

Music Similarity Ranking: Case Study Fairphonic

Rayhan Titho Kharisma¹, Denny²

rayhan.titho@ui.ac.id¹, denny@cs.ui.ac.id²

^{1,2} Faculty of Computer Science, Universitas Indonesia, Jakarta, Indonesia

Article Information

Received : 12 Jul 2024
Revised : 26 Jul 2024
Accepted : 8 Aug 2024

Keywords

Fairphonic, HPCP,
Chroma, Rhythm Pattern

Abstract

Fairphonic Pte Ltd is a technology company operating in the music sector. Fairphonic provides services to detect content on social media that is suspected of copyright infringement. Fairphonic utilizes audio features for the detection process. The current algorithm used by Fairphonic requires pairwise comparison, and the content to be compared is collected through a scraping process immediately after the process is run. Fairphonic has hundreds of thousands of music data in their database. Fairphonic desires a more scalable algorithm to compare an input music piece with the entire Fairphonic music catalog. This research uses features such as Harmonic Pitch Class Profile (HPCP), Chroma, and Rhythm Pattern. The study compares previously researched algorithms, namely binary similarity matrix, Euclidean distance, and similarity matrix profile. The results show that the combination of HPCP with the binary similarity matrix yields the highest Mean Average Precision of 0.989. Speed testing by performing comparisons 10 times shows that the combination of Chroma and the similarity matrix profile is 72% faster compared to the combination of HPCP with the binary similarity matrix. The author recommends the Chroma and similarity matrix profile algorithm for music similarity ranking due to its faster process.

A. Introduction

Fairphonic is a Singapore-based start-up with a vision to revolutionize digital rights management [1]. Fairphonic's business process involves collecting social media content that resembles the audio master provided by clients. Social media content deemed similar by Fairphonic's algorithm is reported to the client. Fairphonic's clients are typically music publishers. After receiving the list of content considered similar, clients can claim royalties on those contents.

Fairphonic's algorithm utilizes pairwise comparison, which is accurate but time-consuming. To enhance scalability, Fairphonic seeks a more efficient algorithm, as it already maintains a substantial MP3 database. Instead of comparing each input individually, Fairphonic aims for an algorithm that returns a list of similar outputs from a single input. Consequently, the author proposes a music similarity ranking system to address this need.

Music similarity ranking is the process of ranking music based on its similarity. This process can be carried out using various algorithms. Music similarity ranking algorithms can be used for a wide range of purposes, such as music recommendation, plagiarism detection, and music analysis[2][3].

Implementations of music similarity ranking are frequently encountered in everyday life. For example, music streaming services like Spotify use music similarity ranking with deep learning to recommend music to users based on the music they have previously enjoyed [4]. Maheshwari [4] calculates similarity based on the tempo, energy, and intensity of the music to perform similarity ranking.

Given the lack of a widely accepted definition for overall musical similarity, an effective approach is to explore models that address specific aspects of music, such as pitch, rhythm, dynamics, and timbre. These models provide tools for tasks related to melody, harmony, rhythm, dynamics, and timbre, among others. It is crucial to emphasize that the concept of music similarity is entirely dependent on the context of the specific retrieval task, which is generally determined by the type of annotations available in the dataset.

The literature on audio-based music similarity showcases various methods, including traditional information retrieval techniques such as feature extraction from audio recordings and calculating distances within a vector space [5,6,7,8,9]. Additionally, it involves modeling the distributions of these extracted features and comparing the resulting statistical models [10,11,12,13,14,15], as well as employing feature learning [16,17], metric learning [18,19], and deep neural networks [20,21].

This paper aims to determine which algorithm is most effective for music similarity ranking for Fairphonic. It will concentrate on calculating the distances between extracted feature vectors, following the first approach mentioned. The study will compare three algorithms [3,22,23], selected for their ease of implementation. Each algorithm will be tested on two datasets: Covers80 and Fairphonic's internal production dataset. The evaluation of these algorithms will be conducted using metrics such as mean rank, mean reciprocal rank, and mean average precision. Based on the evaluation results, the most effective algorithm will be recommended to Fairphonic.

B. Research Method

The overall flow of this experiment can be seen in Figure 1. Both datasets will be formatted as depicted in Figure 1. After extracting features from each song in the datasets, we will retrieve the most similar songs based on the original input, followed by calculating the evaluation metrics. This section will explain the selected algorithms and the evaluation metrics used.

