

# PAPER TEMPLATE FOR ISMIR 2015

**First author**

Affiliation1

author1@ismir.edu

**Second author**

**Retain these fake authors in**

**submission to preserve the formatting**

**Third author**

Affiliation3

author3@ismir.edu

## ABSTRACT

The abstract should be placed at the top left column and should contain about 150-200 words.

## 1. INTRODUCTION

This template includes all the information about formatting manuscripts for the ISMIR 2015 Conference. Please follow these guidelines to give the final proceedings a uniform look. Most of the required formatting is achieved automatically by using the supplied style file ( $\LaTeX$ ) or template (Word). If you have any questions, please contact the Program Committee (ismir2015-papers@ismir.net). This template can be downloaded from the ISMIR 2015 web site (<http://ismir2015.ismir.net>).

## 2. PAPER LENGTH

**The regulations for the maximal paper length have changed.** Instead of the strict limit of six pages (as used for ISMIR 2014), we adopt a “(6+1)-page policy” for ISMIR 2015. This means, the paper may have a maximum of 6 pages for technical content including figures and possible references with one additional optional 7th page containing only references. Note that this is a strict requirement. **The seventh page (if used at all) must not contain any other material except for references.**

## 3. SHORT FAST FOURIER TRANSFORM (SFFT)

A transformada discreta de fourier (DFT ou FFT) muito utilizada para a construçao das características cromáticas (chroma feature) do audio. A funcao dessa transformada eh traduzir informacoes que estao em dominio temporal para dominio frequencial de tal forma a projetar, em bases ortonormais, o valor de cada componente senoidal presente no sinal tratado. Essa projecao se da pelo somatorio do produto interno das senoides (exponenciais complexas) pelo sinal.

Visto que essa transformada somente oferece informacoes em termos de frequencias, surgiu-se a necessidade de adapta-la para a visualizacao das variacoes das frequencias

ao longo do tempo. Essa adaptacao se denomina transformada de fourier janelada (SFFT). Esse processo de janelamento se resume a dividir o sinal em partes, para serem analisadas individualmente no dominio da frequencia, ou seja, a cada instante de tempo referente a cada trecho eh possivel obter uma analise frequencial.

Porem essa tecnica possui as seguintes limitacoes:

- surgimento de frequencias fantasmas (aliasing) que podem dificultar a identificacao das notas musicais que foram verdadeiramente tocadas;
- problemas de vazamento gerados pelo truncamento do sinal, podendo gerar frequencias que dificultam a visualizacao do espectro (leakage);
- dificuldades na determinacao do tamanho da janela pois trata-se de um parametro de amostragem que pode limitar a banda de frequencia a ser analisada.

## 4. CHROMA CONVOLUTION METHOD (CCM)

Em vista das limitacoes apresentadas da SFFT, foi desenvolvido um metodo no dominio do tempo que preserva as caracteristicas originais do sinal. O chroma convolution method (CCM) objetiva a projecao dos trechos do sinal sobre cada um dos sinais das notas musicais procuradas. Observa-se que quando o trecho do sinal eh projetado sobre o sinal de uma nota musical, atraves da convolucao, esta nota eh amplificada e as demais so suprimidas.

Seja uma nota musical formada a partir de uma senoide monocromatica. Quando esse sinal eh convoluido com um trecho de sinal de audio eh possivel extrair como resultado o nivel de energia relacionado a esta frequencia. Tal energia eh mensurada a partir da seguinte equacao Eqn (1):

$$E = \sum [f(x) * g(x)]^2 \quad (1)$$

Neste contexto as características cromáticas (chroma feature) foram construidas a partir dos seguintes procedimentos:

1. dividiu-se o audio a ser analisado em sinais com duracao pre-determinada;
2. funcoes senoidais monocromaticas foram utilizadas para construir notas musicais no dominio do tempo, correspondendo a escala cromatica musical;
3. convoluiu-se cada parte do sinal do sinal de audio com as notas musicais;

