ENSEMBLE SPEAKER AND SPEAKING ENVIRONMENT MODELING APPROACH WITH ADVANCED ONLINE ESTIMATION PROCESS

Yu Tsao, Jinyu Li, and Chin-Hui Lee
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0250, USA
{yutsao, jinyuli, chl}@ece.gatech.edu

ABSTRACT

Recently, we proposed an ensemble speaker and speaking environment modeling (ESSEM) framework to characterize speaker variability and speaking environments. In contrast to multi-style training, ESSEM uses single-style training to prepare multiple sets of environment-specific acoustic models. The ensemble of these acoustic models forms a prior structure of the environment for flexible prediction of unknown environment during testing. In this study, we present methods to further improve the precision for model characterization. We first study a weighted *N*-best information technique to well utilize the *N*-best transcription hypothesis in an unsupervised adaptation manner. Next, we introduce cohort selection and environment space adaptation techniques to online improve the resolution and coverage of the prior structure. With an integration of the proposed methods, we further improve the ESSEM performance over our previous study. On the Aurora-2 task, ESSEM achieves an average word error rate (WER) of 4.64%, corresponding to a 15.64% relative WER reduction over our best baseline result (5.50% to 4.64% WER) obtained with multi-condition training.

***Index Terms*—**noise robustness, ensemble speaker and speaking environment modeling, *N*-best transcription

1. INTRODUCTION

A key issue that limits the current applicability of automatic speech recognition (ASR) is the inevitable mismatch between the training and testing conditions. The sources of the mismatch may come from speaker variability and speaking environment distortions. The exact mismatch is usually an unknown combination of these sources. Although some parametric functions have been developed to well characterize particular distortions, the exact form of the desired combination of speaker and speaking environment distortions can be complex and hard to specify.

Many approaches have been proposed to deal with the mismatch issue. Among them, a category of approaches adjusts parameters of the original hidden Markov model (HMM) set to match the testing conditions. Maximum a posteriori (MAP) [1] and maximum likelihood linear regression (MLLR) [2] are two well-known and widely used approaches. More recently, some approaches prepare prior knowledge to facilitate the characterization of the unknown testing condition. The prior knowledge is usually obtained from multiple sets of hidden Markov models (HMM) prepared in the offline. The HMM sets are trained on the available training data. During testing, another transformation is estimated based on the prior knowledge to generate a new HMM set that matches the testing data. Examples include reference speaker weighting (RSW) [3], cluster adaptive training (CAT) [4], and eigenvoice [5].

In the mid-90s, a stochastic matching (SM) approach [6] was proposed to improve the ASR performance under mismatched conditions. The effects of speaker variability and environment distortions are characterized by a mapping structure. The nuisance parameters in the mapping structure are estimated based on the testing utterances. Finally, the acoustic model from the training condition is compensated by the mapping structure to match the testing utterances. More recently, we extended the original SM framework by including the abovementioned prior knowledge and proposed an ensemble speaker and speaking environment modeling (ESSEM) approach [7, 8]. With the prior knowledge, ESSEM estimates a new set of HMMs based on the stochastic matching criterion [6]. From our previous studies, we verified that ESSEM can significantly improve the ASR performance robustness under mismatched conditions [7, 8].

In this paper, we investigate two directions to further improve the ESSEM performance. First, we address the problem that the decoded transcription of unsupervised adaption may not be correct to guide the adaptation; second, we attempt to improve the resolution and coverage of the prior knowledge. For the first issue, we study a weighted *N*-best information technique; for the second issue, we introduce two techniques—cohort selection and environment space adaption. With the combined enhancement, we achieve an average word error rate (WER) of 4.64%, corresponding to a 15.64% relative WER reduction over our best baseline of 5.50% WER obtained with multi-condition training.

2. REVIEW OF THE ESSEM FRAMEWORK

First, we briefly review the two phases of ESSEM—offline environment preparation and online super-vector estimation.

2.1. Offline Environment Preparation

In the offline phase, we can collect or use the Monte Carlo (MC) methods to obtain speech data from a wide range of different speaker and speaking environments. With P sets of speech data, we can train P sets of HMMs, Λ_p, p=1… P. For ease of modeling, the entire set of mean parameters within a set of HMMs is concatenated into a super-vector, V_p, p=1,…, P. These P super-vectors form an ensemble speaker and speaking environment space (ESS space), $\Omega_{\mathbf{V}}$, where $\Omega_{\mathbf{V}}=\{\mathrm{V}_1\,\mathrm{V}_2\ldots\,\mathrm{V}_P\}$.

