QUANTITATIVE ANALYSIS OF MELODIC SIMILARITY IN MUSIC COPYRIGHT INFRINGEMENT CASES

Saebyul Park[1] **Halla Kim**[1] **Jiye Jung**[2]
Juyong Park[1] **Jeounghoon Kim**[1] **Juhan Nam**[1]
[1]Graduate School of Culture Technology, KAIST, Korea
[2]Heinrich Heine University Düsseldorf, Germany
{saebyul_park, kimhalla, juyongp, miru, juhannam}@kaist.ac.kr
jujiy100@uni-duesseldorf.de

ABSTRACT

This study aims to measure the similarity of melodies objectively using natural language processing (NLP) techniques. We utilize Mel2word which is a melody tokenization method based on byte-pair encoding to facilitate the semantic analysis of melodies. In addition, we apply two word weighting methods: the modified Tversky measure for word salience and the TF-IDF method for word importance and uniqueness, to better understand the characteristics of each melodic element. We validate our approach by comparing song vectors calculated from an average of Mel2Word vectors to the ground truth in 108 cases of music copyright infringement, sourced from an extensive review of legal documents from law archives. The results demonstrate that the proposed approach is more in accordance with court rulings and perceptual similarity.

1. INTRODUCTION

Since the landmark case of *Millett v. Snowden* [1] in 1844, music plagiarism has been a contentious issue for over a century. The term "plagiarism" refers to the subcategory of copyright infringement that involves the false designation of authorship and other unattributed uses of copyrighted material [1]. In determining plagiarism, courts have traditionally considered three major aspects of music infringement lawsuits: 1) copyright ownership, 2) accessibility, and 3) substantial similarity [2]. "substantial similarity", which is the most crucial yet debatable factor, lacks a complete definition with no general agreement [3, 4, 5] due to the varying requisite level from case to case [5]. Court analyses are inconsistent within the same circuit, making it more a matter of quality than quantity [6, 7].

Melodic similarity is usually the determining element in assessing whether or not two musical works are substantially similar [8, 6]. Melody is the most memorable and characteristic feature of music [9, 10], and many cases involve the plagiarism of the melody of an original work [11, 9, 12]. Although numerous studies have developed various quantitative measures of melodic similarity [13, 12, 14, 15], it still remains unclear what constitutes substantial similarity. While a high degree of melodic similarity may suggest plagiarism, it does not necessarily indicate plagiarism. Instead, substantial parts of an existing work that are considered essential and worthy of protection can be crucial in determining plagiarism. For example, in the case of *Hawkes & Sons v. Paramount Film Services* (1934, as cited by [16] and [17]), twenty seconds (of 4 minutes) of a musical work without permission was deemed infringement. Therefore, the use of any "recognizable" parts may establish infringement, even if the overall similarity of the pieces is questionable [17].

This study aims to develop a novel approach for quantitatively evaluating the substantial similarities of melodies by employing natural language processing (NLP) techniques. Due to the shared characteristics between music and language [18, 19, 20], various NLP approaches have been applied to music analysis in different ways [21, 22, 23, 24]. The primary focus of the proposed approach is to define the individual elements of melody using NLP-based methods. To achieve this, we employ Mel2word [25], a novel method for melody segmentation using NLP tokenization techniques to represent melodies as word-like units and capture semantic information through word embeddings. In addition, two word weighting methods are proposed to understand the characteristics of individual melodic elements: a modified Tversky measure for *word salience* and the TF-IDF method for *word importance* and *word uniqueness*. The method is evaluated on 108 plagiarism cases with court decisions and perceptual similarity as ground truth, compiling data from diverse sources to represent one of the most extensive symbolic melodic datasets available. This study provides detailed case analyses, showcasing the numerical and graphical representation of the proposed method and its practical applications. By doing so, we aim to provide empirical and quantitative

[1] *Millett v. Snowden*, available at: https://blogs.law.gwu.edu/mcir/case/millett-v-snowden/

evidence for the qualitative aspects of substantial similarities in music.

2. LITERATURE REVIEW

There have been numerous studies on plagiarism detection based on melodic similarities, which can be broadly categorized into two types of approaches: (1) audio-based and (2) text-based.

Audio-based approaches employ music signal processing to develop plagiarism detection tools that can identify similar parts of music [26, 27, 28, 29]. While they use advanced audio-based analysis techniques to determine the level of similarity between songs, they mainly focus on identifying similarities rather than explaining how the degree of plagiarism is related to the level of similarity. The audio-based approaches are particularly useful in plagiarism cases involving unauthorized sampling or use of musical works. However, for research purposes related to artistic analysis, notated music provides more useful information than audio-based analysis [15].

