

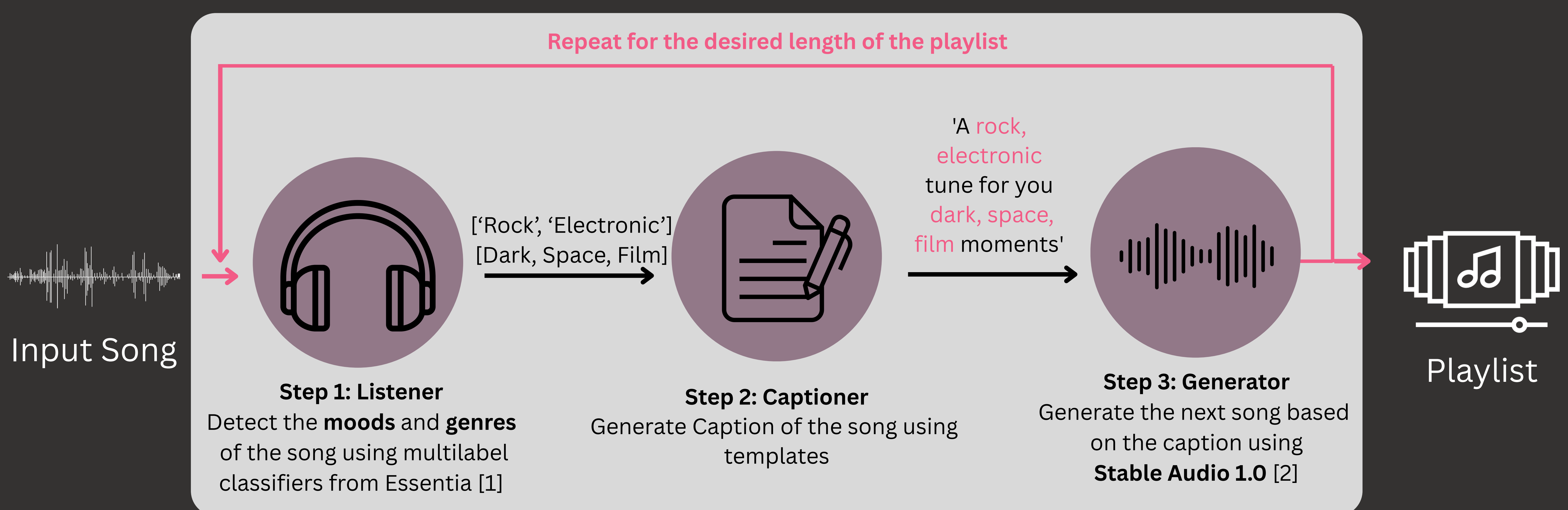
Vibe Sorcery: Integrating Emotion Recognition with Generative Music for Playlist Curation

