

Recepción: 18 de mayo de 2025

Aprobación: 15 de Noviem de 2025

Evaluación de la reforma de la enseñanza de la dirección coral universitaria en la era de los nuevos medios mediante un algoritmo estadístico de reconocimiento de melodías corales

Assessing College Chorus Conducting Teaching Reform in the New Media Era via a Statistical Choral Melody Recognition Algorithm

Avaliação da reforma do ensino de regência coral universitária na era das novas mídias por meio de um algoritmo estatístico de reconhecimento de melodias corais

Daniel Roberto Vega Torres

Doctor en Ciencias Sociales Aplicadas Fundación Universitaria Juan de Castellanos (Tunja, Colombia)dvega@jdc.edu.co

¿Cómo citar este artículo?

Daniel Roberto Vega Torres (2025). Evaluación de la reforma de la enseñanza de la dirección coral universitaria en la era de los nuevos medios mediante un algoritmo estadístico de reconocimiento de melodías corales. *Cultura Científica*, (23), pp. 1–11.

RESUMEN

La educación universitaria en dirección coral recibe una atención cada vez mayor y, debido a su marcado carácter práctico, exige una reforma pedagógica que vaya más allá de los métodos tradicionales centrados en el docente, hacia enfoques que integren mejor la teoría con la práctica de ensayo. Para apoyar dicha reforma, este estudio propone un método de reconocimiento de melodías corales para escenarios de enseñanza en nuevos medios y lo utiliza para evaluar los resultados de la reforma docente. El método se formula dentro de un marco de modelado estadístico adaptado del reconocimiento continuo del habla, lo que ofrece una mayor robustez en comparación con los enfoques convencionales no estadísticos. En la etapa de modelado acústico, se adoptan coeficientes cepstrales de alto orden para reducir la degradación del rendimiento causada por las decisiones de sonoridad y los errores en la estimación de la frecuencia fundamental. Para incorporar aún más conocimiento previo musical y mejorar la generalización entre tonalidades, se entrena un modelo de cuater-

niones independiente de la tonalidad. El reconocimiento final de la melodía se obtiene mediante decodificación de Viterbi. Los resultados experimentales indican que el algoritmo propuesto mantiene un buen rendimiento de reconocimiento en condiciones ruidosas, mientras que su precisión en grabaciones limpias es cercana a la de los sistemas con mejor desempeño entre las líneas base consideradas. A partir de estos resultados, el artículo examina direcciones clave para la reforma de la enseñanza de la dirección coral universitaria en la era de los nuevos medios—como la retroalimentación basada en datos, la evaluación de ensayos asistida por tecnología y el diseño instruccional híbrido—ofreciendo referencias prácticas para mejorar la pedagogía de la dirección coral en la educación superior.

Palabras clave: Dirección coral universitaria, pedagogía de la dirección coral, reforma de la enseñanza, entorno de nuevos medios, reconocimiento de melodías, modelado estadístico.

ABSTRACT

College choral conducting education is increasingly emphasized and, due to its strong practical nature, calls for instructional reform that moves beyond traditional teacher-centered methods toward approaches that better integrate theory with rehearsal practice. To support such reform, this study proposes a chorus melody recognition method for new-media teaching scenarios and uses it to assess teaching reform outcomes. The method is formulated within a statistical modeling framework adapted from continuous speech recognition, yielding improved robustness compared with conventional non-statistical approaches. In the acoustic modeling stage, high-order cepstral coefficients are adopted to reduce performance degradation caused by voicing decisions and fundamental-frequency estimation errors. To further embed musical prior knowledge and enhance generalization across keys, a key-independent quaternion model is trained. Final melody recognition is produced via Viterbi decoding. Experimental results indicate that the proposed algorithm maintains strong recognition performance under noisy conditions, while its accuracy on clean recordings is close to the best-performing systems among the baselines considered. Building on these results, the paper examines key directions for college chorus conducting teaching reform in the new media era—such as data-informed feedback, technology-assisted rehearsal evaluation, and blended instructional design—providing practical references for improving choral conducting pedagogy in higher education.

Keywords: College choral conducting, chorus conducting pedagogy, teaching reform, new media environment, melody recognition, statistical modeling.