Text-based approaches analyze symbolic musical representations, such as notated music. The study by [12] is a remarkable attempt to quantitatively model court decisions in plagiarism cases. This study compared several similarity calculation algorithms and investigated how melodic similarity calculated by text-based algorithms relates to court decisions based on a sample of US copyright cases from 1970. The study unveiled that an algorithm rooted in statistical methods, notably Tversky's similarity measures [30], outperformed in predicting court decisions. This finding was further corroborated by research conducted by [31]. Percent Melodic Identity (PMI) also stands out as another major measure in this context. Drawing from automatic sequence alignment algorithms in the field of molecular genetics, [32] introduced the PMI method to quantify melodic similarity, which was further utilized by [33, 34] to successfully predict plagiarism. Recent advancements in music research have demonstrated significant progress, particularly in utilizing vectorized representations. These include fuzzy vector-based approaches [8], CNN-based methods [35], and hybrid approaches [36].

While previous studies have explored the quantitative similarities between melodies, the specific elements contributing to plagiarism and the underlying reasons remain unclear. This gap highlights the need for further investigation into what exactly constitutes melodic plagiarism and "*why*" these particular elements are implicated. To address this, we propose an NLP-based approach to define individual melodic elements by defining words as the basic unit of text to reconstruct a melody as a sentence of meaningful word units. We also introduce a function that combines psychological and NLP models, particularly Tversky and TF-IDF approaches, aiming to provide a comprehensive framework for understanding melodic similarity.

3. METHODS

The proposed method involves three steps: 1) segmenting melodies using Mel2Word [25], 2) vectorizing melodies

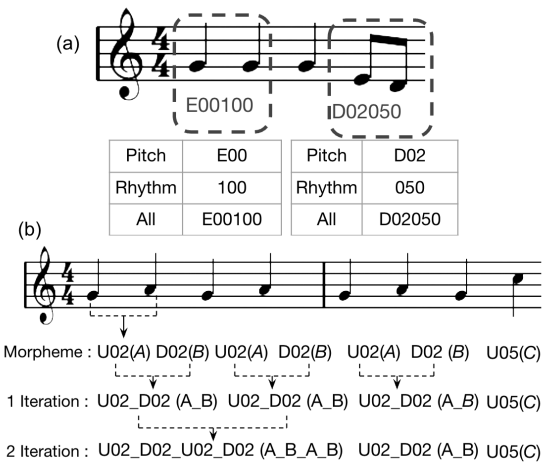

Figure 1: Example of (a) the Mel2Word representation with (b) Byte Pair Encoding (BPE) process.

using the Word2Vec [37] algorithm, and 3) applying word-by-word weighted measures to determine word salience, importance, and uniqueness.

3.1 Textual Representation

Mel2word is a novel text-based representation method to segment melodies into word-like units [25]. In this textual representation, each note is translated into a pitch feature indicating the interval's direction and size, alongside a rhythm feature denoted by the inter-onset interval (IOI) between consecutive notes. Specifically, pitch features are represented by the first character indicating the melody's direction ("U" for upward,"D" for downward, and "E" for no change), followed by a two-digit number specifying the interval size. Rhythm features are depicted with three-digit numbers, obtained by multiplying the IOI by 100, assuming a quarter note equals one beat with a 16th note quantization. This unit, as depicted in Figure 1-(a), composed of two notes, form "morphemes" utilized for constructing a melody word dictionary using Byte Pair Encoding, a commonly used tokenization technique in the field of NLP.

3.2 Byte-Pair Encoding

Mel2Word represents melodies as word-like units using Byte Pair Encoding (BPE), a data-driven NLP method. BPE is a bottom-up method that builds a vocabulary for computational text analysis by replacing frequently occurring byte pairs with a single and less frequently used byte [38]. Originally developed for data compression, BPE has found widespread adoption due to its successful application in word segmentation for NLP tasks [39]. The utilization of BPE in music, as implemented in Mel2word has been effectively adopted for melody analysis, classifying folk song families and jazz artists [25,40]. Similar to its application in language, this method involves creating subwords or *tokens* based on the frequency of consecutive pairs. In other words, it identifies the most frequent consecutive pairs in the melody and merges them into a single unit. As a result, the most frequent pairs are combined using an underscore ('_') symbol. Figure 1-(b) illustrates the

basic BPE process and the resulting token outcomes. [2]

3.3 Word Embedding

Word embedding is a vector representation that captures the meaning and relationships of words by representing them as dense, distributed, and fixed-length vectors based on their context in text. Built on the distributional hypothesis [41], it maps words onto a high-dimensional space, placing similar words close together. In music information retrieval, word embeddings have been used to analyze and model relationships between melodic elements. Specifically, the Word2Vec model [37] has been successfully employed in previous studies to represent notes [42, 43], chords [44, 45] or motifs [46, 47] in a distributed vector space. To capture the semantic analysis of melodic elements, we utilize the Word2Vec model in this study.

