June 16, 2023

Make Your Favorite Music Curative: --- --- Artificial Intelligence Generated Music for Anxiety Reduction

Fiona Yan Liu
yan.liu@polyu.edu.hk
Department of Computing
The Hong Kong Polytechnic University

Background

- Anxiety is becoming more prevalent
  - Nearly 500,000,000 individuals are affected by anxiety problems [Baghaei, 2021]
  - The situation is even worse recently
  - Result in severe consequences such as palpitation, insomnia, and pain [Woo, 2010]
  - An effective treatment is needed
  - Music therapy has been used for more than forty years in clinical practice
- Music therapy shows distinguished effectiveness and few side effects in emotion regulation
  - Is a health profession to help patients improve and maintain their mental and physical health [Bunt, 2014]
  - A commonly accepted theory is that music acts as a distractor, drawing patient's attention to the melody of music rather than his or her own negative feelings [Nilsson, 2008]
  - Music decreased anxiety level of physiologically unhealthy patients such as heart diseases and cancer [Sendelbach, 2006]
  - A positive relationship between music listening and anxiety reduction on normal persons [Khan, 2018]

Music for Therapy

- The methods of music therapy
  - Receptive: Listen to the music and respond to the experience silently, verbally, or in another modality
  - Re-creation: play or sing along to a pre-composed song in a manner that supports identified goals
  - Improvisation: Involve spontaneous music making using simple instruments, body percussion, or the voice
  - Composition: Create own music
- Aim of our work
  - AI-powered therapeutic music generation
- Objectives
  - What are the music factors that essentially evoke human emotions?
  - How to generate therapeutic music with better user engagement?

What Strikes the Strings of Your Heart?

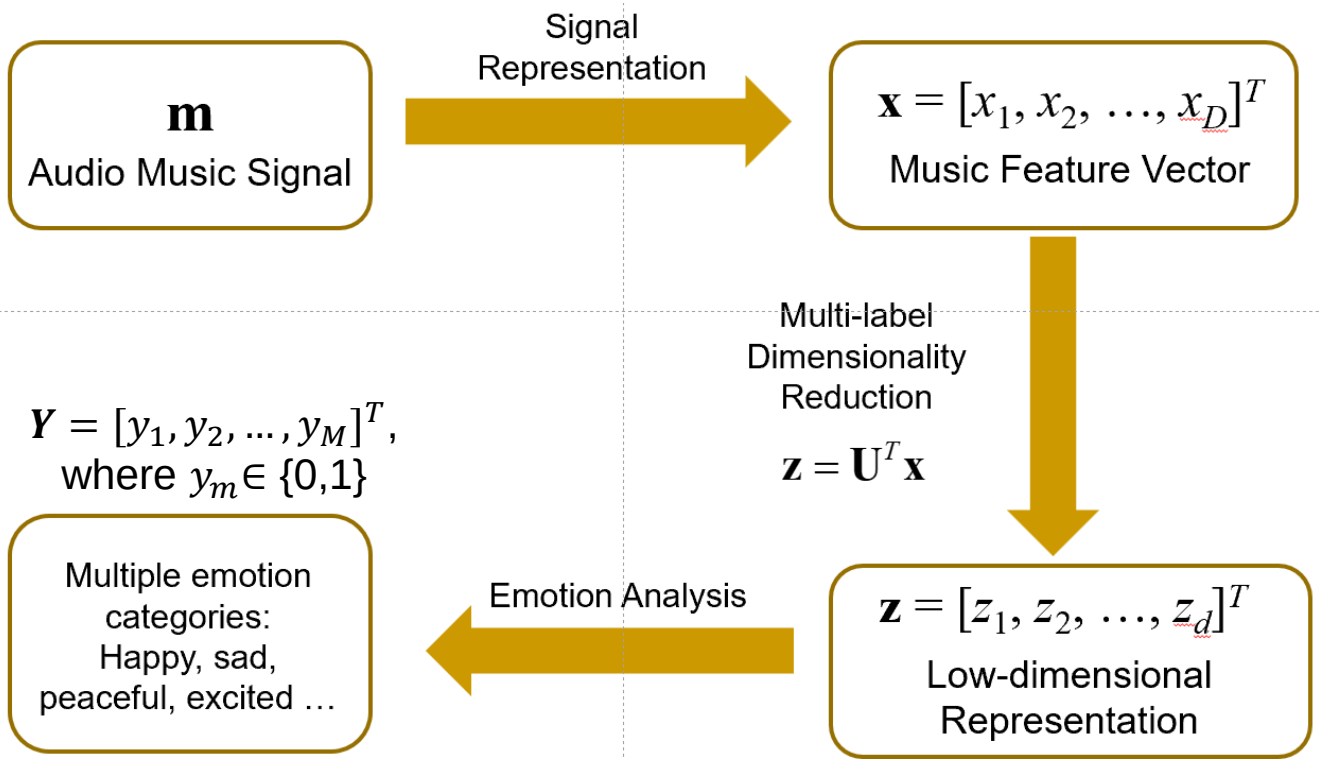
--- --- Feature Mining for Music Emotion Analysis

