

Obtendo estatísticas normalizadas para características de croma

Álvaro Almeida Freire Stivi

astivi@ime.usp.br

Grupo de Computação Musical

<http://compmus.ime.usp.br>

Instituto de Matemática e Estatística – Universidade de São Paulo

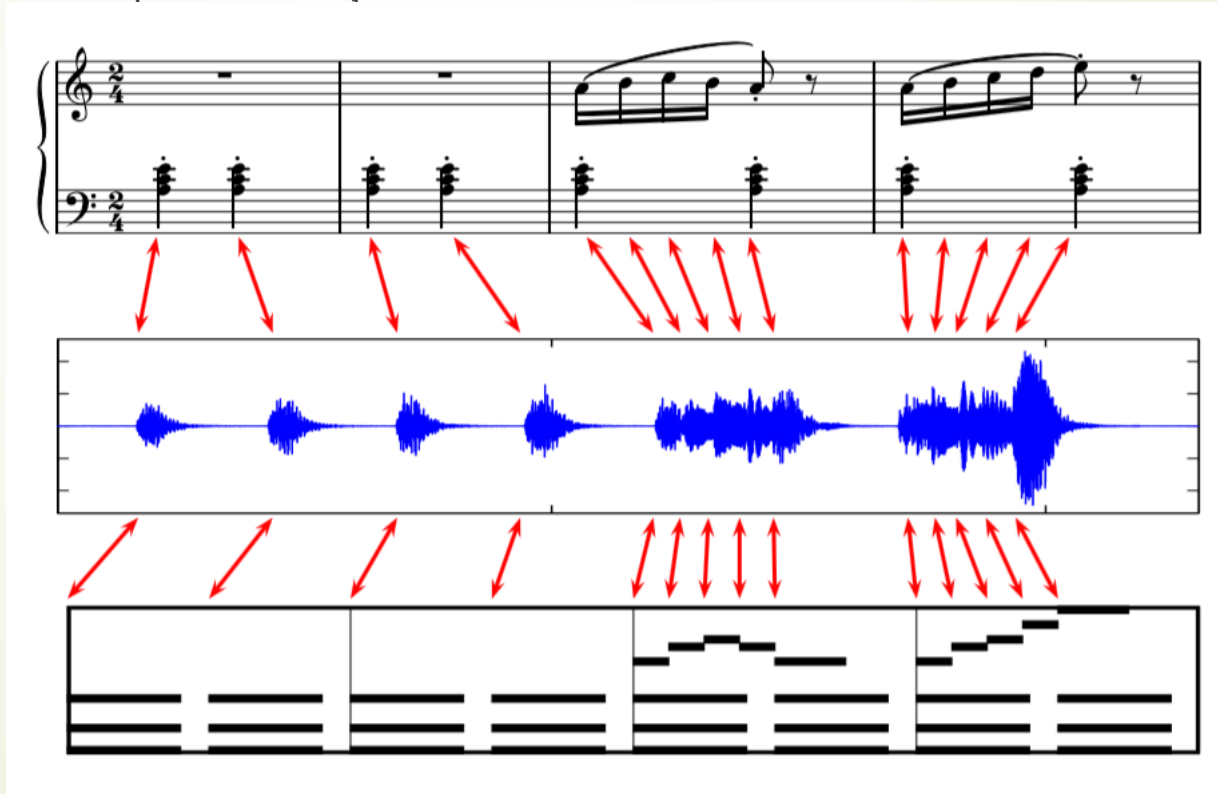


Motivação

- ▶ Obter uma representação eficiente das características de coroma.
- 

Motivação

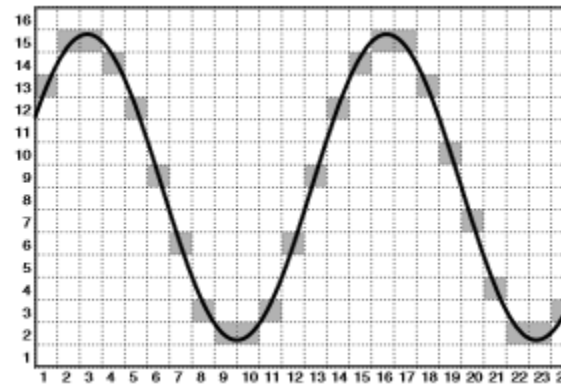
- Obter uma representação eficiente das características de croma.



Motivação

- Obter uma representação eficiente das características de croma.

```
<note>
  <pitch>
    <step>E</step>
    <alter>-1</alter>
    <octave>4</octave>
  </pitch>
  <duration>2</duration>
  <type>half</type>
</note>
```



Ticks	Message	Ch.	MNN	Vel
60	NOTE ON	1	67	100
0	NOTE ON	2	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	2	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE ON	2	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	2	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE ON	2	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	2	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	63	100
0	NOTE ON	2	51	100
0	NOTE ON	2	39	100
240	NOTE OFF	1	63	0
0	NOTE OFF	2	51	0
0	NOTE OFF	2	39	0



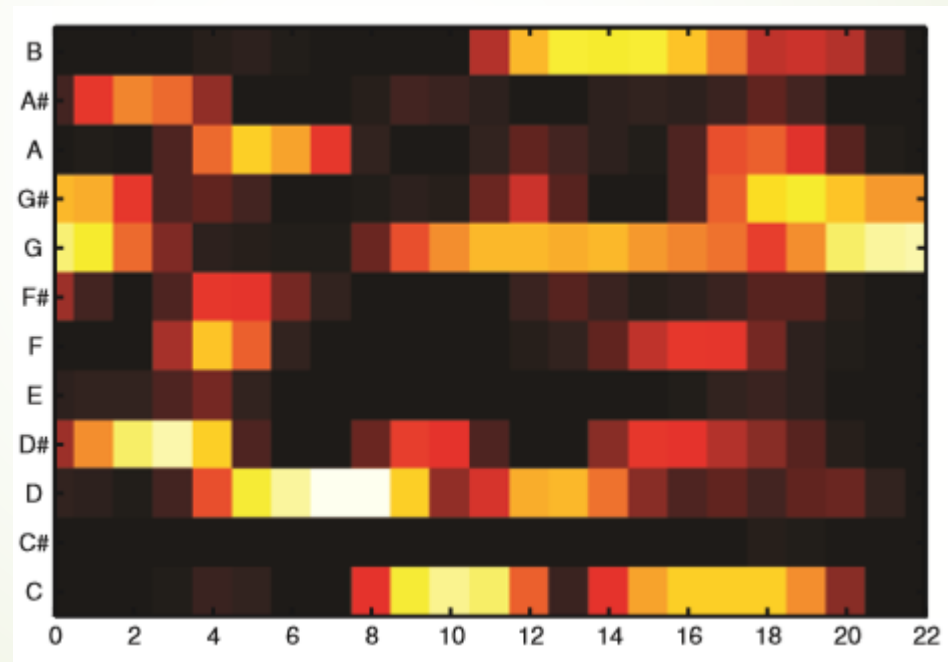
Motivação



Características robustas?