3.4 Toward the Substantiality of Melody

In determining the substantiality of music, the court has considered the "distinctive characteristics" of the subject matter as a crucial factor [48]. To evaluate the distinctive features of a melody, we propose two methods drawn from the fields of psychology and NLP: 1) assessing the *salience* of a word or how noticeable it is, and 2) evaluating the *importance* and *uniqueness* of a word or how important and rare it is.

3.4.1 Word Salience

The Tversky ratio is a formula for similarity proposed by Amos Tversky, a cognitive psychologist who suggested that human perceptions and judgments of similarity are based on the number of features two objects have in common and the salience of these features [30]. Tversky's formula is given by:

$$s(A,B) = \frac{|A \cap B|}{(|A \cap B| + \alpha|A \setminus B| + \beta|B \setminus A|)} \quad (1)$$

where A and B are sets, $|A \cap B|$ is the number of common elements in A and B, $|A \setminus B|$ is the number of elements in A that are not in B, $|B \setminus A|$ is the number of elements in B that are not in A. The parameters α and β adjust the impact of the unique elements of A and B respectively, with higher $s(A,B)$ indicating stronger similarity. In the context of melody, features and elements could refer to components such as note pitch or inter-onset interval.

Since the original Tversky model does not account for the individual salience of specific components, we introduce a modified measure specifically designed to evaluate the significance of individual melodic elements. This adaptation evaluates the significance of each melodic element by considering its prevalence in two melodies and its distribution within each, providing a refined perspective on their commonality and relative frequency. To evaluate the significance of elements shared between two melodic sequences A, B and an element x of A, we propose a salience measure $TV_{A,B}(x)$. When a_x and b_x represent the counts of element x in sequences A and B respectively, along with the lengths l_A and l_B of the sequences, the formula for $TV_{A,B}(x)$ is given by:

$$TV_{A,B}(x) = \frac{\frac{a_x}{l_A}}{\frac{a_x}{l_A} + \alpha\left(1 - \frac{a_x}{l_A}\right) + \beta\left(1 - \frac{b_x}{l_B}\right)} \quad (2)$$

Here, α and β are coefficients designed to adjust for the lengths of A and B, calculated as $\alpha = \frac{l_A}{l_A + l_B}$ and $\beta = \frac{l_B}{l_A + l_B}$.[3] The $TV_{A,B}(x)$ measure evaluates the salience of element x from the perspective of sequence A, taking into account both shared and unique elements. This method allows for a balanced evaluation across sequences, aligning with Tversky's concept of asymmetrical similarity. By incorporating α and β, the measure provides a nuanced assessment of each element's salience, considering its frequency within the sequences and the overall sizes of the melodies. This approach ensures a standardized measure, assigning a salience score ranging from 0 (indicating no shared elements) to 1 (indicating fully shared), facilitating equitable comparisons regardless of sequence length.

3.4.2 Word Importance and Uniqueness

TF-IDF is a widely used algorithm in NLP that measures the importance and uniqueness of a term in a document compared to a collection of documents. It takes into account the frequency and rarity of each term in the document and the corpus, respectively. The TF component considers the relevance of a term proportional to its frequency in the document, while the IDF component measures its rarity in the corpus. If a term is frequently used in the corpus, it is considered less representative of a specific document, and if it is rare, it is considered more relevant to a specific document. The TF-IDF value is obtained by multiplying the TF and IDF scores of a term in a document. The formula is as follows:

$$\text{TF-IDF}(t,d) = \text{TF}(t,d) \times \text{IDF}(t) \quad (3)$$

where t represents a term or word in a document, d, $\text{TF}(t,d)$ represents the term frequency of t in d, and $\text{IDF}(t)$ represents the inverse document frequency of t in a collection of documents. In this study, each token in a melody is treated as a single unit of text and the entire melody as a single document. To determine the importance and uniqueness of each melodic element, the TF and IDF scores are utilized as weightings for the word embedding vectors, respectively.

4. EXPERIMENTS

This section describes our empirical investigation of analyzing copyright infringement cases.

[2] Figures 1-(a) and (b) are sourced from [25]. More details on the Mel2word representation and the BPE process, including the subsequent steps of Dictionary Generation (Section 4.2) and Tokenization (Section 4.3) are found at [25].

[3] The constant 1 is derived from $\frac{l_A}{l_A}$ for sequence A and $\frac{l_B}{l_B}$ for sequence B.

4.1 The Dataset

We collected copyright infringement cases from various sources, including previous research and law school databases. We prioritized the data provided by [34], who used a similar sampling approach to [12] and included updated metadata with perceptual data. [4] Additionally, we extensively reviewed cases from the Music Copyright Infringement Resource (MCIR) [5] and Lost in Music by Westminster Law School [6] to compile a comprehensive analysis, aiming to consider as many legal cases as possible. After the landmark case of *Arnstein v. Porter*, which established the concept of substantial similarity, our analysis delved into an extensive repository of legal documents and accompanying materials, up to 314 cases from 1946 to 2023. Following an in-depth review, we excluded cases lacking audio or sheet transcription, those not involving similar musical elements (e.g., licensing, sampling, arrangements, rap lyrics, etc.), and those without relevant expert commentary and opinions for the rulings. As a result, we collected and transcribed into MIDI data on 116 cases (Infringed N=32, Denied N=66, Settled N=18), encompassing 232 songs. We included settlement cases collected in our database; however, for evaluation analysis, we included only settlements with official payments or public records of royalties or credit. In total, we analyzed 108 cases in this study.