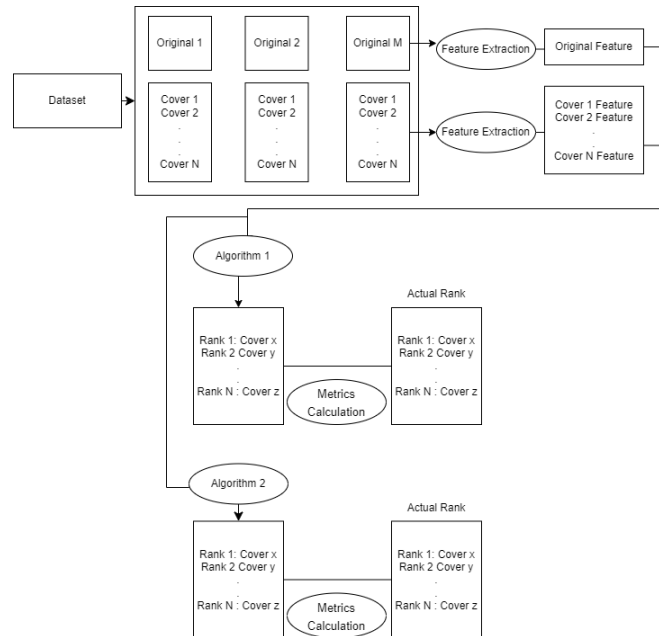


Figure 1. Overall Flow of Experiment

2.1 Chroma & HPCP

Chroma or Chromagram represents the tonal content in an audio signal [24]. In music, chroma features correlate with the 12 musical notes. Essentially, a Chromagram is a Mel-Spectrogram that does not use the mel scale, but instead divides the frequency range into 12 musical notes. There is a variant of the Chromagram that does not use the Fourier transform when transforming from the time domain to the frequency domain, but uses the constant Q-transform (CQT), known as Chroma CQT. CQT is considered more suitable for musical signals compared to the Fourier transform, although CQT computation is more demanding. Another variant is known as Chroma Energy Normalized Statistics (CENS), which is a normalized Chroma CQT, and is widely used for audio matching applications [25].

Another name for Chromagram is Pitch Class Profile (PCP). If PCP only considers pitch, there is a variant that includes harmonic components in the PCP calculation, called the Harmonic Pitch Class Profile (HPCP). Harmonic components are integer multiples of a base frequency. For example, if the base frequency is 50 Hz, the first harmonic component is 100 Hz, the second harmonic component is 150 Hz, and so on. In the PCP calculation, these harmonic components are not considered.

2.2 Rhythm Pattern

According to Lidy and Rauber [23], Rhythm Pattern or Fluctuation Pattern is an audio feature that captures information about the strength of amplitude fluctuation changes within a certain frequency range [23]. The frequency range in the rhythm pattern uses the bark scale. The rhythm pattern can be visualized with the x-axis representing the rate of noise changes, the y-axis representing the frequency range on the bark scale, and the color indicating the strength of the noise changes.

2.3 Similarity Matrix Profile: SiMPle-Fast

Simple-Fast quickly calculates the similarity matrix profile by comparing subsequences from two time series data such as audio data [3]. The similarity matrix profile is formed by calculating the distance between subsequences. This distance calculation is performed using Mueen's Algorithm for Similarity Search (MASS). MASS utilizes the Fourier transform algorithm and convolution theory [26]. The distance is determined using the median of the similarity matrix profile. This method ensures efficient and accurate comparisons, making it suitable for large datasets. By leveraging the strengths of Fourier transforms and convolution, MASS provides a robust approach for identifying similar patterns within complex audio signals.

2.4 Euclidean Distance

Euclidean distance calculates the shortest distance between two points [27]. Euclidean distance also accommodates calculations in higher dimensions. In the example of two-dimensional Cartesian coordinates, the formula for calculating the distance between two points p and q in Figure 2 using Euclidean distance is Equation 1. Euclidean distance for more than two dimensions is described in equation 2.

$$d(p, q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

$$d(p, q) = \sqrt{\begin{matrix} (x_1 - y_1)^2 + (x_2 - y_2)^2 + \\ (x_3 - y_3)^2 + \dots \\ + (x_n - y_n)^2 \end{matrix}} \quad (2)$$

Euclidean distance is computationally efficient and straightforward, making it a popular choice for applications in clustering, classification, and similarity measurement. Despite its simplicity, Euclidean distance assumes that the feature space is isotropic and that all dimensions contribute equally to the distance, which might not always be the case in real-world data. However, its ease of implementation and clear interpretation make it a foundational tool in many analytical and scientific endeavors.