Music Emotion Analysis

- Pioneer research in music and emotions
  - Validate the existence of relationship between music and emotion Aristotle [Lord, 1984];
- Some specific characters of music can evoke certain kinds of emotions
  - Flowing rhythm patterns may be more indicative of happiness, gracefulness, and serenity while firm rhythm may be perceived as expressing sadness, dignity and vigor [Juslin 2001]
  - Simple and consonant harmony is associated with expression such as serenity and happiness while complex and dissonant harmony with excitement and sadness [Hevner 1936]
  - Ascent and descent of melody may be associated with different emotions [Hevner 1936]
- Our work intends to answer three generally interested questions in data driven approach
  - What are the intrinsic features embedded in music signal that essentially evoke human emotions?
  - To what extent these features influence human emotions?
  - Whether the findings from computational models are consistent with the existing research results from psychology?

Problem Formulation

- Find the intrinsic factors in music that evoke human emotions
  - Motivates us to conduct dimensionality reduction techniques
  - Dimensionality reduction can generate a lower dimensional equivalence to the original high-dimensional feature for given target
- Music may evoke more than one emotion [Zentner, 2008]
  - Motivates us to formulate the problem as multi-label learning
  - In multi-label classification, each data point might be associated with multiple labels



Computational Model

- Idea
  - if two pieces of music express similar emotions
  - the low-dimensional representations of them should be close
- Optimization procedure

$$\mathbf{U} = \arg\min_{\mathbf{U}} \sum_{i=1}^{n} \sum_{j=1}^{n} ||\mathbf{U}^T\mathbf{x}_i - \mathbf{U}^T\mathbf{x}_j||^2 \cdot \hat{S}_{ij}$$

$\mathbf{x}_i$: Original high-dimensional representation of the *i*-th music
$\mathbf{U}$: Transformation matrix
$\hat{S}_{ij}$: Similarity coefficient which represents the emotion similarity between the *i*-th music and the *j*-th music

Algorithm 1: Multi-Emotion Similarity Preserving Embedding (ME-SPE)

Input: Training dataset: $\{(\mathbf{x}_1, \mathbf{y}_1), ..., (\mathbf{x}_n, \mathbf{y}_n)\}$; the dimension of the subspace: d
Output: Transformation matrix: $\mathbf{U}$

1 **for** $i = 1, ..., n$ **do**
2 **for** $j = 1, ..., n$ **do**
3 $\hat{S}_{ij} \leftarrow \langle \mathbf{y}_i/||\mathbf{y}_i||, \mathbf{y}_j/||\mathbf{y}_j|| \rangle$;
4 **for** $i = 1, ..., n$ **do**
5 $D_{ii} \leftarrow \sum_{j=1}^{n} \hat{S}_{ij}$;
6 $\mathbf{L} \leftarrow \mathbf{D} - \hat{\mathbf{S}}$;
$\mathbf{X} \leftarrow [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n]$;
for $i = 1, ..., d$ **do**
 Solve $\mathbf{X}\mathbf{L}\mathbf{X}^T\mathbf{u}_i = \lambda_i\mathbf{X}\mathbf{D}\mathbf{X}^T\mathbf{u}_i$;
10 $\mathbf{U} \leftarrow [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_d]$;

