

Comparación de Representaciones Interválicas Hansonianas para Recuperación de Información Musical

Johnny F. Serrano^{*αβ*} y José M. Iñesta^{*α*}

^{*α*} Departamento de Lenguajes y Sistemas Informáticos
Universidad de Alicante, España

^{*β*} Centro de Investigación en Informática Aplicada
Universidad Latina, Costa Rica
johnny@prais.net, inesta@dlsi.ua.es

Resumen

La similitud melódica es un concepto importante a considerar en la recuperación de información musical. Algunas de las posibles aplicaciones son sistemas basados en contenido desarrollados para administración de derechos de autor, detección de plagio de ideas ya expuestas por un artista en el pasado, la asistencia a la composición, etc. Existen varias técnicas expuestas en similitud melódica que utilizan diversos análisis estadísticos y probabilísticos. El objetivo en este trabajo es establecer un equivalente de la notación musical a palabras de texto utilizando tres representaciones basadas en simples relaciones interválicas, rítmicas y de duración, y evaluar tres de las técnicas de recuperación de información textual aplicadas a estas representaciones, además de proponer cambios para mejorar el rendimiento de sistema.

Palabras clave: Recuperación de Información Musical, Similitud Melódica.

1. Introducción

La representación de música mediante descriptores como duración, altura, contorno, intervalos, etc., es muy habitual en sistemas de análisis y recuperación de información musical [Marsden92, Hewlett01]. Las dos dimensiones en las que se codifica una partitura (vertical para las alturas y horizontal para las duraciones) como registro simbólico de la música han sido representadas de muy diferentes maneras para luego realizar análisis por medios probabilísticos como cadenas ocultas de Markov [Hu02], máxima entropía [Stuart04] y estadísticos a través de correlaciones [Stuart04], entre otros [Typke04].

Aquí se va a explorar el análisis interválico como unidad fundamental para la tarea similitud melódica, donde se analizan los fragmentos musicales por medio de sus relaciones interválicas en una escala temperada [Hanson60]. Se ha determinado al intervalo como la unidad básica de análisis, dejando de tomar en cuenta como descriptores directos la altura y la tonalidad. Algunos trabajos anteriores han utilizado este concepto haciendo una notación cartesiana de los intervalos [Clausen01] o determinando la correlación de las progresiones interválicas [Engelbrecht02]. Un conocido trabajo que hace especial hincapié en las relaciones entre intervalos contiguos es el de [Narmour90]. El autor establece una serie de grupos de intervalos contiguos (patrones de implicación - resolución) que

permiten analizar melodías mediante su transformación en símbolos. Un importante inconveniente de este enfoque es que ignora las relaciones de duraciones, lo que es útil para la recuperación de información

En el presente trabajo se explora la conversión a palabras de texto de los intervalos extraídos de documentos MIDI. Este enfoque no es totalmente novedoso en el sentido de que otros autores han utilizado previamente técnicas provenientes del análisis de textos para resolver problemas relacionados con el procesamiento de la información musical. Una posible codificación es la propuesta por [Downie99] o [Doraisami01] que han utilizado modelos basados en n-gramas para la identificación de melodías. Por otra parte, en [Pérez04] los autores utilizan modelos basados en palabras para identificar estilos musicales.

Además, se han utilizado tres técnicas de recuperación de información textual: Coseno [Frakes92], Coseno según [Kaszkiel99] y Coseno Pivotado [Singhal96] para comprobar la efectividad en la determinación de la recuperación de información relevante a través de similitud melódica. Además, se proponen algunas modificaciones a las técnicas mencionadas para mejorar el desempeño.