RESUMO

A educação universitária em regência coral tem sido cada vez mais valorizada e, devido à sua forte natureza prática, requer uma reforma pedagógica que vá além dos métodos tradicionais centrados no professor, avançando para abordagens que integrem melhor a teoria com a prática de ensaio. Para apoiar essa reforma, este estudo propõe um método de reconhecimento de melodias corais para cenários de ensino em novas mídias e o utiliza para avaliar os resultados da reforma do ensino. O método é formulado em um arcabouço de modelagem estatística adaptado do reconhecimento contínuo de fala, oferecendo maior robustez em comparação com abordagens convencionais não estatísticas. Na etapa de modelagem acústica, adotam-se coeficientes cepstrais de alta ordem para reduzir a degradação de desempenho causada por decisões de sonoridade e erros na estimativa da frequência fundamental. Para incorporar ainda mais conhecimento prévio musical e aumentar a generalização entre tonalidades, treina-se um modelo de quatérnions independente de tonalidade. O reconhecimento final da melodia é obtido por meio de decodificação de Viterbi. Os resultados experimentais indicam que o algoritmo proposto mantém forte desempenho de reconhecimento em condições ruidosas, enquanto sua precisão em gravações limpas se aproxima da dos sistemas de melhor desempenho entre as linhas de base consideradas. Com base nesses resultados, o artigo examina direções-chave para a reforma do ensino de regência coral universitária na era das novas mídias—como feedback orientado por dados, avaliação de ensaios assistida por tecnologia e desenho instrucional híbrido—oferecendo referências práticas para aprimorar a pedagogia da regência coral no ensino superior.

Palavras-chaves: Regência coral universitária, pedagogia da regência coral, reforma do ensino, ambiente de novas mídias, reconhecimento de melodias, modelagem estatística.

1. INTRODUCCIÓN

Algunos docentes han intentado integrar Internet y las tecnologías de nuevos medios en el aula (Zhang and Zhang, 2022a; Chew, 2022); sin embargo, todavía no se ha establecido un sistema de enseñanza sistemático. Por un lado, las instalaciones de enseñanza basadas en nuevos medios en colegios y universidades no están plenamente generalizadas, y muchas instituciones carecen de equipos y plataformas de difusión para la docencia con nuevos medios, permaneciendo en una fase exploratoria. Por ello, el profesorado universitario debe aprovechar al máximo los recursos y plataformas disponibles en la era de los nuevos medios, incorporar nuevas tecnologías en la enseñanza en el aula, explorar modelos modernos de instrucción en dirección coral y buscar métodos más eficaces para mejorar los resultados de aprendizaje y fomentar una mayor innovación (Vizcaíno-Verdú and Aggraded, 2022; Shi, 2023).

En el contexto de la formación en dirección en instituciones de educación superior, el ritmo de toda la actividad es altamente dinámico, lo que exige la integración de enlaces de transición (Zhang and Zhang, 2022b). Esto mejora el efecto global y ofrece al público una mejor experiencia (Weng and Chen, 2020b). La relación entre el coro y el director es muy estrecha, y ambos son mutuamente dependientes. No obstante, algunos docentes presentan una comprensión superficial del curso de dirección musical, centrándose únicamente en explicar técnicas de dirección sin abordar aspectos más amplios de la dirección desde la perspectiva coral. Esta brecha de comprensión afecta la eficacia de la enseñanza (Guan et al., 2024; Qi and Liu, 2022).

Al enseñar dirección, el profesorado debe innovar continuamente sus métodos pedagógicos para garantizar una formación de alta calidad (Lu et al., 2022). Sin embargo, algunos docentes siguen apoyándose en modelos tradicionales de enseñanza, lo que vuelve pasivo el aprendizaje del estudiantado y genera un ambiente de aula monótono, conduciendo en última instancia a resultados docentes deficientes.

Para estudiantes de música, el ritmo es un elemento crucial (Weng and Chen, 2020a; Yang, 2022). Por tanto, los docentes deben priorizar este aspecto y fortalecer el sentido del ritmo musical mediante métodos de entrenamiento específicos. Al animar al estudiantado a comprender el ritmo musical a través del movimiento corporal, pueden asimilarlo con rapidez y contribuir a una interpretación coral más cohesionada (Schindler, 2020). Los directores corales deben poseer una sólida capacidad de apreciación musical, que les permita captar la esencia de una obra durante la interpretación y reinterpretarla creativamente para aumentar el atractivo y el impacto global del desempeño (Zheng and Dai, 2022). En consecuencia, también es esencial desarrollar las habilidades de apreciación musical del estudiantado

(Schuller et al., 2022). Por ejemplo, los docentes pueden exponer de manera regular a los estudiantes a piezas musicales diversas para potenciar su capacidad de apreciación.

El reconocimiento de melodías corales consiste en identificar la secuencia de notas a partir de una señal coral y se utiliza principalmente como etapa frontal de un sistema de consulta coral (Sun, 2022). Muchos sistemas actuales de reconocimiento coral necesitan analizar la secuencia de notas presente en la señal para realizar la comparación de melodías (Xiaoyu et al., 2023). Así, la precisión del reconocimiento melódico incide directamente en el rendimiento del sistema de consulta coral. Otras aplicaciones del reconocimiento de melodías incluyen la entrada conveniente de música MIDI sin dispositivos electrónicos de entrada musical, así como sistemas de puntuación para karaoke.