Musical Features Evoke Human Emotions

Datasets	Songs	Emotion Labels	Dimension
EMOTIONS	593	6	72
CAL-500	502	18	$51,143\ (257 \times 199)$

Detailed information of EMOTIONS and CAL-500 datasets

Datasets	Emotion Labels
EMOTIONS	angry/aggressive, quiet/still, happy/pleased, sad/lonely, amazed/surprised, relaxing/calm
CAL-500	happy, sad, calming, arousing, pleasant, cheerful/Festive, tender/soft, powerful/strong, oving/romantic, carefree/lighthearted, exciting/thrilling, emotional/passionate, positive/optimistic, touching/loving, light/playful, angry/aggressive, laid-back/mellow, bizarre/weird

Emotion labels of EMOTIONS and CAL-500 datasets

Performance Comparison on Music Emotion Classification

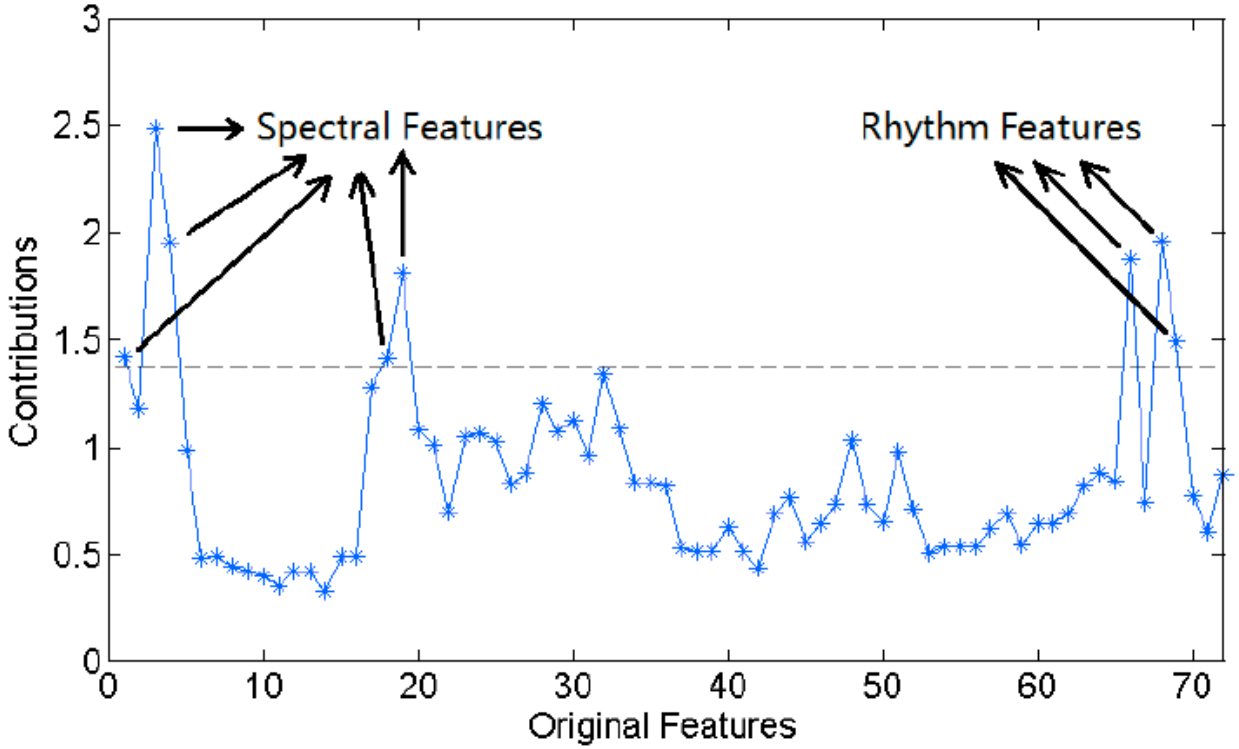
Methods \ Criteria	Average precision	Hamming loss	One-error	Ranking loss
Original 72-D	0.699 ± 0.015	0.275 ± 0.019	0.401 ± 0.023	0.275 ± 0.024
PCA	0.712 ± 0.004	0.262 ± 0.004	0.385 ± 0.009	0.260 ± 0.003
LPP	0.721 ± 0.005	0.258 ± 0.004	0.373 ± 0.009	0.248 ± 0.005
HSL	0.790 ± 0.007	0.211 ± 0.004	0.296 ± 0.015	0.174 ± 0.005
ML-LDA	0.715 ± 0.006	0.261 ± 0.002	0.384 ± 0.014	0.256 ± 0.005
ML-OPLS	0.784 ± 0.006	0.211 ± 0.004	0.301 ± 0.013	0.180 ± 0.007
ME-SPE	$\mathbf{0.814 \pm 0.003}$	$\mathbf{0.193 \pm 0.003}$	$\mathbf{0.257 \pm 0.010}$	$\mathbf{0.149 \pm 0.003}$