Motivação

- Representação eficiente e robusta.

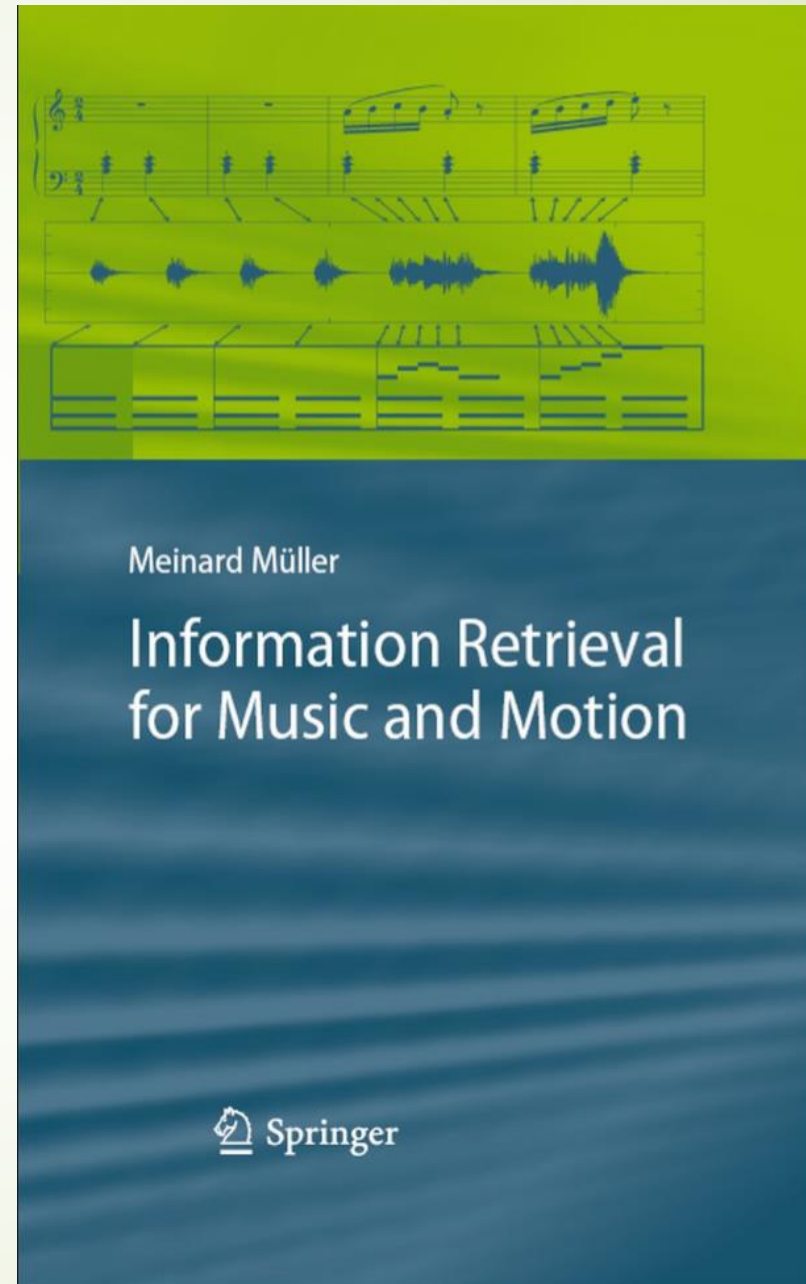


Information Retrieval for Music and Motion

Capítulo 3

Meinard Müller
Springer, 2007

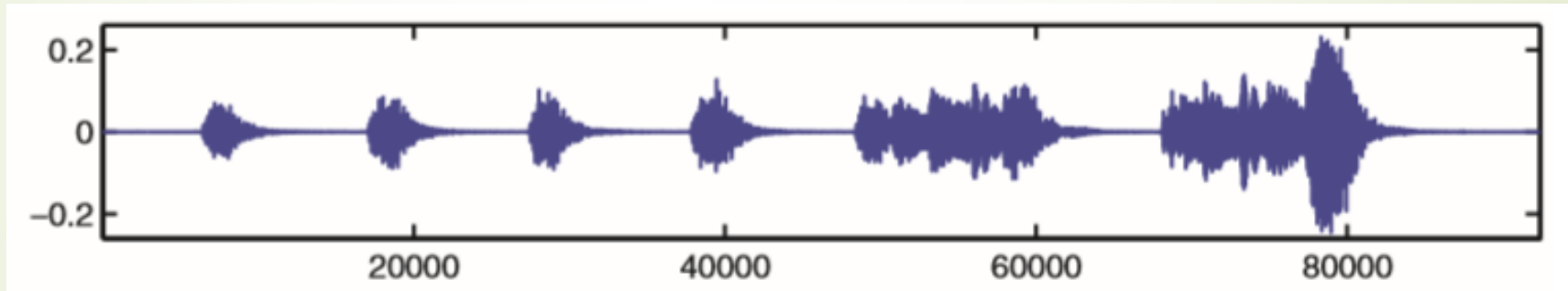
ISBN 978-3-540-74047-6



Op. 100, No. 2 de Friedrich Burgmüller
4 compassos



Op. 100, No. 2 de Friedrich Burgmüller
4 segundos (22050Hz)



