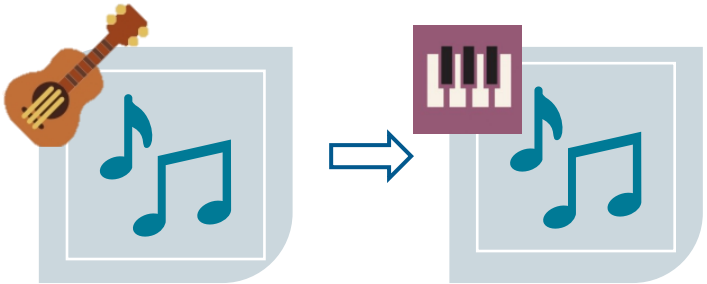
Unsupervised Musical Timbre Transfer for Notification Sounds



Jing Yang[1], Tristan Cinquin[1], Gábor Sörös[2]

[1]Department of Computer Science, ETH Zurich, Switzerland
[2]Nokia Bell Labs, Budapest, Hungary

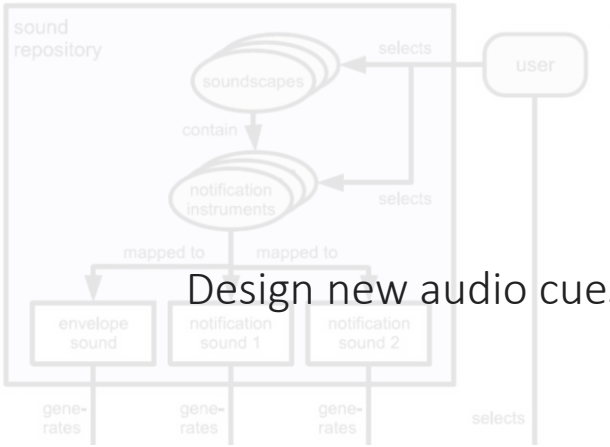
IEEE ICASSP 2021

ETH zürich

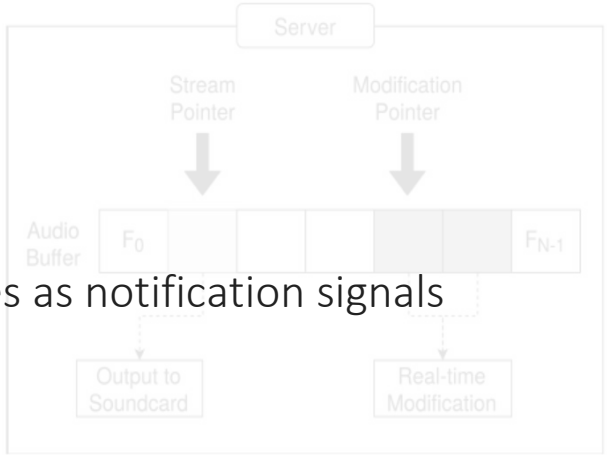
NOKIA Bell Labs

Useful but distracting notifications

Deliver **auditory** notifications in a less intrusive way

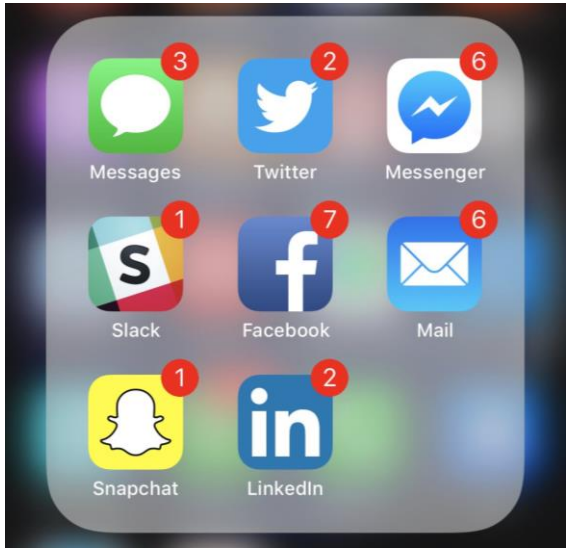


[2008] Jung – Ambience for Auditory Displays Embedded Musical Instruments as Peripheral Audio Cues