2. Metodología

2.1. Conversión a palabras de texto

2.1.1. Representación utilizando intervalos y sus duraciones

Varias representaciones basadas en intervalos y sus duraciones serán probadas para estudiar su rendimiento en la recuperación de información musical. A partir de un fragmento musical (ver Figura 1) se toma la primera nota y se calcula la distancia en semitonos con respecto a la siguiente nota. Esta distancia (intervalo) se representa por medio de las letras d-s-n-m-p-t del sistema hansoniano [Hanson60] más las cinco vocales a-e-i-o-u para juntas formar los intervalos de dos octavas y media¹[Flores79](ver Figura 2) De esta forma, los intervalos quedarán codificados mediante caracteres de esta manera:

¹ Estas sílabas se han utilizado por muchos años en la enseñanza del solfeo interválico sistemático.

Intervalo en semitonos: 1 2 3 4 5 6 7...
Codificación: da sa na ma pa ta pe...



Figura 1. Análisis de Intervalos. El código de la parte superior corresponde la base de datos RISM(ver sección 2.3)



Figura 2. Ejemplos de sílabas del solfeo interválico sistemático. En la parte superior del pentagrama se indican las sílabas correspondientes a la codificación del intervalo entre la nota intermedia y la superior de cada acorde y en la inferior las que codifican e intervalo entre la nota más baja y la intermedia

El trabajo original de Hanson estaba dirigido al análisis de la música moderna, por lo que es conveniente introducir algunas modificaciones para adaptar dicha codificación a los fines que persigue este trabajo. De esta manera, a la representación se le agrega una tercera letra: “l” para representar un intervalo descendente y “r” para un intervalo ascendente. La representación de un intervalo de cero semitonos no estaba considerada originalmente en el análisis de Hanson pero ha sido aquí para obtener una representación más completa de las melodías. Para ello se ha utilizado la palabra “loo”.

Además, se considera como la duración del intervalo la duración de las dos notas que lo conforman en unidades según su figura musical, donde para una nota negra se le da el valor de 1. De esta forma, la duración puede tomar valores fraccionarios. Un par de ejemplos de la representación interválica hansoniana más duración (IH+D) de la melodía se pueden apreciar en la Figura 3. Esta representación viene a ser una notación musical sencilla y comprensible de un fragmento musical, donde definiendo una nota base cualquiera puede reproducirse en diferentes tonalidades manteniendo su similitud [Flores79] De esta manera, al deslizar la tonalidad de la

melodía se simplifica el análisis y se enfoca a la relación entre los intervalos y sus duraciones.

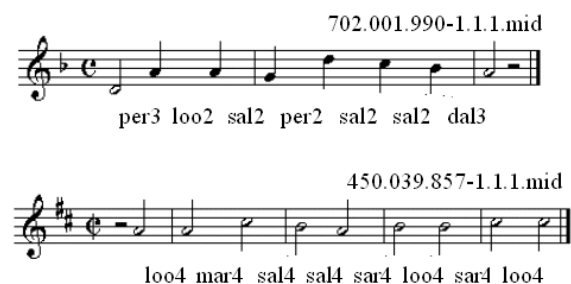


Figura 3. Resentación Interválica Hansoniana de dos fragmentos musicales.

2.1.2. Representación utilizando intervalos, análisis rítmico simple y sus duraciones

La representación IH+D da énfasis a las relaciones interválicas y generaliza las relaciones rítmicas que componen los intervalos, la segunda representación integrará a una cuarta letra donde “A” representa que la segunda nota del intervalo es de mayor duración que la primera, “D” representa que la segunda nota en de menor duración que la primera e “I” que representa que ambas notas que conforman el intervalos son de igual duración, además igual que IH+D se considera la duración del intervalo.

2.1.3 Representación utilizando intervalos, análisis rítmico detallado y sus duraciones

La tercera representación utilizará la palabra interválica hansoniana de IH+D y se agregará dos letras adicionales ([a-z][a-z]), para diferenciar todas las posibles relaciones de duración entre la notas que componen el intervalo, así la representación será interválica hansoniana con análisis rítmico detallado más duración(IHRC+D), que a diferencia de IH+D no generaliza la conformación rítmica del intervalo. Un ejemplo comparativo de las tres representaciones puede verse en la figura 4.



Figura 4. Comparación de los tres tipos de representación.