Características de tom

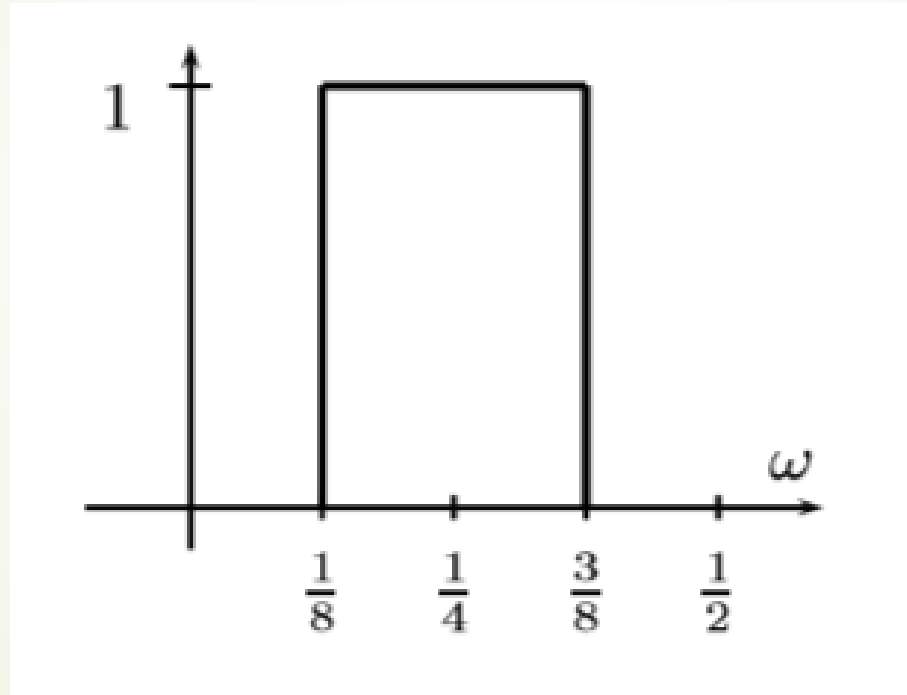
Notas/MIDI/Frequência

$$f(p) = 2^{\frac{p-69}{12}} \cdot 440$$

Note	p	$f(p)$ (CT)
A3	57	220.00
A [#] 3	58	233.08
B3	59	246.94
C4	60	261.63
C [#] 4	61	277.18
D4	62	293.66
D [#] 4	63	311.13
E4	64	329.63
F4	65	349.23
F [#] 4	66	369.99
G4	67	392.00
G [#] 4	68	415.30
A4	69	440.00
A [#] 4	70	466.16