4.2 Dictionary Generation

We utilized the Meertens Tune Collection - Folk Song dataset (MTC-FS) to train our BPE model, consisting of over 18,000 monophonic melodies from Dutch sources spanning five centuries [49]. The MTC-FS is one of the largest monophonic datasets, offering a rich repository of melodies that have influenced both classical and modern music. We selected this dataset for its diverse range rooted in oral transmission across generations, providing a strong foundation for analyzing copyright infringement cases across various eras and styles. With BPE applied to the MTC-FS dataset, we initially constructed a base dictionary, which serves as the primary resource for tokenization in subsequent analyses. We constructed the base dictionary using the Mel2word code by [25] [7], applying BPE to extract words with a minimum frequency of 10 occurrences and limiting the maximum unit size to 11 to prevent redundancy.

4.3 Melody Tokenization

For tokenization, we utilized subsets of the base dictionary to enable tokenization with dictionaries of varying sizes for different levels of segmentation. The subsets were selected based on the most frequent tokens in the base dictionary. For instance, choosing 100 tokens would produce a dictionary with the 100 most frequent entries for tokenization. We relied on statistics from the base dictionary for the maximum length (Mode) and minimum count parameters (Q1, 1st quartile). Consequently, we tokenized melodies from copyright-infringed cases for subsequent analyses using dictionaries of sizes N=100, 500, 1000, and *Full-token* [8], which indicates the maximum number of words available with the parameter settings. Figure 2 illustrates an example of the resulting melody tokenization in our dataset.

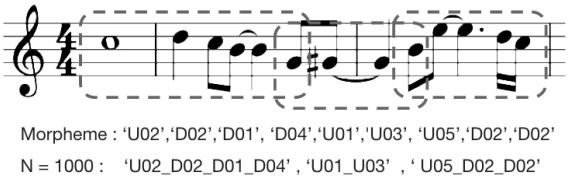

Figure 2: An example of melody tokenization (Dictionary N=1000, pitch feature)

4.4 Melody Embedding

To build semantic word embeddings for melodic tokens, we utilized Word2Vec embedding in our experiment. Using the MTC-FS dataset, we tokenized all songs for different dictionary sizes (N=100, 500, 1000, and Full) and trained the corresponding Word2Vec models for each size. We used the Gensim module [50], a Python implementation of the Word2Vec [9], with a dimension size of 512, a window size of 10, a minimum count of 2, and the skip-gram model option, which is known to better represent sparse words [51].

4.5 Similarity Calculation

Cosine similarity is a widely used measure of similarity between two vectors that quantifies the cosine of the angle between the two vectors in a high-dimensional space. In this study, we used the cosine similarity to quantify the similarity between two songs in infringement cases. In order to determine the essential effectiveness of different methods, we opted to calculate melody vectors by averaging as a baseline approach. Although vector summarization through averaging involves a loss of information, it also brings several advantages, such as simplicity in computation, low storage memory requirement, and faster processing speed [52]. Consequently, to generate the melody vectors for each song, we calculated the average of all word vectors for each word unit using the trained Word2Vec model.

4.6 Weight Functions

To assess individual melodic elements, we employed multiple weight functions. These weights are utilized for each token when calculating the average vector of words to derive the final melody vector. For each weight function,

[4] Except for case 14 (*Vargas v. Pfizer)*, as the supplied MIDI data did not contain a melody.

[5] https://blogs.law.gwu.edu/mcir/

[6] https://www.lostinmusic.org/

[7] https://github.com/saebyulpark/Mel2word

[8] With dictionary sizes of N=2399 for pitch, N=1184 for rhythm, and N=3112 for both pitch and rhythm

[9] https://radimrehurek.com/gensim/

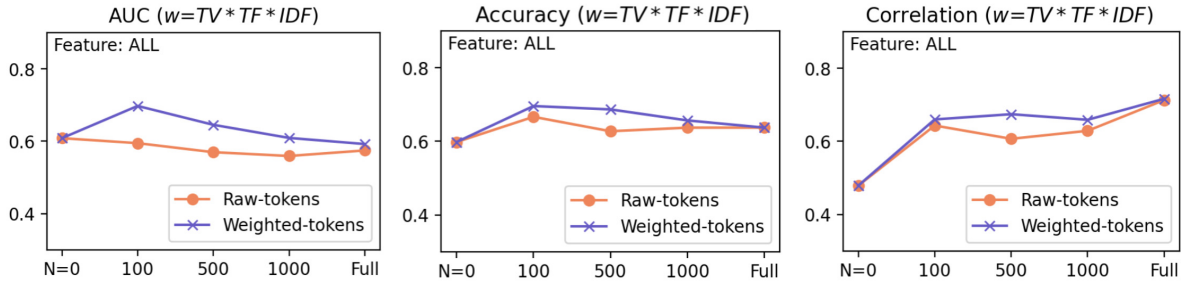

Figure 3: Summary of results with various dictionaries and weight presence.

