Blind Estimation of the Subband Reverberation Time

Shrishti Saha Shetu

Final Presentation of Research Internship
Chair of Multimedia Communications and Signal Processing

Jan. 14, 2020

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

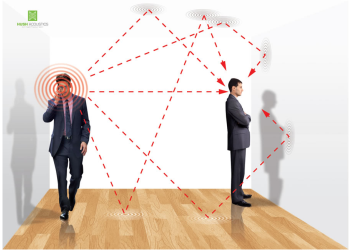
LMS

Outline

- Introduction
- Common Model-Based Approaches
- New Method for Non-Blind Subband T60 Estimation
- Model-Based Approaches for Fullband and Subband T60 Estimation
- Conclusions

Introduction

- **Reverberation time (T60)**
  - Time required for sound energy to decay by 60dB after switching off the excitation
  - Acoustical property of a room
- Blind T60 estimation needed for various applications such as dereverberation, ASR, etc.



Picture taken from [1]

Introduction

- **Fullband T60**
  - Estimated in the time-domain (broadband T60)
- **Subband T60**
  - Estimated in the frequency-domain

Introduction

- **Fullband T60**
  - Estimated in the time-domain (broadband T60)
- **Subband T60**
  - Estimated in the frequency-domain
- Subband T60 estimation is much less explored than fullband T60 estimation, cf. [2]

Goal:

Model-based approach to estimate single-channel subband T60 and comparison with existing approaches.

Introduction

Goal:

Model-based approach to estimate single-channel subband T60 and comparison with existing approaches.

Approach:

- Comparison of different subband T60 estimators
- Creating dataset for analysis and training of different methods
- Investigation of different model-based approaches

Common Model-Based Approaches

- Method based on Maximum Likelihood (ML) Estimation [8]
  - Based on a statistical decay model
  - T60 obtained by ML estimation and statistical post-processing

Common Model-Based Approaches

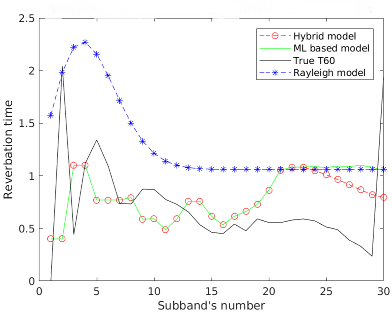
- **Method based on Maximum Likelihood (ML) Estimation [8]**
  - Based on a statistical decay model
  - T60 obtained by ML estimation and statistical post-processing
- **Rayleigh Model-based Method [9]**
  - Estimates for the upper subbands are extrapolated from the more reliable estimates of the lower subbands by a Rayleigh model function

Common Model-Based Approaches

- Method based on Maximum Likelihood (ML) Estimation [8]
  - Based on a statistical decay model
  - T60 obtained by ML estimation and statistical post-processing
- Rayleigh Model-based Method [9]
  - Estimates for the upper subbands are extrapolated from the more reliable estimates of the lower subbands by a Rayleigh model function
- A Hybrid Method [6]
  - T60 decreases monotonically from 4 to 20 kHz due to the material and air absorption
  - 2nd-order polynomial model function for estimating the T60 from mid- to high-frequencies used

Common Model-Based Approaches

- **Comparison:**
  - RIRs from ACE challenge dataset, cf. [2]
  - Octave filterbank used for subband decomposition

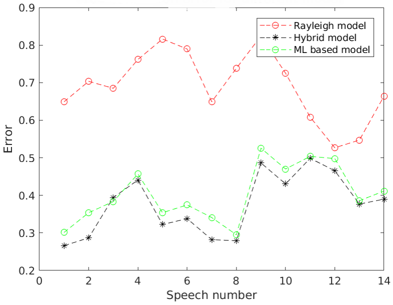


Common Model-Based Approaches

- **Comparison:**
  - 14 RIRs used for evaluation
  - Average absolute error taken as performance measure