2.4 Binary Similarity matrix

This approach [22] involves four main stages: pre-processing, similarity matrix creation, dynamic programming local alignment (DPLA), and post-

processing. In the pre-processing stage, Harmonic Pitch Class Profiles (HPCP) matrices are extracted, and then these HPCP matrices are averaged to become global HPCP vectors for each song.

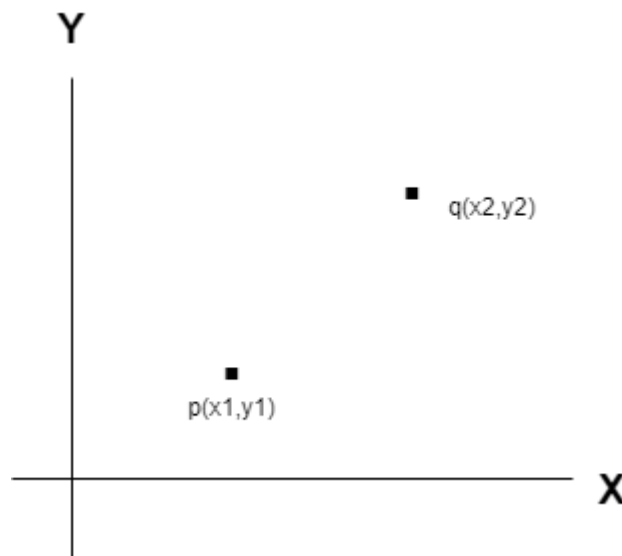


Figure 2. Two points cartesian coordinates

One song is then transposed to match the key of the other song. The transposition process first calculates the optimal transposition index based on the global HPCP vector. The similarity matrix creation involves calculating the binary similarity function for two HPCP vectors, one of which has already been transposed. If HPCP vector A has dimensions (300, 12) and HPCP vector B has dimensions (400, 12), the resulting binary similarity matrix will have dimensions (300, 400), with each element being binary, either 1 or -1. DPLA is then implemented on the binary similarity matrix to account for drastic changes in song structure, and finally, post-processing produces the similarity metrics between the two songs. This method is said to improve accuracy by avoiding intermediary techniques such as key estimation or beat tracking and focusing on structural changes within the songs. By emphasizing structural modifications, this approach aims to provide more reliable similarity assessments, making it particularly useful for diverse musical genres and complex compositions. Additionally, the method's robustness in handling key transpositions and structural variations allows it to capture deeper musical relationships that simpler models might miss, ensuring a comprehensive analysis of musical similarity. This can be particularly advantageous in applications such as music recommendation systems, plagiarism detection, and musicological research, where nuanced understanding of musical similarity is essential.

2.5 Mean Rank

Mean Rank is the average rank of the first cover song (the song considered similar) that appears in the ranking results [28]. For example, Figure 3 shows a dataset of original and cover song pairs. Then, Figure 4 shows the computed similarity ranking results based on their similarity levels.

Based on Figure 4, for original song 1, the correct cover song is ranked 2nd (Cover Song 1c), for original song 2, it is ranked 1st (Cover Song 2b), for original song 3, it is ranked 1st (Cover Song 3a), and for original song 4, it is ranked 3rd (Cover Song 4a). Therefore, the mean rank of this scheme is $(2 + 1 + 1 + 3)$ divided by 4, which equals 1.75. This metric provides a straightforward way to evaluate the effectiveness of similarity ranking algorithms by focusing on the rank positions of the first relevant items. A lower mean rank indicates better performance, as it shows that the correct cover songs appear closer to the top of the ranking list. Thus, Mean Rank is a crucial metric for assessing and comparing different algorithms in music similarity ranking tasks..