La mayoría de los algoritmos existentes para el reconocimiento de melodías corales se basan en métodos no estadísticos (Lim et al., 2023), utilizando algoritmos de extracción de frecuencia fundamental basados en procesamiento de señales y diversos procesos de segmentación de notas (Thayyib et al., 2023). Estos métodos suelen carecer de robustez frente a distintos hablantes y entornos ruidosos. En consecuencia, algunos investigadores han explorado modelos estadísticos para el reconocimiento de melodías. En (Wenxuanzi and Li, 2020) se propuso un algoritmo de reconocimiento estadístico de dos etapas. Primero, la segmentación de notas se realiza mediante un modelo oculto de Markov entrenado con características MFCC, dividiendo la señal coral en segmentos etiquetados como “nota” o “silencio”. Luego, el tono de cada segmento se etiqueta usando un modelo de mezclas gaussianas entrenado con características de frecuencia fundamental. Para asegurar una segmentación precisa, dicho algoritmo asume que la señal coral contiene únicamente ciertas sílabas, como “da” o “ta”. El algoritmo de reconocimiento melódico propuesto por (Pachet et al., 2021) y (Thayyib et al., 2023) integra un modelo de trayectoria de tono basado en Markov oculto y un modelo musical basado en una gramática binaria, asumiendo que el tempo de la melodía coral es conocido.

El algoritmo propuesto en este artículo se apoya en marcos probabilísticos y modelos provenientes de sistemas avanzados de reconocimiento continuo del habla. En el algoritmo, cada nota (y el silencio) se trata como una palabra independiente. Al entrenar el modelo acústico, se emplean coeficientes cepstrales de alto orden como características, lo que evita decisiones rígidas sobre sonoridad y estimación de la frecuencia fundamental, reduciendo el impacto negativo de estos errores en el rendimiento del sistema. Simultáneamente, se entrena un modelo de lenguaje cuaternario independiente de la tonalidad para reflejar la distribución de probabilidad a priori de diferentes secuencias de notas en la canción. Finalmente, los resultados de reconocimiento melódico se obtienen mediante el algoritmo de búsqueda global de Viterbi.

2. ALGORITMO DE RECONOCIMIENTO DE MELODÍAS CORALES

En un sistema de reconocimiento continuo del habla, sea $X = X_1 X_2 \dots X_n$ la secuencia de características del habla. La secuencia de palabras $\hat{W} = w_1 w_2 \dots w_m$ generada por el sistema maximiza la probabilidad a posteriori $P(W | X)$, lo cual puede expresarse como:

$$\hat{W} = \arg \max_W P(W | X) = \arg \max_W \frac{P(W)P(X | W)}{P(X)}. \quad (1)$$

Para aplicar la tecnología de reconocimiento continuo del habla al reconocimiento de melodías, la secuencia de palabras puede interpretarse como una secuencia de notas. Mediante el entrenamiento del modelo acústico y del modelo de lenguaje correspondientes, el resultado del reconocimiento melódico puede obtenerse usando el marco del reconocimiento continuo del habla.

2.1. MODELO ACÚSTICO

La mayoría de los sistemas existentes de reconocimiento de melodías utilizan la frecuencia fundamental como característica. Aunque existe una relación directa entre las características de la frecuencia fundamental y la altura (pitch) de la nota, los errores en la estimación de la frecuencia fundamental (y en la determinación de sonoridad) no pueden evitarse, incluso con algoritmos avanzados. En un entorno ruidoso, este problema se vuelve más pronunciado. Si la frecuencia fundamental se usa como característica, los errores en su estimación afectarán negativamente los resultados del reconocimiento melódico. Por ello, para mejorar la robustez, este trabajo adopta coeficientes cepstrales de alto orden como características acústicas.

La Figura 1 muestra, respectivamente, las características cepstrales de una trama de una señal con una altura correspondiente al do central (C4) y de una trama de señal silenciosa. Para mayor claridad, se ha filtrado la parte de bajo orden del cepstrum, que no está relacionada con la información de frecuencia fundamental.

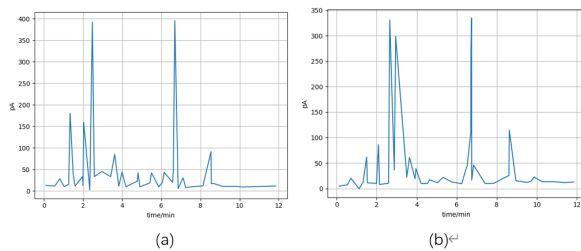


Figura 1. Características cepstrales de una trama de (a) una señal con la altura del do central (C4) y (b) una señal silenciosa. La parte de bajo orden ha sido filtrada

La frecuencia de muestreo es 16 kHz. Puede observarse en la Figura 1 que existen picos evidentes en el cepstrum de la señal sonora, mientras que el cepstrum de la señal silenciosa es relativamente plano. El retardo de orden N del cepstrum y su frecuencia correspondiente satisfacen la siguiente relación: $\text{lag } N \times f(\text{lag } N) = \text{frecuencia de muestreo}$.