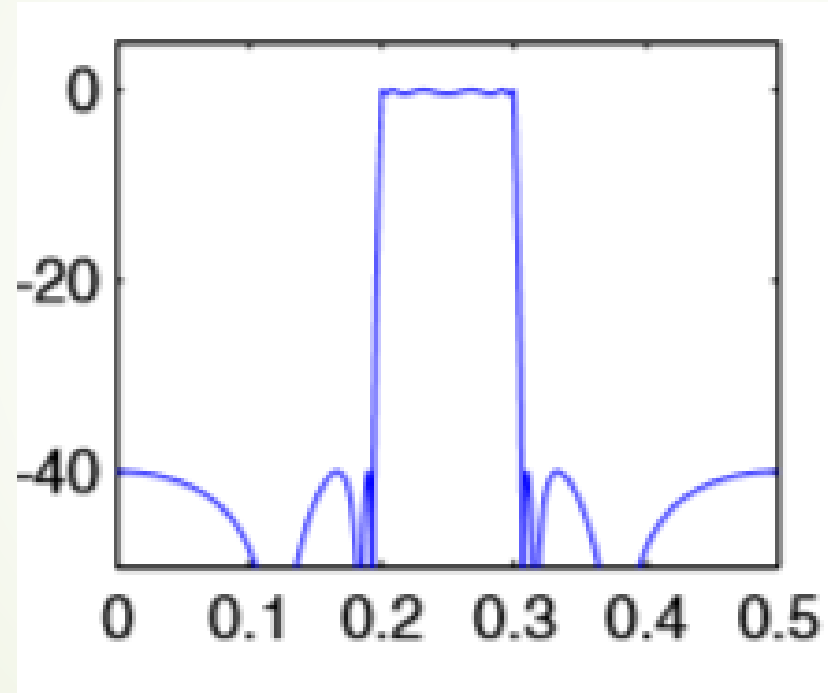
Filtros passa-banda

Filtro Ideal

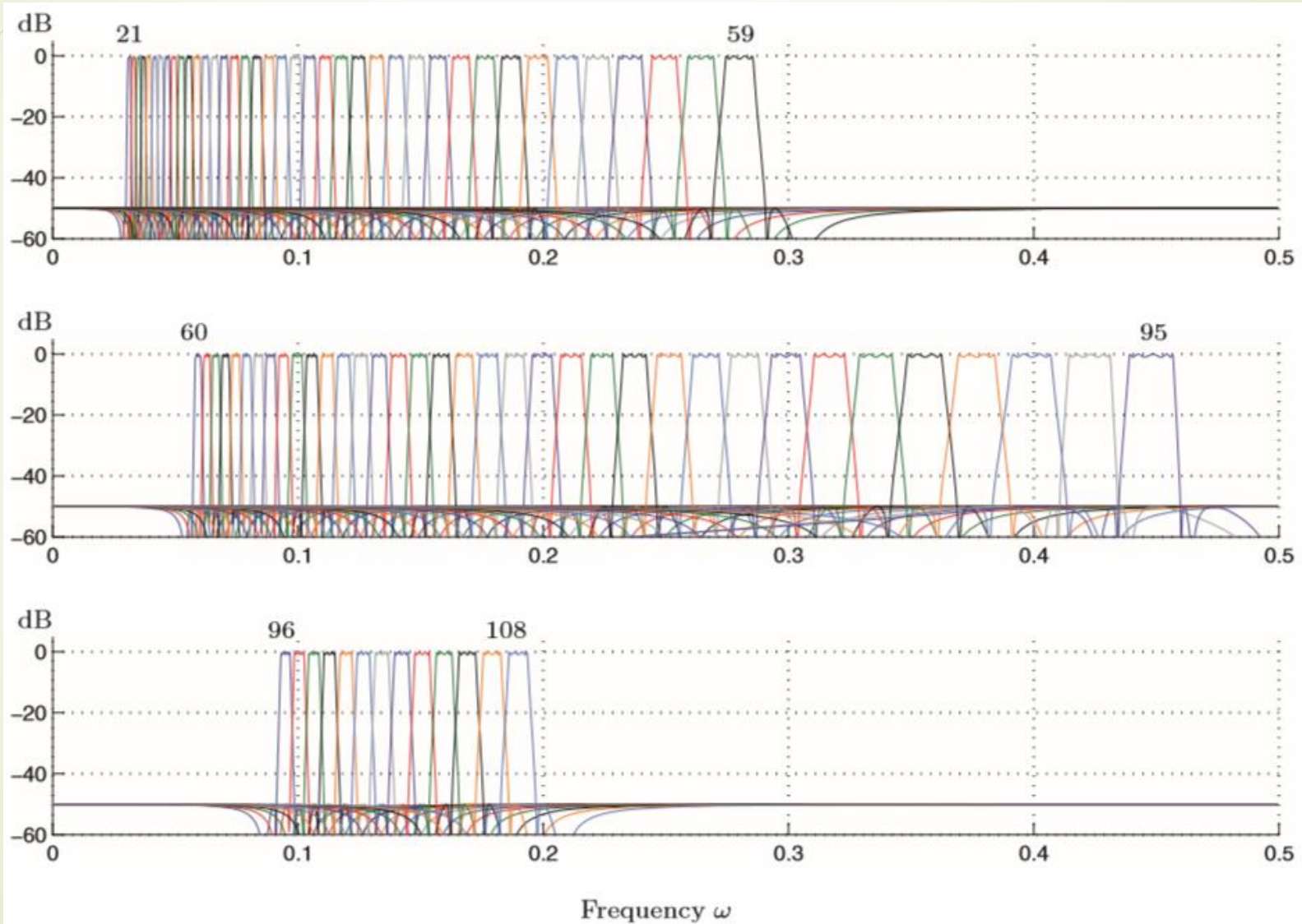


Filtros passa-banda

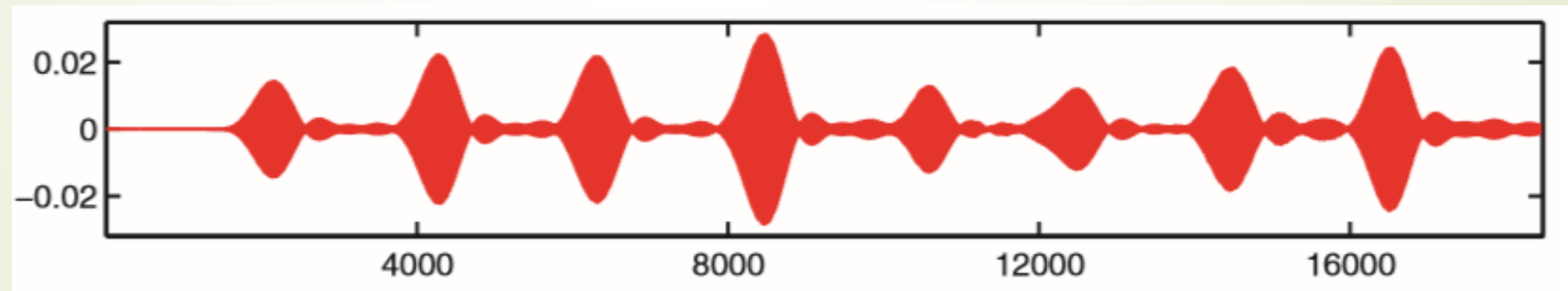
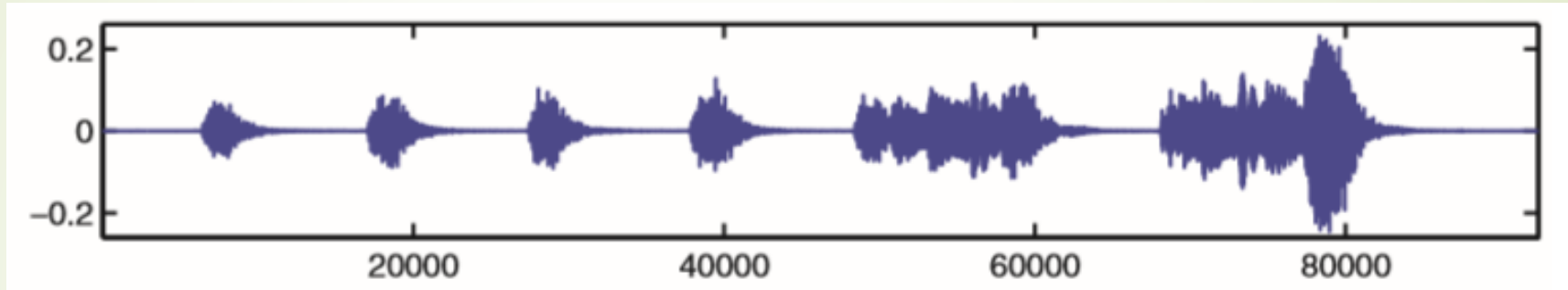
Filtro elíptico de ordem 12