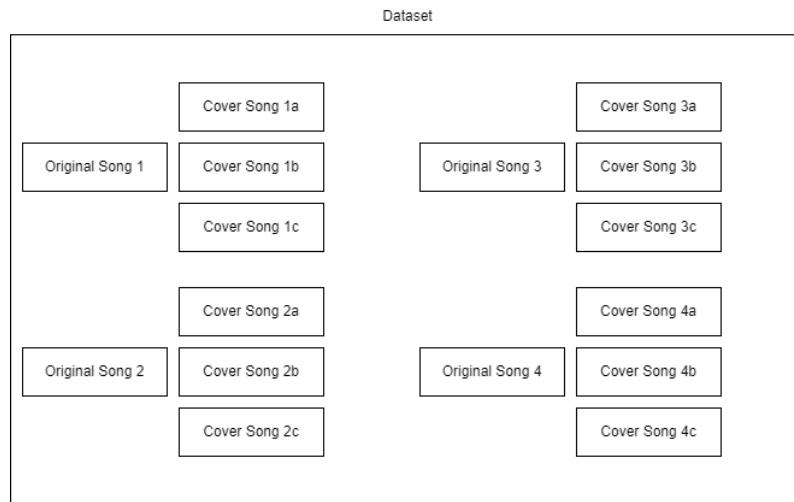


Figure 3. Example original song – cover song pairs

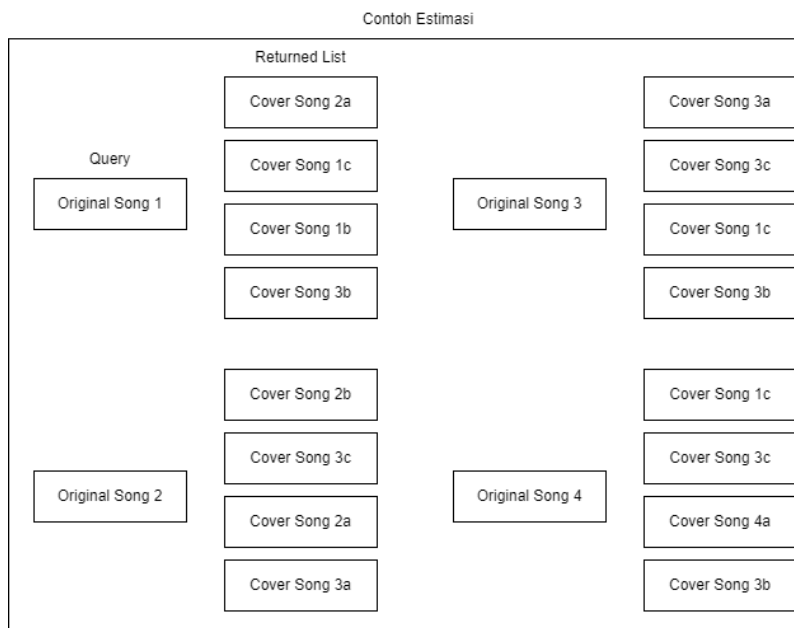


Figure 4. Example of similarity ranking results

2.6 Mean Reciprocal Rank

Mean Reciprocal Rank (MRR) transforms the index position into a range from 0 to 1, with higher values indicating better performance in the returned ranks [28]. Essentially, a higher MRR value signifies more accurate ranking results. The MRR calculation process is illustrated in Figure 5. This metric provides a clear and concise measure of ranking effectiveness, making it a valuable tool for evaluating the performance of similarity ranking algorithms in various applications, including music similarity analysis and recommendation systems. By using MRR, one can easily assess and compare the precision of different algorithms.

2.7 Mean Average Precision

Mean Average Precision (MAP) is a widely used metric for evaluating the performance of similarity ranking systems [28]. MAP provides a single-figure measure of quality across all ranking positions, making it a comprehensive metric for performance evaluation. It is the average of the Average Precision (AP) values calculated for each query or task. Precision@K, a related metric, measures how many relevant items are present among the top K recommended items. However, Precision@K does not account for the actual ranking positions of these relevant items.

To address this, Average Precision@K modifies Precision@K by incorporating the ranks of the relevant items into the calculation, ensuring that items appearing earlier in the ranking contribute more to the score. This approach offers a more nuanced understanding of ranking effectiveness by emphasizing the importance of higher-ranked relevant items. An illustration of MAP is provided in Figure 6, which demonstrates how this metric helps in assessing the overall effectiveness of similarity ranking results.