- The larger the better
  - Average precision
- The small the better
  - Hamming loss
  - One-error
  - Ranking loss

PCA: principal component analysis
LPP: locality preserving projections
HSL: hyper-graph spectral learning
ML-LDA: multi-label linear discriminant analysis
ML-OPLS: multi-label orthonormalized partial least squares
ME-SPE: multi-emotion similarity preserving embedding

To what Extent These Features Influence Human Emotions

- The contributions of each features in the classification task



Whether the Findings of Computational Model are Consistent with Existing Research Results

- Mean of spectral flux
  - Measures the change of the frequency of current frame compared to that of the previous one
  - Reflect how fast the pitch of a song changes
  - Contribute to classify all six emotions
  - Match the argues that harmony and melody features are closely related to emotion expressions
- Mean of the first MFCC coefficient
  - The weighted sum of all the log-energies
  - An overall measure of the signal loudness
  - Contribute to differentiating “Amazed” and “Happy”
  - Support the claim that the loudness of the music contributes to convey emotions
- Beat histograms
  - Periods of the first and second beat histogram peaks in bpm (beats per minute) respectively
  - Reflect the tempo of a song
  - Neither of these two dimensions improves the performance individually
  - The combination of them make special contributions in classifying “Angry” and “Amazed”, as well as “Sad” and “Quiet”
  - Consistent with the findings that rhythm features are important in emotional expression

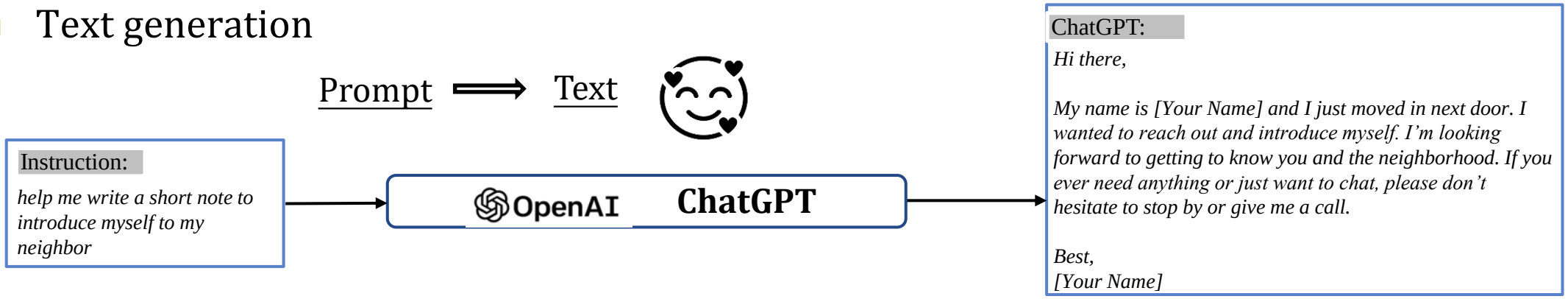
Make Your Favorite Music Curative: --- --- Generate Music For Anxiety with Better User Engagement