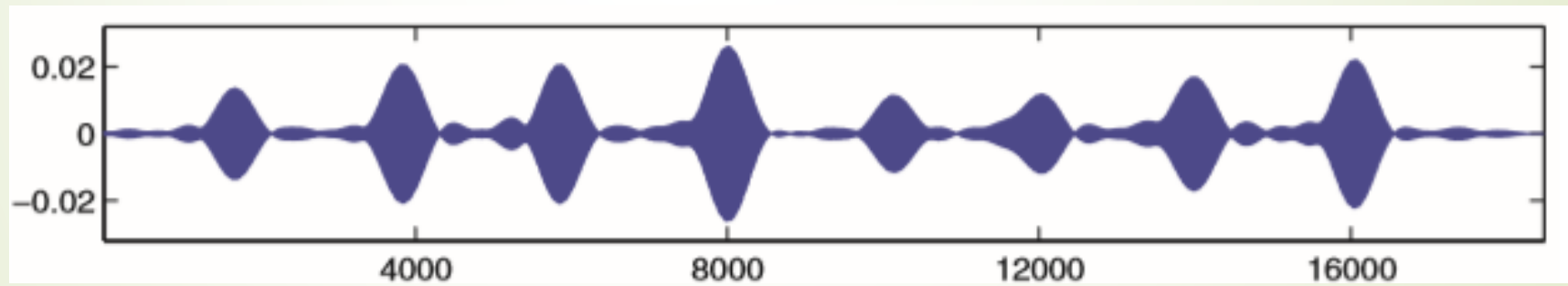
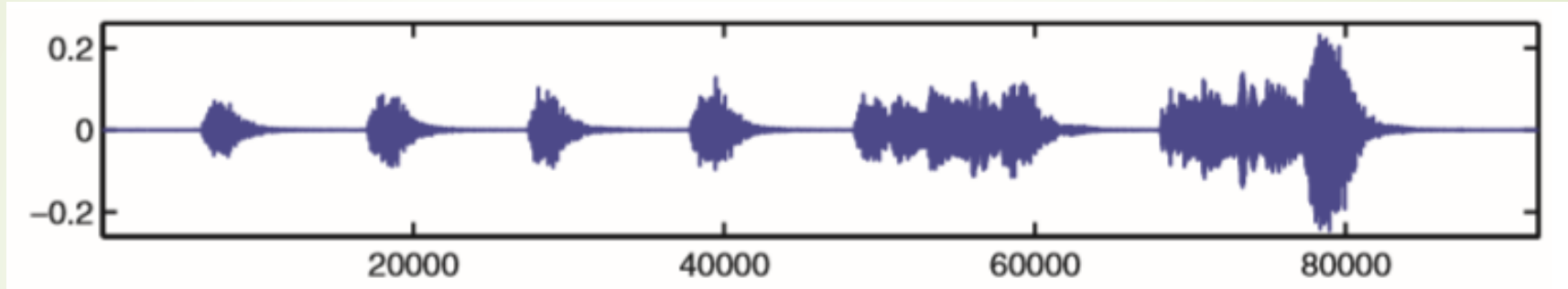
Filtros passa-banda



Filtros passa-banda



Filtros passa-banda

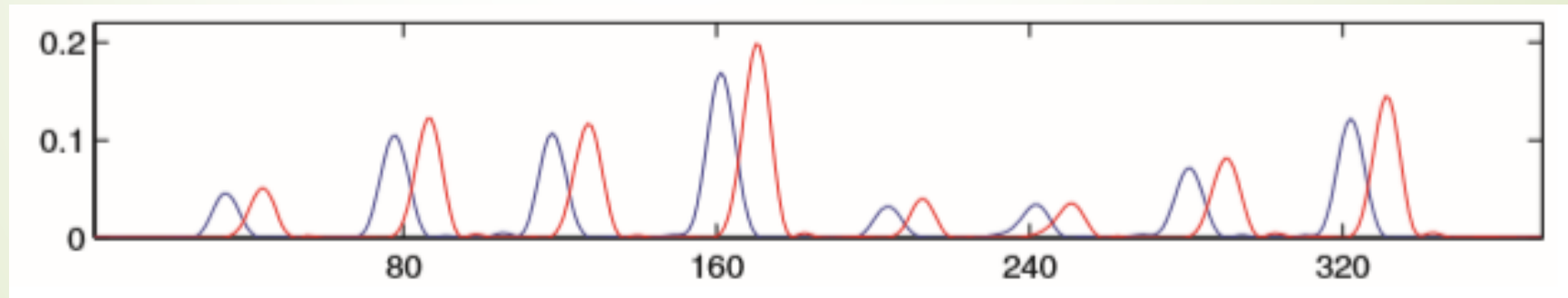
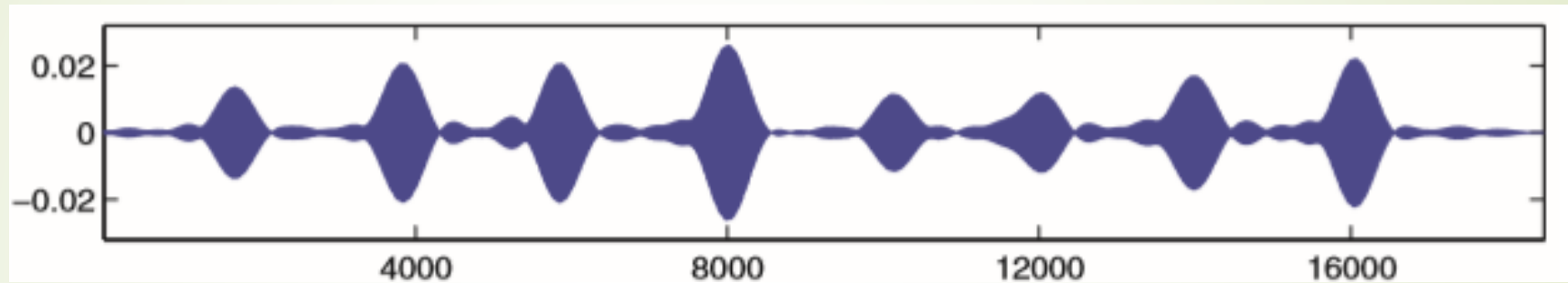


Short-time mean-square power (STMSP)

$$\sum_{k \in [n - \lfloor \frac{w}{2} \rfloor : n + \lfloor \frac{w}{2} \rfloor]} |x(k)|^2$$

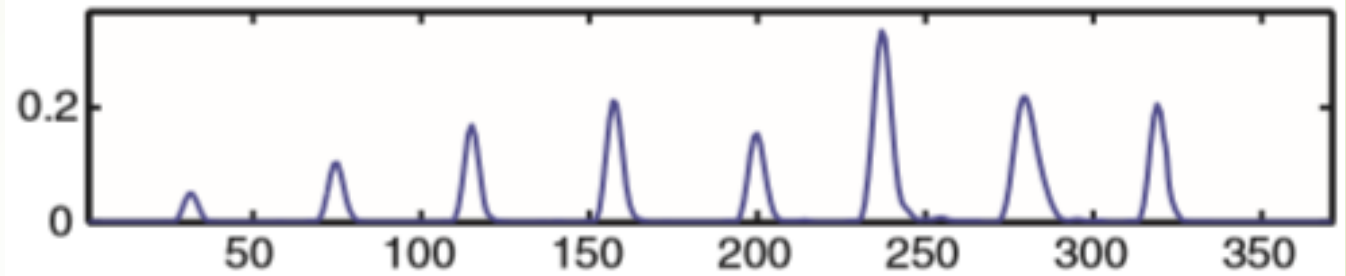
Calculado a cada d amostras.