2.2. Online Super-vector Estimation

In the online phase, we estimate the target super-vector, V_{Y}, for the testing environment through a mapping function, $\mathbf{G}_{\varphi}$:

$$\mathrm{V}_{\mathrm{Y}} = \mathbf{G}_{\varphi}(\Omega_{\mathrm{V}}), \tag{1}$$

with

$$\hat{\varphi} = \underset{\varphi}{argmax}\ P(F_{\mathrm{Y}} \mid \Omega_{\mathbf{V}}, \varphi, W), \tag{2}$$

where $\hat{\varphi}$ represents the nuisance parameters in the mapping function, and W is the transcription corresponding to the testing

utterances, F_Y. The nuisance parameters are only used in the mapping procedure but not involved in the recognition procedure. We estimate the nuisance parameters based on the expectation-maximization (EM) algorithm in an unsupervised adaptation (compensation) style, which uses the decoded transcription as the guide. With the estimated target super-vector, V_Y, we can have the set of acoustic models, Λ_Y, for the testing condition.

3. WEIGHTED N-BEST INFORMATION

In an unsupervised adaptation style, we can use the best decoded transcription for stochastic matching (W in Eq-(2)). However, the decoded best transcription may not be the ground truth, especially in the severed noisy conditions. Using N-best transcriptions from the decoder is a good way to address this issue, since the ground truth may be embedded in other candidates. By introducing the N-best transcriptions, we can rewrite Eq-(2) as:

$$\hat{\varphi} = \underset{\varphi}{argmax} \sum_{n \in N} \lambda_n \, P(F_Y \mid \Omega_{\mathbf{V}}, \varphi, W_n) \ , \tag{3}$$

where W_n and λ_n are the decoded transcription and weight for the n-th hypothesis, respectively. Based on a study about unsupervised speaker adaptation [9], we adopt the following equation to dynamically determine λ_n:

$$\lambda_n = \frac{\exp[(L_n - L_1)/\eta]}{\sum_{m \in N} \exp[(L_m - L_1)/\eta]} , \tag{4}$$

where L_n is the log-likelihood of the n-th hypothesis, and η is a parameter that determines the confidence of the hypotheses. Eq-(4) is a general formulation for using N-best transcription. When setting $\eta \to \infty$, an equal weighting is applied to the N-best hypotheses; when setting $\eta = 0$, only the 1-best hypothesis is used in Eq-(3) (this case is the same as using Eq-(2)). If we use a small value of η, the best candidate's likelihood will dominate Eq-(4). Therefore, we prefer a value around 12-15 used in discriminative training [10] to scale down the dominating likelihood.

4. ONLINE ESS SPACE CONSTRUCTION

Next, we present two techniques to construct the ESS space that provides better resolution and coverage to model the test conditions—cohort selection and environment space adaptation.

4.1. Cohort Selection

In our previous study, we have demonstrated that using a succinct environment space with higher resolution helps ESSEM to better characterize unknown testing environments, especially when only limited adaptation data is available [7]. In this section, we study a cohort selection technique [11] to construct a space with good resolution in the online phase. The concept of cohort selection resembles that of the family of subset selection methods [12] that find a subset of components from the entire set of components to model a signal of interest. In the implementation aspect, cohort selection can be seen as an extension of the best first function [8]. However, instead of locating one most matched environment, cohort selection finds N training environments (cohorts) that are closest to the testing environment.

In this study, we use the likelihood to measure the closeness. With the selected cohort environments, we build a cohort ESS space, $\Omega_{\mathbf{V}_{CH}}$. Finally, we use the stochastic matching algorithm to estimate the target super-vector, V_Y, for the testing condition:

$$V_Y = \mathbf{G}_{\varphi}(\Omega_{\mathbf{V}_{CH}}) . \tag{5}$$

The nuisance parameters can be estimated with the EM algorithm.

4.2. Environment Space Adaptation

As mentioned earlier, ESSEM prepares the ESS space using the available training data in the offline. Thus, the ESS space may have a poor coverage for the testing environments that contain distortions not included in the training set. This poor coverage limits the ESSEM performance. Here, we propose an environment space adaptation (ESA) technique to online generate a new ESS space that provides better coverage for the testing conditions.

