Automatic DJ Transitions with Differentiable Audio Effects and Generative Adversarial Networks

Bo-Yu Chen[1], Wei-Han Hsu[1], Wei-Hsiang Liao[2],
Marco A. Martínez-Ramírez[2], Yuki Mitsufuji[2] and Yi-Hsuan Yang[1]

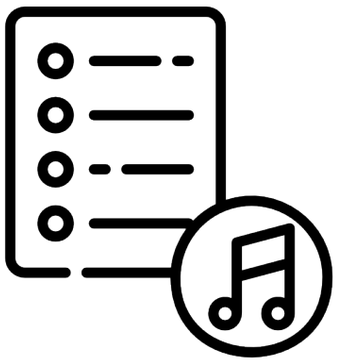
[1] Academia Sinica, Taiwan, [2] Sony Group Corporation, Japan

DJ

Auto DJ (1/3)

Auto DJ (2/3)

Track Selection



Track Ordering

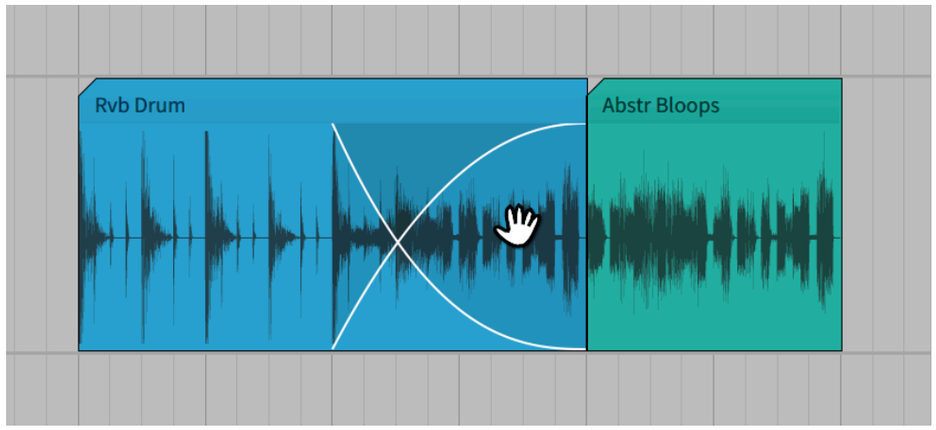
Mixing

Auto DJ (3/3)


Equalizer
Control the frequency property of the track
Fader
Control the volume of the track

How to do Auto DJ Mixing ?

Rule based Approach



CrossFade

Limited EQ Settings & Pattern

[1] Len Vande Veire and Tijl De Bie, “From raw audio to a seamless mix: Creating an automated DJ system for drum and bass,” EURASIP Journal on Audio, Speech, and Music Processing, vol. 2018, no. 1, pp. 1–21, 2018
[2] Rachel M. Bittner et al., “Automatic playlist sequencing and transitions,” in Proc. International Society for Music Information Retrieval Conference (ISMIR), 2017.

