

Voicexer

TFGAN Vocoder - Training - Time Domain Losses



Losses

Function:

$$\boldsymbol{L}_{syn} = \boldsymbol{L}^T + \boldsymbol{L}^F + \lambda_1 \boldsymbol{L}^D$$



Times

Donna

Losses:



$$L^T = \sum_k L_k^t$$



$$L_k^t(\hat{s}, s) = \lambda_5 L_k^{energy}(\hat{s}, s) + \lambda_6 L_k^{phase}(\hat{s}, s) + \lambda_7 L_k^{time}(\hat{s}, s)$$

k	1	2	3	4
frame-length	1	240	480	960
hop-length	1	120	240	480

Table.3 Window parameter for each k

Dommarin

Training

TFGAN

Time





I

o

s

s

e

s

Wonder



Times



LOSSS

$$L_{syn} = L^T + L^F + \lambda_1 L^D$$

$$L^T = \sum_k L_k^t$$

$$L_k^t(\hat{s}, s) = \lambda_5 L_k^{energy}(\hat{s}, s) + \lambda_6 L_k^{phase}(\hat{s}, s) + \lambda_7 L_k^{time}(\hat{s}, s)$$

enervoy

time

sample:



IOS: ■■

function

Losses,



Tinne

effect:

o

L

o

s

s

netairic

or **Renov**es

nn

es

er

nn

and

I

O

S

S

enquiry

o Captures



t

n

e

phases



information station

$$L^{time}(\hat{s}, s) = \left\| v(\hat{s}) - v(s) \right\|_1$$

$$L_{energy}(\hat{s}, s) = \left\| v(\hat{s}_w^2) - v(s_w^2) \right\|_1$$

$$L^{phase}(\hat{s}, s) = \left\| \Delta v(\hat{s}_w^2) - \Delta v(s_w^2) \right\|_1,$$

$$v(s)_{1 \times w} = (n(s_0), n(s_1), \dots, n(s_w))$$

m(.)

paranmeter

earcn

Windowing

Table. 3

for





Trine

Donnair

Functorion.

Losses:



LOSS