Based on the testing utterances, ESA generates a new ESS space online by compensating the parameters of the original ESS space. The stochastic matching criterion is used for the compensation process with a mapping function, $\mathbf{G}_{\boldsymbol{\theta}}(\bullet)$:

$$\Omega_{\mathbf{V}_{ESA}} = \mathbf{G}_{\boldsymbol{\theta}}(\Omega_{\mathbf{V}}) , \tag{6}$$

where $\Omega_{\mathbf{V}_{ESA}}$ is the compensated ESS space, $\boldsymbol{\theta}$ denotes the nuisance parameters of the mapping function. The new space provides a better coverage for the testing condition. Finally, we estimate the target super-vector, V_Y, through stochastic matching:

$$V_Y = \mathbf{G}_{\varphi}(\Omega_{\mathbf{V}_{ESA}}) . \tag{7}$$

Similarly, the set of nuisance parameters, φ, of the mapping function can be estimated by the EM algorithm as shown in Eq-(2).

5. EXPERIMENTS

We conducted experiments on the Aurora-2 database [13]. The multi-condition training set was used to obtain environment-specific HMMs and to build the ESS spaces. We tested ESSEM in a per-utterance unsupervised compensation mode on a gender dependent (GD) system [7, 8]. For the GD system, two GD HMM sets were first trained. Then, 17 environment-specific HMM sets for each gender were obtained by adapting mean vectors from that GD HMM set to specific environments. Accordingly, two GD ESS spaces along with two GD HMM sets were prepared. An automatic gender identification (AGI) unit was used to determine the gender identity of each testing speaker. The full evaluation set was used to test the ESSEM performance. In this paper, we only report the results of 50 conditions (10 types of noise, 0dB to 20dB SNRs). A modified ETSI advanced front-end (AFE) was used for feature extraction, and we followed a complex back-end topology as presented in [14] to train HMMs. More details about the experimental setup can be found in our previous study [7, 8].

In the following experiments, we tested ESSEM performance on our current optimal ESS space. This ESS space was refined by soft margin estimation (SME) [10] and minimum classification error (MCE) training [7, 15] in the offline phase. SME is to increase discrimination within each super-vector, and MCE is to increase the distance between each pair of super-vectors. Based on our preliminary experiments, this ESS space always gives the best performance for a same online method.

5.1 Baseline

First, we report two baseline results in Table-1. For "Baseline (AGI)" in Table-1, we directly used the AGI unit to identify speaker's gender for each testing utterance. Then, the HMM set for the identified gender is used to decode the same testing utterance.

For "Baseline(EC)" in Table-1, we followed our previous study [7] and adopted an environment clustering (EC) tree to obtain this set of testing results. First, we built a two-layer hierarchical EC tree to structure the 34 environments into seven clusters. In the first layer, the 34 environments were exactly divided into two groups, each corresponding to one gender. In the second layer,

another two groups were classified roughly according to high/low SNR levels. We prepared a representative HMM set for each of these seven nodes. Each set of representative HMMs was trained in a multi-style training manner using the speech data belonging to its corresponding node. During testing, we located one cluster from this EC tree and used its corresponding representative HMM set to recognize the testing utterance. Here, the same AGI unit was used for the first layer of the EC tree to identify speaker's gender. At the second layer, an online cluster selection was conducted to determine one most suitable cluster of speaking environments. More details about the EC-structured baseline can be found in our previous study [7, 8]. To have a fair comparison, we applied the SME criterion [10] to improve each of the GD and representative HMM sets used in the two baseline experiments.

By comparing the two baselines in Table-1, we observe that "Baseline(EC)" provides better performance than "Baseline(AGI)". This result confirms that in addition to two genders, a speaking environment clustering process can give us a better baseline system. In the following discussions, we will use "Baseline(EC)" as an additional set of baseline to compare with the proposed techniques.

