

UNTREF - Ingeniería de Sonido

Fundamentos de Procesamiento Musical usando Cuadernos Python

Alejandro Grons kis

Año 2020

Objetivo del curso

Este curso se engloba en el contexto de la recuperación de la información musical (Music Information Retrieval, MIR). El principal objetivo del mismo reside en proporcionar una introducción a esta creciente área de investigación interdisciplinaria la cual se apoya en el procesamiento de señales, recuperación de información, aprendizaje automático, ingeniería multimedia, musicología y psicoacústica. Este curso pretende, por medio de un entorno contributivo de acceso libre que denominamos [cuadernos Jupyter](#) (archivos `.ipynb` que son cargados y procesados por algún entorno de ejecución como [Cocalc](#) o [Colab](#)) poner a disposición de los alumnos una herramienta web que dé solución a problemas presentados en el área MIR.

Resultados de aprendizaje (Objetivos formativos)

- Introducir el problema de la extracción de información musical
- Conocer los formatos y parámetros involucrados en la digitalización del sonido y su uso en la recuperación de información sonora
- Adquirir la capacidad de desarrollar sistemas que utilicen técnicas y algoritmos aplicables a estos problemas.

Incumbencias del curso

Capacita para desempeñarse en las siguientes áreas de aplicación:

- Técnicas de recuperación de información
- Separación de pistas y reconocimiento de instrumentos
- Conversión de un registro de audio a notación simbólica. Este proceso implica varias tareas que incluyen:
 - ▶ detección de tonos distintos
 - ▶ detección del inicio de una nota musical
 - ▶ estimación del tamaño de la ventana temporal
 - ▶ extracción de información rítmica
- Técnicas de aprendizaje automático (*machine learning*).

Metodología

El empleo de cuadernos FMP [2] provee de un material útil para la enseñanza y el aprendizaje de los fundamentos del procesamiento musical. La ventaja de su empleo radica en facilitar la introducción al campo MIR con un interés particular puesto en el análisis de audio y técnicas de recuperación de información. En concordancia con el libro de texto [1], los cuadernos abordan tareas cuyo desarrollo implica tratar los algoritmos y métodos requeridos. Complementando los [paquetes de software código abierto](#) para aplicaciones MIR y procesamiento de señales de audio, los cuadernos FMP combinan explicaciones en formato de libros de texto con ejemplos escritos en código **Python**. Entrelazando conceptos técnicos, detalles matemáticos, ejemplos de código, ilustraciones y archivos de sonido en un entorno interactivo **Jupyter Notebooks** permite al alumno acortar la brecha entre clases Teóricas y Prácticas. A fin de expandir con nuevas capacidades los sistemas tradicionales de gestión del aprendizaje (como Moodle) se pretende introducir a los alumnos en la elaboración de actividades prácticas y tareas usando la herramienta [Github Education](#) en el aula [3].

Estructura de los cuadernos

Parte 0: provee enlaces a paquetes de software.

Parte B: provee introducción al entorno **Jupyter**, el lenguaje de programación **Python** y conceptos técnicos involucrados en el uso de los cuadernos.

Partes 1 a 6: tratan los diversos escenarios MIR y tareas de procesamiento musical.

CONTEXTO TÉCNICO

Los cuadernos correspondientes a la **Parte B** describen las principales herramientas empleadas en el desarrollo de las **FMP Notebooks**, a saber:

- Instalación de paquetes
- Integración de objetos multimedia (audio, imagen y video) en el entorno **Jupyter**
- Introducción a **Python** resumiendo los tipos de datos, estructuras de control y funciones. Diversos ejemplos de código ilustran cómo crear figuras, procesar archivos de audio y programar gráficos interactivos.