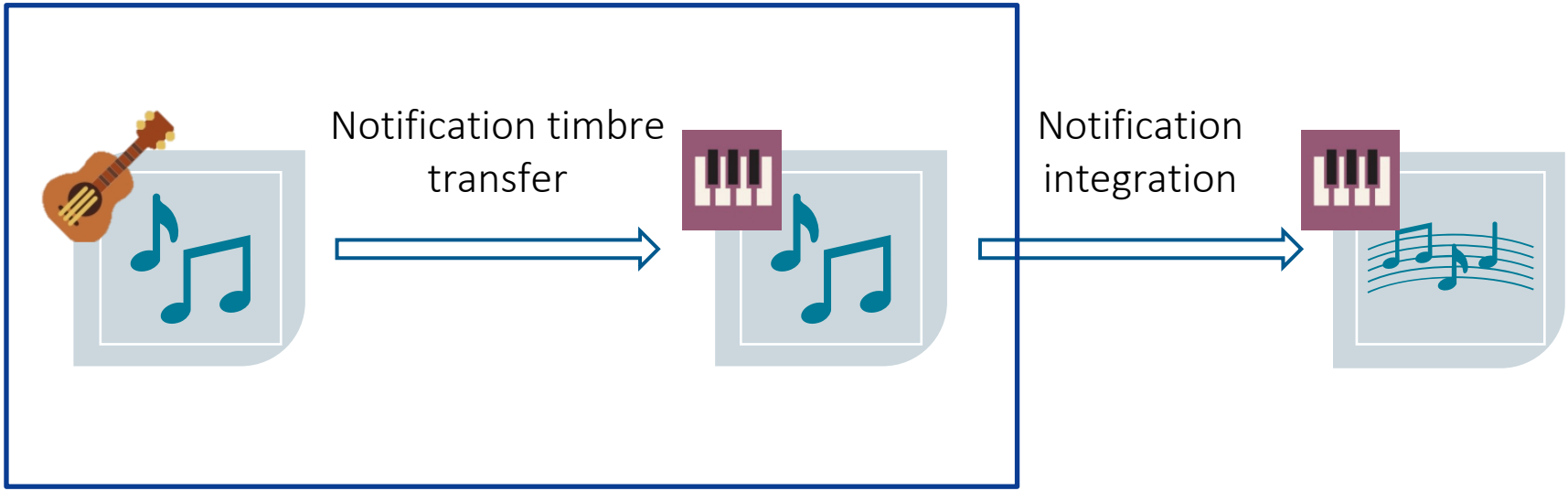
[2019] Ananthabhotla and Paradiso – SoundSignaling: Realtime, Stylistic Modification of a Personal Music Corpus for Information Delivery

Design new audio cues as notification signals

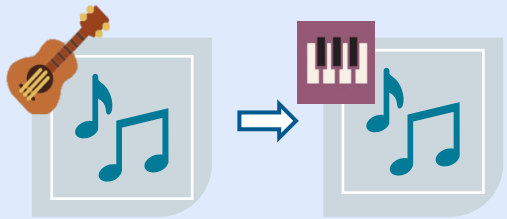


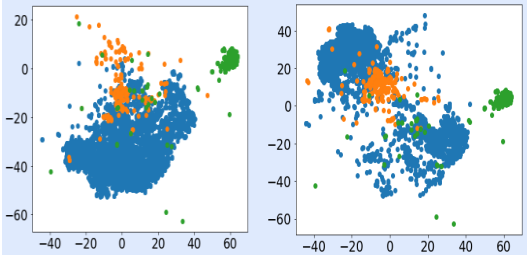
Deliver commonly used notification sounds less intrusively by developing audio style transfer techniques.