Isabel Urrego-Gómez, Simon Colton, Iran Roman
Queen Mary University of London

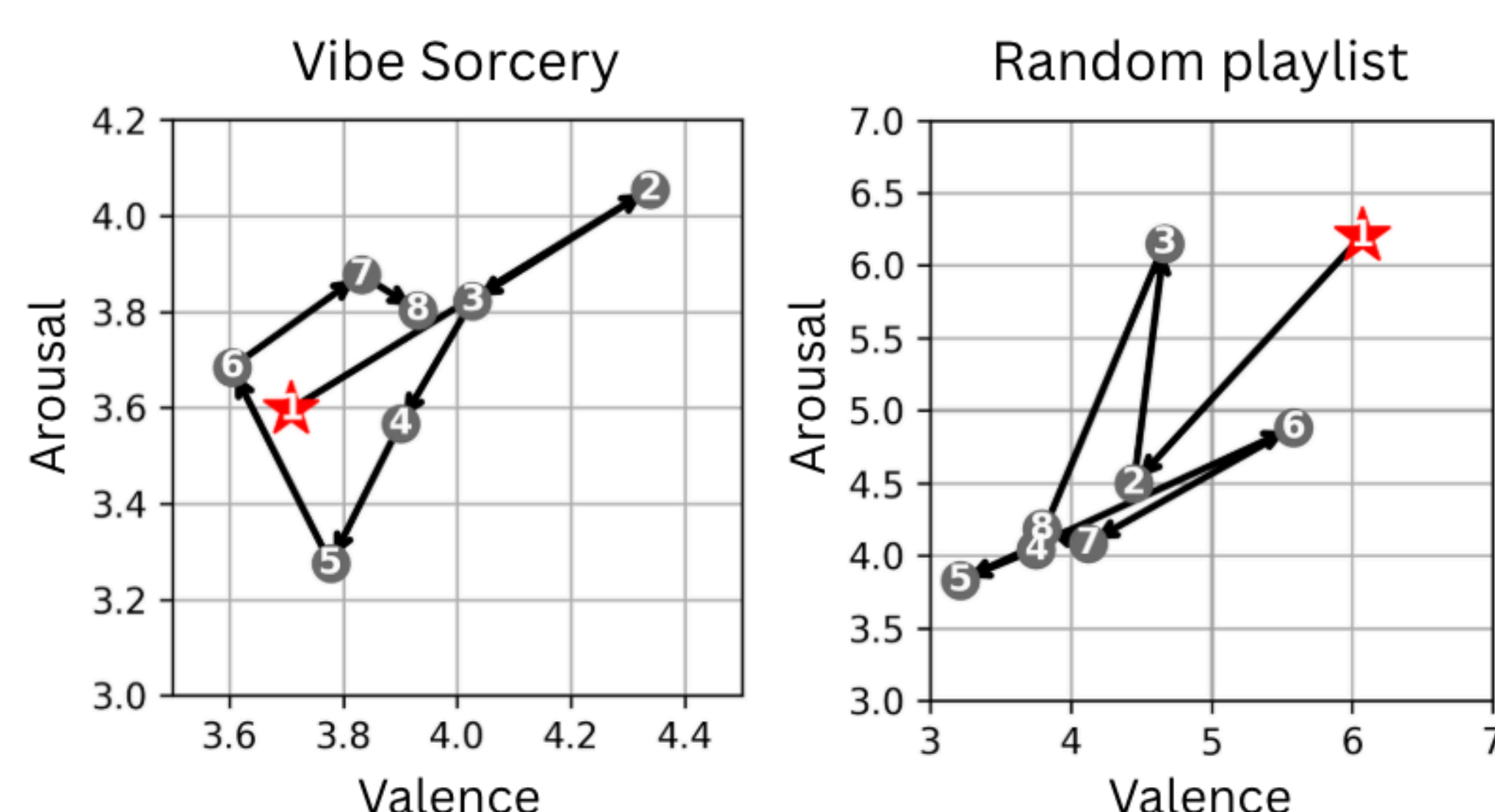
MOTIVATION

- Conventional playlist generators **cluster** songs by similarity (artist, tempo, timbre), overlooking progression.
- We model playlists as a **Markov process**, where each track depends only on the previous one.
- By combining emotion recognition with language-prompted generative music, we generate adaptive playlists that follow emotional cues.

METHODOLOGY



RESULTS

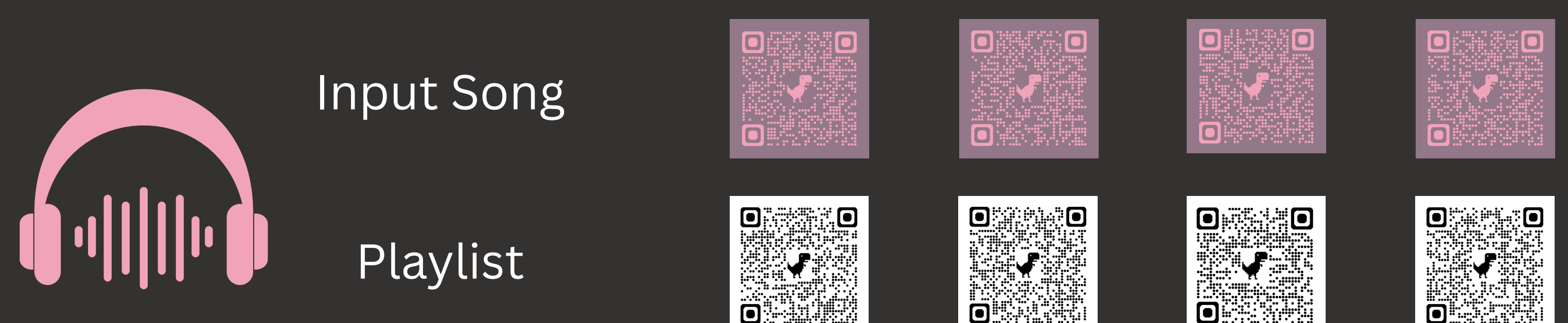


Comparison of playlist trajectories in Arousal-Valence plane [3]. **Vibe Sorcery** (left) produces gradual, coherent emotional shifts, while the **random baseline** (right) exhibits more abrupt transitions.

- Baseline:** 100 songs were created by sampling mood-genre pairs independently, with consecutivity defined only by generation order.
- Evaluation Metric:** Average euclidean distance in Arousal-Valence (AV) plane.

Method	Avg. AV Distance
Random Baseline	2.4
Vibe Sorcery	0.82

SOME EXAMPLES



CONCLUSIONS

- Generative music avoids contextual constraints and associations of existing songs.
- Vibe Sorcery produces adaptive, coherent emotional playlists.

FUTURE WORK

- Use LLMs to improve captions
- Optimize thresholds & parameters.
- Human listener evaluation.

REFERENCES

- [1] Pablo Alonso-Jiménez et al. "Tensorflow Audio Models in Essentia". In: International Conference on Acoustics, Speech and Signal Processing (2020).
- [2] Zach Evans et al. "Stable Audio Open". In: IEEE International Conference on Acoustics, Speech and Signal Processing. 2025, pp. 1–5. DOI: 10.1109/ICASSP49660.2025.10888461.
- [3] James A Russell. "A circumplex model of affect." In: Journal of personality and social psychology (1980).