Gap

No golden rule in DJing

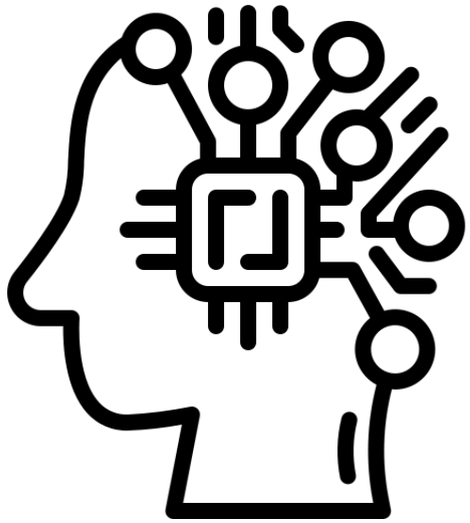
Rule Based ⇢ Data Driven

Challenge

Dataset

DJ mixes

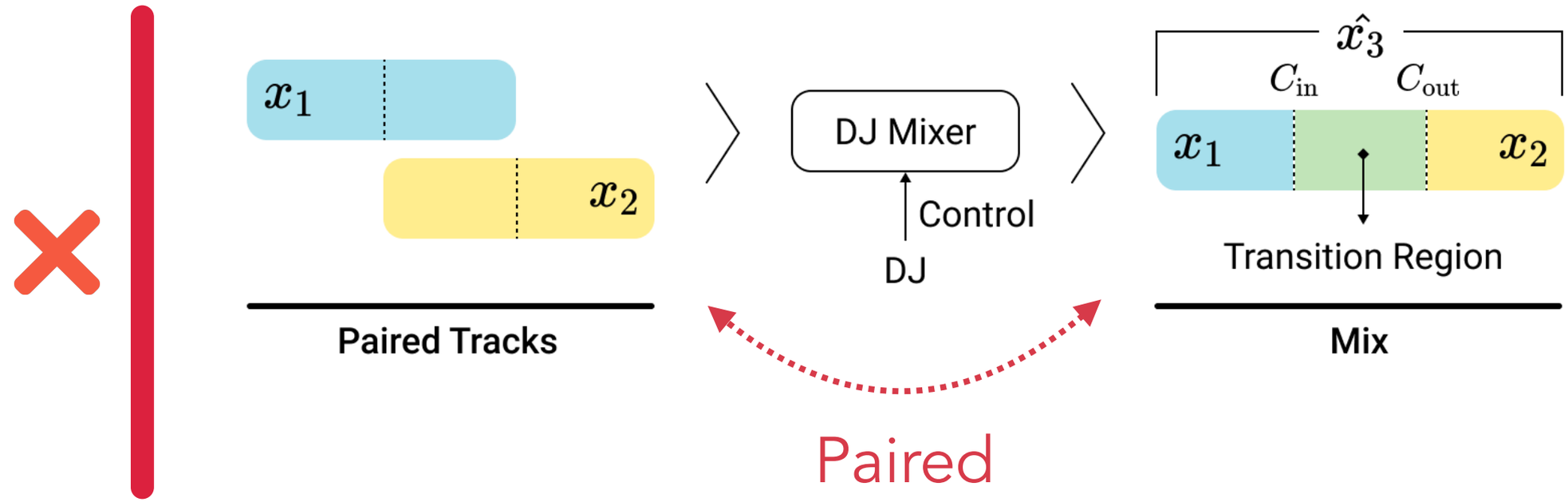
Model

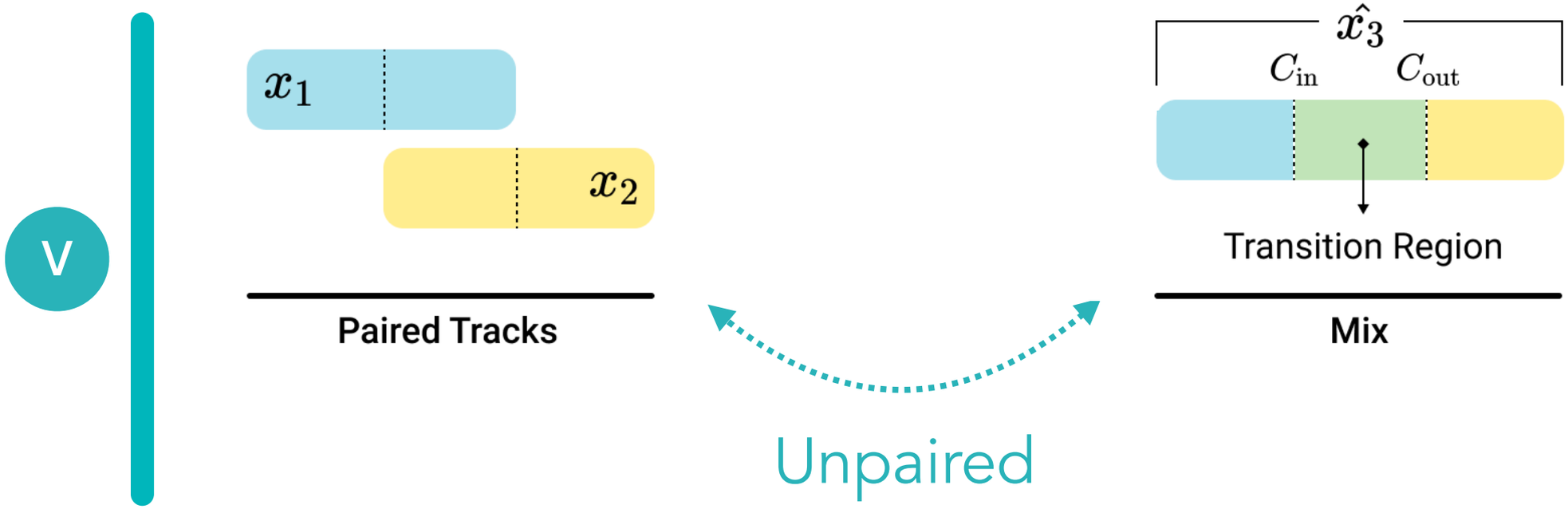


Mixing prior

How to prepare data ?

Data Requirement





Real DJ Mix

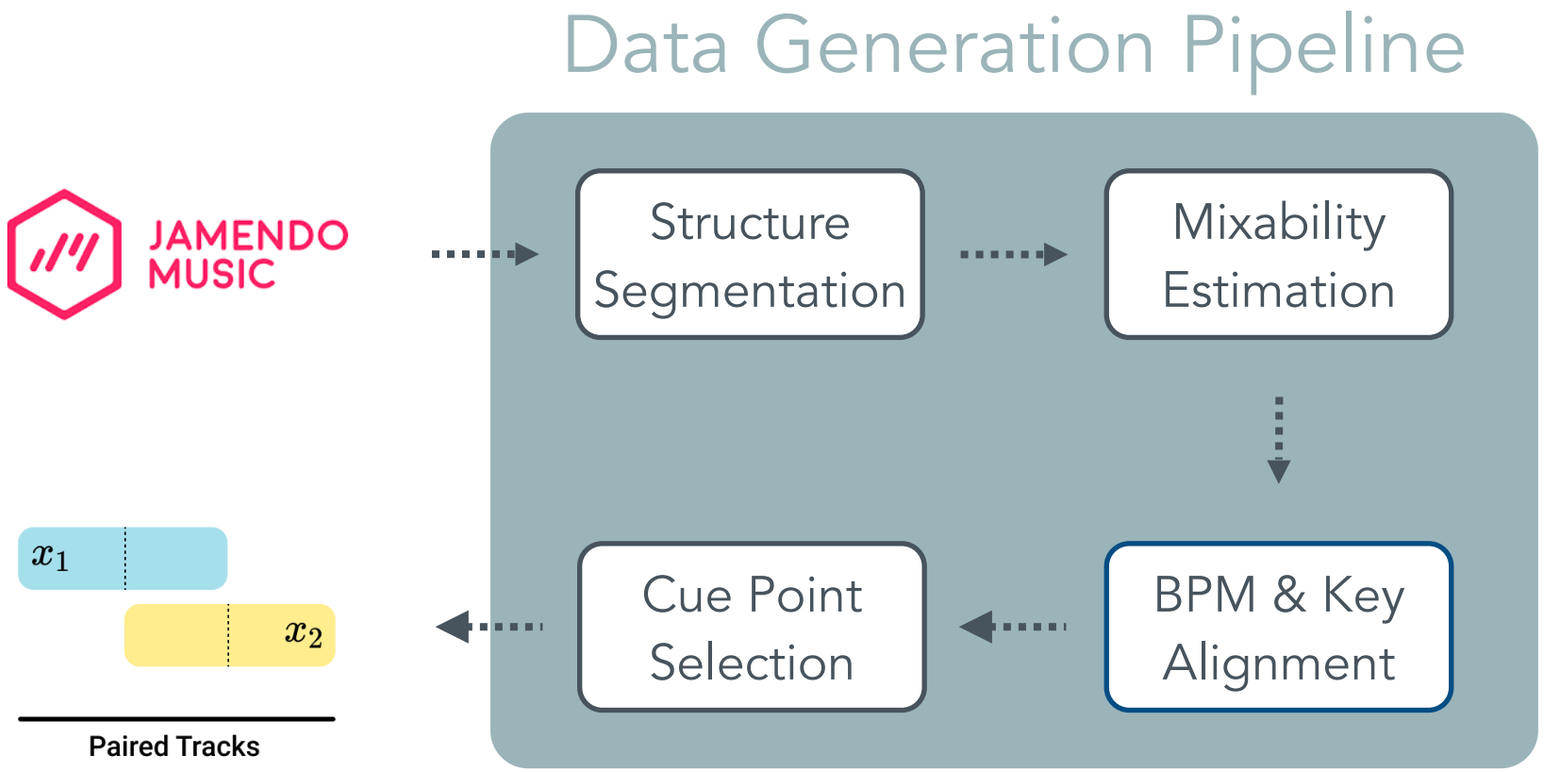
DJ Mixes

Timestamps

7,064 DJ mix sections totally

[1] "Livetracklist," https://www.livetracklist.com/

Paired tracks



9,148 paired tracks totally

[1] Len Vande Veire and Tijl De Bie, "From raw audio to a seamless mix: Creating an automated DJ system for drum and bass," EURASIP Journal on Audio, Speech, and Music Processing, vol. 2018, no. 1, pp. 1–21, 2018.
[2] Dmitry Bogdanov et al., "The MTG-Jamendo dataset for automatic music tagging," in Proc. Machine Learning for Music Discovery Workshop, International Conference on Machine Learning (ICML), 2019.