Weight Function(w)	Foundational Method
TV	Tversky Model
TF	TF-IDF
IDF	TF-IDF
$TF * IDF$	TF-IDF
$TV * TF$	Tversky + TF-IDF
$TV * IDF$	Tversky + TF-IDF
$TV * TF * IDF$	Tversky + TF-IDF

Table 1: Summary of Weight Functions

we defined specific parameters, as summarized in Table 1. We calculated term frequency (TF) and inverse document frequency (IDF) values using the *TfidfVectorizer* module from the sklearn library [10] with the default settings. We also computed the Tversky value (TV) using the formula described in Section 3.4.1 (Equation 2). To avoid zero-multiplication for hybrid variables (e.g., $TV * TF * IDF$), we added one to each variable and multiplied it by the subsequent value. Finally, all weights were experimented with various normalization methods.

4.7 Evaluation

We conducted three types of evaluations, following the previous work [32, 33, 34]. First, we assessed how well the similarities evaluated by the algorithm corresponded to the court's decision. To measure this, we computed the Area Under the ROC Curve (AUC), a commonly used method to evaluate binary classification performance. Second, we utilized AUC to determine TPR and FPR at different thresholds, identifying the threshold with the highest accuracy (ACC). Finally, we measured how well the similarities correlated with human perceptual data provided by [34] [11], for which we computed the Pearson coefficient only for the subset of songs with perceptual data available.

5. RESULTS

5.1 Overall Result

Figure 3 presents an overview of the results considering different dictionaries and the presence of weights, with combined feature (Pitch + Rhythm) and $TV * TF * IDF$ weights achieving the highest scores. While the tokenization method has a minor impact on AUC and ACC metrics, it notably influences the correlation with perceptual data, showing better performance across all dictionary sizes. Additionally, the adoption of weights generally enhances performance across most cases (except for morpheme-level and Full-level).

Regarding measures related to legal decisions (AUC and ACC), as previously discussed about performance limits [34], once again, we found that the proposed method was more effective in correlating with perception than with court decisions. This is likely due to courts considering various factors such as lyrics, arrangement, and other musical elements, as well as the worthiness of a melody to be protected (e.g., *Intersong-USA v. CBS*). Additionally, they consider the possibility of subconscious copying (e.g., *Francis Day & Hunter v. Bron*), and proof of access to the original work (e.g., *Ellis v. Diffie*), even when similarities between melodies exist. Since our study targeted all possible cases involving melody, there may be various confounding variables.

Interestingly, we noticed that the weight function underperformed when analyzing melodies at the morpheme-level ($N = 0$), possibly due to the high number of randomly shared features at this level. Performance also decreased at the Full-level, likely because longer words led to a decrease in shared features. Additionally, we found that applying all weights multiplied by the Tversky model improved performance, while the default TF and IDF weights tended to reduce performance. This result supports the basic assumption of Tversky's model that we perceive similarity based on how many features are shared, which is consistent with previous research [53, 12, 31] showing a strong association of the Tversky model with infringement decisions and perceptual similarity.

5.2 Comparison Results with Previous Studies

Table 2 compares evaluations conducted on subsamples of 17 (N=17) [12] and 39 (N=39) [13] cases each, facilitating comparison with existing literature. As observed, our method performed remarkably well, achieving the highest scores for both sets. [14] These subsets consist of cases with

[10] https://scikit-learn.org/

[11] This data consists of a similarity scale ranging from 0 to 5 points, where 0 represents dissimilarity and 5 represents similarity, available at: https://github.com/comp-music-lab/music-copyright-expanded

[12] Based on [33], which includes 14 songs from [32].

[13] While [34] included 40 cases; we analyzed 39, excluding *Vargas v. Pfizer* due to the absence of melody.

[14] Pitch feature, N=100, with quantile Gaussian normalization for 17 cases; pitch + rhythm, N=100 (AUC) and Full (ACC) with quantile Gaus-

Cases		Savage [32]	Yuan1 [33]	Yuan2 [34]	Proposed
N=17	AUC	0.69	0.61	0.61	0.94
	ACC	0.80	0.71	0.71	0.94
N=39	AUC	N/A	N/A	0.73	0.79
	ACC	N/A	N/A	0.75	0.79

Table 2: Comparison Results with Previous Studies

a significant indication of melodic similarity from the outset, making them subjects of a number of previous studies [12, 31, 32, 33, 34]. Therefore, they exhibited effective discrimination based solely on the melody itself, compared to our overall findings. Given that our study examined the entire melodies obtained from the archive, we anticipate that further investigation focusing on specific parts or cases emphasizing the melody will yield even more intriguing results.