Short-time mean-square power (STMSP)

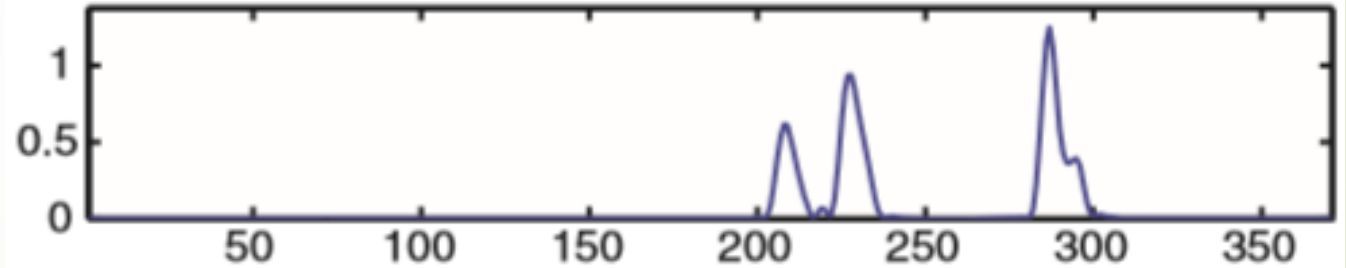


Short-time mean-square power (STMSP)

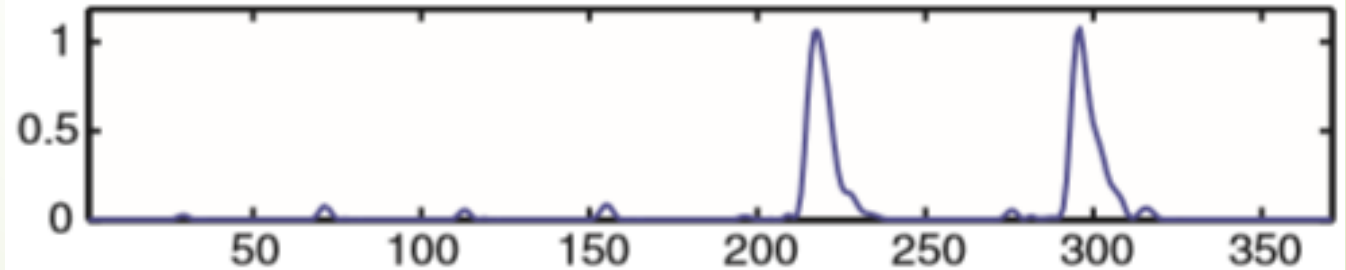
A4



B4

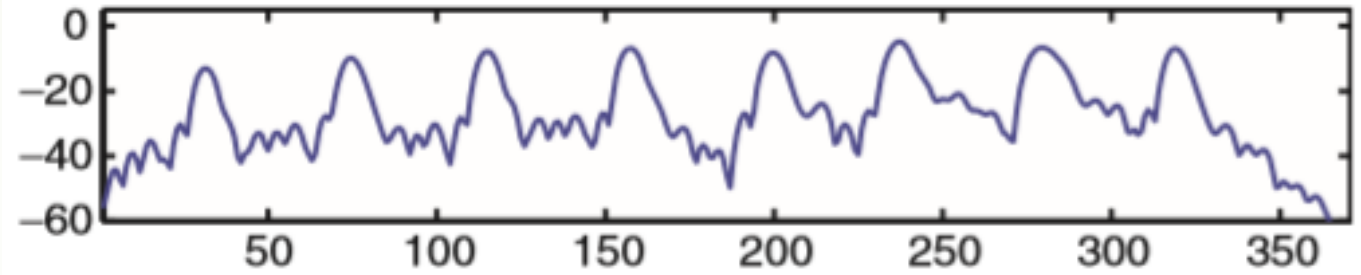


C5

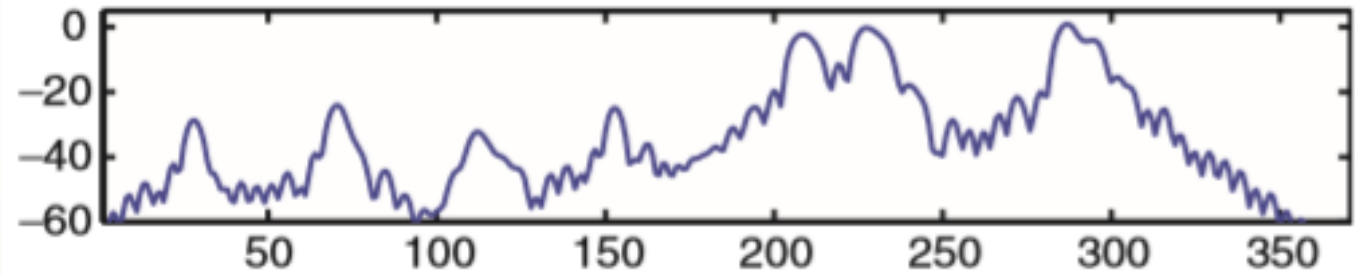


Short-time mean-square power (STMSP)