How to model DJ transition ?

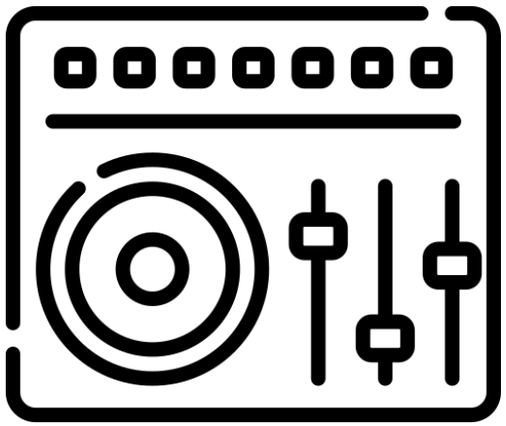
Idea

GAN

DJ

Unpaired data

DDSP



DJ Mixer

Mixing Prior

[1] Jesse Engel et al., "DDSP: Differentiable digital signal processing," in Proc. International Conference on Learning Representations (ICLR), 2020.

[2] Ian Goodfellow et al., "Generative adversarial nets," Proc. Advances in neural information processing systems, vol. 63, no. 11, pp. 139–144, 2014.

DDSP (1/4)

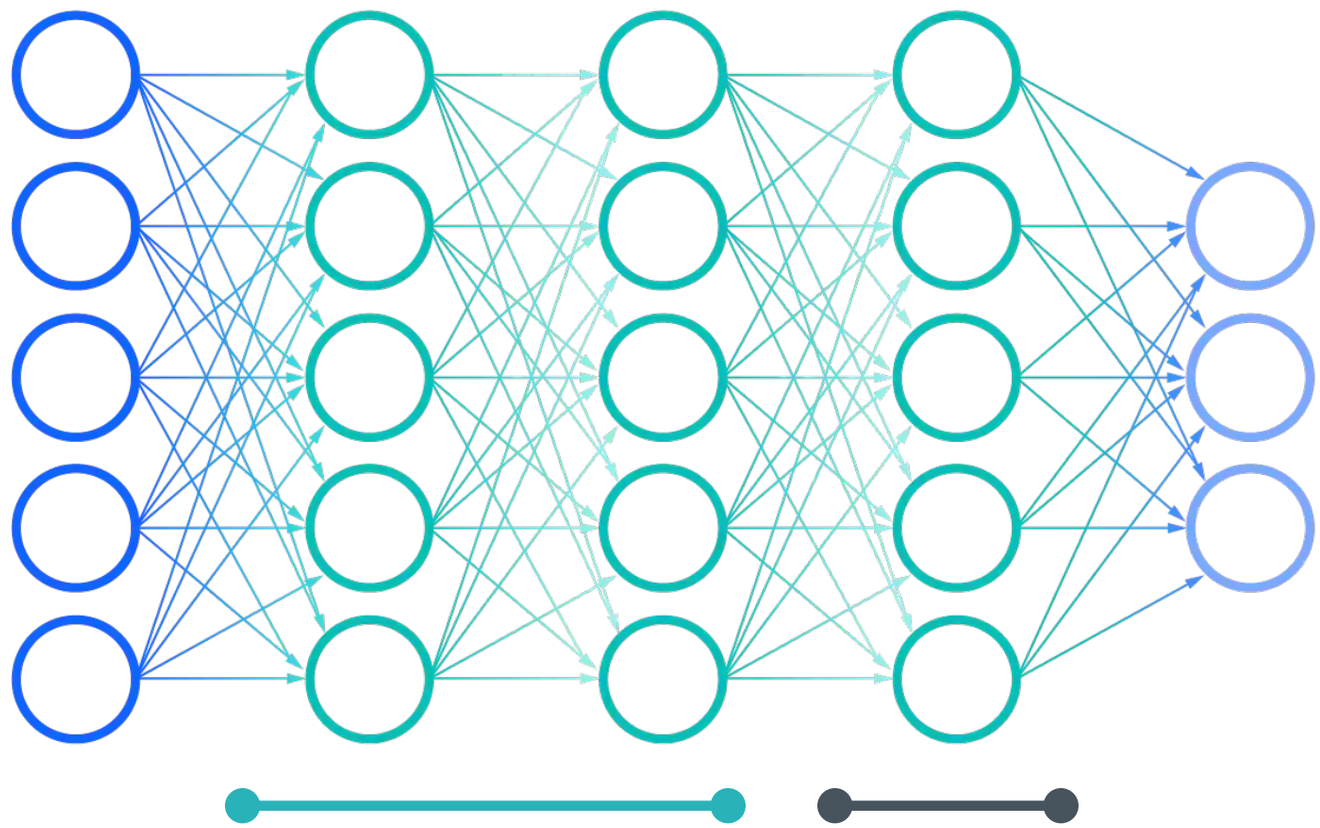
Trainable		Pre-defined
1D CNN without activation function	→	FIR Filter
RNN without activation function	→	IIR Filter

↓

Differentiable

DDSP (2/4)