Las características cepstrales mostradas en la Figura 1 se procesan como sigue antes de utilizarse para entrenar el modelo acústico. Primero, se recorta del cepstrum completo la porción correspondiente al rango de frecuencia fundamental de la voz humana. En los experimentos, para voces masculinas este rango va desde el retardo $N = 48$ hasta el retardo $N = 240$ (con muestreo a 16 kHz), lo que corresponde al rango de alturas C2 a E4 en un teclado de piano; para voces femeninas, el rango va desde $N = 24$ hasta $N = 120$, correspondiente a C3 a E5 en un teclado de piano. Luego, la parte recortada del cepstrum se normaliza a una longitud fija promediando los valores cepstrales de órdenes adyacentes. En los experimentos, dicha longitud fija es 24. Cabe señalar que, en este proceso, la señal cepstral de voz masculina se comprime el doble que la de voz femenina, lo cual equivale a duplicar el valor de frecuencia fundamental contenido en la señal masculina.

En los experimentos, los datos utilizados para entrenar el modelo acústico consistieron en tres hablantes (dos hombres y una mujer). Los datos de cada hablante incluyen dos tipos: el Tipo 1 contiene coros con la sílaba 'da', y el Tipo 2 contiene letras (lyrics). La cantidad total de datos de Tipo 1 es de aproximadamente 45 minutos, mientras que la cantidad total de datos de Tipo 2 es de aproximadamente 1 hora. Los datos se grabaron en un entorno de oficina con una frecuencia de muestreo de 16 kHz. Todos los datos corales se etiquetan como un semitono en el teclado del piano según la altura real de las notas.

Cada semitono (y el silencio) entre E3 y D5 corresponde a un HMM. Cada modelo HMM contiene tres estados, y la función de densidad de probabilidad de cada estado se modela mediante un GMM. De manera similar al proceso de entrenamiento del modelo acústico para el reconocimiento continuo del habla, todos los modelos HMM se entrenan simultáneamente usando el algoritmo forward-backward. La Figura 2 muestra el vector medio de los modelos GMM de todos los estados en los modelos HMM de dos notas, C4 y C5, tras el entrenamiento.

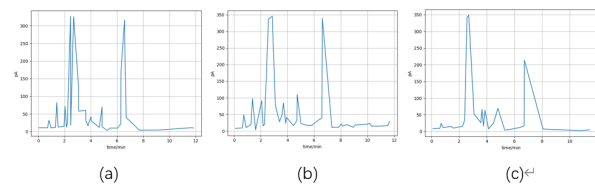


Figura 2. Vector medio del modelo GMM de todos los estados en el modelo HMM de las notas (a) C4 y (b) C5

2.2. MODELO DE LENGUAJE

En teoría musical, cada tonalidad contiene siete notas, y cada octava consta de doce semitonos. Esto implica que la probabilidad de aparición de una nota en una canción no es uniforme; en su lugar, existe una distribución de probabilidades a priori. Por tanto, para el reconocimiento de melodías se entrenan modelos de lenguaje sobre distribuciones a priori de secuencias de notas.

Los datos de entrenamiento del modelo de lenguaje provienen de la base de datos EsAC e incluyen 522 melodías. El entrenamiento del modelo de lenguaje en nuestros experimentos es independiente de la tonalidad. Primero, todas las melodías se normalizan a una misma tonalidad, luego se transponen a las otras once tonalidades y, finalmente, se combinan para formar una gran base de datos para entrenar el modelo de lenguaje. La ventaja de este enfoque es que elimina la necesidad de determinar explícitamente la tonalidad de la melodía coral. Tras una comparación experimental, se seleccionó un modelo de lenguaje cuaternario.

Para evaluar los resultados de reconocimiento melódico en términos del rendimiento global del sistema de consulta coral, desarrollamos un sistema base (baseline) de consulta coral. El sistema convierte tanto la melodía de la biblioteca como las señales de consulta coral en secuencias de notas y luego realiza el emparejamiento (matching) basado en estas secuencias. Finalmente, se genera una lista de melodías coincidentes.

Dado que la melodía de la biblioteca y la melodía coral consultada pueden estar en tonalidades distintas, durante el emparejamiento las notas se expresan como alturas relativas, que representan la diferencia entre las alturas de dos notas adyacentes. En el sistema base actual no se utiliza la información de duración de las notas. El proceso de emparejamiento de melodías puede dividirse en tres pasos: (1) emparejamiento rápido, donde se identifican con rapidez posibles puntos candidatos de coincidencia basándose en el índice de la secuencia de notas n-gram; (2) emparejamiento grueso, que utiliza un algoritmo de emparejamiento de cadenas de menor complejidad para acotar el rango de melodías candidatas; (3) emparejamiento fino, que utiliza un algoritmo de programación dinámica para determinar el resultado final del emparejamiento.

