

Voice Synthesis, Transformations, and Sound Effects: A portfolio

Erkut, Cumhur

Creative Commons License
CC BY 4.0

Publication date:
2025

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Erkut, C. (2025). *Voice Synthesis, Transformations, and Sound Effects: A portfolio*. Abstract from DAS | DAGA 2025 - 51st Annual Meeting on Acoustics, Copenhagen, Denmark.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Voice Synthesis, Transformations, and Sound Effects: A portfolio

Cumhur Erkut,

with help of Christie Laurent, Juan Alonso, David Südholt, and Anders R Børgum

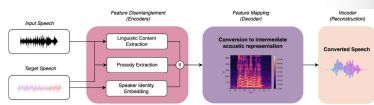
Intro

- Started physically on aging formants
- Benefited in DL disentanglements in DDSP, RAVE -> SRAVE, back to music

Methods

- Tunable Voice Aging Model
- MFCC playback using Z in DDSP
- DDSP cross-synthesis, plugins
- Full embeddings in S-RAVE

Results



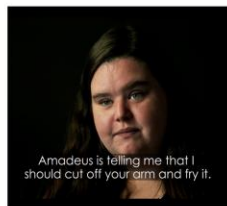
- Significant performance: RT from 1 to 15, transformation data from 20 minutes to 30 seconds
- Concerns: Voice spoofing

Discussion

- Diverse portfolio (Hearing Aids, XR, Robotics, Embedded AI)
- PINNS <-> Transformers
- Active Inference (Bayesian)
- Evaluation: Audio Aesthetics?

Interplay between sound, music, and speech. Transformers for embeddings, diffusion for real-time synthesis.

PROBLEM



- 30% of patients with schizophrenia treatment-resistant Meltzer HV. (1997)
- those whose symptoms get better from medication experience terrible side-effects and often discontinue taking their meds Lieberman, J. A., et al. (2005)
- This is the most expensive disease to treat in psychiatry Kadakia A. et al. 2022, Christensen, M.K. et al. 2022, FosseehA. et al. 2018,

Amadeus is telling me that I should cut off your arm and fry it.

AVATAR THERAPY

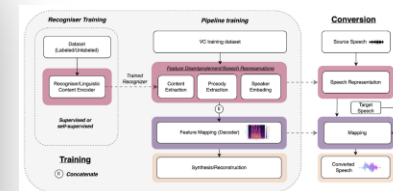


The patient together with the therapist create the avatar of the malevolent voice



Scan to hear more at cerkut.github.io, click on photos to read more about talented Christie Laurent, Juan Alonso, David Südholt, and Anders R Børgum

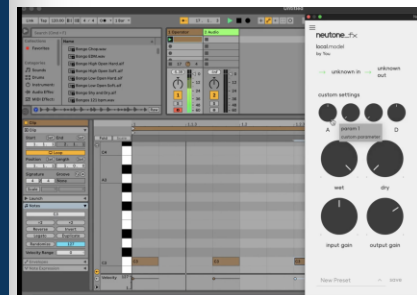
Recent Work



General Voice Conversion pipeline: S-RAVE



Commercial deployment possible: BabyAudio Humanoid



Back to musical instruments