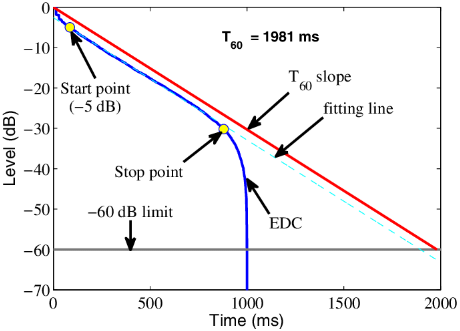
$$\mathrm{AE} = \frac{\sum_{i=1}^{N_{f_c}} \mid \tilde{T}_{60}(f_{c,i}) - T_{60}(f_{c,i}) \mid}{N_{f_c}}$$

- **Result:**
  - Hybrid method achieves lowest error



Schroeder Method

- Energy decay curve approximated by a linear LS fit [10]
- Measure the time taken for the energy to decay by 60dB



Picture taken from [7]

Schroeder Method

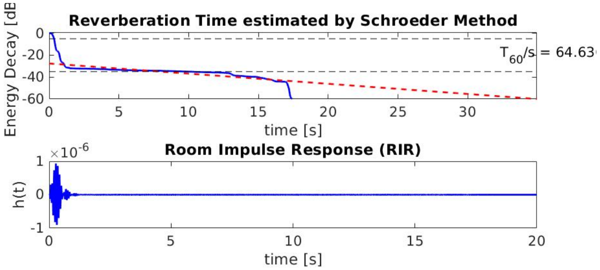
- **Problem:**
  - Determining the limits for the LS fit of the Schroeder method is a tedious task for large databases
  - Can only be used for subbands above the Schroeder frequency [4]

$$f_s = 2000\sqrt{\frac{T}{V}}$$

  - Ringing effect and sharp energy decay at the beginning of time frames

Schroeder Method

- **Sharp Energy Decay:**
  - Example from ChurchMargaret RIR dataset
  - Corresponds to the 2nd subband with center frequency 31.6228 Hz using a 30 channel octave filtebank
- **Problem:**
  - Infeasible T60 estimate obtained



New Method for Non-Blind Subband T60 Estimation

Problem:

Many databases do not contain ground-truth data for subband T60.

New Method for Non-Blind Subband T60 Estimation

Problem:

Many databases do not contain ground-truth data for subband T60.

Goal:

Automatic Schroeder Method for obtaining ground truth data for evaluating and analysing model-based approaches.

New Method for Non-Blind Subband T60 Estimation

Approach:

Modify the Schroeder method for automatically calculating ground-truth suband T60.

New Method for Non-Blind Subband T60 Estimation

Approach:

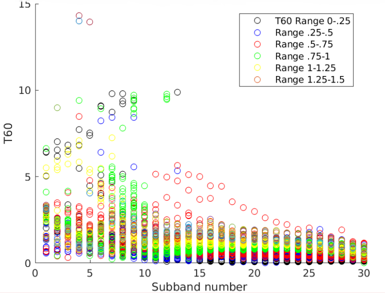
Modify the Schroeder method for automatically calculating ground-truth suband T60.

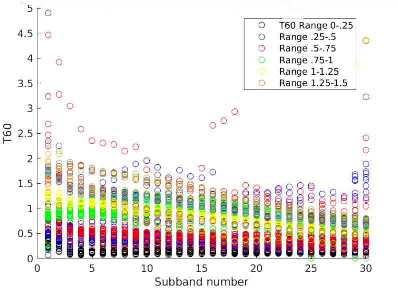
- **Steps:**
  - Set threshold for subband T60 based on fullband T60
  - Specify different energy decay range for LS fit
  - Iteration over those ranges
  - Check threshold
  - Minimum estimation if T60 not found at the end of all iterations

New Method for Non-Blind Subband T60 Estimation

▶ Subband T60 Estimation using Automatic Schroeder Method:

- Octave and DCT filterbank used
- RIRs grouped in different fullband T60 ranges





Subband T60 Model of Jeub

- In [3], two models for the subband RT obtained by a DCT filter bank are proposed
  - Model based on two or three 'reliable' suband estimates
  - Parameters

Center frequency	f_1	f_2	f_3
RT Model 1	0.8KHz	7.2 KHz	-
RT Model 2	0.8KHz	3.2 KHz	7.2 KHz

Evaluation of Jeub's Subband T60 Model

- Average absolute error:

	11 DCT Subbands	30 DCT Subbands
Model 1	0.0534 per subband	0.0648 per subband
Model 2	0.1447 per subband	0.1703 per subband

- Conclussion: RT Model 1 performs better than RT Model 2.