Music Generation for Anxiety Reduction

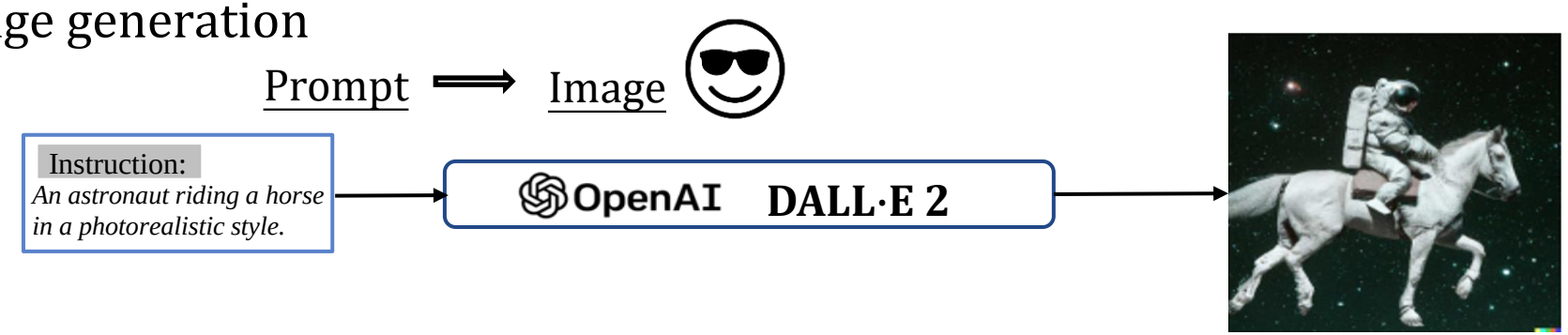
- Music for receptive music therapy to reduce anxiety [Grocke, 2006]
  - Therapeutic music
    - Relaxation response
    - Emotional regulation
  - User preferred music
    - Emotional connection
    - Personalized expression
- Artificial Intelligence Generated Content (AIGC)
  - Involve the creation of digital content, such as images, natural language, and music, through AI models

Typical Products of AIGC

- Text generation



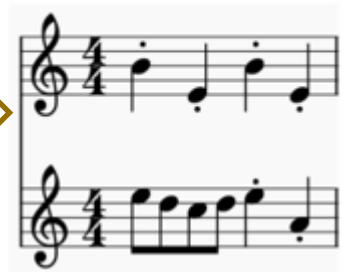
- Image generation



AI-Powered Music Generation

Score Generation (Composition)

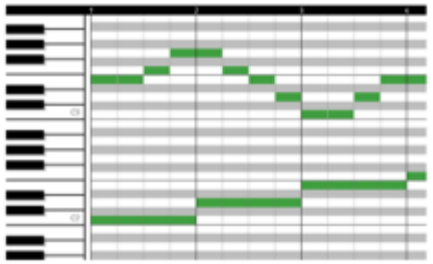
Score



Composer creates the musical score

Performance Generation

Performance

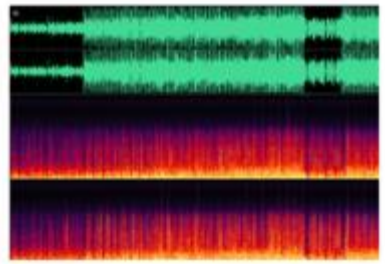


Performer interprets and plays the score.

Audio engineer synthesizes the performance into a digital audio format.