Higher MAP values indicate better performance of the ranking system, reflecting a higher proportion of relevant items appearing at the top of the ranking list. By using MAP, researchers and practitioners can gain insights into how well their similarity ranking algorithms perform, allowing them to make informed decisions about algorithm improvements and optimizations. This metric is particularly useful in applications such as information retrieval, recommendation systems, and any domain where the quality of ranked results is crucial.

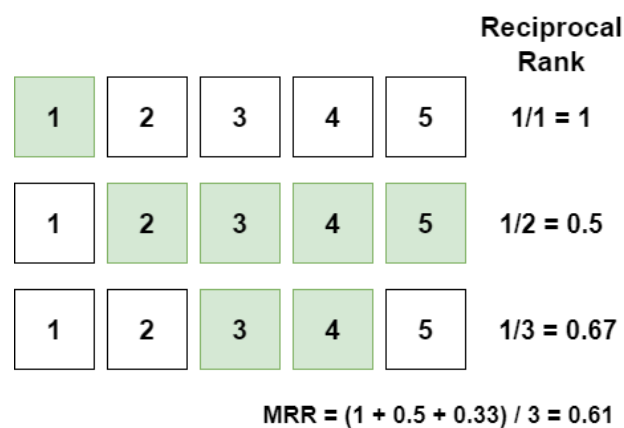


Figure 5. MRR illustration.

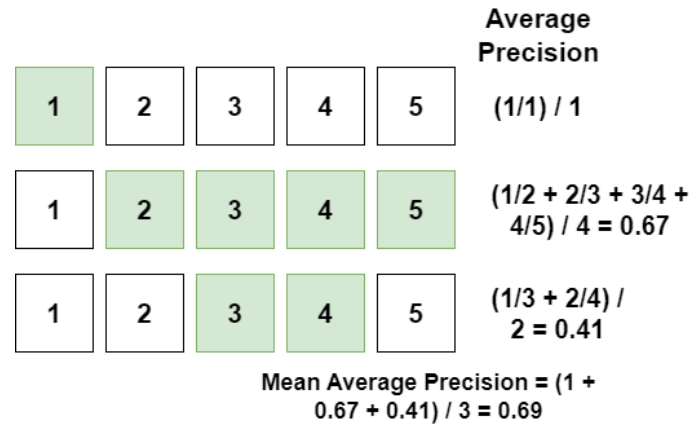


Figure 6. Mean Average Precision example

C. Result and Discussion

This section discusses a comprehensive analysis of the performance metrics used to evaluate the algorithms. This includes the scores for Mean Rank, Mean Reciprocal Rank (MRR), and Mean Average Precision (MAP), alongside the velocity of the processing. The Mean Rank scores provide insight into the average position of the first correctly identified similar song in the ranking results, while the MRR scores transform these positions into a range between 0 and 1, indicating higher values for better performance. MAP scores offer an average precision value that considers the ranking positions of relevant items, providing a nuanced measure of algorithm accuracy. Additionally, the velocity of the process, or the time taken to compute these metrics, is evaluated to ensure that the algorithms are not only accurate but also efficient. These metrics collectively highlight the strengths and weaknesses of each algorithm, guiding the recommendation for the most suitable algorithm for Fairphonic's needs.

The experiment tested on two datasets, Covers80 and Fairphonic's internal production dataset. The Covers80 dataset comprises 39 original songs and 80 cover songs, while the production dataset contains 80 original songs and 160 cover songs. The diversity and size of these datasets ensure a thorough evaluation of the algorithms, reflecting real-world scenarios. This comprehensive testing framework allows for a robust assessment of each algorithm's performance, providing valuable insights into their practical applicability and reliability in various contexts.

The evaluation metric calculations for the three algorithms: Chroma CENS - Simple Fast, Rhythm Pattern - Euclidean Distance, and HPCP - Binary Similarity Matrix can be seen in Table 1. Based on Table 1, the most accurate algorithm is HPCP - Binary Similarity Matrix, outperforming the other two algorithms across both datasets. The second-best algorithm is Chroma CENS - Simple Fast. The Rhythm Pattern - Euclidean Distance algorithm performed poorly on the Covers80 dataset.