2.2. Sistema de Recuperación de Información

Uno de los objetivos del presente trabajo es exportar técnicas habitualmente utilizadas en extracción de información textual al ámbito de la información musical. De esta manera vamos a considerar que los documentos con los que trabajamos son fragmentos melódicos monofónicos y los términos de los que se componen estos documentos son las codificaciones textuales de cada uno de los intervalos presentes en dichos documentos.

Dada una base de datos de fragmentos melódicos se indexará mediante el siguiente procedimiento². Primero, los fragmentos musicales son convertidos a texto según la metodología descrita. Después, de cada fragmento se extrae un índice para el que se calcula la frecuencia de cada intervalo como la sumatoria de las duraciones de todos los intervalos del mismo tipo hallados en dicho fragmento. De esta manera, se está asimilando el concepto “frecuencia” propio del ámbito de la extracción de información a la duración de las notas en cada fragmento melódico. Para ello se determina:

- La frecuencia (es decir, la duración) de los términos (intervalos de un mismo tipo) en el documento (f_i) (ver Tabla 1). Por ejemplo, en el primer documento, el intervalo “sal” aparece en 3 ocasiones y con una duración de 2 unidades en cada intervalo, por lo que el valor del f_i del intervalo “sal” es 6.

Tabla 1. Indexación por frecuencias de términos correspondientes a los fragmentos de la Figura 3.

| Fragmento Musical | Intervalo | f_i |
|-----------------------|-----------|-------|
| 702.001.9901.1.1.mid | dal | 3 |
| | loo | 2 |
| | per | 5 |
| | sal | 6 |
| 450.039.857-1.1.1.mid | loo | 12 |
| | mar | 4 |
| | sal | 8 |
| | sar | 8 |

- La frecuencia de los documentos de la colección (f_d) que contienen cada uno de los términos (ver Tabla 2).

² El procedimiento es el mismo para los tres tipos de representación

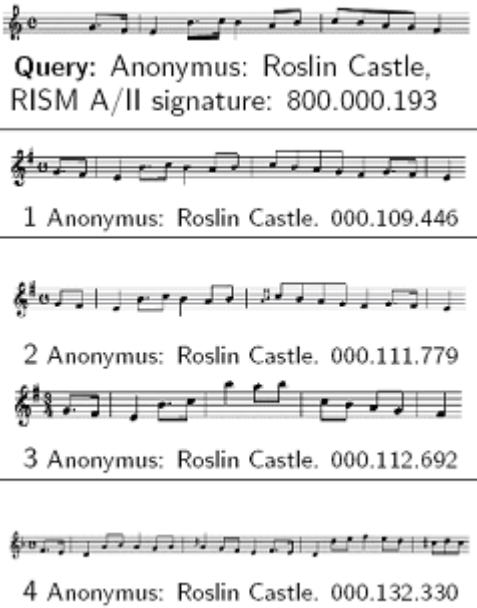
Tabla 2. Indexación por frecuencias de documentos de los dos fragmentos musicales de la Figura 3. Sólo hay dos documentos en la colección, por lo tanto $f_d \leq 2$.

| Intervalo | f_d |
|-----------|-------|
| dal | 1 |
| loo | 2 |
| mar | 1 |
| per | 1 |
| sal | 2 |
| sar | 1 |

2.3. Corpus

El corpus utilizado para la experimentación está compuesto de 581 fragmentos melódicos con un promedio de 14 intervalos por melodía. El formato de las melodías es MIDI. Estas melodías conforman las respuestas a 11 consultas generadas por el equipo de Rainer Typke de Utrecht University [Typke05a], siendo los documentos más similares (relevantes) evaluados y ordenados por 35 expertos en el campo musical a partir de la colección RISM A/II de más de 456000 melodías.

Las 11 consultas tienen alrededor de 50 melodías similares, éstas además están ordenadas según un grado de similitud y en grupos de igual similitud. En la Figura 5 se visualiza un ejemplo de consulta y un extracto de las melodías que resultan más relevantes tras ella.