Audio Generation

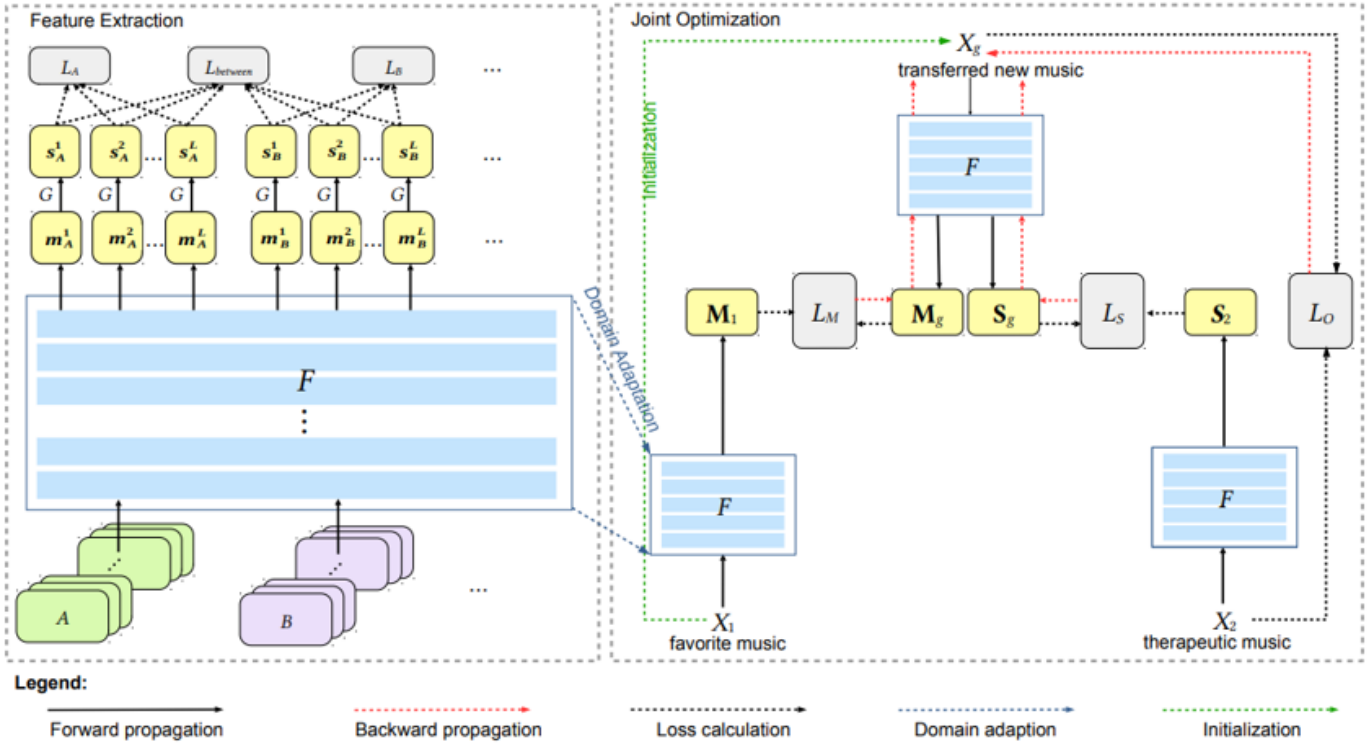
Sound



Audience listens to the finalized audio.