The superior performance of Chroma CENS - Simple Fast and HPCP - Binary Similarity Matrix over Rhythm Pattern - Euclidean Distance may be due to the fact that song similarity is more influenced by pitch features (Chroma CENS and HPCP) rather than rhythm features (Rhythm Pattern). Additionally, the Euclidean Distance algorithm tends to be ineffective for calculating distances in high-dimensional spaces (the curse of dimensionality). This phenomenon arises because, as the number of dimensions increases, the distance between points becomes less

meaningful due to the dilution of data density. Thus, the Rhythm Pattern - Euclidean Distance algorithm may not capture the intricate nuances of musical similarity as effectively as the pitch-based algorithms. The findings suggest a need for more sophisticated methods to handle high-dimensional data in music similarity tasks.

Table 1. Metrics Calculation Result

Algorithm	Evaluation Metrics	Production Dataset	Covers 80
Chroma Cens - Simple Fast	MR	42.205	4.525
	MRR	0.886	0.897
	MAP	0.865	0.83
Rhythm Pattern - Euclidean Distance	MR	117.346	55.812
	MRR	0.613	0.068
	MAP	0.558	0.059
HPCP - Binary Similarity Matrix	MR	2.576	2.068
	MRR	1.0	0.979
	MAP	0.989	0.889

This study calculates the time required for the three algorithms to perform 100 comparisons. The experiment was conducted 30 times, and the average time in seconds to perform 100 comparisons and its standard deviation were calculated. The calculation results can be seen in Table 2.

The Rhythm Pattern - Euclidean Distance algorithm recorded a much faster time compared to the other two algorithms. This is due to the relatively simple computation, which only involves calculating the Euclidean distance. The second fastest algorithm is Chroma CENS - Simple Fast, which recorded a time almost four times faster than HPCP - Binary Similarity Matrix. These findings highlight the trade-off between speed and accuracy, emphasizing the need for balancing computational efficiency and performance in practical applications.

Table 2. Processing Speed Results

Algorithm	Mean (seconds)	Standard Deviation
Chroma CENS - Simple Fast	6.536	0.143
Rhythm Pattern - Euclidean Distance	0.216	0.057
HPCP - Binary Similarity Matrix	22.963	0.115

D. Conclusion

Based on the experimental research and evaluation metrics testing conducted, the values of Mean Rank, Mean Reciprocal Rank, and Mean Average Precision were calculated and recorded in Table 1. The most accurate algorithm, according to these evaluation metrics, is HPCP - Binary Similarity Matrix. However, the fastest algorithm is Rhythm Pattern - Euclidean Distance. The Chroma CENS - Simple Fast algorithm falls in the middle, being reasonably fast but not as accurate as HPCP - Binary Similarity Matrix. Specifically, Chroma CENS - Simple Fast is 72 percent faster (six seconds) than HPCP - Binary Similarity Matrix (22 seconds).

The HPCP - Binary Similarity Matrix algorithm excels in accuracy due to its robust handling of harmonic and pitch class features, which are crucial for precise music similarity assessment. Its comprehensive analysis of musical structure allows

for more reliable similarity determinations, making it ideal for applications requiring high accuracy. Despite its accuracy, the computational cost and processing time are significant, which may limit its practicality in real-time applications or large-scale datasets.

In contrast, the Rhythm Pattern - Euclidean Distance algorithm, while less accurate, offers a much faster processing time. Its simplicity in calculating Euclidean distances between rhythm patterns makes it efficient, suitable for scenarios where speed is prioritized over precision. This algorithm's efficiency can be particularly beneficial in applications where quick responses are necessary, albeit with a trade-off in similarity detection accuracy.

The Chroma CENS - Simple Fast algorithm strikes a balance between speed and accuracy. It provides a reasonable level of precision in identifying similar music tracks while maintaining a faster processing time compared to HPCP - Binary Similarity Matrix. This makes it a versatile option for applications needing a compromise between computational efficiency and accuracy.

In summary, the choice of algorithm depends on the specific requirements of the application, whether it be the need for high accuracy, fast processing time, or a balance of both.

E. Acknowledgment

The author wishes to express special thanks to hi supervisor, Denny, S.Kom., M.I.T., Ph.D. for his excellent guidance, supervision, encouragement, and invaluable advice throughout a period of study and in writing this paper.