En el emparejamiento rápido, las melodías de la biblioteca se indexan previamente según la secuencia de n-gramas (altura relativa). Al realizar el emparejamiento, se consulta el índice a partir de la secuencia de alturas relativas contenida en la melodía coral. Dado que pueden existir errores en los resultados de reconocimiento, este trabajo adopta una estrategia de emparejamiento difuso (fuzzy matching) para expandir la secuencia de notas identificada a partir de la señal coral en una red, como se muestra en la Figura 3. Cada nodo representa una nota producida por el algoritmo de reconocimiento melódico, y la arista entre dos nodos

representa la altura relativa entre las notas. Además de la altura relativa obtenida directamente del resultado de reconocimiento, también se añade a la red una altura relativa que difiere en un semitono. Todas las secuencias de alturas relativas contenidas en la red se utilizan para recuperar el índice.

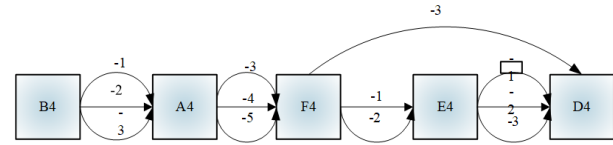


Figura 3. Red de melodía coral utilizada en el emparejamiento rápido

Los puntos candidatos de coincidencia devueltos por el emparejamiento rápido se representan como vectores ternarios QkP , donde P es la posición inicial de la secuencia n-gram en la red de melodía coral, y k y Q son, respectivamente, la posición inicial de la secuencia n-gram correspondiente en la melodía candidata y el número de serie de la melodía en la biblioteca.

Al determinar el valor de n , si dicho valor es demasiado pequeño, el número de puntos candidatos devueltos por el emparejamiento rápido puede ser excesivo y el rango de búsqueda no puede acotarse eficazmente. Por el contrario, si el valor es demasiado grande, la cantidad de información contenida en la secuencia n-gram es relativamente alta y puede acotar el rango de búsqueda de forma efectiva, pero el recall del emparejamiento rápido disminuirá debido a posibles errores en el reconocimiento de melodías. Mediante experimentos, este trabajo fija $n = 4$ (4 notas, lo que implica 3 alturas relativas). La biblioteca de melodías utilizada en el experimento incluye 1325 melodías.

Tanto el emparejamiento grueso como el emparejamiento fino emplean algoritmos de emparejamiento de cadenas. En la etapa de emparejamiento grueso se utiliza un algoritmo de baja complejidad. Sea $a_1 a_2 \dots a_m$ la secuencia de alturas relativas identificada a partir de la señal coral, y $b_1 b_2 \dots b_n$ la secuencia de alturas relativas contenida en la melodía candidata. Si se asume que, en cierto punto del proceso, a_i coincide con b_j , denotado como p_{ij} , entonces el siguiente punto de coincidencia en la trayectoria de emparejamiento es:

$$\text{Next}(p) = \arg \min \begin{cases} d(a_{i+1}, b_{j+1}) \\ d(a_i, b_{j+1}) + d(\phi, b_{j+1}) \\ d(a_{i+1}, b_j) + d(a_{i+1}, \phi) \end{cases} \quad (2)$$

En comparación con el algoritmo de dynamic time warping (DTW), este algoritmo conserva únicamente una trayectoria durante el proceso de emparejamiento, lo que lo hace más eficiente que DTW. Tras calcular el coste de emparejamiento grueso de todos los puntos candidatos devueltos por el emparejamiento rápido, se ordenan y se devuelven los

N mejores puntos candidatos al módulo de emparejamiento fino.

El módulo de emparejamiento fino adopta el algoritmo DTW. El algoritmo calcula la matriz de costes $D_{0...m,0...n}$ con las condiciones iniciales:

$$\begin{cases} D_{0,0} = 0, \\ D_{0,j} = \text{INF}, & 1 \leq j \leq n, \\ D_{i,0} = \text{INF}, & 1 \leq i \leq m. \end{cases} \quad (3)$$

3. RESULTADOS EXPERIMENTALES

En la plataforma de pruebas, se llevaron a cabo experimentos de rendimiento para evaluar el cálculo de la distancia de edición, los métodos DTW y OSCM en cinco casos, con volúmenes de datos de 4500, 9000, 18 000, 36 000 y 72 000 canciones, respectivamente. También se realizaron pruebas simuladas con 1500 consultas que contenían 3 errores aleatorios.

3.1. EXPERIMENTOS SOBRE LA SELECCIÓN DE PARÁMETROS DEL ALGORITMO OSCM

Dado que el algoritmo OSCM permite que el usuario seleccione los parámetros, en esta sección se exploran los efectos de diferentes selecciones de parámetros en el rendimiento del algoritmo.

En el algoritmo OSCM, la base de datos utiliza el código binario de cada segmento como código hash para el índice de contorno melódico. La longitud n del segmento de características tiene un impacto significativo en el rendimiento del algoritmo. Un n mayor, es decir, un segmento más largo, da como resultado menos registros en la base de datos correspondientes al elemento de índice de ese segmento, lo que produce menos resultados intermedios de la consulta y, por lo tanto, velocidades de consulta más rápidas. Los resultados experimentales que se muestran en la figura 4 confirman esta expectativa.

