F$, i.e.

$$\min_{\hat{\mathbf{S}},\hat{\mathbf{P}}^1,\cdots,\hat{\mathbf{P}}^K} \sum_{k=1}^{K} ||\hat{\mathbf{S}^k} - \hat{\mathbf{P}}^k\hat{\mathbf{S}}||_F, \tag{9}$$

subject to

$$\begin{cases} \hat{\mathbf{P}}^k_{ij}\left(1 - \hat{\mathbf{P}}^k_{ij}\right) = 0 \\ \quad \sum_i \hat{\mathbf{P}}^k_{ij} = 1 \\ \quad \sum_j \hat{\mathbf{P}}^k_{ij} = 1 \end{cases}, k = 1, \cdots, K. \tag{10}$$

We propose an alternating optimization method for the problem (9)-(10)[3]. In the first step, referred to as the matching step, $\hat{\mathbf{S}}$ is assumed to be known, and we try to update $\hat{\mathbf{P}}^k$ for $k = 1, \cdots, K$. Similarly in the second step, referred to as the consensus step, $\hat{\mathbf{P}}^k$ for $k = 1, \cdots, K$ is assumed to be known and we update $\hat{\mathbf{S}}$. These two steps are elaborated in the next subsections.

[3]Similar constrained alternating optimization methods can be found in [14, 15].

3.1. Matching step

Since $\hat{\mathbf{P}}^k$ is a permutation matrix, it is trivial to show that the optimization problem (9)-(10) can be reformulated as

$$\min_{\mathbf{P}^1,\cdots,\mathbf{P}^K} \sum_{k=1}^{K} \sum_{i,j} \mathbf{Q}^k_{ij}\hat{\mathbf{P}}^k_{ij}, \tag{11}$$

subject to the constrains (10), where $\mathbf{Q}^k$ is the Euclidean distance matrix between $\hat{\mathbf{S}}$ and $\hat{\mathbf{S}}^K$ obtained as

$$\mathbf{Q}^k_{ij} = ||\hat{\mathbf{S}^k}_i - \hat{\mathbf{S}}_j||_2, \tag{12}$$

where $||\cdot||_2$ denotes the vector 2-norm and where $\hat{\mathbf{S}^k}_i$ and $\hat{\mathbf{S}^k}_j$ are the $i-th$ and $j-th$ row of $\hat{\mathbf{S}^k}$ respectively.

Assuming $\hat{\mathbf{S}}$ is known, the minimization of (11) over $\hat{\mathbf{P}}^k$ for $k \in 1, \cdots, K$, depends on $\hat{\mathbf{S}^k}$ only. Therefore, the minimization of (11) decouples into K independent minimizations as follows

$$\min_{\hat{\mathbf{P}}^k} \sum_{i,j} \mathbf{Q}^k_{ij}\hat{\mathbf{P}}^k_{ij}, \tag{13}$$

subject to

$$\begin{cases} \hat{\mathbf{P}}^k_{ij}\left(1 - \hat{\mathbf{P}}^k_{ij}\right) = 0 \\ \quad \sum_i \hat{\mathbf{P}}^k_{ij} = 1 \\ \quad \sum_j \hat{\mathbf{P}}^k_{ij} = 1 \end{cases}. \tag{14}$$

The optimization problem (13)-(14) is a so-called linear assignment problem, which is a well-known sub-class of linear programming problems. The optimal solution to this problem can be obtained using methods such as the Hungarian algorithm [16] and the auction algorithm [17].

In this method, the permutation matrix of each sub-network $\hat{\mathbf{P}}^k$ is locally calculated within the sub-network, i.e. there is no cooperation between the sub-networks in the matching step.

3.2. Consensus step

In the consensus step, we update $\hat{\mathbf{S}}$ assuming $\hat{\mathbf{P}}^k$ is available in all sub-networks.

Since the loss function (9) is convex, $\hat{\mathbf{S}}$ is obtained by setting the derivative of the loss function (9) with respect to $\hat{\mathbf{S}}$ to 0, i.e.

$$-\sum_{k=1}^{K}\left[(\hat{\mathbf{P}}^k)'\hat{\mathbf{S}^k} - \underbrace{(\hat{\mathbf{P}}^k)'\hat{\mathbf{P}}^k}_{\mathbf{I}}\hat{\mathbf{S}}\right] = 0. \tag{15}$$

Since $(\hat{\mathbf{P}}^k)'\hat{\mathbf{P}}^k = \mathbf{I}$, we obtain $\hat{\mathbf{S}}$ as

$$\hat{\mathbf{S}} = \frac{1}{K}\sum_{k=1}^{K}\tilde{\mathbf{S}}^k \tag{16}$$

$$\tilde{\mathbf{S}}^k = (\hat{\mathbf{P}}^k)'\hat{\mathbf{S}^k}. \tag{17}$$

As (16) implies, the estimation of $\hat{\mathbf{S}}$ results in an average of $\tilde{\mathbf{S}}^k$ over all sub-networks. We can calculate this averaging in a distributed fashion using a consensus averaging protocol as explained in [18]. It is noted that unlike the matching step which is performed locally in each sub-network, the consensus step is performed by a cooperation of all sub-networks in a distributed fashion.