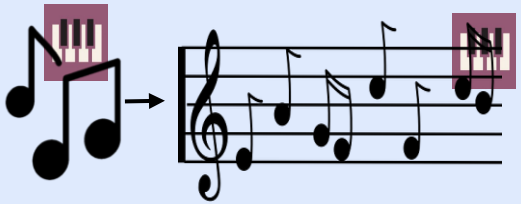
Our concept



Our contributions







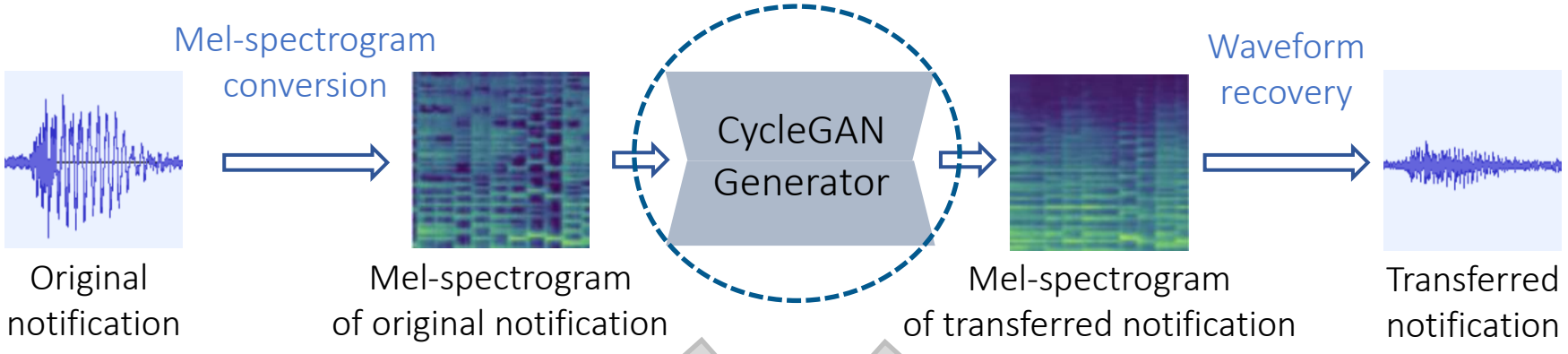
- Musical timbre transfer for artificial notification sounds
- Objective and subjective evaluations of our notification timbre transfer method
- Application of less intrusive notification delivery

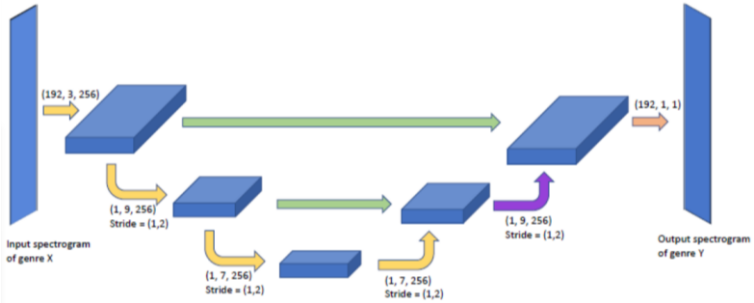
Major challenges of notification timbre transfer

- Insufficient training data of notification sounds
- Complete lack of pairs of notifications and their corresponding music sequences

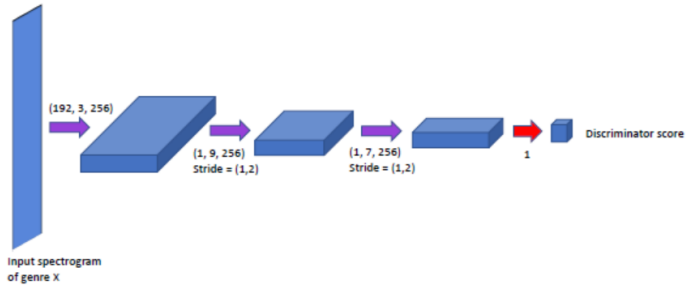
Ambiguous timbre of notification sounds

Our method for notification timbre transfer





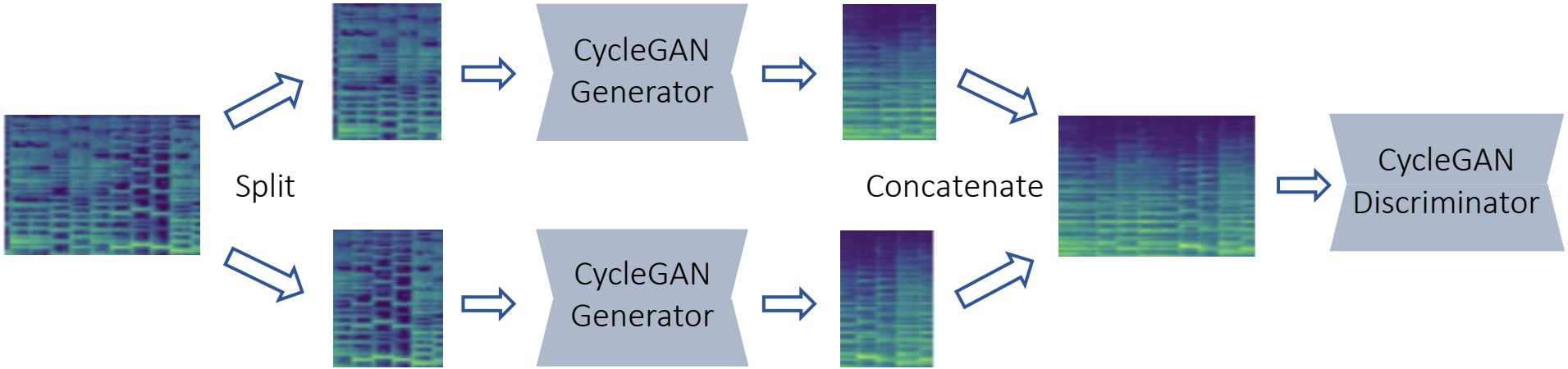
Generator: fully convolutional U-Net architecture