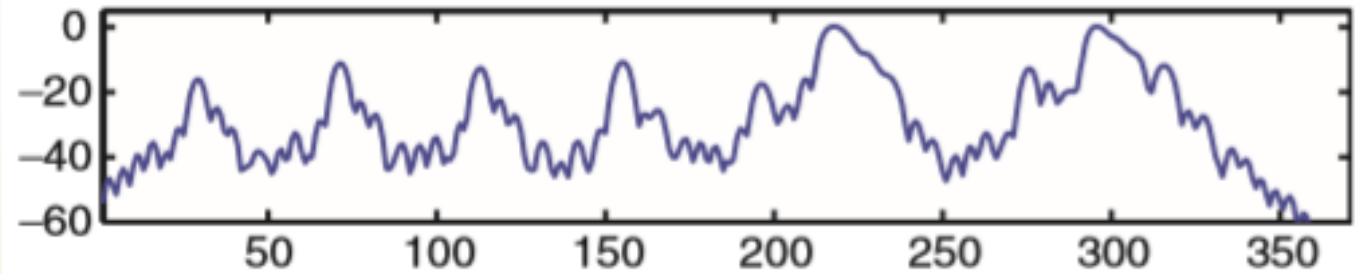
A4



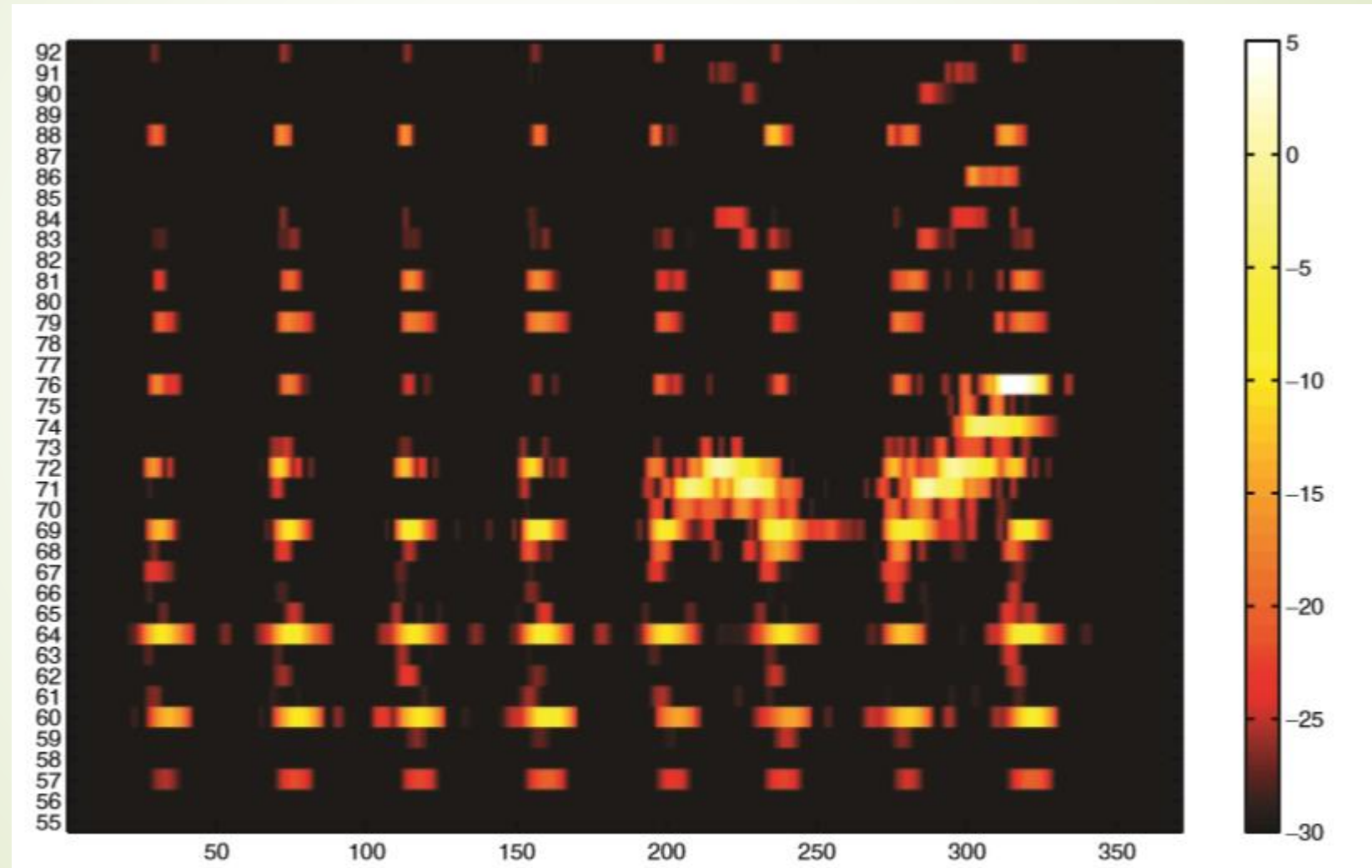
B4



C5



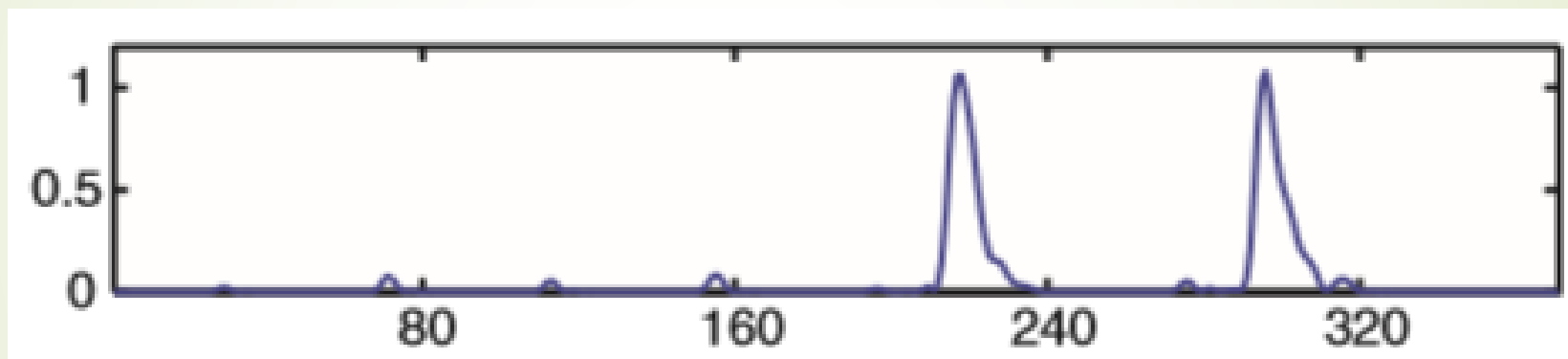
STMSP para cada tom



Reconhecem as estruturas?