Transfer Learning

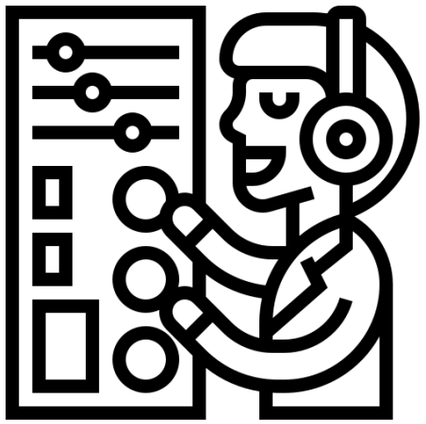


Freeze

FIR or IIR Filter

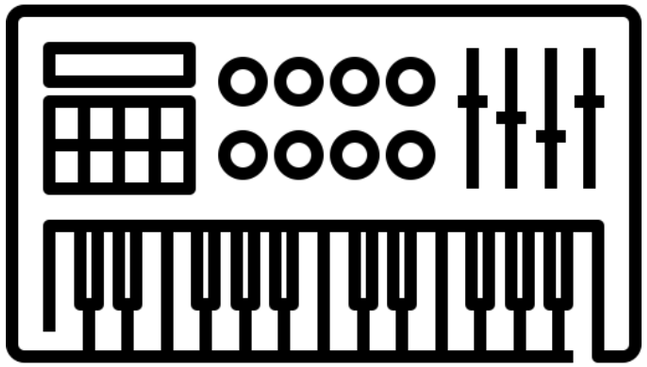
(non-trainable -> only provide gradient)

DDSP (3/4)



Producer

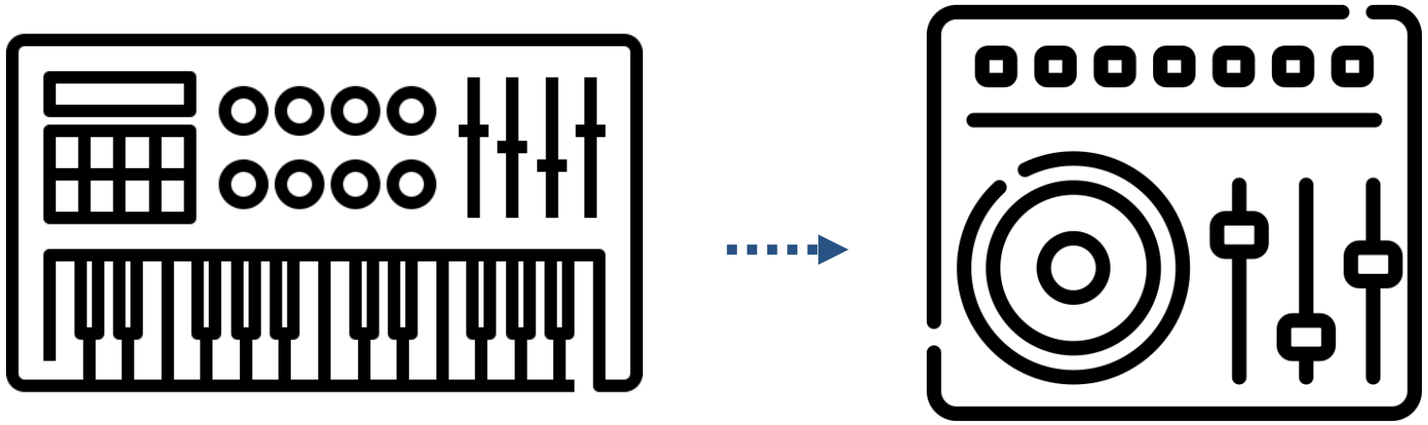
Trainable



Synthesizer

Differentiable

DDSP (4/4)

