

Experiments

Misalignment of VioFixer

Models	SiSNR	LSD	SSIM	PESQ	SiSPNR	MOS
Unprocessed	4.44	2.00	0.64	1.94	7.20	2.38
Oracle	-22.31	0.91	0.74	2.52	11.73	3.74
Target	128.00	0.01	1.00	4.64	110.55	3.95



Although

NOS

Score

O

f

Oracle

and

Target

are

close,

onrade

SisNR



significantly

Lower.



Beaurs

O

f

the

discriminative

training

and

multiple

l

o

s

s

e

s

used



Wonder.



The

misalignment



usually

shout

1

~

3

0

0

samples,

still

with in

the

10ms

winndown.

Only frequency domain metrics can be used!

Table.8 Metrics on GSR Evaluation Set, ALL-GSR

Although *MO Score* and *Target* are close, *SiNR* is significantly lower.

• Because the discriminative training and multiple loss used in Vocoder

• The mis-alignment is usually about 1~300 samples, still within the 10ms window

Experiments

Mis-alignment of VoiceFixer

Table.8 Metrics on GSR evaluation set, ALL-GSR

Models	SiSNR	LSD	SSIM	PESQ	SiSPNR	MOS
Unprocessed	4.44	2.00	0.64	1.94	7.20	2.38
Oracle	-22.31	0.91	0.74	2.52	11.73	3.74
Target	128.00	0.01	1.00	4.64	110.55	3.95

- Although MOS Score of *Oracle* and *Target* are close, *oracle* SiSNR is significantly lower.
 - Because of the discriminative training and multiple losses used in Vocoder.
 - The mis-alignment is usually about 1~300 samples, still within the 10ms window.

Only frequency domain metrics can be used!

Experiments

Overview