First Order Polynomial Model

Motivation:

- Previous observations show that the subband T60s largely depend on the fullband T60
- Monotonic decay of the subband T60

First Order Polynomial Model

Motivation:

- Previous observations show that the subband T60s largely depend on the fullband T60
- Monotonic decay of the subband T60

Goal:

Creating two first order linear regression models for only upper subbands and all subbands respectively.

First Order Polynomial Model

- Approach:
  - Dataset (DCT clean) grouped based on different fullband T60 ranges(state in steps of 100ms starting with 0.1s)
  - Linear regression is realized by a first order polynomial fitting
  - Evaluate both the models for the specified dataset
  - Evaluation results obtained by averaging for all fullband T60 ranges

First Order Polynomial Model

Evaluation Criteria:

$$\mathsf{mean}_{\mathsf{AE}} = \frac{\sum_{N=1}^{N} \frac{\sum_{N_s=1}^{N_s} \mathsf{AE}}{N_s}}{N}$$

First Order Polynomial Model

Evaluation Criteria:

$$\text{mean}_{\text{AE}} = \frac{\sum_{N=1}^{N} \frac{\sum_{N_s=1}^{N_s} \text{AE}}{N_s}}{N}$$

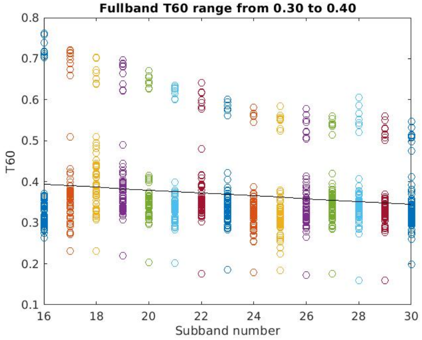
- Results:

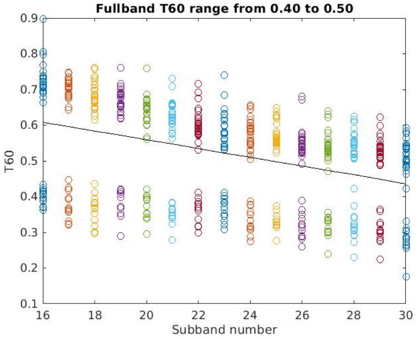
Model Name	Mean Average Error
Model for upper subband	0.0400 per subband
Model for all subband	0.1010 per subband

First Order Polynomial Model

▶ Test Results for the Upper Subband Model:

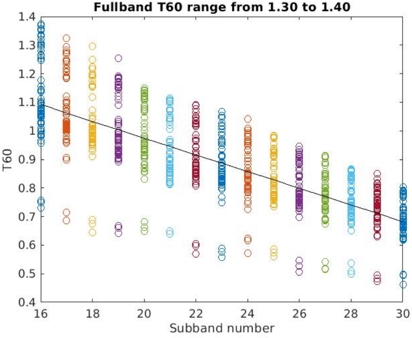
- RIRs from DCT clean dataset with no outliers
- Fullband T60 ranges from 0.3 to 0.5