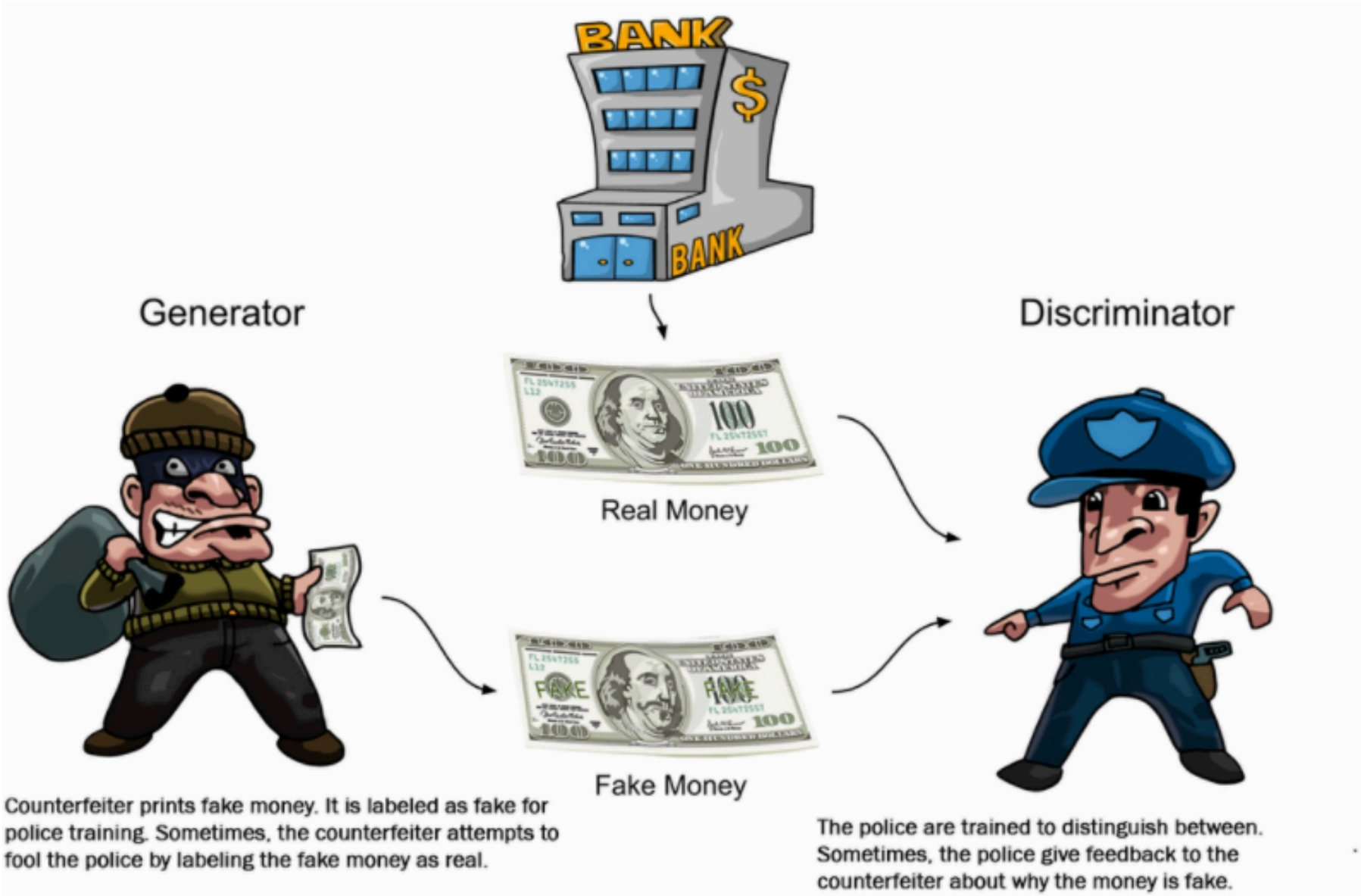


Synthesizer

DJ Mixer

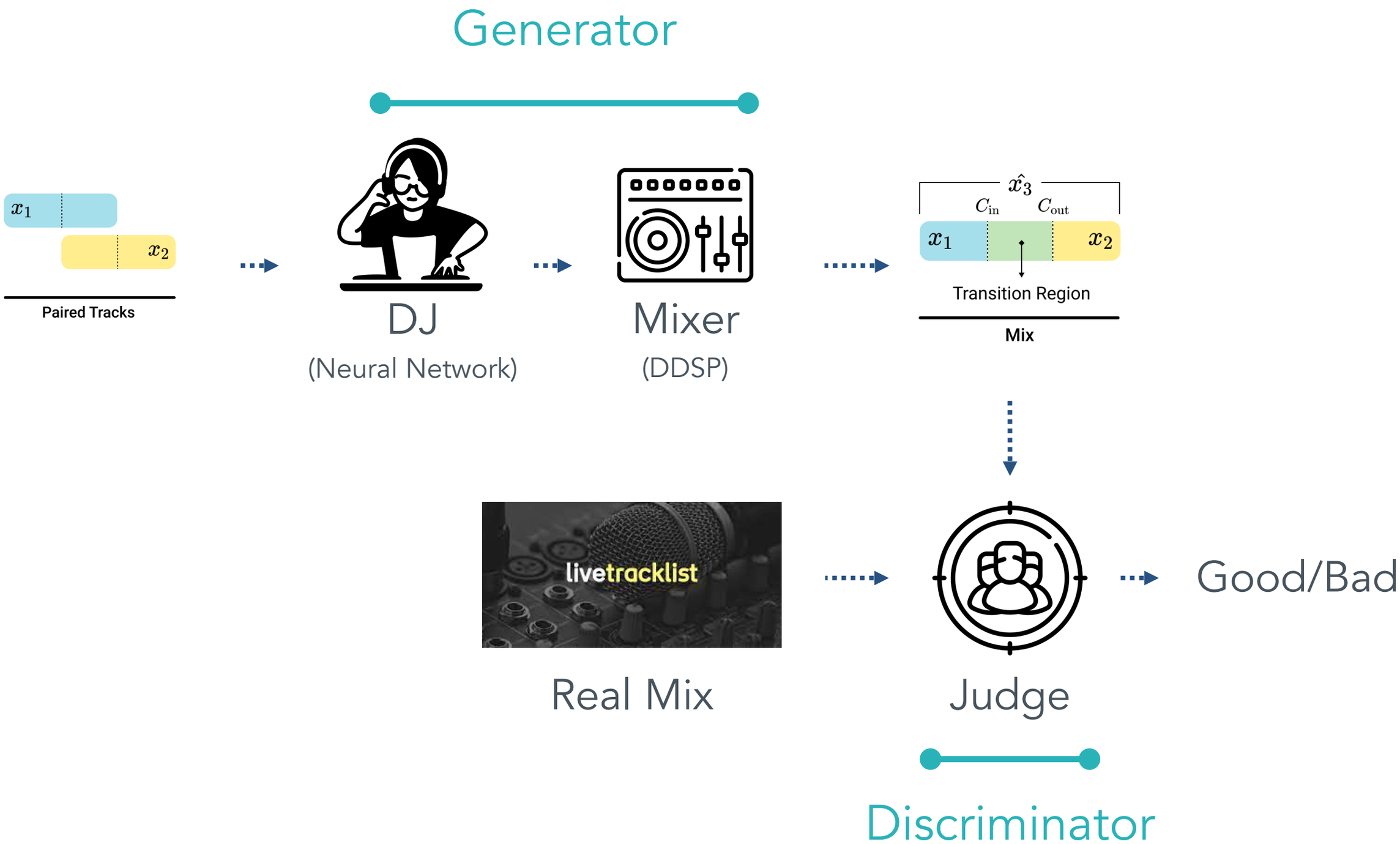
[1] Jesse Engel et al., "DDSP: Differentiable digital signal processing," in Proc. International Conference on Learning Representations (ICLR), 2020.

GAN (1/2)



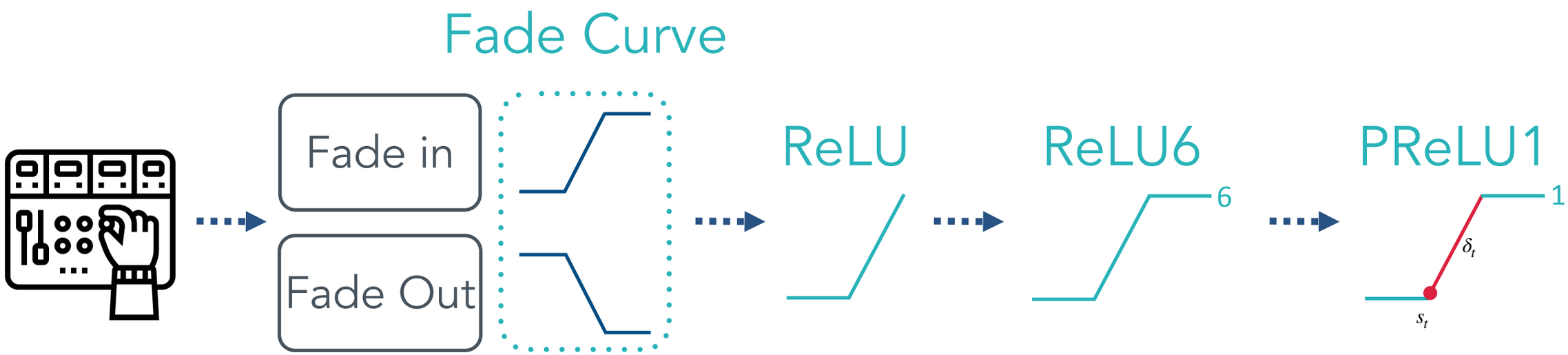
[1] Ian Goodfellow et al., "Generative adversarial nets," Proc. Advances in neural information processing systems, vol. 63, no. 11, pp. 139–144, 2014.