Query: Anonymus: Roslin Castle,
RISM A/II signature: 800.000.193

1 Anonymus: Roslin Castle. 000.109.446

2 Anonymus: Roslin Castle. 000.111.779

3 Anonymus: Roslin Castle. 000.112.692

4 Anonymus: Roslin Castle. 000.132.330

Figura 5. Ejemplo de consulta (parte superior) con los 4 documentos que resultaron más relevantes, en orden de relevancia.

2.4. Evaluación

En esta sección se presentan las medidas propuestas para evaluar la precisión, cobertura y capacidad de los algoritmos utilizados en este trabajo para recuperar los documentos en un orden de relevancia adecuado. Para esta evaluación se utilizarán dos métodos:

- Evaluación según las competiciones del TREC³ [Rasmussen03], considerando todas las respuestas con el mismo grado de relevancia, donde se mencionarán el promedio de la precisión (no interpolada), R -precisión, la precisión a los 5 candidatos y gráficas de precisión frente a cobertura interpoladas.
- Evaluación ADR (average dynamic recall) [Typke05b], “promedio de cobertura dinámica” que toma en consideración además un grado de relevancia de cada respuesta, donde los grados de relevancia son determinados por grupos.

Considerando R como el número de documentos relevantes para una consulta:

- Se toma los primeros R candidatos devueltos por el sistema de recuperación de información.
- En cada división de grupo, se calcula la precisión.
- Se integra sobre la curva de precisión y se divide por R .

El resultado es un valor entre 1 y 0, donde un 1 equivale a recuperar la totalidad de documentos relevantes, en su orden correcto.

2.5. Experimentación

En la experimentación se utilizaron las técnicas de recuperación de información del Coseno, Kaszkiel y Coseno Pivotado. La evaluación inicial que se hizo utilizando las técnicas originales demostró en el caso de Kaszkiel y Coseno Pivotado los resultados no eran más favorables que la técnica del Coseno. Y se determinó que la normalización del logaritmo natural afectaba negativamente ya que atenuaba de forma excesiva las frecuencias, por lo que se reemplazó por la raíz cuadrada de las frecuencias. Asimismo en el caso del Coseno Pivotado la normalización se considera con respecto a las duraciones de los candidatos y no a su longitud en

³ TREC: Text Retrieval Conference.

Bytes. A continuación se muestran las fórmulas para el cálculo del peso de la consulta, peso del candidato y su similitud.

2.5.1. *Coseno:*

Dados dos documentos, uno de consulta, Q , y otro candidato, D , seleccionado de entre una base de datos en la que hay N documentos, se define la similitud entre ambos como:

$$sims(Q, D) = \frac{\sum_{i=1}^k q_i \cdot d_i}{\|Q\| \cdot \|D\|}, \quad (1)$$

donde $\|Q\| = \sqrt{\sum_{i=1}^k q_i^2}$, $\|D\| = \sqrt{\sum_{i=1}^k d_i^2}$ y

k es el número de términos en Q .

Los pesos de cada uno de los términos se definen como $q_i = ft_{q,i} \cdot \log_e\left(\frac{N}{fd_i}\right)$ para los términos de

la consulta y $d_i = ft_{d,i} \cdot \log_e\left(\frac{N}{fd_i}\right)$ para los de cada candidato. (2)

En estas expresiones, $ft_{d,i}$ representa la frecuencia del término i -ésimo del candidato y $ft_{q,i}$ la frecuencia del término i -ésimo de la consulta.