5.2 Weighted *N*-best Information

Next, we present the results of weighted *N*-best information on the ESSEM framework. The environment clustering (EC) technique was applied on ESSEM to enhance performance [7]. We adopted the same hierarchical EC tree as stated in the previous section to prepare seven clusters. Environments belonging to a same cluster then formed an EC sub-space, $\Omega_{\mathbf{V}^{(c)}}$, c=1,2,…C (here C=7). In the online stage, ESSEM selected a cluster (for example, the t-th cluster) and located its corresponding sub-space ($\Omega_{\mathbf{V}^{(t)}}$). With the selected sub-space, we estimated the target super-vector, $\mathrm{V_Y}$, by:

$$\mathrm{V_Y} = \mathbf{G}_{\varphi}(\Omega_{\mathbf{V}^{(t)}}), \tag{8}$$

with

$$\hat{\varphi} = \underset{\varphi}{argmax} \sum_{n\in N} \lambda_n P(F_{\mathrm{Y}} \mid \Omega_{\mathbf{V}^{(t)}}, \varphi, W_n) . \tag{9}$$

In this set of experiments, we used a linear combination (LC) function [8] as the mapping structure. Therefore, Eq-(8) becomes:

$$\mathrm{V_Y} = \sum_{p=1}^{P^{(t)}} \hat{w}_p \mathrm{V}_p , \tag{10}$$

with

$$\{\hat{w}_p\}_{p=1}^{P^{(t)}} = arg \max_{\{w_p\}_{p=1}^{P^{(t)}}} \sum_{n\in N} \lambda_n P(F_{\mathrm{Y}} \mid \sum_{p=1}^{P^{(t)}} w_p \mathrm{V}_p, W_n), \tag{11}$$

where $\hat{w}_p$ is the p-th weighting coefficient of the LC function, and $P^{(t)}$ is the total number of super-vectors of the t-th EC sub-space.

We used a 8-best list (N=8). Table-1 lists results of ESSEM plus EC with setting $\eta\to\infty$ and η=0. By testing many different values, we found η=15 gave the best performance, and we listed the results in Table-1. In the following ESSEM experiments, we integrated the weighted *N*-best information technique with η=15.

Table-1 Average word error rates (in %) from 0dB to 20dB.

Test Condition	SetA	SetB	SetC	Overall
Baseline(AGI)	5.09	5.32	6.69	5.50
Baseline(EC)	5.05	5.31	6.31	5.41
ESSEM+EC($\eta\to\infty$)	4.58	4.88	5.51	4.89
ESSEM+EC(η=15)	4.53	4.78	5.46	4.82
ESSEM+EC(η=0)	4.53	4.89	5.54	4.88

5.3 Online Cohort Selection and ESA

Next, we tested the two online methods, cohort selection and ESA. For cohort selection, we located 15 environments closest to the testing condition in the original ESS space. For the ESA technique, we integrated it with EC (named EC-ESA) and used the EC tree introduced in Section 5.1. Similar to the original EC algorithm, a cluster was first selected based on the testing utterances. Then, ESA compensated the parameters in the selected EC sub-space, $\Omega_{\mathbf{V}^{(t)}}$, to the EC-ESA space, $\Omega_{\mathbf{V}_{ESA}^{(t)}}$, through stochastic matching:

$$\Omega_{\mathbf{V}_{ESA}^{(t)}} = \mathbf{G}_{\boldsymbol{\theta}}(\Omega_{\mathbf{V}^{(t)}}) . \tag{12}$$

In this paper, we adopted a simple mapping process for $\mathbf{G}_{\boldsymbol{\theta}}(\bullet)$. We compensated each super-vector to match the testing condition individually. Accordingly, Eq-(12) becomes:

$$\mathrm{V'}_p = \mathbf{G}_{\theta_p}(\mathrm{V}_p) , p=1\ldots P^{(t)}, \tag{13}$$

where $\mathrm{V'}_p$ and V_p, are the compensated and original super-vectors for the p-th environment, and $\mathbf{G}_{\theta_p}(\bullet)$ is the mapping structure for the p-th super-vector. Finally, we obtain the EC-ESA space by:

$$\Omega_{\mathbf{V}_{ESA}^{(t)}} = \{\mathrm{V'}_1\ \mathrm{V'}_2 \ldots \mathrm{V'}_{P^{(t)}}\} . \tag{14}$$

Here, we used diagonal MLLR [2] for $\mathbf{G}_{\theta_p}(\bullet)$ to compensate each super-vector in an unsupervised manner. With the EC-ESA space, $\Omega_{\mathbf{V}_{ESA}^{(t)}}$, ESSEM used the LC function as shown in Eq-(10) for the mapping structure to estimate the target super-vector, $\mathrm{V_Y}$. Table-2 lists the cohort selection and EC-ESA results as "ESSEM+cohort (LC)" and "ESSEM+EC-ESA(LC)". For ease of comparison, Table-2 also lists the results of ESSEM with EC as "ESSEM+ EC(LC)"; which is the same set of results to the η=15 in Table-1.