GAN (2/2)



[1] Ian Goodfellow et al., "Generative adversarial nets," Proc. Advances in neural information processing systems, vol. 63, no. 11, pp. 139–144, 2014.

Differentiable Fader



$$PReLU1(\vec{v_t}, s_t, \delta_t) = \min(\max(0, \min(\max(0, \vec{v_t} - s_t), 1) * \delta_t), 1)$$

[1] Andrew G Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.

Differentiable Equalizer (1/2)

Time Domain

FIR Filter

IIR Filter

Frequency Domain

Frequency Sampling

⋮

Computational Inefficient

[1] Boris Kuznetsov, Julian D Parker, and Fabian Esqueda, "Differentiable iir filters for machine learning applications," in Proc. Int. Conf. Digital Audio Effects, 2020.

[2] Jesse Engel et al., "DDSP: Differentiable digital signal processing," in Proc. International Conference on Learning Representations (ICLR), 2020.

[3] Shahan Nercessian, "Neural parametric equalizer matching using differentiable biquads," in Proc. Int. Conf. Digital Audio Effects (eDAFx-20), 2020, pp. 265–272.

Differentiable Equalizer (2/2)

Time-Frequency Domain

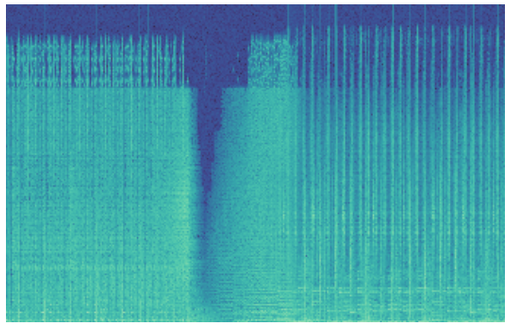
PReLU1 in f axis

=

H_{lpf}

Low Pass Filter

X



Spectrogram

$$H_{lpf}(\vec{v_f}, \theta_f) = 1 - PReLU1(\vec{v_f}, s_f, \delta_f)$$

[1] Haohe Liu et al., "Channel-wise subband input for better voice and accompaniment separation on high resolution music," in Proc. Interspeech 2020, 2020, pp. 1241–1245.