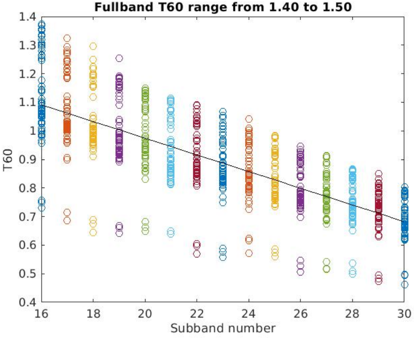




First Order Polynomial Model

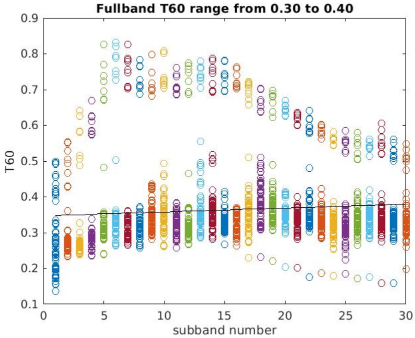
- **Test Results for the Upper Subband Model:**
  - RIRs from DCT clean dataset with no outliers
  - Fullband T60 ranges from 1.3 to 1.5

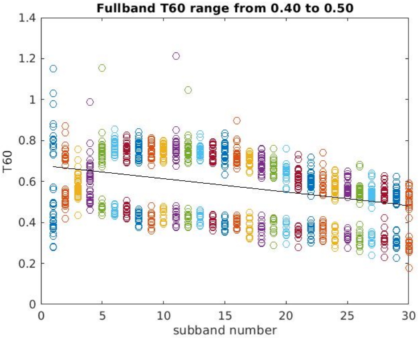




First Order Polynomial Model

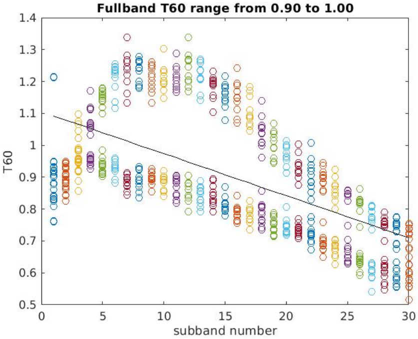
- **Test Results for the All Subband Model:**
  - RIRs from DCT clean dataset with no outliers
  - Fullband T60 ranges from 0.3 to 0.5

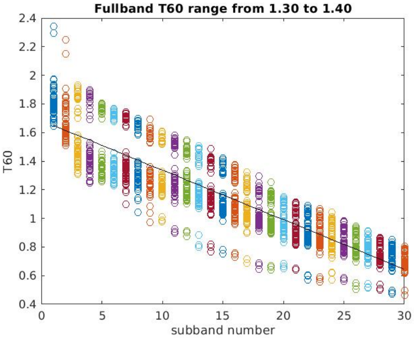




First Order Polynomial Model

- **Test Results for the All Subband Model:**
  - RIRs from DCT clean dataset with no outliers
  - Fullband T60 ranges from 0.9 to 1.0 and 1.3 to 1.4





First Order Polynomial Model

- Observations:
  - Different subband T60 'patterns' in different fullband T60 ranges
  - Motivates to estimate fulland T60 from subband estimates using the following model

$$RT_{T60} = \frac{\frac{\sum_{n=1}^{30} sub_{T60}(n)}{30} + \frac{\sum_{n=1}^{10} sub_{T60}(n)}{10} + \frac{\sum_{n=1}^{5} model_{T60}(n)}{5}}{3} - error_{model} \quad (1)$$

First Order Polynomial Model

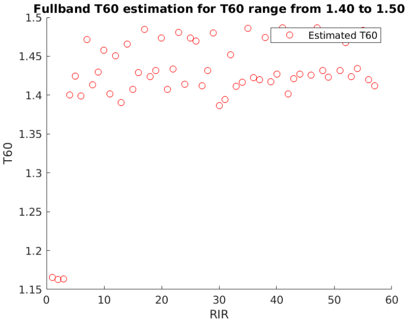
- **Observations:**
  - Different subband T60 'patterns' in different fullband T60 ranges
  - Motivates to estimate fulland T60 from subband estimates using the following model

$$RT_{T60} = \frac{\frac{\sum_{n=1}^{30} sub_{T60}(n)}{30} + \frac{\sum_{n=1}^{10} sub_{T60}(n)}{10} + \frac{\sum_{n=1}^{5} model_{T60}(n)}{5}}{3} - error_{model} \quad (1)$$

  - Estimate based on true subband T60 and linear regresson model estimations
  - Subtract the model error as a regularizing parameter

First Order Polynomial Model

- **Fullband T60 Estimation:**
  - Evaluated for DCT clean dataset
  - Fullband T60 ranges from 1.40 to 1.50 used
  - Same model function from the previous slide used for estimation
  - Results in 0.0485 mean average error



DNN Model for Fullband T60 using the Subband Estimates

Rationale:

Previous observations show dependency of fullband T60 estimation on subband T60s.