Table-2 Average word error rates (in %) from 0dB to 20dB.

Test Condition	SetA	SetB	SetC	Overall
ESSEM+EC(LC)	4.53	4.78	5.46	4.82
ESSEM+cohort(LC)	4.53	4.71	5.49	4.79
ESSEM+EC-ESA(LC)	4.41	4.75	4.97	4.66

By comparing Table-1 and Table-2, we can see that the three ESSEM results are clearly better than the two baseline results. Next from Table-2, we observe that "ESSEM+cohort(LC)" can give slightly better performance than "ESSEM+EC(LC)". Please note that both EC and cohort selection resemble the subset selection methods [12]. EC online selects one sub-space from many prepared EC-structured sub-spaces, while cohort selection online collects super-vectors to construct a cohort ESS space. The testing results actually confirm that online cohort selection can provide relatively better resolution to model the testing condition.

We also observe that "ESSEM+EC-ESA(LC)" achieves clearly better performance than "ESSEM+EC(LC)". Especially for test Set C, where an additional channel diction is added, EC-ESA gives a clear improvement of 8.97% (5.46% to 4.97% WER) relative WER reduction over EC alone. This result suggests that ESA can online generate an ESS space that has better coverage to characterize the testing conditions, especially for those containing distortions not included in the training set.

5.4 Integration with Complex Online Mapping Structure

Next, we compare the same three techniques—EC, cohort selection, and EC-ESA—with a more complex mapping structure, linear

combination with a correction bias (LCB) [8]. When applying EC and the LCB function along with the weighted *N*-best information technique on ESSEM, Eq-(8) now becomes:

$$\mathrm{V_Y} = \sum_{p=1}^{P^{(t)}} \hat{w}_p \mathrm{V}_p + \hat{b}\,, \tag{15}$$

with

$$\{\hat{w}_p, \hat{b}\}_{p=1}^{P^{(t)}} = arg \max_{\{w_p, b\}_{p=1}^{P^{(t)}}} \sum_{n \in N} \lambda_n P(F_\mathrm{Y} \mid \sum_{p=1}^{P^{(t)}} w_p \mathrm{V}_p + b, W_n)\,, \tag{16}$$

where $\hat{b}$ is a global correction bias [8].

Table-3 lists the results of EC as "ESSEM+EC(LCB)". We used the same procedure as described in the previous section to test cohort selection and EC-ESA. The corresponding results are listed as "ESSEM+cohort(LCB)" and "ESSEM+EC-ESA(LCB)". As shown in Table-3, the results for all three sets are similar. However, the bottom row (EC-ESA) still gives slightly better performance than the other two techniques and provides WER reductions of 15.64% and 14.23%, respectively, over "Baseline(AGI)" (5.50% to 4.64% WER) and "Baseline(EC)" (5.41% to 4.64% WER). In Table-4, we list the detailed EC-ESA results (in accuracy %) at each testing condition of the Aurora-2 task.

Table-3 Average word error rates (in %) from 0dB to 20dB.

Test Condition	SetA	SetB	SetC	Overall
ESSEM+EC(LCB)	4.43	4.74	4.98	4.66
ESSEM+cohort(LCB)	4.48	4.74	4.92	4.67
ESSEM+EC-ESA(LCB)	4.41	4.73	4.92	4.64

6. CONCLUSION

In this paper, we study techniques to enhance online estimation in the ESSEM framework. We first incorporated weighted *N*-best information in an unsupervised compensation mode. Next, we introduced cohort selection and environmental space adaptation from environment clustering (EC-ESA) to construct the ESS spaces online. Cohort selection and EC-ESA, respectively, enable ESSEM to enhance the resolution and coverage of the environment space to better characterize the testing conditions. When compared with the multi-style trained baseline, ESSEM with an integration of these online techniques achieves a significant 15.64% word error reduction (5.50% to 4.64% WER) on the Aurora-2 task.

The important issue of enhancing the precision and coverage of the ESS space was only discussed briefly in Section 5. SME was used to improve intra-environment precision while MCE was applied to increase the inter-environment separation, and hence enlarge the coverage of the ESS space with a small number of environment-specific super-vectors. Based on our experiments, this enhanced ESS space always improved performance over configurations without using any discriminative training for ESS enhancement. A detailed report will be given in a future paper.