2.5.2. *Coseno según Kaszkiel:*

De forma equivalente, esta medida de similitud entre la consulta Q y los candidatos D se define como:

$$sims(Q, D) = \frac{\sum_{i=1}^k q_i \cdot d_i}{\|Q\| \cdot \|D\|}, \quad (3)$$

donde $\|Q\| = \sqrt{\sum_{i=1}^k q_i^2}$, $\|D\| = \sqrt{\sum_{i=1}^k d_i^2}$,

y los pesos de los términos: $d_i = \sqrt{ft_{d,i}}$,

$$q_i = \sqrt{ft_{q,i}} \cdot \log_e\left(\frac{N}{fd_i} + 1\right) \quad (4)$$

2.5.3. *Coseno Pivotado:*

La tercera de las medidas utilizadas para evaluar la similitud entre la consulta y los candidatos, se define como:

$$sims(Q, D) = \sum_{i=1}^k \frac{q_i \cdot d_i}{W_d}, \quad (5)$$

y los pesos: $d_i = 1 + \sqrt{ft_{d,i}}$

$$, q_i = 1 + \sqrt{ft_{q,i}} \cdot \log_e\left(\frac{N+1}{fd_i}\right) \quad (6)$$

El factor de normalización que aparece en la ecuación (6) sería:

$$W_d = (1 - slope) + slope \cdot \frac{L_I}{L_I}, \quad (7)$$

donde L_I es la duración total de los intervalos del candidato y $\overline{L_I}$ el promedio de la longitud total de los intervalos de los candidatos.

3. Resultados

3.1. Resultados Preliminares

La primera prueba fue evaluar el rendimiento aplicando los tres métodos y los tres tipos de representaciones, utilizando como peso de la consulta y peso del documento las duraciones respectivas. En el caso del Coseno se normalizaron las frecuencias entre 0 y 1 con respecto a la máxima frecuencia encontrada dentro del fragmento analizado. Y en el caso del Coseno Pivotado se utilizó la norma con respecto al promedio de la longitud de intervalos de los candidatos y un *slope* de 0.4. Las adaptaciones propuestas para la normalización de los pesos en los casos de Kaszkiel y coseno pivotado mejoraron los resultados respecto a los obtenidos en las pruebas preliminares con las ecuaciones originales. Los resultados de la evaluación ADR y de precisión a los 5 candidatos de las 11 consultas se aprecian en las Tabla 3 y 4.

Tabla 3. Resultados Preliminares según ADR⁴

| Representación | Coseno | Kaszkiel | Coseno Pivotado |
|----------------|--------|----------|-----------------|
| IH+D | 0.4569 | 0.4436 | 0.5518 |
| IHRS+D | 0.4573 | 0.4769 | 0.5372 |
| IHRC+D | 0.4210 | 0.4420 | 0.5198 |

⁴ ADR: Average Dynamic Recall

Tabla 4. Resultados Preliminares según Prec. at 5⁵

| Reprentación | Coseno | Kazkiel | Coseno Pivotado |
|--------------|--------|---------|-----------------|
| IH+D | 0.6364 | 0.5455 | 0.7091 |
| IHRS+D | 0.7091 | 0.6727 | 0.7818 |
| IHRC+D | 0.6364 | 0.7091 | 0.8182 |

Se puede apreciar que la técnica del Coseno Pivotado da los mejores resultados con respecto a las dos técnicas de Coseno y Kaszkiel, y que la representación IH+D proporciona el mejor rendimiento, con un 55.18% de ADR en promedio de las 11 consultas. Aunque las representaciones IHRS+D y IHRC+D tienen peores resultados, un aspecto importante es que alcanzan una precisión mayor que IH+D, esto quiere decir que el incluir el análisis rítmico ayuda al ordenamiento de los documentos, pero perdiendo cobertura.

3.2. Cambio de Normalización

Como el objetivo es encontrar los documentos más relevantes según la consulta, en el caso de similitud melódica es importante hacer un análisis más estricto de la duración de los intervalos. Para ello se establece el factor Rft_i como la relación entre la frecuencia en el candidato y la frecuencia en la consulta, este valor es calculado dividiendo el valor menor entre el mayor, y se obtiene un factor numérico entre 1 y 0 que indica el grado de similitud entre la duración de los intervalos del mismo tipo del candidato con respecto a la consulta:

$$\text{si } ft_{q,i} > ft_{d,i}, Rft_i = \frac{ft_{d,i}}{ft_{q,i}} \quad (8)$$

$$\text{si } ft_{q,i} < ft_{d,i}, Rft_i = \frac{ft_{q,i}}{ft_{d,i}} \quad (9)$$

A continuación se evalúa el Coseno utilizando como peso del candidato el factor Rft_i :

$$d_i = Rft_i \cdot \log_e\left(\frac{N}{fd_i}\right) \quad (10)$$

En el caso de Kaszkiel, el peso del candidato sería el factor Rft_i multiplicado por la frecuencia del término de la consulta:

$$d_i = \sqrt{Rft_i \cdot ft_{q,i}} \quad (11)$$

Y en el Coseno Pivotado el peso del candidato sería el factor $Rft_i + 1$ multiplicado por la frecuencia del candidato. Además que la norma será con respecto a la duración total de los intervalos de la consulta y no al promedio de la duración de los candidatos:

$$d_i = 1 + \sqrt{ft_{d,i} \cdot (Rft_i + 1)}, \quad (12)$$

Se modifica además la normalización del utilizando una relación entre duración y cantidad de intervalos que componen el candidato de la siguiente forma:

$$W_d = (1 - slope) \cdot \frac{\sqrt{L_1}}{\sqrt{L_{qI}}} + slope \cdot \frac{\sqrt{I_1}}{\sqrt{I_{qI}}} \quad (13)$$

donde L_{qI} es la duración total de los intervalos de la consulta, I_1 es la cantidad de intervalos que tiene el candidato y I_{qI} la cantidad de intervalos que tiene la consulta.

Los resultados de la evaluación según ADR y Prec. At 5 con un *slope* de 0.4, de las 11 consultas se aprecian en las Tablas 5 y en la Tabla 6. Se nota que la utilización del factor Rft_i mejora el rendimiento de todos los métodos. En método del Coseno Pivotado junto con la representación IH+D devuelve los mejores resultados, el incremento en la evaluación ADR mejora en un 9%, y la precisión en un 8%.

Tabla 5. Resultados según ADR utilizando Rft_i

| Reprentación | Coseno | Kazkiel | Coseno Pivotado |
|--------------|--------|---------|-----------------|
| IH+D | 0.5495 | 0.5684 | 0.6424 |
| IHRS+D | 0.5399 | 0.5465 | 0.5686 |
| IHRC+D | 0.5345 | 0.5220 | 0.5420 |

Tabla 6. Resultados según Prec. at 5 utilizando Rft_i

| Reprentación | Coseno | Kazkiel | Coseno Pivotado |
|--------------|--------|---------|-----------------|
| IH+D | 0.7730 | 0.8121 | 0.7818 |
| IHRS+D | 0.7940 | 0.8312 | 0.8545 |
| IHRC+D | 0.8213 | 0.8510 | 0.8545 |

⁵ Prec. at 5: Precisión a los 5 documentos

3.3. Combinación de representaciones.

De forma similar la otras representaciones mantienen el comportamiento de baja cobertura y alta precisión, con respecto a la IH+D. Utilizando la técnica del Coseno Pivotado, el sistema de recuperación de información se aplica en dos etapas, una utilizando la representación IH+D de mayor cobertura, para luego aplicar en una segunda pasada la representación IHRC+D de mayor precisión. En la tabla 7 y en la gráfica 1, se pueden apreciar los resultados utilizando IH+D aplicados a las tres técnicas, agregando además los resultados de la técnica del Coseno Pivotado aplicada en un primera etapa de máxima cobertura para IH+Dy una segunda etapa aplicando la representación IHRC+D para mejorar la precisión.

