ConferencingSpeech 2022 Challenge Ranking Announcement

In this challenge, we calculated root mean squared error (RMSE) and RMSE_MAP for evaluating the accuracy and the pearson correlation coefficient (PCC) for measuring linearity. For more details:

1. PCC: measures the linear relationship between a model's estimation and the subjective data. It can be obtained through:

$$PCC = \frac{\sum_{i=1}^{N}(Xi - \overline{X}) \times (Yi - \overline{Y})}{\sqrt{\sum(Xi - \overline{X})^2} \times \sqrt{\sum(Yi - \overline{Y})^2}}$$

Where Xi denotes the subjective score MOS and Yi the objective one (MOSp). N represents the total number of speech samples considered in the analysis.

2. root mean squared error (RMSE): evaluate the accuracy between a model's estimation and the subjective data. It can be obtained through:

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}-1}\sum_{\mathrm{i}=1}^{\mathrm{N}}(Xi - \mathrm{Yi})^2}$$

3. RMSE_MAP: due to bias or offsets, different gradients and different qualitative rank orders are always present in subjective evaluations. The statistical uncertainty always exists in the collected MOS. Therefore, a third-order polynomial function is applied to compensate for the possible variance between several subjective experiments in this challenge. For each model, we create one mapping function per test dataset. Then the mapped predictions were used to calculated RMSE_MAP.

The final ranking is shown in the following table, which is determined by RMSE_MAP

Ranking	Team	PCC	RMSE	RMSE_MAP
1	LCN	0.812	0.344	0.298
2	kuaishou	0.797	0.436	0.31
3	ks1024	0.782	0.464	0.324
4	UNISOC_audio	0.792	0.458	0.329
5	ByteDance	0.781	0.474	0.332
6	Speech_Stars	0.746	0.518	0.339
7	BIT_MI	0.757	0.486	0.342
8	TeamWHU	0.77	0.58	0.345
9	DTAQ	0.741	0.427	0.356
10	YXAudiolab	0.743	0.49	0.357
11	NTUEE SPML Lab	0.798	0.502	0.36
	Baseline2	0.724	0.543	0.374
12	CCAT	0.714	0.548	0.378
13	cto-logi	0.692	0.593	0.393
14	MIR@NTU	0.7	0.454	0.4

15	xiaoban	0.699	0.547	0.401
16	SDSound	0.616	0.573	0.403
17	xaddwell	0.501	0.691	0.450
18	chair	0.581	0.592	0.47
	Baseline1	0.551	0.745	0.475

For more details, please follow our subsequent article on arxiv. (The specific link will be notified by email later.)