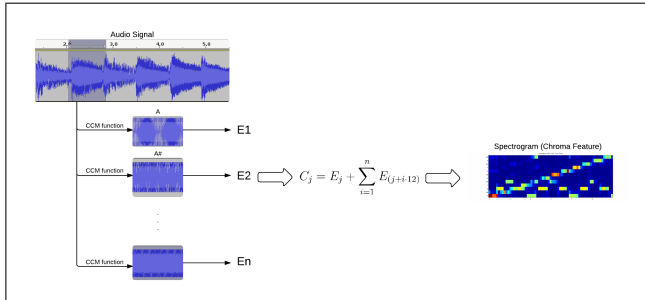


© First author, Second author, Third author.

Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** First author, Second author, Third author. “Paper Template For ISMIR 2015”, 16th International Society for Music Information Retrieval Conference, 2015.

4. a energia da convolucao foi extraida a partir da equacao Eqn (1);
5. por fim cada energia foi somada com suas respectivas oitavas, originando o chroma feature.

A figura abaixo representa uma visao esquematica do procedimento descrito. Ao final do processo as caracteristicas cromaticas (chroma feature) podem ser visualizadas a partir de um espectrograma.



**Figure 1.** Schematic of process to build chroma feature using CCM.

## 5. EXPERIMENTS AND RESULTS

Com o intuito de demonstrar a eficacia do CCM em relacao ao tradicional metodo SFFT [1] no que diz respeito identificacao das notas tocadas a partir de suas caracteristicas cromaticas (chroma feature), foram realizados 2 experimentos usando a linguagem de programacao MATLAB. Como entrada de dados usou-se sinais de audio gravados por instrumentos musicais reais<sup>1</sup>.

### 5.1 Experiment 1: Identification of Musical Notes

O primeiro experimento se trata de identificacao de notas musicais numa melodia de piano. Tais notas foram tocadas segundo a Figure 2:



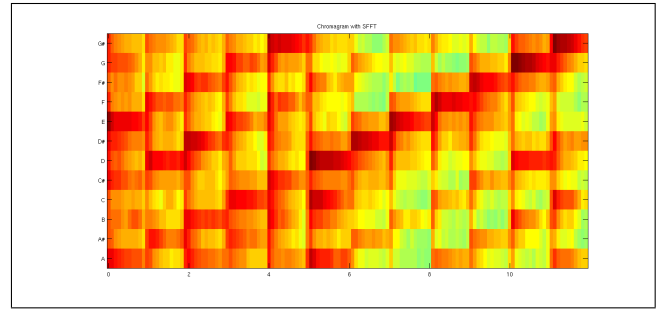
**Figure 2.** Notes that were played and recorded on piano.

Utilizou-se convolucao com 96 tons puros de diferentes frequencias para gerar o chroma feature. As janelas de ambas as propostas foram configuradas no tamanho de 0.120 segundos e a taxa de amostragem do audio foi de 44.1 kHz.

#### 5.1.1 SFFT Results

Segue chroma feature encontrado utilizando o metodo tradicional da SFFT em Figure 5. Observa-se que a escala cromatica foi identificada, porem com erros em relacao ao momento exato em que as notas foram tocadas. Nota-se, por exemplo, que uma nota E foi identificada no comeco quando na verdade deveria ter sido identificada nos instantes proximos de 7 segundos.

<sup>1</sup> Code and files available in [https://github.com/josepedro/ismir\\_article](https://github.com/josepedro/ismir_article).



**Figure 3.** Chroma feature of audio using SFFT.

Foi tambem feito uma analise de eficiencia das notas gravadas no audio com a sequencia de notas identificas no chroma feature. As notas identificadas foram seleccionadas atraves da maior intensidade em cada instante de tempo. Segue tabela comparativa:

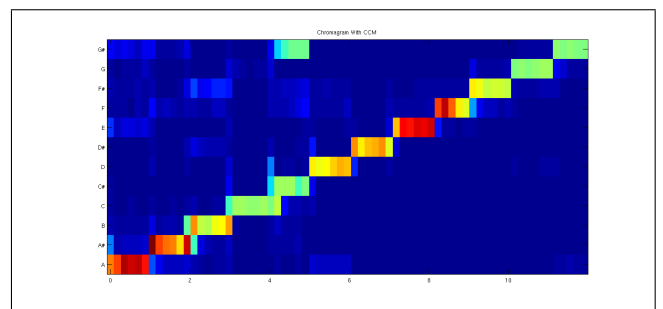
Played Notes	Identified Notes
A	E
A#	D
A#	D#
B	B
B#	E
C	D
C#	G#
D	D
D#	D#
E	E
F	F
F#	F#
G	G
G#	G#

**Table 1.** Comparison between played notes and identified notes.

Dado os dados da tabela Table 1, houve 9 de 14 notas tocadas em instantes corretos, ou seja, 64.2857% de fidelidade com a melodia original.

#### 5.1.2 CCM Results

Segue chroma feature encontrado utilizando o CCM:



**Figure 4.** Chroma feature of audio using CCM.

Observa-se na figura Figure 4 que a melodia apresentou uma maior clareza. Uma tabela comparativa das notas to-

cadadas e identificadas foi feita tambem atraves da estrategia dos maximos:

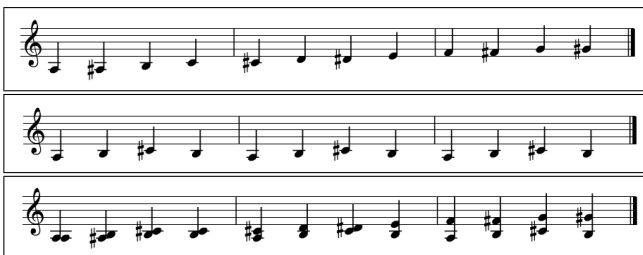
Played Notes	Identified Notes
A	A
A#	A#
B	B
C	C
C#	C#
C#	G#
C#	C#
D	D
D#	D#
E	E
F	F
F#	F#
G	G
G#	G#

**Table 2.** Comparison between played notes and identified notes from chroma feature.

Em vista dos resultados de Table 2, houve 13 de 14 notas tocadas em instantes corretos, ou seja, 92.8571% de fidelidade com a melodia original.

## 5.2 Experiment 2: Identification of Musical Notes of Different Instruments

O segundo experimento diz respeito a identificao de notas musicais de diferentes instrumentos. Para tal foram tocadas as seguintes melodias no piano e violao juntas:



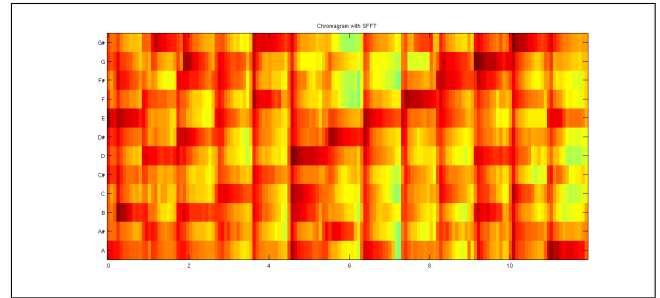
**Figure 5.** The first sheet music is only piano, the second sheet music is only acoustic guitar and the third sheet music is the both together.

Utilizou-se convolucao com 12 tons de violao e 12 tons de piano de diferentes frequencias gerando dois chroma feature respectivamente, um para cada instrumento. As janelas de ambas as propostas foram configuradas no tamanho de 0.120 segundos e a taxa de amostragem do audio foi de 44.1 kHz.

### 5.2.1 SFFT Results

Segue chroma feature encontrado utilizando o metodo tradicional da SFFT:

Observa-se na figura Figure 6 que a escala cromatica do piano foi parcialmente identificada. As notas do violao sao praticamente imperceptiveis.