El valor de n también afecta a la tolerancia al error del algoritmo. Si la entrada no tiene errores, o contiene errores coherentes con el contorno de la melodía, la tasa de aciertos del algoritmo es del 100%. Si la entrada contiene errores opuestos al contorno de la melodía, el algoritmo demuestra una buena tolerancia a los fallos. Según el método de cálculo de la coincidencia continua unilateral, la tasa de aciertos de los tres primeros resultados seguirá siendo alta. Sin embargo, si la entrada incluye dos o más errores en dirección opuesta al contorno de la melodía, estos errores afectan significativamente a la tasa de aciertos del algoritmo. Si la distancia entre dos errores adyacentes en dirección opuesta al contorno de la melodía es inferior a n , el algoritmo calcula una similitud baja, lo que da lugar a

una tasa de aciertos más baja. Según el modelo de error de coro de usuarios en (Weng and Chen, 2020b), la distancia entre dos errores de coro de usuarios adyacentes que son opuestos al contorno de la melodía es generalmente mayor que 4. Por lo tanto, cuando n es 3, el algoritmo debería mostrar una buena tolerancia a fallos. Se seleccionaron para la prueba los casos en los que n es 3, 4 y 5. Como se muestra en la figura 5, cuando $n = 3$, la tasa de aciertos de la consulta inicial, con volúmenes de datos variables, es entre un 15% y un 20% superior a la de las 10 consultas principales con $n = 5$.

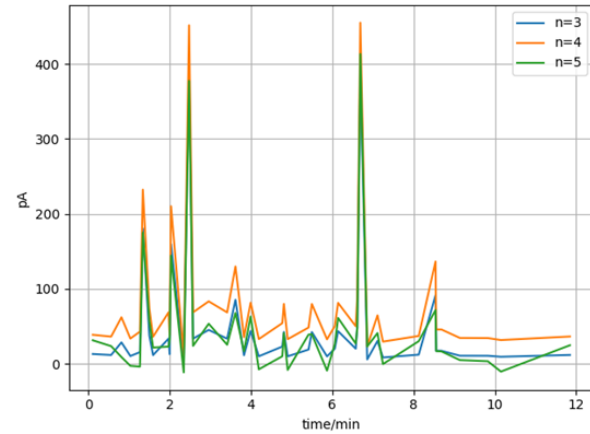


Figura 4. La influencia de la granularidad en el tiempo de respuesta de la consulta del algoritmo oscmAlgorithm

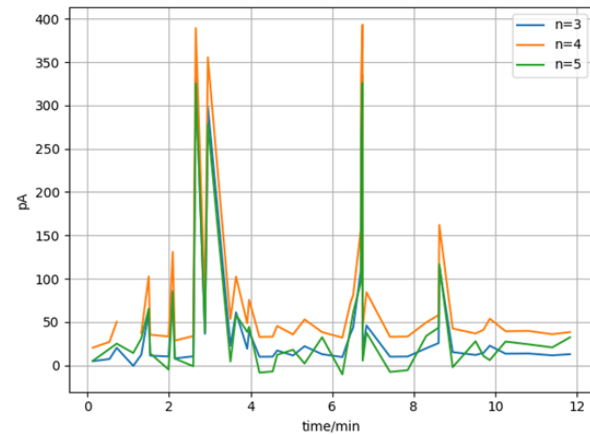


Figura 5. La influencia de la granularidad en las tasas de éxito de las consultas del algoritmo OSCM

En el algoritmo OSCM, el número de niveles del valor de la característica del contorno melódico también tiene un impacto significativo en el rendimiento del algoritmo. Cuando el número de niveles es grande, el número de registros correspondientes a cada elemento del índice de segmento disminuye, lo que da lugar a menos resultados intermedios de la consulta y, por lo tanto, a velocidades de consulta más rápidas. Los resultados experimentales que se muestran en la figura 6 son coherentes con este análisis. Cuando la base de datos contiene 72 000 canciones, el tiempo de respuesta de la consulta utilizando el método de contorno melódico

8 Cultura Científica

de 3 niveles es aproximadamente 2,5 veces más rápido que el que utiliza el método de contorno melódico de 5 niveles.

Del mismo modo, un aumento en el número de niveles da como resultado una representación más precisa de la entrada de la consulta, lo que mejora la tasa de aciertos de la consulta. Además, dado que la representación de la característica sigue siendo un contorno melódico, conserva una buena tolerancia a fallos. Los resultados experimentales que se muestran en la figura 7 demuestran que la tasa de aciertos de la consulta del algoritmo varía con el número de piezas de la base de datos musical para diferentes estrategias de clasificación del contorno melódico. Se puede observar que, cuando se emplea la estrategia de clasificación del contorno melódico de 5 niveles, la tasa de aciertos de las 10 primeras posiciones del algoritmo de consulta es aproximadamente un 20% mayor en comparación con cuando se utiliza la estrategia de clasificación del contorno melódico de 3 niveles.