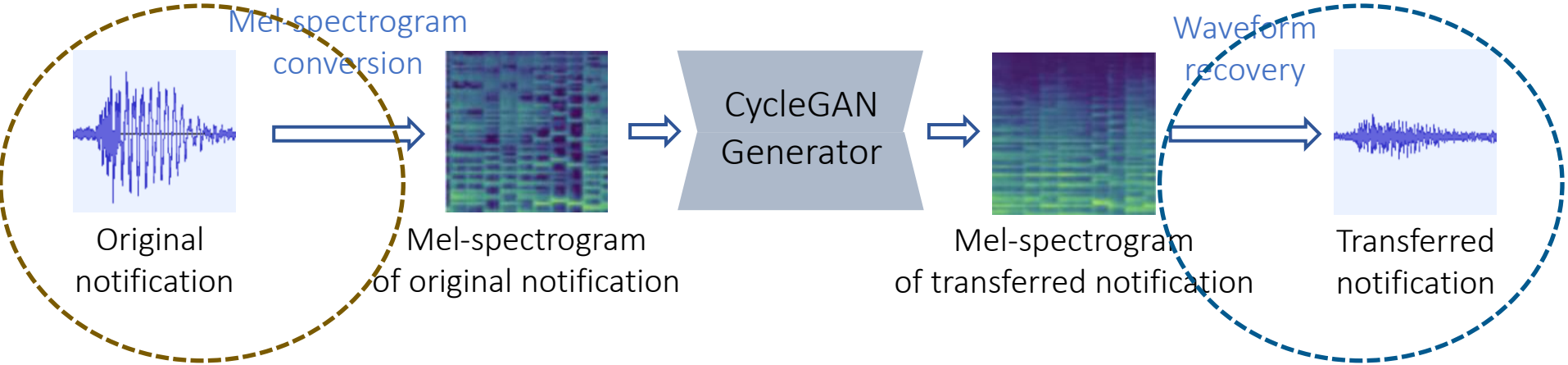
Discriminator: convolutional PatchGAN discriminator

Splitting-and-concatenation mechanism for arbitrary notification length

[2019] Pasini – MelGAN-VC: Voice Conversion and Audio Style Transfer on Arbitrarily Long Samples using Spectrograms



Training data



Input domain (?):

- Sufficient training data
- Similar timbre features as the notification sounds

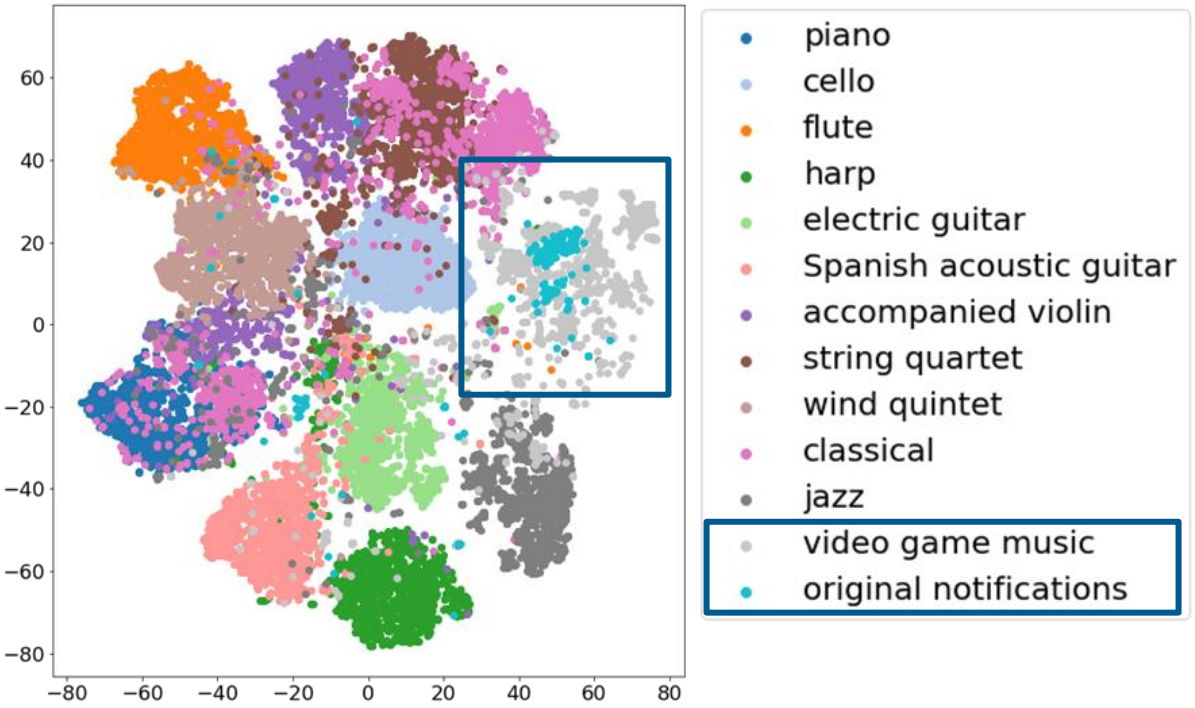
Output domain:

- MusicNet dataset
- YouTube

Video game music as the training data for input domain

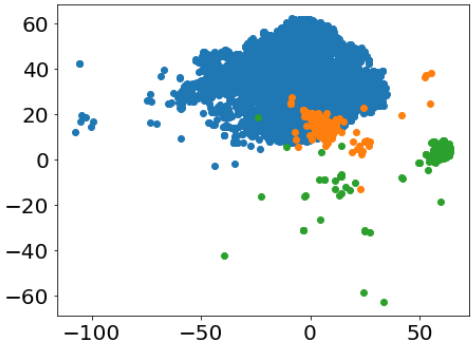
VGGish feature vectors [2017] Hershey et al. – CNN Architectures for Large-Scale Audio Classification

- Recognizing different musical instruments and soundtracks

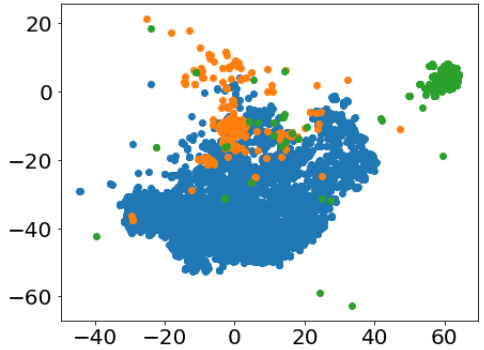


Evaluation – timbre (using VGGish feature vectors)

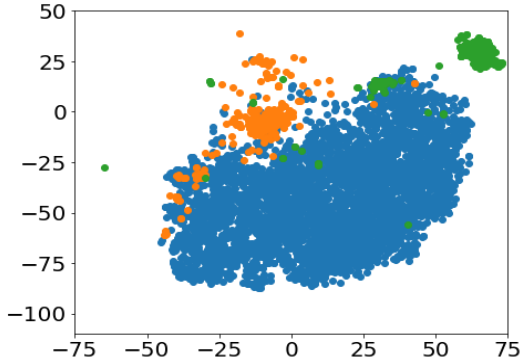
Original notifications Timbre-transformed notifications Target timbre



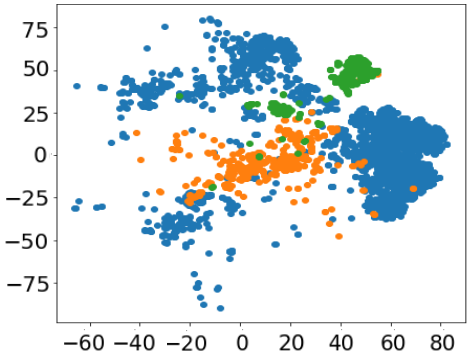
(a) Cello



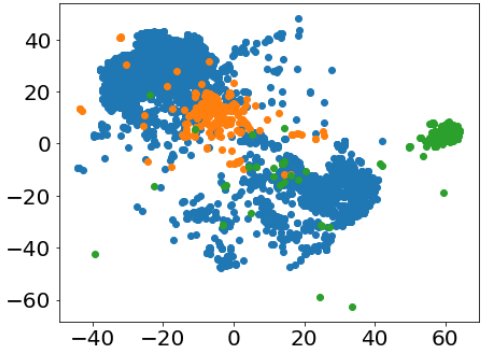
(b) Guitar



(c) Piano



(d) Classical music



(e) Piano-accompanied violin

Evaluation – melody (pitch and rhythm)