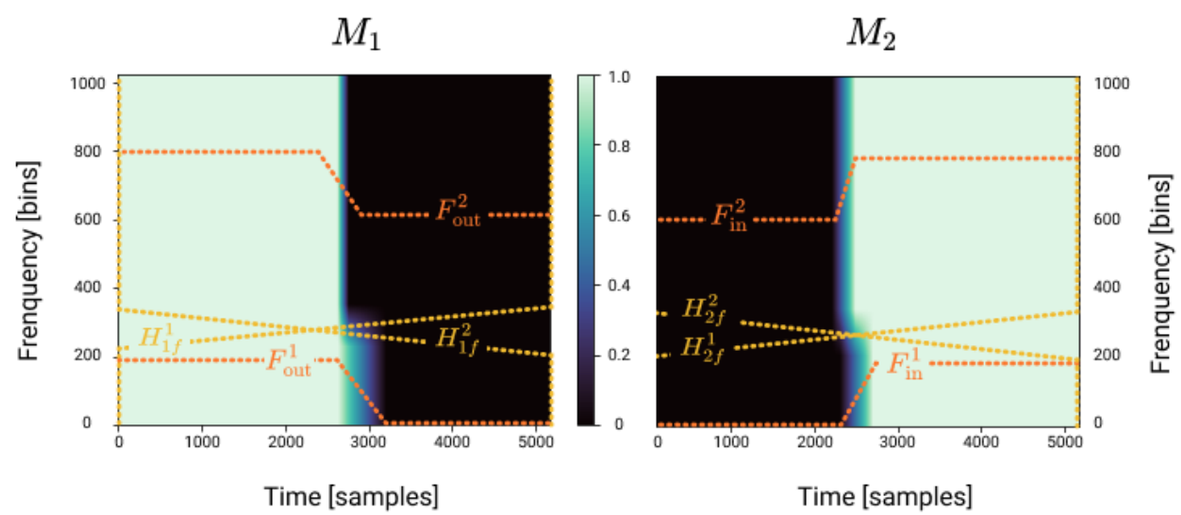
Differentiable DJ Mixer

Differentiable Mixer

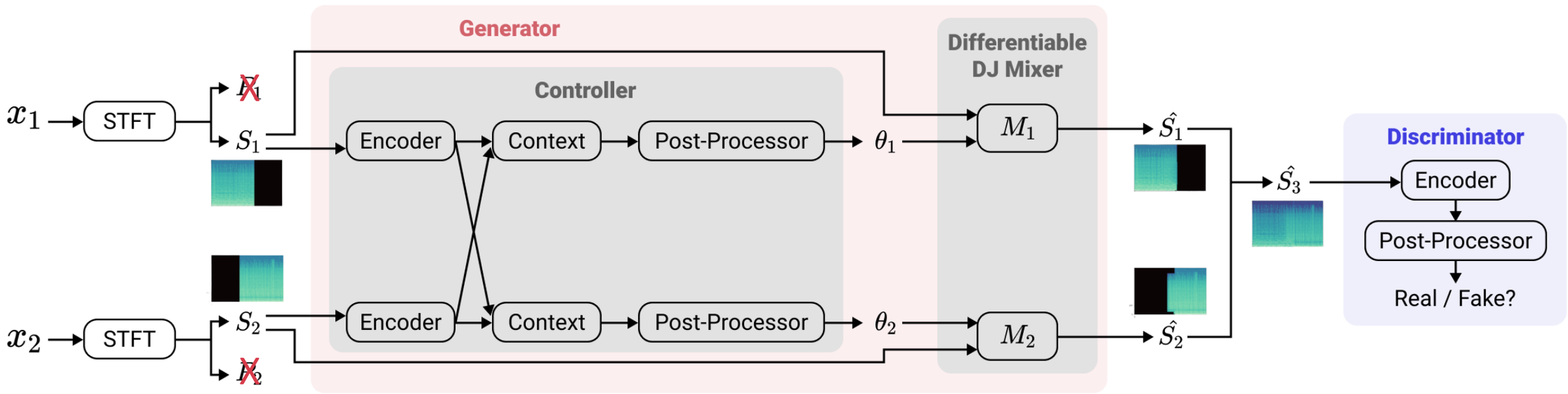
Differentiable Fader

Differentiable Equalizer

PReLU in both t and f axis

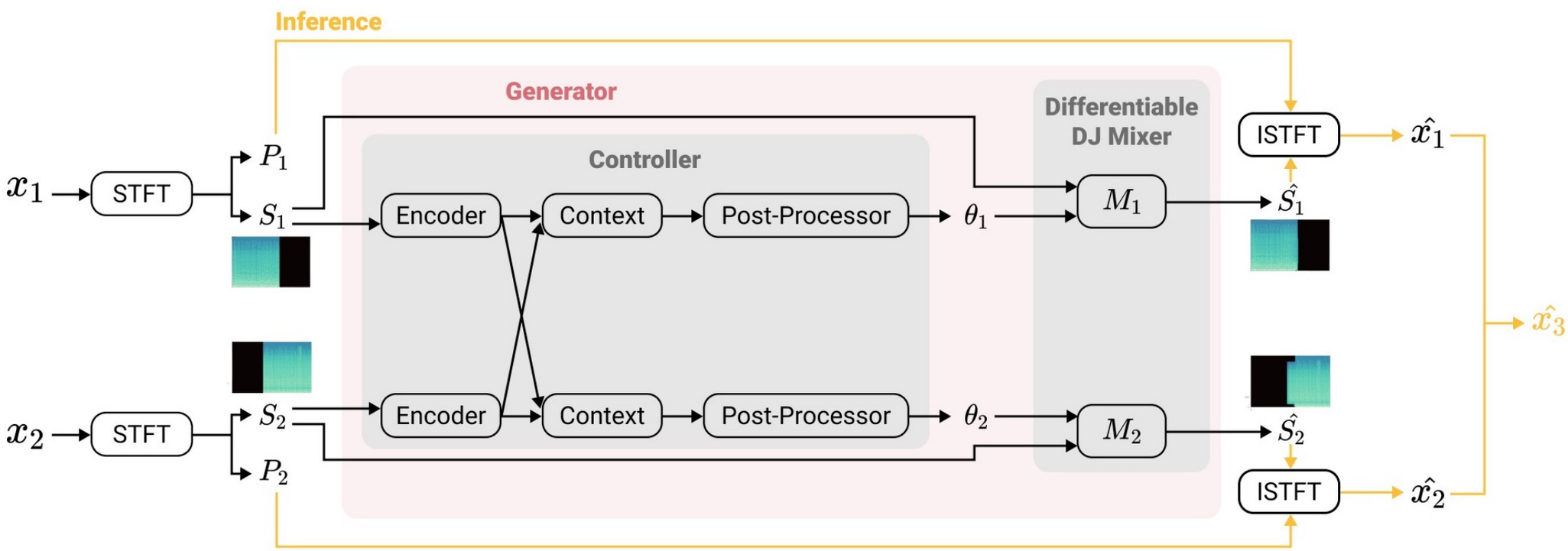


Model Architecture



[1] Christian J Steinmetz et al., "Automatic multitrack mixing with a differentiable mixing console of neural audio effects," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, 2021.

Inference


Inference
Generator
Controller
Differentiable DJ Mixer
x_1
STFT
P_1
S_1
Encoder
Context
Post-Processor
θ_1
M_1
$\hat{S}_1$
ISTFT
$\hat{x}_1$
$\hat{x}_3$
x_2
STFT
S_2
P_2
Encoder
Context
Post-Processor
θ_2
M_2
$\hat{S}_2$
ISTFT
$\hat{x}_2$