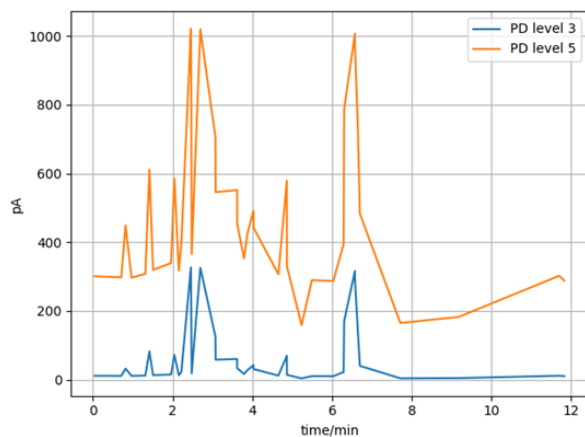


Figura 6. La estrategia de calificación de las características del contorno melódico frente a la tasa de éxito de las consultas del método o scm

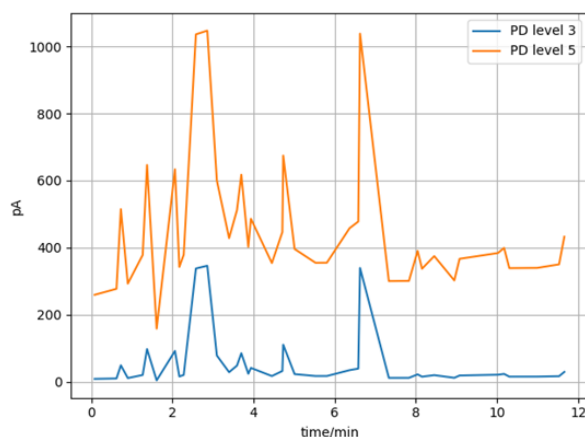


Figura 7. La estrategia de calificación de las características del contorno melódico frente a la tasa de éxito de la consulta del método SCM

En los algoritmos OSCM incluidos en la comparación, n es 3 y el contorno melódico adopta una estrategia de clasi-

ficación de 5 niveles. La figura 7 muestra las tres tasas de acierto de consulta más altas de los tres algoritmos en diferentes conjuntos de datos. La figura 8 muestra los tres algoritmos en diferentes tasas de acierto de consulta más altas en el conjunto de datos.

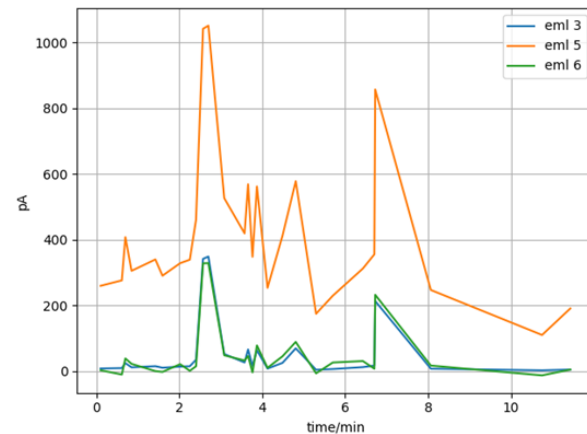


Figura 8. Número de errores de consulta frente a tasa de éxito de consulta del método O SCM

Como se muestra en las figuras 7 y 8, de los tres algoritmos, el algoritmo DIW tiene las tasas de aciertos de las 3 y las 10 consultas principales más altas. Cuando aumenta el número de canciones en la base de datos, la tasa de aciertos de la consulta del algoritmo DTW disminuye más lentamente. La tasa de aciertos de la consulta es casi la misma para volúmenes de datos de 18 000 y 72 000, y de manera similar, el volumen de datos no afecta significativamente a la tasa de aciertos entre 36 000 y 72 000. Esto indica que el algoritmo DTW tiene la mejor escalabilidad entre los tres algoritmos.

Por lo general, las fluctuaciones emocionales de los alumnos son significativas, pero su pensamiento racional mejora con la edad. Los profesores pueden aprovechar esta característica adaptando sus métodos de enseñanza a la edad y los rasgos de personalidad de los alumnos para estimular su interés por el aprendizaje. Entre las estrategias más comunes se encuentran los efectos psicológicos y la enseñanza rítmica. Por ejemplo, si los alumnos muestran falta de interés por las clases de música y prefieren dormir durante la clase, los profesores pueden orientarlos para que estudien al aire libre y disfruten de la naturaleza en los parques locales. Los profesores también pueden participar en estas actividades, lo que puede relajar enormemente a los alumnos, mejorar las relaciones con los profesores y fomentar el intercambio de puntos de vista y opiniones, mejorando así su capacidad general.

Teniendo en cuenta la tasa de aciertos y la velocidad de consulta, el algoritmo OSCM se considera el mejor de los tres algoritmos.