DNN Model for Fullband T60 using the Subband Estimates

Rationale:

Previous observations show dependency of fullband T60 estimation on subband T60s.

Approach:

Using a Convolutional Neual Network (CNN) to estimate fullband T60 from subband estimates.

DNN Model for Fullband T60 using the Subband Estimates

- Parameters:

	Clean Dataset	ACE Dataset	DCT Dataset(Full)	Octave Dataset(Full)
Train Samples	253	348	532	532
Test Samples	109	14	216	216
Filterbank	DCT	DCT	DCT	Octave
Outliers	No	NO	Yes	Yes
Conv. layer	1	-	-	-
Dense layer	3	-	-	-
Regularization	Dropout	-	-	-

DNN Model for Fullband T60 using the Subband Estimates

- Results:

	Clean Dataset	ACE Dataset	DCT Dataset(Full)	Octave Dataset(Full)
MAE	0.0557	0.0841	0.0584	0.0609
MSE	0.0048	0.0099	0.0027	0.0062

DNN Model for Fullband T60 using the Subband Estimates

- Results:

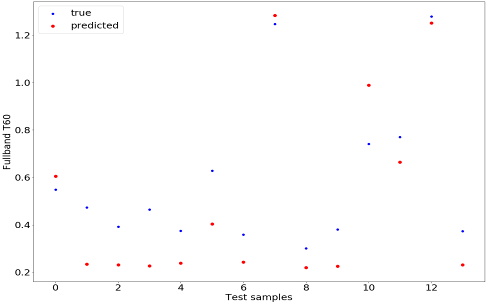
	Clean Dataset	ACE Dataset	DCT Dataset(Full)	Octave Dataset(Full)
MAE	0.0557	0.0841	0.0584	0.0609
MSE	0.0048	0.0099	0.0027	0.0062

- CNN model is able to estimate the fullband T60 with very low MSE and MAE
- Test MSE for ACE challenge RIRs was lower than the best performing algorithm reported in [2]
- Model also performs well with the dataset containing outliers

DNN Model for Fullband T60 using the Subband Estimates

▶ Test Results for ACE Dataset:

- 14 test samples with no outliers and DCT filterbank used for subband decomposition
- MSE $= 0.0099$



DNN Model for Subband T60 Estimation

Motivation:

- Analysis of different model-based approaches for subband T60 estimation
- Deep learning based approach to estimate multi-channel fullband T60 [5]

DNN Model for Subband T60 Estimation

Motivation:

- Analysis of different model-based approaches for subband T60 estimation
- Deep learning based approach to estimate multi-channel fullband T60 [5]

.

Approach:

- Using a Convolutional Neual Network (CNN) to create an universal model for subband T60 estimation

.

DNN Model for Subband T60 Estimation

- Parameters:

	DCT full dataset
Input features	Spectogram of 2s sequence of given RIR
Input shape	257×251
Training Samples	800
Test Samples	343
Filterbank	DCT
Outliers	No
Conv. layer	3
Dense layer	3
Regularization	Dropout

DNN Model for Subband T60 Estimation

- Results:

	DCT full dataset
MAE	0.1207
MSE	0.0320

DNN Model for Subband T60 Estimation

- Results:

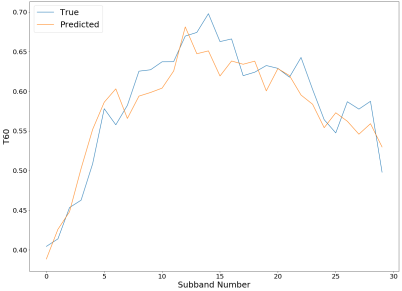
	DCT full dataset
MAE	0.1207
MSE	0.0320

- MSE and MAE is higher than the DNN fullband T60 estimation approach
- MAE is better than the 2nd subband model of Jeub and close to the 1st model

DNN Model for Subband T60 Estimation

▶ Test Result of DNN Subband Model:

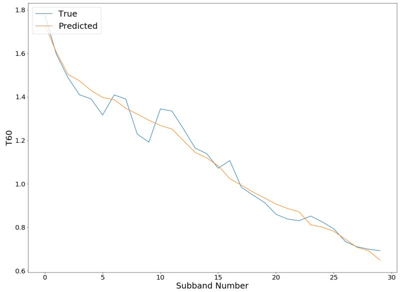
- Sample containing no outliers
- DCT filterbank used for subband decomposition



DNN Model for Subband T60 Estimation

▶ Test Result of DNN Subband Model:

- Sample containing no outliers
- DCT filterbank used for subband decomposition



Conclusions

- Algorithm for non-blind subband T60 estimation method devoloped

Conclusions

- Algorithm for non-blind subband T60 estimation method devoloped
- First order polynomial model cannot approximate well the subband T60

Conclusions

- Algorithm for non-blind subband T60 estimation method devoloped
- First order polynomial model cannot approximate well the subband T60
- DNN model shows promising result for subband T60 estimation

Conclusions

- Algorithm for non-blind subband T60 estimation method devoloped
- First order polynomial model cannot approximate well the subband T60
- DNN model shows promising result for subband T60 estimation
- A 1-D CNN model is also developed for the estimation of fullaband T60 from subband estimates

Bibliography I

[1] https://www.hushacoustics.co.uk/.

[2] James Eaton et al. "ESTIMATION OF ROOM ACOUSTIC PARAMETERS: THE ACE CHALLENGE". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (June 2016).

[3] Marco Jeub. "JOINT DEREVERBERATION AND NOISE REDUCTION FOR BINAURAL HEARING AIDS AND MOBILE PHONE". In: (2012), pp. 41–46.

[4] S. M. KHEZRI and P. J. SHALKOUHI. "THE SCHROEDER FREQUENCY OF FURNISHED AND UNFURNISHED SPACES". In: *Graduate School of the Environment and Energy, Science and Research Branch, Islamic Azad University, Tehran, Iran* (2012), pp. 1–4.

[5] Myungin Lee and Joon Hyuk Chang. "DEEP NEURAL NETWORK BASED BLIND ESTIMATION OF REVERBERATION TIME BASED ON MULTI-CHANNEL MICROPHONES". In: *ACTA ACUSTICA UNITED WITH ACUSTICA* 104 (2018), pp. 486–495.

[6] Song Li, Roman Schlieper, and Jürgen Peissig. "A HYBRID METHOD FOR BLIND ESTIMATION OF FREQUENCY DEPENDENT REVERBERATION TIME USING SPEECH SIGNALS". In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2019), pp. 1–5.

[7] Amaro A de Lima et al. "FEATURE ANALYSIS FOR QUALITY ASSESSMENT OF REVERBERATED SPEECH". In: *2009 IEEE International Workshop on Multimedia Signal Processing*. IEEE. 2009, pp. 1–5.

[8] Heinrich W Löllmann et al. "AN IMPROVED ALGORITHM FOR BLIND REVERBERATION TIME ESTIMATION". In: *In Proceedings of International Workshop on Acoustic Echo and Noise Control (IWAENC)* (2010), pp. 1–4.

[9] Heinrich W Löllmann et al. "SINGLE-CHANNEL MAXIMUM-LIKELIHOOD T60 ESTIMATION EXPLOITING SUBBAND INFORMATION". In: *Proceedings of the ACE Challenge Workshop* (2015), pp. 1–5.