F. References

- [1] Palisady, "Music tech start-up Fairphonic appoints Gavin Parry as Chairman," Palisady, 2021. [Online]. Available: <http://palisady.co/news/music-tech-start-up-fairphonic-appoints-gavin-parry-as-chairman/>. [Accessed: Jun. 3, 2024]..
- [2] H. Redelinghuys and Z. Wang, "Evaluating audio features for speech/non-speech discrimination," 2022 First International Conference on Artificial Intelligence Trends and Pattern Recognition (ICAITPR), 2022, pp. 1-6, doi: 10.1109/ICAITPR51569.2022.9844226.
- [3] D. F. Silva, C.-C. M. Yeh, Y. Zhu, G. E. A. P. A. Batista, and E. Keogh, "Fast Similarity Matrix Profile for Music Analysis and Exploration," IEEE Transactions on Multimedia, vol. 21, no. 1, pp. 29-38, Jan. 2019, doi: 10.1109/TMM.2018.2849563.
- [4] C. Maheshwari, "Music Recommendation on Spotify using Deep Learning," arXiv preprint arXiv:2312.10079, 2023. [Online]. Available: <https://arxiv.org/abs/2312.10079>. [Accessed: Jul. 11, 2024].
- [5] Tao Li and M. Ogihara, "Content-based music similarity search and emotion detection," 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, QC, Canada, 2004, pp. V-705, doi: 10.1109/ICASSP.2004.1327208.
- [6] K. Seyerlehner and M. Schedl, "Block-Level Audio Features for Music Genre Classification," in Proceedings of the 2009 International Conference, 2009.

- [Online]. Available: <https://api.semanticscholar.org/CorpusID:17208112>. [Accessed: Jul. 11, 2024].
- [7] X. Yu, J. Zhang, J. Liu, W. Wan and W. Yang, "An audio retrieval method based on chromagram and distance metrics," 2010 International Conference on Audio, Language and Image Processing, Shanghai, China, 2010, pp. 425-428, doi: 10.1109/ICALIP.2010.5684543.
 - [8] Peter Knees and Markus Schedl. 2013. A survey of music similarity and recommendation from music context data. *ACM Trans. Multimedia Comput. Commun. Appl.* 10, 1, Article 2 (December 2013), 21 pages. <https://doi.org/10.1145/2542205.2542206>
 - [9] C. J. Tralie, "Early MFCC and HPCP Fusion for Robust Cover Song Identification," arXiv preprint arXiv:1707.04680, 2017. [Online]. Available: <https://arxiv.org/abs/1707.04680>. [Accessed: Jul. 11, 2024].
 - [10] B. Logan and A. Salomon, "A music similarity function based on signal analysis," IEEE International Conference on Multimedia and Expo, 2001. ICME 2001., Tokyo, Japan, 2001, pp. 745-748, doi: 10.1109/ICME.2001.1237829.
 - [11] J.-J. Aucouturier and F. Pachet, "Music Similarity Measures: What's the use?" in *Proceedings of the International Society for Music Information Retrieval Conference*, 2002. [Online]. Available: <https://api.semanticscholar.org/CorpusID:14783916>. [Accessed: Jul. 11, 2024].
 - [12] J.-J. Aucouturier and P. F., "Improving Timbre Similarity: How high is the sky?," *Journal of Negative Results in Speech and Audio Sciences*, vol. 1, no. 1, 2004.
 - [13] Adam Berenzweig, Beth Logan, Daniel P.W. Ellis, Brian Whitman; A Large-Scale Evaluation of Acoustic and Subjective Music-Similarity Measures. *Computer Music Journal* 2004; 28 (2): 63-76. doi: <https://doi.org/10.1162/014892604323112257>
 - [14] R. Typke, F. Wiering, and R. C. Veltkamp, "Evaluating the Earth Mover's Distance for Measuring Symbolic Melodic Similarity," in *MIREX-ISMIR 2005: 6th International Conference on Music Information Retrieval*, Citeseer, 2005.
 - [15] D. Schnitzer, A. Flexer, and G. Widmer, "A fast audio similarity retrieval method for millions of music tracks," *Multimedia Tools and Applications*, vol. 58, pp. 23-40, 2012. [Online]. Available: <https://doi.org/10.1007/s11042-010-0679-8>. [Accessed: Jul. 11, 2024].
 - [16] M. Panteli, E. Benetos, and S. Dixon, "Learning a Feature Space for Similarity in World Music," in *Proceedings of the International Society for Music Information Retrieval Conference*, 2016. [Online]. Available: <https://api.semanticscholar.org/CorpusID:217471943>. [Accessed: Jul. 11, 2024].
 - [17] Z. Yu, X. Xu, X. Chen and D. Yang, "Learning a Representation for Cover Song Identification Using Convolutional Neural Network," *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, pp. 541-545, doi: 10.1109/ICASSP40776.2020.9053839.
 - [18] M. Slaney, K. Q. Weinberger, and W. White, "Learning a Metric for Music Similarity," in *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR 2008)*, J. P. Bello, E. Chew, and D. Turnbull, Eds.,