5.3 Exploratory Result Analysis

Beyond the performance, the strength of our method lies in its ability to numerically represent the characteristics of each melodic element within a song. For example, Figure 4 illustrates a copyright infringement case, *Three Boys Music v. Michael Bolton*, where distinctive and shared melodic features are quantified using TV, TF, IDF, and $TVTFIDF$. In this manner, by examining these melodies, we can observe the numerical values of their importance, uniqueness, and the degree to which they are shared for each melodic element. Moreover, this can play a crucial role when melodies are tokenized into more meaningful units, potentially enhancing their interpretability. For example, Figure 5 presents a cross-scape plot visualization, which provides a hierarchical analysis of the similarities between two songs, indicating where and how they are similar [54]. The left side represents the infringed case (*Three Boys Music v. Michael Bolton*), while the right side represents the denied case (*Baxter v. MCA, Inc.*). On the left, (a) depicts the morpheme-level, while on the right, (b) showcases the token-level melody with weighting applied. [15] As observed, at the morpheme level, segmentation of each note leads to overall similarity across all parts due to the frequently shared elements at the note level. However, in the weighted tokenized songs, certain crucial phrases in the infringed case stand out notably darker (i.e., more similar). This visually demonstrates how our approach highlights specific parts that contribute to a stronger similarity between two pieces of music. In this way, by providing a quantitative method to identify the individual characteristics of melody elements, our research can be of significant help in practical applications such as legal analysis, as well as various fields of music research.

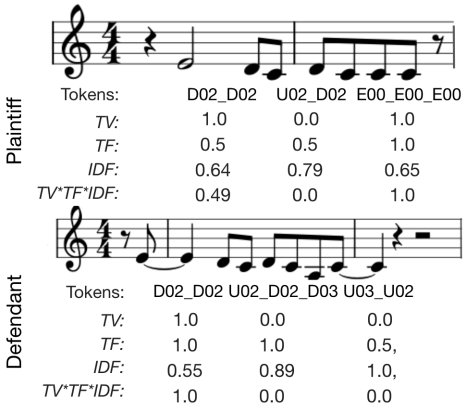

Figure 4: An example of melody weighting values (*Three Boys Music v. Michael Bolton*, N=100, pitch feature)

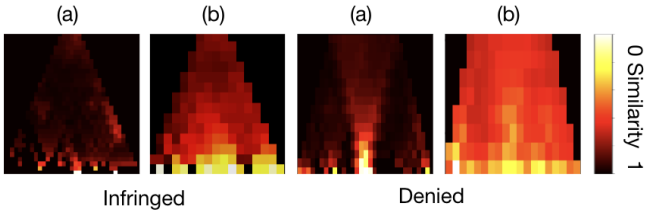

Figure 5: Cross-scape plots: (a) Word2Vec at morpheme-level, (b) Word2Vec at token-level with $TV * TF * IDF$.

6. CONCLUSION

In this study, we employed natural language processing (NLP) techniques to objectively grasp the substantial similarity of melodies, thereby making notable contributions in several key areas: First, after an extensive review of legal documents, we compiled one of the most extensive public datasets, the Music Copyright Infringement Collection (MCIC). [16] Although it is not big data, this dataset is significant given the limited number of legal cases, as it includes MIDI transcriptions, sheet music, and metadata on legal issues and decisions, forming the crucial groundwork for future studies on music similarity and copyright issues. Second, we encoded melodies into word-like units using Mel2word to analyze melodic similarity for the music plagiarism study. This approach extends semantic analysis beyond the note- or n-gram level, surpassing conventional analysis methods. Third, we introduced the modified-Tversky measure to evaluate the salience of each melodic element. Derived from a prominent psychological theory, this refined measure offers potential applications beyond music, exhibiting its general versatility. Moreover, by incorporating traditional NLP-based weighting algorithms, we conducted an in-depth analysis of individual features to comprehensively grasp substantial similarity. Thus, by integrating computational methods, psychological models, data-driven techniques, and rule-based approaches, we performed a detailed exploration of melodic similarity.

sian normalization for 39 cases.

[15] The original plot was modified to compare song similarities using word vectors. Details and base code for the cross-scape plot are at [54] and `https://github.com/saebyulpark/cross_scapeplot`.

[16] `https://github.com/saebyulpark/MCIC`. This site includes supplementary materials with comprehensive experimental details, including transcriptions, statistics, normalization methods, additional results, and full and sub-dataset lists.

7. ETHICS STATEMENT

This research adheres to ethical guidelines, ensuring data integrity and confidentiality. All sources are credited, and only publicly available data were used.