[10] M R. Schroeder. " NEW METHOD OF MEASURING REVERBERATION TIME". In: *The Journal of the Acoustical Society of America* 37 (Mar. 1965), pp. 409–412.

Thank you for your attention!

New Method for Non-Blind Subband T60 Estimation

Reliability Measurement:

Goal: Creating a clean dataset using automatic Schoreder method without any unreliable or outlier estimates.

New Method for Non-Blind Subband T60 Estimation

Reliability Measurement:

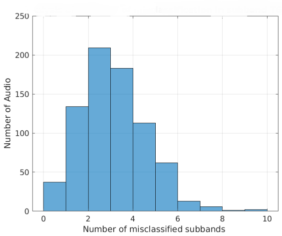
Goal: Creating a clean dataset using automatic Schoreder method without any unreliable or outlier estimates.

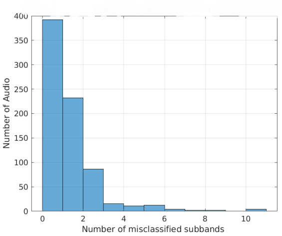
- **Approach:**
  - Group subbands in different regions
  - Set local threshold for T60 and LS error
  - Marking an estimate as reliable or not

New Method for Non-Blind Subband T60 Estimation

Analysis for DCT and Octave Filterbanks:

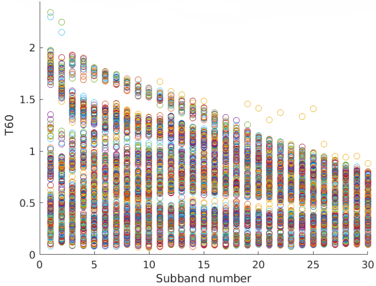
- Subband T60 estimation using both filterbanks
- Reliability measurement





New Method for Non-Blind Subband T60 Estimation

- **Creating Clean Dataset:**
  - Based on analysis results of previous slide
  - Creating DCT clean dataset by only considering RIRs having no outliers
  - Observing the tendency of subband T60 decreasing for higher frequencies

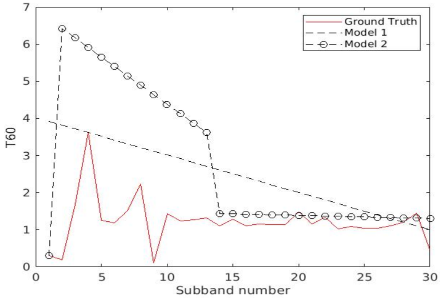


Subband T60 Model of Jeub

- Observations:
  - Totally depends on the assumption that the true T60 of two or three subbands are already known
  - Results in higher error, if the RT of these subbands are miscalculated

Subband T60 Model of Jeub

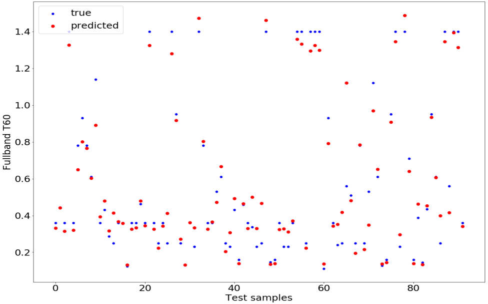
- **Observations:**
  - Totally depends on the assumption that the true T60 of two or three subbands are already known
  - Results in higher error, if the RT of these subbands are miscalculated
    - In figure an example is shown, where the subband T60 of .8khz band (4th subband) is miscalculated to almost 4s
    - So the models result with higher mean absolute error of 1.18 and 1.06 respectively



DNN Model for Fullband T60 Using the Subband Estimates

Test Results for Clean Dataset:

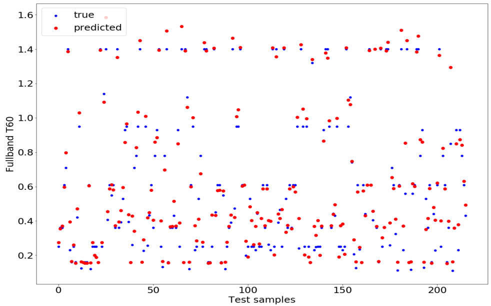
- 109 test samples with no outliers and DCT filterbank used for subband decomposition
- MSE $= 0.0048$