- Pitch similarity score [2010] Urbano et al. – Melodic Similarity through Shape Similarity
- Rhythm similarity score [2018] Tomczak et al. – Audio Style Transfer with Rhythmic Constraints

0 (different) - 1 (identical)

Baselines: Notifications reconstructed using Griffin-Lim algorithm GL-5 GL-200

	Piano	Cello	Guitar	Acc. violin	Classical	GL-5	GL-200
Pitch similarity score	0.458 ± 0.014	0.453 ± 0.016	0.449 ± 0.015	0.451 ± 0.013	0.445 ± 0.015	0.523 ± 0.023	0.831 ± 0.028
Rhythm similarity score	0.357 ± 0.038	0.330 ± 0.040	0.354 ± 0.047	0.345 ± 0.043	0.353 ± 0.046	0.479 ± 0.037	0.837 ± 0.042

Evaluation – user perception study on both timbre and melody

- 10 arbitrary notification sounds transferred into piano
- 53 participants (20 female, 33 male, age ϵ [20,52], average age = 28.132)
- 5-point Mean Opinion Score

1	Bad (very different from piano timbre / original melody)
2	Poor (different from piano timbre / original melody)
3	Fair (perceptible difference, but acceptable)
4	Good (slightly perceptible difference, can be recognized as piano / the original melody)
5	Excellent (imperceptible difference, it is piano / original melody)

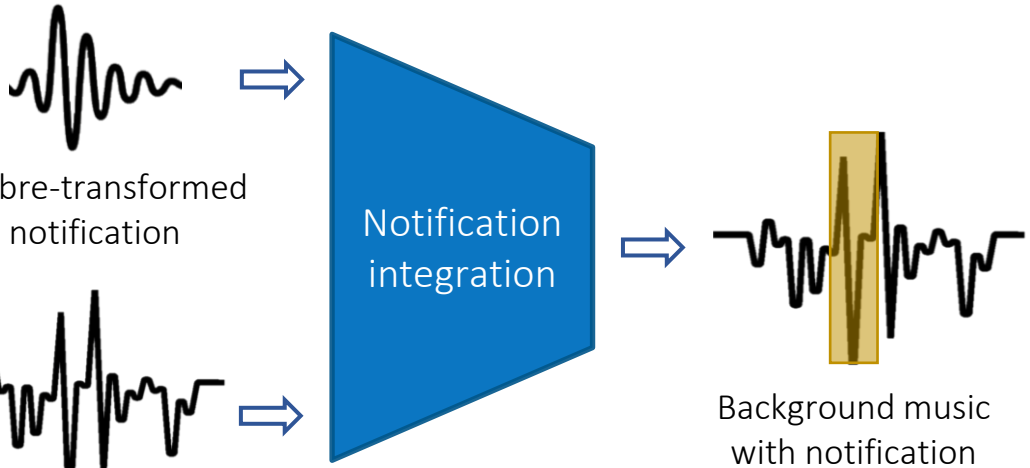
Timbre transfer: 3.345 ± 0.861
Melody preservation: 3.720 ± 0.261

Original notification

Transferred notification

More examples: https://gladys0313.github.io/notification-timbre-transfer/

Potential application – less distracting notification delivery



1. Amplitude adjustment
2. Tempo adjustment
3. Fade-in and fade-out effects

Original notification

Timbre-transformed notification

Music with notification (at 5s)

More examples: https://gladys0313.github.io/notification-timbre-transfer/

Conclusions and future work

Conclusions

- ✓ Notification timbre transfer
- ✓ CycleGAN-based model trained in an unsupervised manner
- ✓ Video game music as the source domain
- ✓ Efficacy and limitation
- ✓ Less intrusive information delivery

Future work

- ➢ Transferring notifications into a style that contains multiple timbre tracks
- ➢ Study to explore the usability of our notification delivery method

Thank you

Unsupervised Musical Timbre Transfer for Notification Sounds

 Jing Yang[1]

 Tristan Cinquin[1]

 Gábor Sörös[2]

[1]Department of Computer Science, ETH Zurich, Switzerland
[2]Nokia Bell Labs, Budapest, Hungary

IEEE ICASSP 2021

https://gladys0313.github.io/notification-timbre-transfer/

ETH zürich

NOKIA Bell Labs