8. REFERENCES

[1] A. Keyt, "An improved framework for music plagiarism litigation," *Cal. L. Rev.*, vol. 76, 1988.

[2] S. J. Jones, "Music copyright in theory and practice: An improved approach for determining substantial similarity," *Duq. L. Rev.*, vol. 31, 1992.

[3] A. B. Cohen, "Masking copyright decisionmaking: The meaninglessness of substantial similarity," *UC DAViS l. reV.*, vol. 20, 1986.

[4] M. F. Sitzer, "Copyright infringement actions: the proper role for audience reactions in determining substantial similarity," *S. Cal. L. Rev.*, vol. 54, 1980.

[5] C. A. Tschider, "Automating music similarity analysis in "sound-alike" copyright infringement cases," *Entertainment, Arts and Sports Law Journal*, vol. 25, no. 2, 2014.

[6] S. N. Hamilton, D. Majury, and D. Moore, *Sensing Law*. Taylor & Francis, 2016.

[7] I. Stav, "Musical plagiarism: A true challenge for the copyright law," *DePaul J. Art Tech. & Intell. Prop. L*, vol. 25, 2014.

[8] R. De Prisco, D. Malandrino, G. Zaccagnino, and R. Zaccagnino, "Fuzzy vectorial-based similarity detection of music plagiarism," in *FUZZ-IEEE*. IEEE, 2017.

[9] E. Selfridge-Field, "Conceptual and representational issues in melodic comparison," *Melodic similarity, concepts, procedures, and applications*, 1998.

[10] R. Typke, "Music retrieval based on melodic similarity," *Ph.D thesis*, 2007.

[11] J. P. Fishman, "Music as a matter of law," *Harv. L. Rev.*, vol. 131, p. 1861, 2017.

[12] D. Müllensiefen and M. Pendzich, "Court decisions on music plagiarism and the predictive value of similarity algorithms," *Musicae Scientiae*, vol. 13, no. 1_suppl, 2009.

[13] D. Müllensiefen and K. Frieler, "Cognitive adequacy in the measurement of melodic similarity: Algorithmic vs. human judgments," *Computing in Musicology*, vol. 13, no. 2003, 2004.

[14] V. Velardo, M. Vallati, and S. Jan, "Symbolic melodic similarity: State of the art and future challenges," *Computer Music Journal*, vol. 40, no. 2, 2016.

[15] J. S. Downie, "Music information retrieval," *Annual Review of Information Science and Technology*, vol. 37, no. 1, 2003.

[16] L.-J. Stark and T. J. Perfect, "The effects of repeated idea elaboration on unconscious plagiarism," *Memory & cognition*, vol. 36, no. 1, pp. 65–73, 2008.

[17] B. Challis, "The song remains the same: A review of the legalities of music sampling," *Wipo Magazine*, vol. 6, 2009.

[18] A. D. Patel, "Language, music, syntax and the brain," *Nature neuroscience*, vol. 6, no. 7, 2003.

[19] B. Maess, S. Koelsch, T. C. Gunter, and A. D. Friederici, "Musical syntax is processed in broca's area: an meg study," *Nature neuroscience*, vol. 4, no. 5, 2001.

[20] S. Koelsch, E. Kasper, D. Sammler, K. Schulze, T. Gunter, and A. D. Friederici, "Music, language and meaning: brain signatures of semantic processing," *Nature neuroscience*, vol. 7, no. 3, 2004.

[21] D. Conklin, "Representation and discovery of vertical patterns in music," in *International Conference on Music and Artificial Intelligence*. Springer, 2002.

[22] J. Wołkowicz, Z. Kulka, and V. Kešelj, "N-gram-based approach to composer recognition," *Archives of Acoustics*, vol. 33, no. 1, 2008.

[23] M. Mongeau and D. Sankoff, "Comparison of musical sequences," *Computers and the Humanities*, vol. 24, no. 3, 1990.

[24] T. Crawford, "String matching techniques for musical similarity and melodic recognition," *Computing in musicology*, vol. 11, 1998.

[25] S. Park, E. Choi, J. Kim, and J. Nam, "Mel2word: A text-based melody representation for symbolic music analysis," *Music & Science*, vol. 7, 2024.

[26] C. Dittmar, K. F. Hildebrand, D. Gärtner, M. Winges, F. Müller, and P. Aichroth, "Audio forensics meets music information retrieval—a toolbox for inspection of music plagiarism," in *2012 Proceedings of the 20th European signal processing conference*. IEEE, 2012, pp. 1249–1253.

[27] K. Suneja and M. Bansal, "Comparison of time series similarity measures for plagiarism detection in music," in *Annual IEEE India Conference*, 2015.

[28] S. De, I. Roy, T. Prabhakar, K. Suneja, S. Chaudhuri, R. Singh, and B. Raj, "Plagiarism detection in polyphonic music using monaural signal separation," *arXiv preprint arXiv:1503.00022*, 2015.