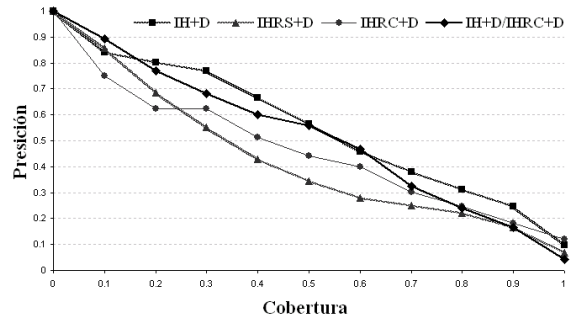
Tabla 7. Resultados para la técnica del Coseno Pivotado, utilizando los tres tipos de representaciones y la combinación en etapas de dos de ellas(IH+D y IHRC+D).

| Representación | AP no int. ⁶ | Prec. at 5 | ADR |
|-----------------|-------------------------|------------|--------|
| IH+D | 0.5467 | 0.7818 | 0.6424 |
| IHRS+D | 0.4265 | 0.8545 | 0.5686 |
| IHRC+D | 0.4047 | 0.8545 | 0.5421 |
| IH+D/ IHRC+D | 0.5160 | 0.8909 | 0.6390 |

En la Gráfica que el Coseno Pivotado genera mejores resultado con un 64% de ADR en promedio de las 11 consultas. Y que la combinación de las representaciones IH+D y IHRC+D disminuyen levemente la cobertura pero aumentan la precisión. Una base de datos de mayor tamaño es necesaria para evaluar mejor la combinación de representaciones. Observando los resultados para cada pregunta individualmente puede notarse cómo afecta el Coseno Pivotado aplicado a IH+D/IHRC+D el resultado de algunas preguntas, ya que favorece a documentos pequeños, pero en otras el resultado sobrepasa del 75%.

En la Figura 6 puede verse el resultado comparativo para una consulta entre los expertos musicales y el sistema de recuperación musical. El sistema de recuperación de información retornó los primeros tres candidatos en el orden exacto, pero no así con los siguientes, sin embargo estos candidatos mantienen una fuerte relación melódica con respecto a la consulta.

⁶ AP not Int.: Precisión promedio no interpolada.



Gráfica 1. Resultados para la técnica del Coseno Pivotado evaluado con diferentes representaciones.

Consulta: 800.000.193/330044

Ordenamiento por Expertos Musicales

1. [Musical notation]
2. [Musical notation]
3. [Musical notation]
4. [Musical notation]
5. [Musical notation]

Sistema de Recuperación de Información Musical

1. [Musical notation]
2. [Musical notation]
3. [Musical notation]
4. [Musical notation]
5. [Musical notation]

ADR at 5: 0.87

Figura 6. Resultados comparativos para una consulta(primeros 5 candidatos relevantes) entre los expertos musicales y el sistema de recuperación de información.

4. Conclusiones

Se ha demostrado que las relaciones interválicas son elementos de criterio para la tarea de similitud melódica y que es posible hacer equivalentes a palabras de texto y utilizar las técnicas de recuperación textual actuales. En el caso de Kaszkiel y Coseno Pivotado fue necesario un cambio en la normalización reemplazando la

función del logaritmo natural por la raíz cuadrada de la frecuencia. Además, las mejoras introducidas a los tres métodos por la relación de frecuencias (Rft_i) y demostraron ser de gran ayuda, y mejoraron los resultados tanto en la evaluación del TREC como ADR.

Además la combinación de dos de las representaciones mejorar los resultados en términos de precisión por medio de un filtrado inicial de los documentos más relevantes mediante una representación de mayor cobertura, para efectuar luego una segunda pasada mediante una representación más precisa, para dar el resultado final. Fue posible además incrementar los resultados utilizando una equilibrada normalización de los fragmentos musicales, utilizando tanto la duración total de los intervalos como la cantidad de intervalos que los componen.

Otra de las tareas pendientes es la de estudiar el comportamiento del sistema en función del tamaño de los fragmentos musicales que se comparan. La disponibilidad de una gran base de datos es necesaria para llevar a cabo estas tareas.