**Figure 6.** Chroma feature of audio using SFFT.

Foi tambem feito uma analise de eficiencia das notas gravadas dos dois instrumentos com a sequencia de notas identificadas no chroma feature, usando a maior intensidade em cada instante de tempo. Segue tabela comparativa:

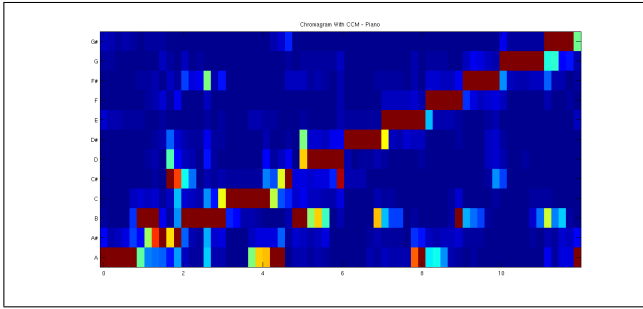
Piano	Acoustic Guitar	Identified Notes
E	A	A
A	A	A
B	A	A
E	A	A
D	A#	B
F	A#	B
D	A#	B
D#	B	C#
F#	B	C#
B	B	C#
F#	B	C#
C	C	B
A	C	B
E	C	B
G#	C#	A
B	C#	A
D	D	B
D#	D#	C#
B	D#	C#
E	E	B
F	F	A
F#	F#	B
F	F#	B
C#	F#	B
F#	F#	B
G	G	C#
B	G	C#
G	G	C#
G#	G#	B
A	G#	B
E	G#	B

**Table 3.** Comparison between played notes of piano, acoustic guitar and identified notes of chroma feature.

Segundo os dados da tabela Table 3, houve 12 de 31 notas de piano tocadas em instantes corretos, ou seja, 38.7096% de fidelidade com a melodia original do piano. Quanto ao violao, somente 1 nota foi tocada em instante certo, ou seja, 3.2258% de fidelidade com a melodia original do violao.

### 5.2.2 CCM Results

Segue chroma feature encontrado utilizando o CCM de tons de piano:



**Figure 7.** Chroma feature of audio using CCM of piano tones.

Observa-se na figura Figure 7 que a melodia apresentou uma maior clareza. Eh visivel tambem a escala cromatica tocada pelo piano e algumas poucas notas do violao. Uma tabela comparativa das notas tocadas identificadas foi feita tambem atraves da estrategia dos maximos:

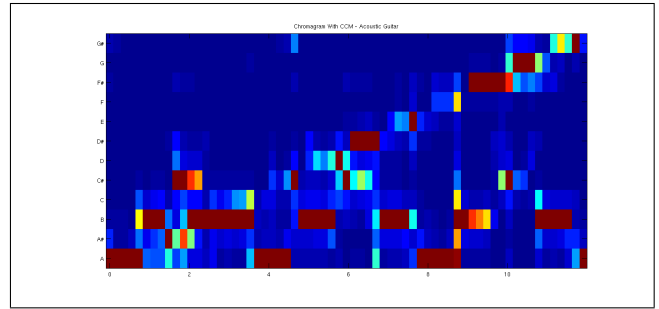
Piano	Acoustic Guitar	Identified Notes
A	A	A
A	B	B
A#	B	A#
A#	B	C#
A#	C#	A#
B	B	B
C	A	C
C#	A	A
C#	B	C#
D	B	B
D	C#	D
D#	B	D#
E	A	E
F	B	F
F#	C#	F#
G	B	G
G#	B	G#
G#	B	A

**Table 4.** Comparison between played notes of piano, acoustic guitar and identified notes of chroma feature with piano tones.

Em vista dos resultados de Table 4, houve 13 de 18 notas de piano em instantes corretos, ou seja, 72.2222% de fidelidade com a melodia original. Quanto ao violao, houve 5 de 18 notas em instantes corretos, ocasionando em 27.7777%. Tal resultado mostra que o CCM usando tons de piano realmente amplificou as suas notas e suprimiu as do violao.