- Drexel University, Philadelphia, PA, USA, Sep. 14-18, 2008, pp. 313-318. [Online]. Available: http://ismir2008.ismir.net/papers/ISMIR2008_148.pdf. [Accessed: Jul. 11, 2024].
- [19] H. Heo, H. J. Kim, W. S. Kim, and K. Lee, "Cover Song Identification with Metric Learning Using Distance as a Feature," in Proceedings of the 18th International Society for Music Information Retrieval Conference (ISMIR 2017), S. J. Cunningham, Z. Duan, X. Hu, and D. Turnbull, Eds., Suzhou, China, Oct. 23-27, 2017, pp. 628-634. [Online]. Available: https://ismir2017.smcnus.org/wp-content/uploads/2017/10/33_Paper.pdf. [Accessed: Jul. 11, 2024].
- [20] M. Mehta, A. Sajani and R. Chapaneri, "Cover Song Identification with Pairwise Cross-Similarity Matrix using Deep Learning," 2019 IEEE Bombay Section Signature Conference (IBSSC), Mumbai, India, 2019, pp. 1-5, doi: 10.1109/IBSSC47189.2019.8973064.
- [21] M. Sheikh Fathollahi and F. Razzazi, "Music similarity measurement and recommendation system using convolutional neural networks," International Journal of Multimedia Information Retrieval, vol. 10, pp. 43-53, 2021. [Online]. Available: <https://doi.org/10.1007/s13735-021-00206-5>. [Accessed: Jul. 11, 2024].
- [22] J. Serra, E. Gomez, P. Herrera, and X. Serra, "Chroma Binary Similarity and Local Alignment Applied to Cover Song Identification," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 16, no. 6, pp. 1138-1151, Aug. 2008, doi: 10.1109/TASL.2008.924595.
- [23] T. Lidy, G. Pözlbauer, and A. Rauber, "Sound Re-synthesis From Rhythm Pattern Features - Audible Insight Into A Music Feature Extraction Process," in Proceedings of the International Computer Music Conference (ICMC), Barcelona, Spain, Sep. 5-9, 2005, pp. 93-96.
- [24] A. S. Nasim, R. H. Chowdory, A. Dey, and A. Das, "Recognizing Speech Emotion Based on Acoustic Features Using Machine Learning," in Proceedings of the 2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2021, pp. 1-7, doi: 10.1109/ICACSIS53237.2021.9631319.
- [25] M. Kattel, A. Nepal, A. K. Shah, and D. Shrestha, "Chroma feature extraction," in Proceedings of the Conference: Chroma Feature Extraction Using Fourier Transform, vol. 20, no. 1, Jan. 2019.
- [26] S. Zhong and A. Mueen, "MASS: distance profile of a query over a time series," Data Mining and Knowledge Discovery, vol. 38, no. 3, pp. 1466-1492, May 2024. [Online]. Available: <https://doi.org/10.1007/s10618-024-01005-2>. [Accessed: Jul. 11, 2024].
- [27] M. Hohmann, K. Devriendt, and M. Coscia, "Quantifying ideological polarization on a network using generalized Euclidean distance," Science Advances, vol. 9, no. 9, pp. eabq2044, 2023. [Online]. Available: <https://www.science.org/doi/abs/10.1126/sciadv.abq2044>. [Accessed: Jul. 11, 2024].
- [28] R. P. P. Bodo, E. Benetos, and M. G. de Queiroz, "A Framework for Music Similarity and Cover Song Identification," in Proceedings of the Conference, 2021. [Online]. Available:

<https://api.semanticscholar.org/CorpusID:244169905>. [Accessed: Jul. 11, 2024].