Remark 3: In the consensus step, $\hat{\mathbf{S}}$ is obtained by cooperation of all sub-networks in the WASN. Estimation of $\hat{\mathbf{S}}$ yields a network-wide consensus on the energy signatures of the sources. The consented signatures will generally have a better estimation accuracy when compared to the initial per-sub-network estimates. This is an advantage when these signatures are further exploited in the processing pipeline, e.g., to perform VAD.

Remark 4: The cost function (9) monotonically decreases in each iteration of the CM algorithm and in all experiments we have carried out, the algorithm was observed to converge.

3.3. Labelling

After convergence of the CM algorithm, we label the sources according to the network-wide consensus signatures such that all sub-network signatures assigned to the first row of $\hat{\mathbf{S}}$ are labelled as 1, all sub-network signatures assigned to the second row of $\hat{\mathbf{S}}$ are labelled as 2, etc.

4. VALIDATION

In this section, the accuracy of the proposed CM algorithm is investigated. The accuracy of the CM algorithm is measured using the labelling error rate calculated as

$$E_{lbl} = \frac{\sum_{k=1}^{K} ||\mathbf{P}^k - \hat{\mathbf{P}}^k||_F^2}{2NK} \times 100, \qquad (18)$$

where $\mathbf{P}^k$ is the permutation matrix obtained by matching the signatures estimated in k-th sub-network with respect to true energy signatures $\mathbf{S}$.

4.1. Experimental Setup

A $20m \times 10m \times 5m$ room with a reflection coefficient of 0.3 at all the walls containing three sources is simulated using the image method [19, 20]. We consider the network depicted in Fig. 1, which consists of 20 nodes clustered in three sub-networks. Each node is equipped with three microphones with a sampling frequency of $f_s = 16kHz$. An uncorrelated additive white Gaussian noise is present in each microphone. The energy of the signals is computed over frames of size $L = 480$.

The CM algorithm is compared with the benchmark method of [4], which is also based on energy signatures, and which is here referred to as k-means.

4.2. Results

Table 1 lists the error rate of the source labelling for different levels of noise variance. The results show that the CM algorithm labels the sources without any error when the variance of the noise is small and that the error rate increases with the variance of the noise. Table 1 also shows that the CM algorithm is more accurate than the k-means.

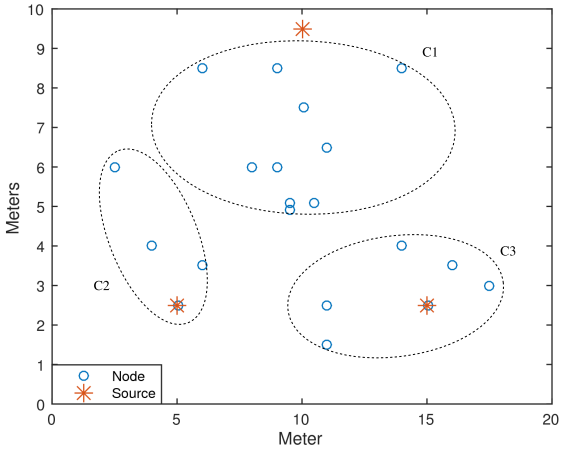

Fig. 1. A WASN of 20 nodes observing 3 speech sources. This network is clustered into three sub-networks.

Table 1. *The error rate of the source labelling using the k-means and CM algorithms (%).*

Noise Variance	k-means	CM
0	11	0
0.01	22	0
0.05	33	0
0.1	56	11
0.5	78	11

Table 2 summarizes the root mean square error (RMSE) of the energy signature estimation for the network-wide consented signatures obtained using the CM algorithm and the signatures obtained using NPCA locally at sub-networks 1, 2 and 3 (C1, C2 and C3). The results of this table show that the RMSE of the consented signatures is smaller than that of each local estimate and hence show the benefit of the applied cooperation between the sub-networks.