Evaluation

Model Comparison

Sum

Rule

Linear

Human

Baselines

GAN

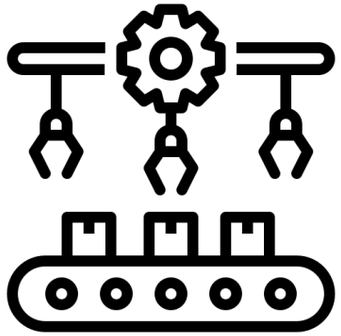
Our Approach

Sum

$$\sum$$

Sum

Rule

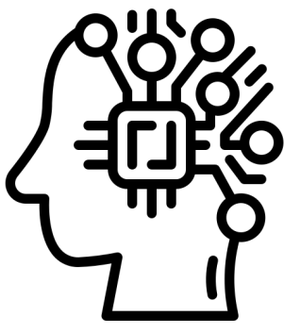


Rule

Linear

Linear

GAN Our Approach



GAN

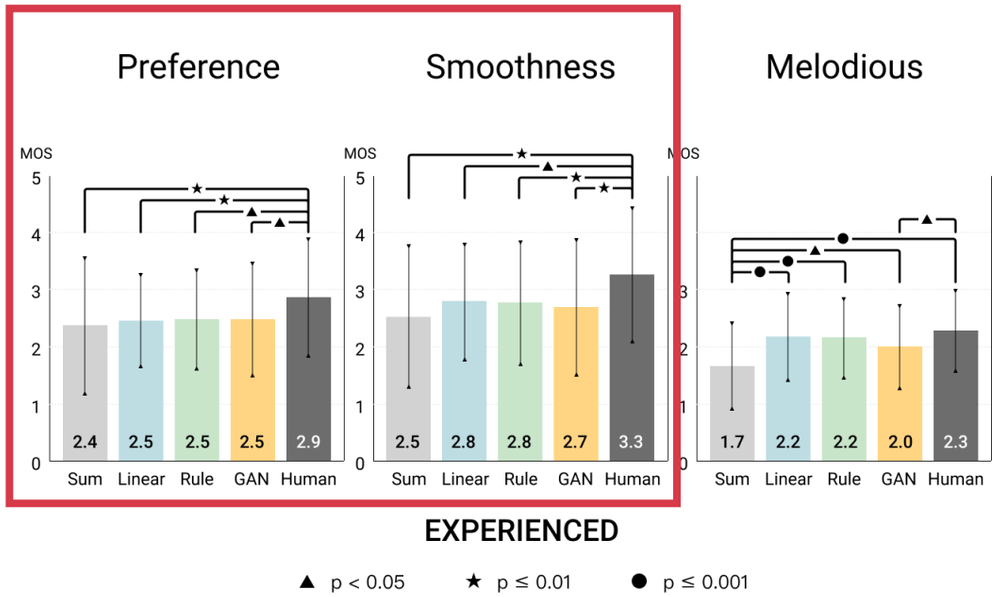
Our Approach

Human

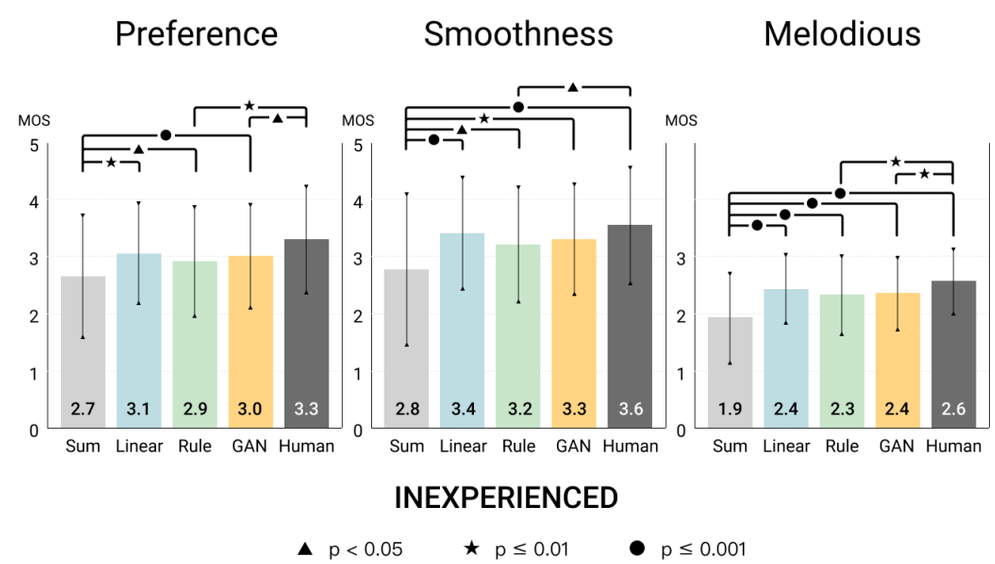
Human

Listening test results (1/2)

Most scores are lower than 3

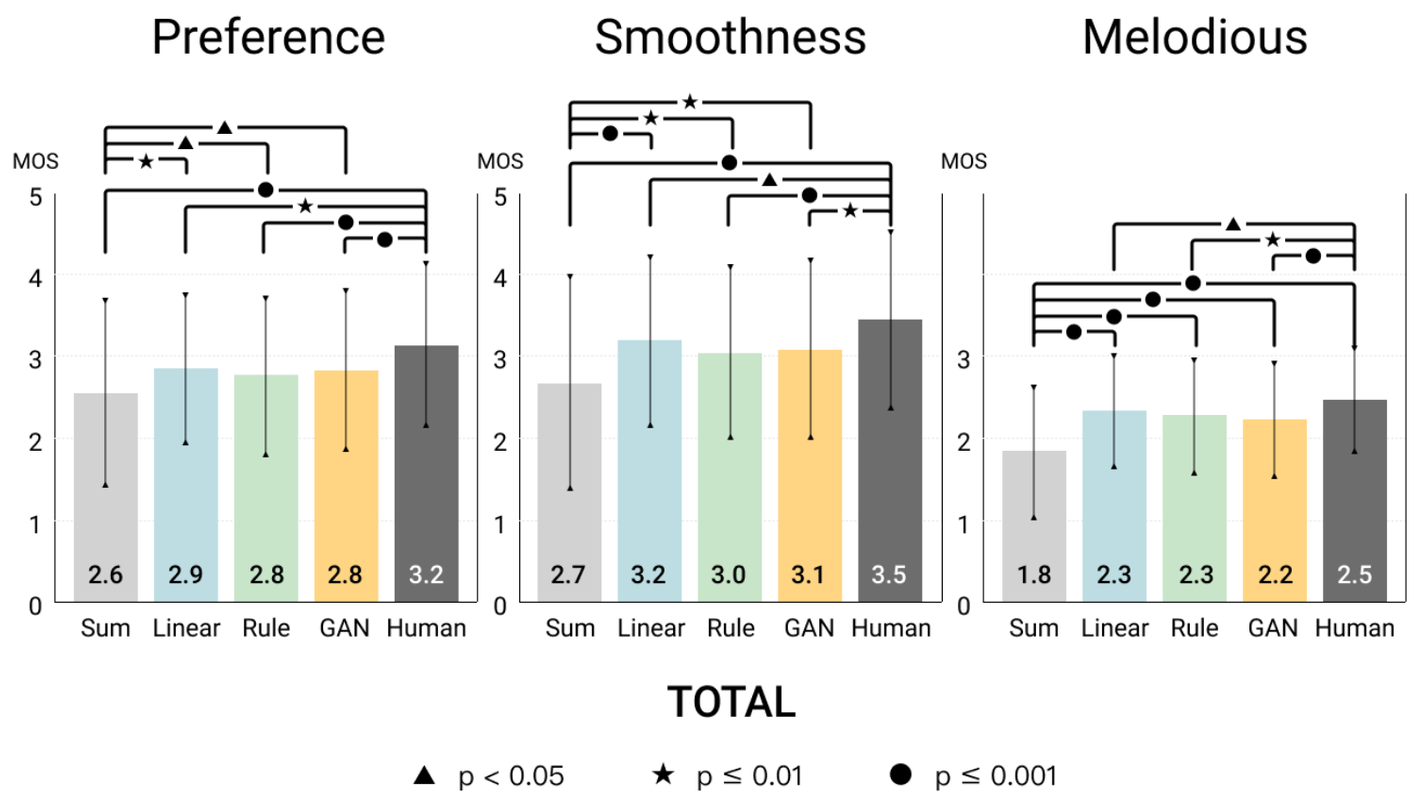


<



Music style of the selected paired tracks is unacceptable to experienced subjects

Listening test results (2/2)



Overall, there is no significant difference among GAN, Linear and Rule.

Comments from experienced subjects

“

They make good use of the filter in the first track and fit with the second track, which takes on an altogether new aspect.

“

This DJ seems to make the transition by vinyl records instead of a modern CDJ or DDJ controller. It sounds like a human playing a vinyl record live, making it organic.

Contributions

Contributions

1. Developed a data generation pipeline to create training data.
2. Propose efficient differentiable DJ mixer in time-frequency domain.
3. Utilize generative adversarial network and differentiable DJ mixer to learn how to generate DJ transition.
4. Conduct a subjective listening test to show that our approach achieve competitive results compared with a number of baselines.

Demo & Code

https://paulyuchen.com/djtransgan-icassp2022

https://github.com/ChenPaulYu/DJtransGAN

Thanks for your Attention

bernie40916@gmail.com @Chen_Paul_u ChenPaulYu