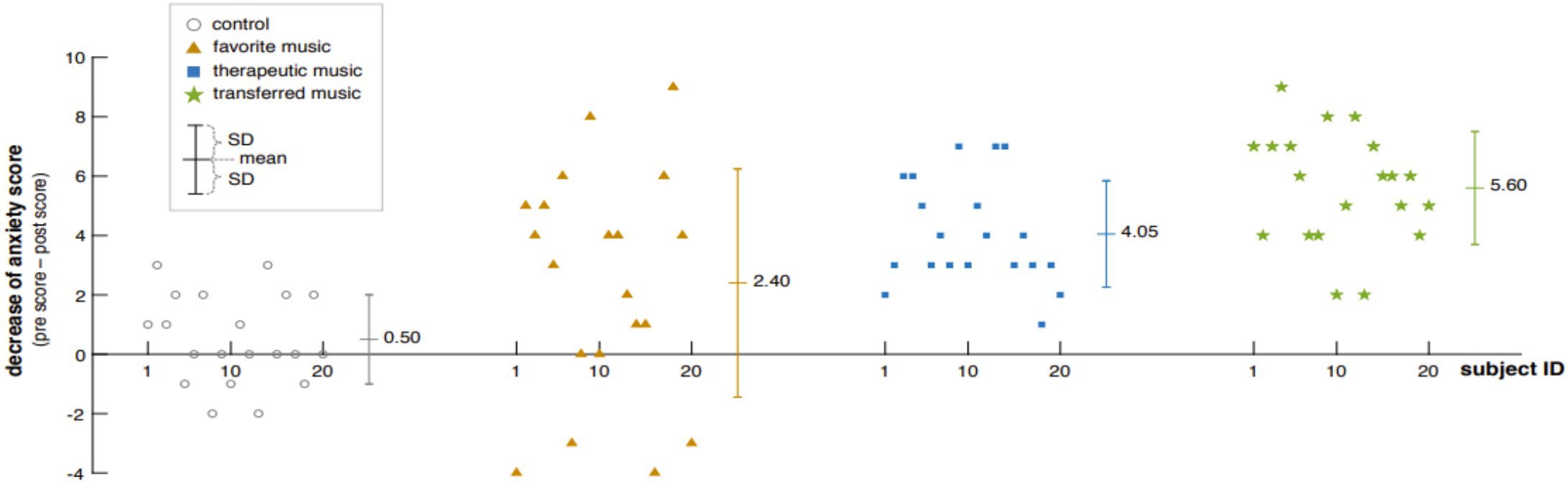
Music Style Transfer

- A novel style transfer model
  - to generate the therapeutic music
  - according to user's preference
- Start from user preferred music
  - Approximate the character of the therapeutic music
  - Emotion features
  - Musicality features
- A new domain adaptation algorithm
  - that transfers the learning result for music genre classification to the music personalization
- Three convolutional neural networks are utilized
  - Minimize the difference in the feature space



Experiment Design

- Paradigm design
  - Subjects: 20 subjects (8 male and 12 female) with mild anxiety experience
  - Stimuli: 3 types of music: : favorite music, therapeutic music, and transferred music
  - Measurement: State-Trait Anxiety Inventory (STAI, Form Y version)
  - Procedure: Four-day trial: control trial, favourite trial, therapeutic trial, and transferred trial in random order.



Conclusion

- Music emotion analysis
  - Dimensionality reduction technique has identified the intrinsic features embedded in music signal that evoke human emotions
- Music generation for anxiety reduction
  - AIGC technique has generated therapeutic music according to user's preference
- Future work
  - Long-term music therapy in more interactive way, such as re-creation or improvisation
    - Seek the cooperation with music therapist
  - Music generation to evoke specific emotion precisely
    - Seek the cooperation with psychologist
  - Computer music in metaverse for healthy aging
    - Seek the cooperation to apply real world problem solving

Q & A

Fiona Yan Liu

yan.liu@polyu.edu.hk

Department of Computing

The Hong Kong Polytechnic University

References

[1] Baghaei, N., Chitale, V., Hlasnik, A., Stemmet, L., Liang, H. N., & Porter, R. (2021). Virtual reality for supporting the treatment of depression and anxiety: Scoping review. *JMIR mental health*, *8*(9), e29681.

[2] Woo, A. K. (2010). Depression and anxiety in pain. *Reviews in pain*, *4*(1), 8-12.

[3] Bunt, L., & Stige, B. (2014). *Music therapy: An art beyond words*. Routledge.

[4] Nilsson, U. (2008). The anxiety-and pain-reducing effects of music interventions: a systematic review. *AORN journal*, *87*(4), 780-807.

[5] Sendelbach, S. E., Halm, M. A., Doran, K. A., Miller, E. H., & Gaillard, P. (2006). Effects of music therapy on physiological and psychological outcomes for patients undergoing cardiac surgery. Journal of cardiovascular nursing, 21(3), 194-200.

[6] Khan, M. A., Chennafi, M., Li, G., & Wang, G. (2018, October). Electroencephalogram-based comparative study of music effect on mental stress relief. In 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI) (pp. 1-5). IEEE.

[7] Lord, C., Aristotle. The Politics. 1st ed. Chicago, IL, USA: Univ. Chicago Press, 1984.

[8] Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework.

[9] Hevner, K. (1936). Experimental studies of the elements of expression in music. *The American Journal of Psychology*, *48*(2), 246-268.

[10] Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, *8*(4), 494.

[11] Grocke, D., & Wigram, T. (2006). Receptive methods in music therapy: Techniques and clinical applications for music therapy clinicians, educators and students. Jessica Kingsley Publishers.