ACKNOWLEDGEMENT

This work was supported by a Texas Instruments Leadership University grant. We also thank Prof. Koichi Shinoda of Tokyo Institute of Technology for helpful discussions and comments.

REFERENCES

[1] J.-L. Gauvain and C.-H. Lee, "Maximum a posteriori estimation for multivariate Gaussian mixture observations of Markov chains," *IEEE Trans. Speech Audio Processing*, vol. 2, pp.291-99, April 1994.

[2] M. J. F. Gales, "Maximum likelihood linear transformations for HMM-based speech recognition," *Computer Speech and Language*, vol.12, no. 2, pp.75-98, 1998.

[3] T. J. Hazen, "A comparison of novel techniques for rapid speaker adaptation," *Speech Comm.*, pp.15-33, 2000.

[4] M. J. F. Gales, "Cluster adaptive training of hidden Markov models," *IEEE Trans. Speech Audio Proc.*, pp. 417-428, 2000.

[5] R. Kuhn, J.-C. Junqua, P. Nguyen, and N. Niedzielski, "Rapid speaker adaptation in Eigenvoice space," *IEEE Trans. Speech Audio Processing*, vol. 8, pp.695-707, Nov. 2000.

[6] A. Sankar and C.-H. Lee, "A maximum-likelihood approach to stochastic matching for robust speech recognition," *Trans. Speech Audio Proc.*, pp.190-202, 1996.

[7] Y. Tsao and C.-H. Lee, "Two extensions to ensemble speaker and speaking environment modeling for robust automatic speech recognition," in *ASRU*, 2007.

[8] Y. Tsao and C.-H. Lee, "Improving the ensemble speaker and speaking environment modeling approach by enhancing the precision of the online estimation process," in *Interspeech 2008*.

[9] P. Nguyen, P. Gelin, J.-C. Junqua, and J.-T. Chien, "N-best based supervised and unsupervised adaptation for native and non-native speakers in cars," *ICASSP'97.*, pp. 257-265, 1997.

[10] J. Li, "Soft margin estimation for automatic speech recognition," Ph.D. Dissertation, School of ECE, Georgia Institute of Technology, 2008.

[11] B. Mak, T.-C. Lai, and R. Hsiao "Improving reference speaker weighting adaptation by the use of maximum-likelihood reference speakers," in *ICASSP 99*, vol. 1, pp. 173-176, 1999.

[12] S. Chen, D. Donoho, and M. A. Saunders, "Atomic Decomposition by Basis Pursuit", in *SIAM J. on Scientific Computing*, vol. 20, No. 1, pp. 33-61, 1998.

[13] D. Pearce and H.-G. Hirsch, "The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions," in *Proc. ISCA ITRW ASR'2000*.

[14] J. Wu and Q. Huo, "Several HKU approaches for robust speech recognition and their evaluation on Aurora connected digit recognition tasks," in *Eurospeech 2003*.

[15] B.-H. Juang, W. Chou, and C.-H. Lee, "Minimum classification error rate methods for speech recognition," *IEEE Trans. Speech Audio Proc.*, pp. 257-265, 1997.

Table-4 ESSEM+EC-ESA(LCB) with the best ESS space on the Aurora-2 task (in word accuracy %).

	Set A					Set B					Set C			Overall
	Subway	Babble	Car	Exhibition	Average	Restaurant	Street	Airport	Station	Average	Subway M	Street M	Average	Average
20 dB	99.60	99.55	99.55	99.48	99.55	99.75	99.30	99.40	99.69	99.54	99.69	99.46	99.58	99.55
15 dB	99.48	99.03	99.37	99.14	99.26	99.36	99.00	99.55	99.48	99.35	99.48	99.24	99.36	99.31
10 dB	98.74	98.43	98.72	97.90	98.45	98.37	97.64	98.48	98.61	98.28	98.28	97.67	97.98	98.28
5 dB	96.25	95.04	96.90	94.57	95.69	95.27	94.47	95.74	95.71	95.30	95.76	94.26	95.01	95.40
0 dB	87.07	80.02	88.70	84.20	85.00	81.61	82.98	86.10	84.88	83.89	85.35	81.56	83.46	84.25
Average	96.23	94.41	96.65	95.06	95.59	94.87	94.68	95.85	95.67	95.27	95.71	94.44	95.08	95.36