5. CONCLUSIONS

A new method for distributed labelling of audio sources in wireless acoustic sensor networks (WASN) has been proposed in this paper. This method uses a hierarchical approach in which first a network clustering algorithm is performed, where in each sub-network, the energy patterns of the sources are estimated using a non-negative principal component analysis (NPCA). Finally the source labels are obtained by an iterative matching algorithm, which performs a consensus step and a matching step in each iteration. In the consensus step, a network-wide consensus is obtained about the signature of the sources. In the matching step, the signatures of each sub-network are labelled according to the consented signature.

Table 2. *The estimation error of the network-wide consented signatures and the signatures estimated in each sub-network locally.*

Noise Variance	Consented	C1	C2	C3
0	8.14	9.69	13.35	11.1
0.01	8.15	9.64	13.43	11.03
0.05	8.17	9.66	13.37	11.01
0.1	8.14	9.85	13.45	13.34
0.5	8.18	9.91	13.05	10.10

6. REFERENCES

[1] A. Bertrand, "Applications and trends in wireless acoustic sensor networks: a signal processing perspective," in *Communications and Vehicular Technology in the Benelux (SCVT), 2011 18th IEEE Symposium on*, 2011, pp. 1–6.

[2] Markovich-Golan et al., "Distributed multiple constraints generalized sidelobe canceler for fully connected wireless acoustic sensor networks," *Audio, Speech, and Language Processing, IEEE Trans.*, vol. 21, no. 2, pp. 343–356, 2013.

[3] M. H. Bahari, A. Bertrand, and M. Moonen, "Blind sampling rate offset estimation based on coherence drift in wireless acoustic sensor networks," in *Signal Processing Conference (EUSIPCO), 2015 23rd European*, Aug 2015, pp. 2281–2285.

[4] S. Chouvardas et al., "Distributed robust labeling of audio sources in heterogeneous wireless sensor networks," in *Acoustics, Speech and Signal Processing, 2015 IEEE Int. Conf.*, 2015, pp. 5783–5787.

[5] J. Plata-Chaves, M. H. Bahari, M. Moonen, and A. Bertrand, "Unsupervised diffusion-based lms for node-specific parameter estimation over wireless sensor networks," in *IEEE 41th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016.

[6] J. Plata-Chaves, A. Bertrand, and M. Moonen, "Distributed signal estimation in a wireless sensor network with partially-overlapping node-specific interests or source observability," in *IEEE 40th Int. Conf. Acoustics, Speech and Signal Processing*, 2015.

[7] E. Weinstein et al., "Multi-channel signal separation by decorrelation," *Speech and Audio Processing, IEEE Transactions on*, vol. 1, no. 4, pp. 405–413, 1993.

[8] E. Oja and M. Plumbley, "Blind separation of positive sources using non-negative PCA," in *Proceedings of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), Nara, Japan*. Citeseer, 2003.

[9] A. Bertrand and M. Moonen, "Energy-based multi-speaker voice activity detection with an ad hoc microphone array," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 85–88.

[10] A. Bertrand and M. Moonen, "Blind separation of non-negative source signals using multiplicative updates and subspace projection," *Signal Processing*, vol. 90, no. 10, pp. 2877–2890, 2010.

[11] A. Bertrand and M. Moonen, "Distributed computation of the Fiedler vector with application to topology inference in ad hoc networks," *Signal Processing*, vol. 93, no. 5, pp. 1106–1117, 2013.

[12] Z. Lu and A. M. Zoubir, "Generalized bayesian information criterion for source enumeration in array processing," *Signal Processing, IEEE Transactions on*, vol. 61, no. 6, pp. 1470–1480, 2013.

[13] P. Forero, A. Cano, G. B. Giannakis, et al., "Distributed clustering using wireless sensor networks," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 4, pp. 707–724, 2011.

[14] M. H. Bahari, N. Dehak, Burget L. Van hamme, H., A. Ali, and J. Glass, "Non-negative factor analysis of gaussian mixture model weight adaptation for language and dialect recognition," *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 22, no. 7, pp. 1117–1129, 2014.

[15] M. H. Bahari and H. Van hamme, "Speaker age estimation using hidden markov model weight supervectors," in *Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on*, July 2012, pp. 517–521.

[16] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics (NRL)*, vol. 52, no. 1, pp. 7–21, 2005.

[17] D. P. Bertsekas, "Auction algorithms for network flow problems: A tutorial introduction," *Computational Optimization and Applications*, vol. 1, no. 1, pp. 7–66, 1992.

[18] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[19] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[20] E. A. P. Habets, "Room impulse response generator," 2006.