4. CONCLUSIÓN

Actualmente, muchos profesores utilizan métodos de enseñanza separados para los alumnos de «coro» y los de «dirección», lo que complica la integración entre ambos y afecta al desarrollo general de cada grupo. Para solucionar esto, los profesores deben crear escenarios interactivos en los que los alumnos puedan cambiar de papel con frecuencia, lo que permite a cada uno de ellos experimentar diferentes perspectivas. En resumen, para mejorar el nivel general de los directores de coro en las universidades, los profesores deben innovar y mejorar los métodos de enseñanza existentes, adaptándolos a las necesidades de los alumnos para aumentar la eficacia de la enseñanza.

FINANCIACIÓN

Este estudio no cuenta con financiación alguna.

REFERENCES

- Elaine Chew. Cosmos: Computational shaping and modeling of musical structures. *Frontiers in Psychology*, 13:527539, 2022.
- X. Guan, J. Li, Z. Zheng, and J. Shi. The experiences of teachers in teaching computer music in selected colleges and universities in hubei china. *Journal of World Englishes and Educational Practices*, 6(1):88–116, 2024.
- J. S. Lim, D. Shin, J. Zhang, S. Masiclat, R. Luttrell, and D. Kinsey. News audiences in the age of artificial intelligence: Perceptions and behaviors of optimizers, mainstreamers, and skeptics. *Journal of Broadcasting & Electronic Media*, 67(3):353–375, 2023.
- M. Lu, D. Pengcheng, and S. Yanfeng. [retracted] research on moral education function of music art in college students based on bayesian learning algorithm. *Security and Communication Networks*, 2022(1):1809364, 2022. Retracted.
- F. Pachet, P. Roy, and B. Carré. Assisted music creation with flow machines: towards new categories of new. In *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*, pages 485–520. 2021.
- L. Qi and N. Liu. [retracted] music singing based on computer analog piano accompaniment and digital processing for 5g industrial internet of things. *Mobile Information Systems*, 2022(1):4489301, 2022. Retracted.
- A. Schindler. Multi-modal music information retrieval: Augmenting audio-analysis with visual computing for improved music video analysis. arXiv preprint, 2020.
- B. W. Schuller, A. Akman, Y. Chang, H. Coppock, A. Gebhard, A. Kathan, others, and F. B. Pokorny. Climate change & computer audition: A call to action and overview on audio intelligence to help save the planet. arXiv preprint, 2022.
- Y. Shi. The use of mobile internet platforms and applications in vocal training: Synergy of technological and pedagogical solutions. *Interactive Learning Environments*, 31(6):3780–3791, 2023.
- J. Sun. Exploring the impact of music education on the psychological and academic outcomes of students: mediating role of self-efficacy and self-esteem. *Frontiers in Psychology*, 13:841204, 2022.
- P. V. Thayyib, R. Mamilla, M. Khan, H. Fatima, M. Asim, I. Anwar, others, and M. A. Khan. State-of-the-art of artificial intelligence and big data analytics reviews in five different domains: a bibliometric summary. *Sustainability*, 15(5):4026, 2023.
- A. Vizcaíno-Verdú and I. Aggraded. Thisis me challenge and music for empowerment of marginalized groups on tiktok. *Media and Communication*, 10(1):157–172, 2022.
- S. S. Weng and H. C. Chen. Exploring the role of deep learning technology in the sustainable development of the music production industry. *Sustainability*, 12(2):625, 2020a.
- S. S. Weng and H. C. Chen. Exploring the role of deep learning technology in the sustainable development of the music production industry. *Sustainability*, 12(2):625, 2020b.
- C. Wenxuanzi and T. Li. Visualized analysis and optimization countermeasures of the current situation of college aesthetic education research: Metrological analysis based on vosviewer. In *Proceedings of the 2020 3rd International Conference on E-Business, Information Management and Computer Science*, pages 268–277, December 2020.
- W. Xiaoyu, X. Ying, L. Shilin, R. Xiubo, Y. Dan, Y. Xinran, others, and X. Qiu Jian. Phonetics as a means of nationalising art songs: a comparative music-phonetics study based on zhao yuanren’s new poetry collection. *Journal of New Music Research*, 52(2-3):202–226, 2023.
- Z. Yang. Data analysis and personalized recommendation of western music history information using deep learning under internet of things. *PLOS ONE*, 17(1):e0262697, 2022.
- H. Zhang and M. Zhang. Innovation and practice of college music teaching in the 5g smart media era. In *Innovative Computing*, pages 439–447. Springer, Singapore, 2022a.
- H. Zhang and M. Zhang. Innovation and practice of college music teaching in the 5g smart media era. In *Innovative Computing: Proceedings of the 4th International Conference on Innovative Computing (IC 2021)*, pages 439–447. Springer Singapore, 2022b.

H. Zheng and D. Dai. Construction and optimization of artificial intelligence-assisted interactive college music performance teaching system. *Scientific Programming*, 2022(1):3199860, 2022.