DNN Model for Fullband T60 Using the Subband Estimates

▶ Test Results for DCT Full Dataset:

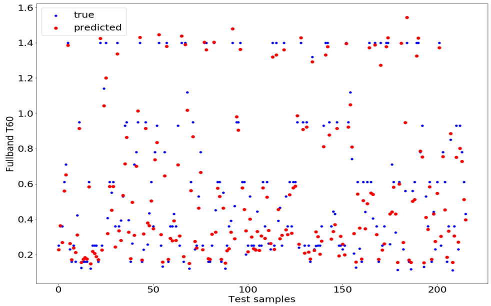
- 216 test samples including outliers and DCT filterbank used for subband decomposition
- MSE $= 0.0027$



DNN Model for Fullband T60 Using the Subband Estimates

▶ Test Results for Octave Full Dataset:

- 216 test samples including outliers and octave filterbank used for subband decomposition
- MSE $= 0.0062$



ACE Challenge Results

T_{60} ESTIMATION ALGORITHM PERFORMANCE IN ALL NOISES FOR ALL SNRS

Ref.	Algorithm	Class	Mic. Config.	Bias	MSE	ρ	RTF
A	QA Reverb [48]	SFM	Single	-0.068	0.0648	0.778	0.4
B	Octave SB-based FB RTE [49]	ABC	Single	-0.104	0.0731	0.738	0.939
C	DCT-based FB RTE [49]	ABC	Single	-0.104	0.0766	0.71	1
D	Model-based SB RTE [49]	ABC	Single	-0.0196	0.0981	0.661	0.451
E	Baseline algorithm for FB RTE [49]	ABC	Single	-0.0432	0.11	0.387	0.0424
F	SDDSA-G retrained [50]	SFM	Single	0.0167	0.0937	0.608	0.0152
G	SDDSA-G [19]	SFM	Single	-0.0423	0.0803	0.6	0.0164
H	Multi-layer perceptron [51]	MLMF	Single	-0.0967	0.104	0.48	0.0578‡
I	Per acoust. band SRMR Section 2.5. [52]	SFM	Single	-0.114	0.109	0.48	0.578
J	NSRMR Section 2.4. [52], [53]	SFM	Single	-0.0646	0.119	0.261	0.571
K	NSRMR Section 2.4. [52], [53]	SFM	Chromebook	0.012	0.116	0.291	1.04
L	NSRMR Section 2.4. [52], [53]	SFM	Mobile	-0.0504	0.0958	0.281	1.58
M	NSRMR Section 2.4. [52], [53]	SFM	Crucif	-0.0516	0.107	0.246	2.62
N	SRMR Section 2.3. [52]	SFM	Single	-0.16	0.144	0.22	0.457
O	SRMR Section 2.3. [52]	SFM	Chromebook	-0.105	0.132	0.221	0.829
P	SRMR Section 2.3. [52]	SFM	Mobile	-0.153	0.12	0.228	1.26
Q	SRMR Section 2.3. [52]	SFM	Crucif	-0.153	0.128	0.225	2.09
R	NIRAv3 [47]	MLMF	Single	-0.192	0.151	0.302	0.899†
S	NIRAv1 [47]	MLMF	Single	-0.184	0.151	0.258	0.899†
T	NIRAv2 [47]	MLMF	Single	-0.179	0.198	-0.0199	0.907†
U	Blur kernel [54]	SFM	Single	0.173	0.15	0.279	8.46
V	Blur kernel with sliding window [55]	SFM	Single	-0.00555	0.139	0.12	0.421
W	Temporal dynamics [18]	SFM	Single	-0.304	0.211	0.269	0.362
X	Improved blind RTE [17]	ABC	Single	-0.0635	0.165	0.166	0.0259
Y	SDD [16]	SFM	Single	0.463	305	0.00158	0.0221