ESCENARIOS MIR Y PROCESAMIENTO MUSICAL

Un fin primordial de los cuadernos FMP consiste en proveer tanto material audiovisual así como ejemplos de código Python los cuales implementen las técnicas computacionales destinadas a ilustrar los diversos conceptos matemáticos. Además, los cuadernos proveen código que permite al usuario experimentar con parámetros y de esta forma adquirir una mejor comprensión de los resultados del cálculo a través de visualizaciones y sonificaciones.

Parte 1. Representaciones musicales

Partituras, representación simbólica y audio. Se provee código Python para estudiar las propiedades acústicas de las señales de audio incluyendo aspectos tales como frecuencia, tono, envolvente dinámica y timbre.

Parte 2. Análisis de Fourier

Se provee código que permite lograr una mejor comprensión de los números complejos y funciones exponenciales, los cuales constituyen la base de la transformada de Fourier discreta (DFT). La transformada rápida de Fourier (FFT) se trata en teoría y práctica. Se introduce además la transformada de Fourier de tiempo corto (STFT). En este contexto, se abordan aspectos tales como el muestreo, relleno de una señal y sistema de coordenadas.

Parte 3. Sincronización musical

El objetivo reside en alinear temporalmente diferentes versiones de la misma pieza musical. En este escenario, se proveen ejemplos de código para extraer la *Secuencia Chroma* de una señal así como para detectar patrones melódicos. Se tratan aspectos de gran interés práctico como afinación, compresión logarítmica y cuestiones de resolución espectral y temporal. Se aborda además una técnica conocida como distorsión de tiempo dinámico (DTW), para tal fin se implementa un algoritmo para medir la similitud entre dos secuencias temporales que pueden variar en velocidad.

Parte 4. Análisis de estructura musical

Dado un registro musical, el objetivo es identificar elementos estructurales críticos y segmentar el registro de acuerdo a estos elementos. En este contexto, se discuten diversos criterios de segmentación basados en repetición, homogeneidad, así como otros criterios novedosos cuyo dominio de aplicación multimedia excede el campo del audio. En particular, se provee código que permite generar, visualizar y modificar las propiedades estructurales de matrices auto-similares. Los cuadernos abarcan además técnicas clásicas tales como la detección de novedades (*Novelty Detection*) y la extracción de una muestra representativa de tamaño reducido a partir de una pieza musical (*Audio Thumbnailing*).

Parte 5. Reconocimiento de acordes

Esta tarea MIR aborda el análisis de propiedades armónicas de una pieza musical estableciendo una progresión de acordes a partir de un registro de audio. En primera instancia, se introducen conceptos básicos referidos a teoría de la armonía, tales como intervalos, acordes y escalas. A continuación, se introduce una aproximación simple basada en un procedimiento de ajuste con plantillas de acordes teóricos y pos-tratamiento por filtrado. Este sistema se extiende finalmente a un modelo oculto de Markov (HMM), siendo este último un concepto de fundamental importancia en el análisis de patrones temporales con aplicaciones en reconocimiento del habla, de gestos y de movimientos corporales.

Parte 6. Tempo y Seguimiento del Pulso (*Beat Tracking*)

Se introducen las ideas básicas acerca de cómo extraer información asociada con el *tempo* a partir de registros de audio. Una primer tarea, conocida como *Detección de Onsets*, procura localizar el momento preciso en el que comienza una nota musical a través de la detección de cambios en el contenido espectral de la señal. A fin de estimar el *tempo* y deducir el período y alineación de los *beats*, la información contenida en cada intervalo entre *onsets* se analiza teniendo en cuenta patrones de periodicidad local. En esta segunda etapa se introducen 2 métodos conceptualmente distintos: uno basado en el análisis de Fourier y el otro basado en autocorrelación.

Referencias

- [1] Meinard Müller, 2015, "Fundamentals of Music Processing", Springer Verlag.
- [2] Meinard Müller and Frank Zalkow, 2019, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing", *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, Delft, The Netherlands.
- [3] Integración de la herramienta [Github Education](#) en el aula.