Segue chroma feature encontrado utilizando o CCM de tons de violao:

Observa-se na figura Figure 8 que a melodia apresentou uma maior clareza. Eh visivel tambem a escala cro-



**Figure 8.** Chroma feature of audio using CCM of acoustic guitar tones.

matica tocada pelo piano suprimida e as notas do violao amplificadas. Uma tabela comparativa das notas tocadas e identificadas foi feita tambem atraves da estrategia dos maximos:

Piano	Acoustic Guitar	Identified Notes
A	A	A
A#	B	B
A#	B	A#
B	C#	C#
C	B	B
C#	A	A
D	B	C#
D#	C#	B
D#	C#	D
D#	C#	C#
E	B	D#
F	A	B
F	A	E
F#	B	A
F#	B	B
G	C#	F#
G	C#	C#
G#	B	G
G#	B	B

**Table 5.** Comparison between played notes and identified notes.

Em vista dos resultados de Table 5, houve 9 de 19 notas de violao em instantes corretos, ou seja, 47.3684% de fidelidade com a melodia original. Quanto ao piano, houve 2 de 19 notas em instantes corretos, ocasionando em 10.5263%. Tal resultado mostra que o CCM usando tons de violao realmente amplificou as suas notas e suprimiu as do piano.

## 6. DISCUSSION

Em vista dos resultados alcançados, ha limitacao e beneficios no uso da convolucao em relacao a transformada de fourier janelada. Uma limitacao encontrada foi o tempo de processamento que no metodo da SFFT foi de 0.100138 segundos enquanto no CCM foi de 13.266711 segundos para o primeiro experimento.

Do ponto de vista dos beneficios, o primeiro experimento mostra que o CCM eh mais eficiente do que o SFFT para caracterizacao de notas. O SFFT obteve um rendimento de 64.2857% enquanto que o CCM obteve 92.8571%, ocasionando em 28.5714% a mais de notas identificadas com o CCM. Tal fato se deve pela reducao de ruidos e intensificacao das frequencias sonoras realmente gravadas pelo instrumento musical.

Do segundo experimento observa-se um beneficio muito importante no uso do CCM, a possibilidade de distinguir instrumentos musicais tocados juntos. Alem da SFFT nao possuir a possibilidade de geracao de um chroma feature para cada instrumento, ela gerou resultados baixos em comparacao com CCM na identificacao das melodias do piano (38.7096%) e violao (3.2258%).

Usando o CCM gerou-se, para cada instrumento, um chroma feature. O chroma feature do piano possuiu identificacao de 72.2222% de notas de melodia de piano e 27.7777% de melodia de violao. Esse fato demonstra que o chroma feature do piano realmente destaca as suas notas de timbre caracteristico. O chroma feature do violao tambem possuiu tal capacidade com 47.3684% de fidelidade para melodia do violao e 10.5263% para melodia de piano. Eh notavel que o CCM possui um potencial maior de identificacao que o SFFT para ambos instrumentos alem de destacar mais diferentes instrumentos.

## 7. CONCLUSIONS AND FUTURE WORKS

Em vista do que foi proposto, foi apresentado uma nova forma de construcao da chroma feature usando o chroma convolution method (CCM). Comparou-se o metodo proposto com o tradicional SFFT e o CCM mostrou-se mais satisfatorio.

A eficiencia do CCM frente ao SFFT nao foi somente pela deteccao de melodias monofonicas, mas tambem pela identificacao de melodias polifonicas com mais de um instrumento. O CCM possui a capacidade diferencial de trabalhar com timbres de instrumentos musicais para a distincao dos mesmos no audio.

Pode-se concluir que melhorias que foram feitas na forma simples da SFFT em trabalhos correlatos podem ser aplicadas no CCM, aumentando assim seu potencial. Tambem eh de significativa importancia demonstrar a otimizacao

dos resultados de sistemas de transcricao automatica de acordes usando o CCM.

Assim como ocorre na SFFT, plausivel de implementacao de um pre-processamento de sinal utilizando filtros no CCM para a caracterizacao de instrumentos em audios mais complexos como o de uma banda ou orquestra.

## 8. REFERENCES

- [1] Columbia University LabROSA. Chroma feature analysis and synthesis, 2015.