Agradecimientos

Este trabajo se realizó gracias al apoyo de Agencia Española de Cooperación Internacional (ref: AEI:B/2098/04). El Ministerio de Ciencia y Tecnología de Costa Rica y al CONICIT (ref: 84-2005), así como de los proyectos CICyT TIC 2003-08496-C04 y GV06/166.

Referencias

- [Clausen01] Michael Clausen, Frank Kurth, Roland Engelbrecht. 'Content-based Retrieval in MIDI and Audio'. ECDL WS Generalized Documents. 2001.
- [Doraisami01] S. Doraisami & S. Rüger. 'Robust polyphonic music retrievals with n-grams'. Journal of Intelligent Information Systems, 21(1): 53-70.
- [Downie99] Stephen Downie. 'Evaluating a simple approach to music information retrieval: conceiving melodic n-grams as text'. Tesis Doctoral. University of Western Notario. 1999.
- [Engelbrecht02] Roland Engelbrecht. 'Statistical Comparison Measures for Searching in Melody Databases'. Tech. Report. University of Bonn. 2002.
- [Flores79] Bernal Flores. 'Music Theory. Systematically Orderly Sonorities'. Universidad de Costa Rica, 1979.
- [Frakes92] W.B. Frakes, D.C. Baeza-Yates, editors. 'Information Retrieval: Data Structures and Algorithms'. Prentice-Hall, 1992.
- [Hanson60] Howard Hanson. 'Music Theory. Harmonic Materials of Modern Music'. ISBN: 0891972072 Publisher: Irvington Pub. 1960.
- [Hewlett01] W.B. Hewlett, E. Selfridge-Field (eds.) 'The Virtual Score. Representation, Retrieval, Restoration'. The MIT Press. 2001.
- [Hu02] Ning Hu, Roger B. Dannenberg, Ann L. Lewis. 'A Probabilistic Model of Melodic Similarity'. Proceedings of the International Computer Music Conference (ICMC), 2002.
- [Kaszkiel99] M. Kaszkiel, J. Zobel and R. Sacks-Davis. 'Efficient Passage Ranking for Document Databases'. ACM Transactions on Information Systems, 17(4), pp 406-439. October 1999.
- [Marsden92] Marsden, A. Pople (eds.) 'Computer Representations and Models in Music'. Academic Press. 1992.
- [Narmour90] Eugene Narmour. 'The Analysis and Cognition of Basic Melodic Structures'. Chicago University Press, 1990.
- [Pérez04] C. Pérez-Sancho, José M. Iñesta, J. Calera-Rubio. 'Style recognition through statistical event models'. Sound and Music Computing (SMC'04) Conference Proceedings, 135-139 2004.
- [Rasmussen03] Edie Rasmussen. 'Evaluation in Information Retrieval'. The MIR/MDL Evaluation Project White Paper Collection #3. Pp. 45-49. 2003.
- [Singhal96] Singhal. A., Buckley, M. Mitra. 'Pivoted document length normalization'. Proceedings of the 19th annual international ACM-SIGIR conference on research and development in information retrieval. 1996.

- [Stuart04] Craig Stuart Sapp, Yi-Wen Liu, Eleanor Selfridge-Field. 'Search-Effectiveness Measures for Symbolic Music Queries in Very Large Databases'. Proceedings of 5th International Conference on Music Information Retrieval, ISMIR. 2004.
- [Typke04] Rainer Typke, Frans Wiering, Remco C. Veltkamp. 'A survey of Music Information Retrieval Systems'. Proceedings of 5th International Conference on Music Information Retrieval, ISMIR. 2004.
- [Typke05a] Rainer Typke, Marc den Hoed, Justin de Nooijer, Frans Wiering, Remco C. Veltkamp. 'A Ground Truth For Half A Million Musical Incipits'. Proceedings of the Fifth Dutch-Belgian Information Retrieval Workshop. Pp.63-70. 2005.
- [Typke05b] Rainer Typke. 'Proposed measure for comparing ranked lists, for MIREX '05 (Symbolic Melodic Similarity) '. 5th Dutch-Belgian Information Retrieval Workshop. 2005.