[29] N. Borkar, S. Patre, R. S. Khalsa, R. Kawale, and P. Chakurkar, "Music plagiarism detection using audio fingerprinting and segment matching," in *Smart Technologies, Communication and Robotics*, 2021.

[30] A. Tversky, "Features of similarity." *Psychological review*, vol. 84, no. 4, 1977.

[31] A. Wolf, D. Müllensiefen *et al.*, "The perception of similarity in court cases of melodic plagiarism and a review of measures of melodic similarity," in *International Conference of Students of Systematic Musicology*, 2011.

[32] P. E. Savage, C. Cronin, D. Müllensiefen, and Q. D. Atkinson, "Quantitative evaluation of music copyright infringement," in *Proceedings of the 8th International Workshop on Folk Music Analysis*. Thessaloniki Greece, 2018.

[33] Y. Yuan, S. Oishi, C. Cronin, D. Müllensiefen, Q. Atkinson, S. Fujii, and P. E. Savage, "Perceptual vs. automated judgments of music copyright infringement," in *Proceedings of the 21st International Society for Music Information Retrieval Conference*, 2020.

[34] Y. Yuan, C. Cronin, D. Müllensiefen, S. Fujii, and P. E. Savage, "Perceptual and automated estimates of infringement in 40 music copyright cases," *Transactions of the International Society for Music Information Retrieval*, vol. 6, no. 1, 2023.

[35] K. Park, S. Baek, J. Jeon, and Y.-S. Jeong, "Music plagiarism detection based on siamese cnn," *Human-centric Computing and Information Sciences*, 2022.

[36] D. Malandrino, R. De Prisco, M. Ianulardo, and R. Zaccagnino, "An adaptive meta-heuristic for music plagiarism detection based on text similarity and clustering," *Data Mining and Knowledge Discovery*, 2022.

[37] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013.

[38] P. Gage, "A new algorithm for data compression," *C Users Journal*, vol. 12, no. 2, 1994.

[39] R. Sennrich, B. Haddow, and A. Birch, "Neural machine translation of rare words with subword units," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 2016.

[40] S. Park and J. Nam, "The language of jazz: A natural language processing-based analysis of the patterns and vocabulary of jazz solo improvisation," in *17th International Conference on Music Perception and Cognition*, 2023.

[41] Z. S. Harris, "Distributional structure," *Word*, vol. 10, no. 2-3, 1954.

[42] C.-H. Chuan, K. Agres, and D. Herremans, "From context to concept: exploring semantic relationships in music with word2vec," *Neural Computing and Applications*, vol. 32, no. 4, 2020.

[43] D. Herremans and C.-H. Chuan, "Modeling musical context with word2vec," in *Proceedings of the First International Conference on Deep Learning and Music*, 2017.

[44] C.-Z. A. Huang, D. Duvenaud, and K. Z. Gajos, "Chordripple: Recommending chords to help novice composers go beyond the ordinary," in *Proceedings of the 21st International Conference on Intelligent User Interfaces*, 2016.

[45] S. Madjiheurem, L. Qu, and C. Walder, "Chord2vec: Learning musical chord embeddings," in *Proceedings of the constructive machine learning workshop at 30th conference on neural information processing systems*, 2016.

[46] A. A. Alvarez and F. Gómez-Martin, "Distributed vector representations of folksong motifs," in *International Conference on Mathematics and Computation in Music*, 2019.

[47] T. Hirai and S. Sawada, "Melody2vec: Distributed representations of melodic phrases based on melody segmentation," *Journal of Information Processing*, vol. 27, 2019.

[48] R. P. Smith, "Arrangements and editions of public domain music: Originally in a finite system," *Case W. Res. L. Rev.*, vol. 34, 1983.

[49] P. Van Kranenburg and M. De Bruin, "The meertens tune collections: Mtc-fs-inst 2.0," *Meertens Online Reports*, vol. 2019, no. 1, 2019.

[50] R. Řehůřek and P. Sojka, "Software Framework for Topic Modelling with Large Corpora," in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 2010.

[51] A. J. Landgraf and J. Bellay, "Word2vec skip-gram with negative sampling is a weighted logistic pca," *arXiv preprint arXiv:1705.09755*, 2017.

[52] M. A. Kharazmi and M. Z. Kharazmi, "Text coherence new method using word2vec sentence vectors and most likely n-grams," in *2017 3rd Iranian Conference on Intelligent Systems and Signal Processing*. IEEE, 2017, pp. 105–109.

[53] D. Müllensiefen and K. Frieler, "Measuring melodic similarity: Human vs. algorithmic judgments," in *Proceedings of the Conference on Interdisciplinary Musicology*, 2004.

[54] S. Park, T. Kwon, J. Lee, J. Kim, and J. Nam, "A cross-scape plot representation for visualizing symbolic melodic similarity," in *Proceedings of the 20th International Society for Music Information Retrieval Conference*, 2019.