Detecção de onsets

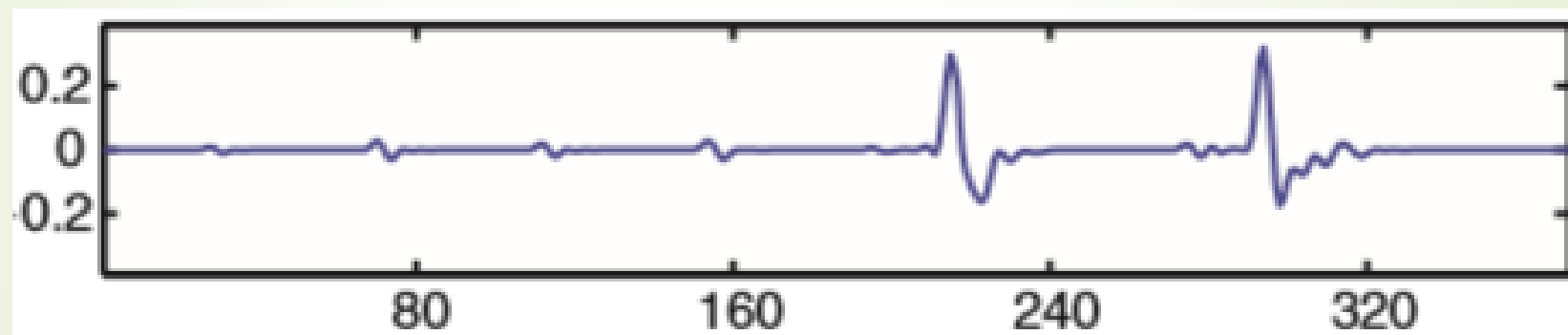




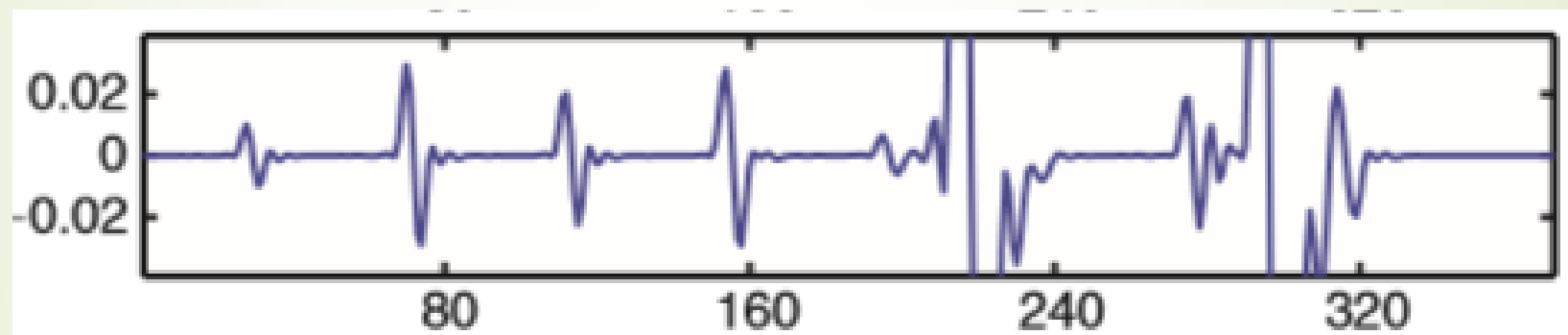
Detecção de onsets

$$x'(n) := x(n) - x(n - 1), n \in \mathbb{Z}$$

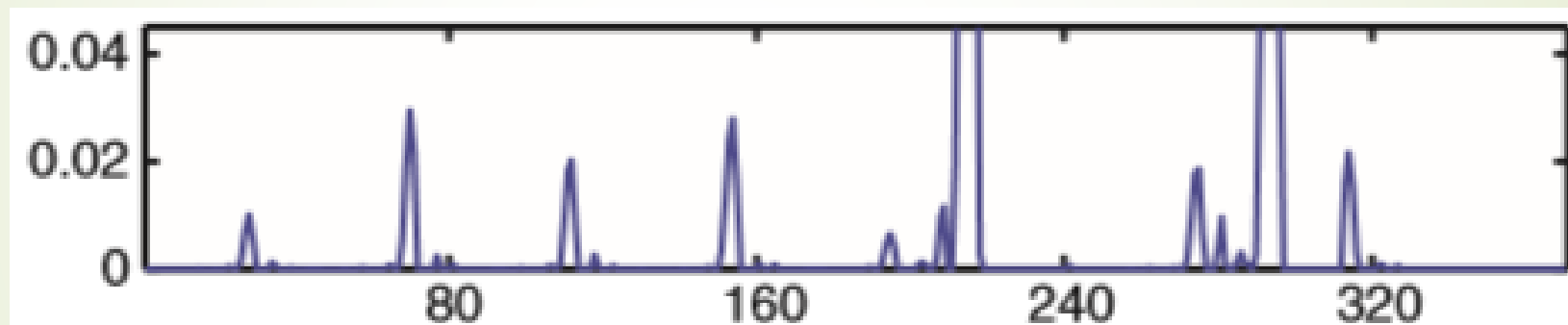
Detecção de onsets



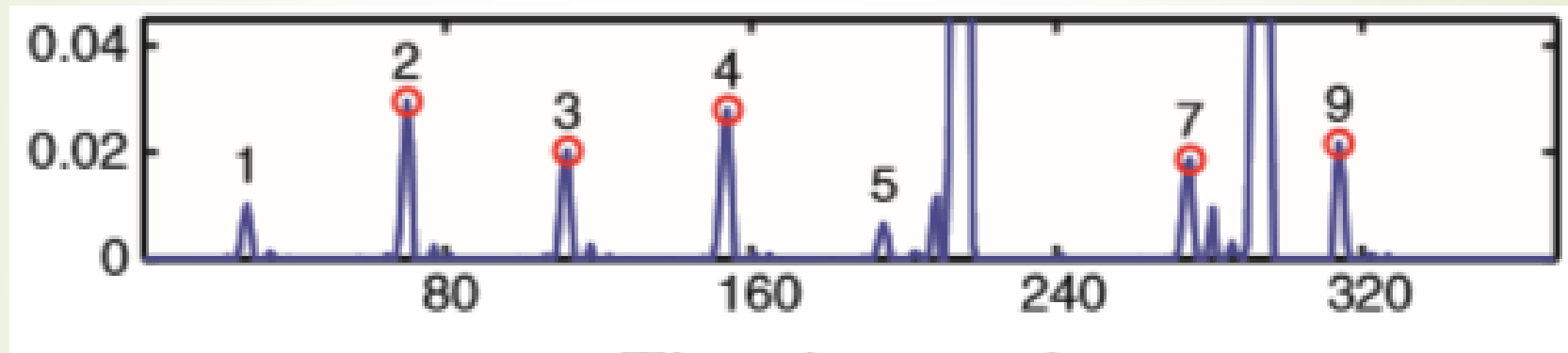
Detecção de onsets



Detecção de onsets



Detecção de onsets

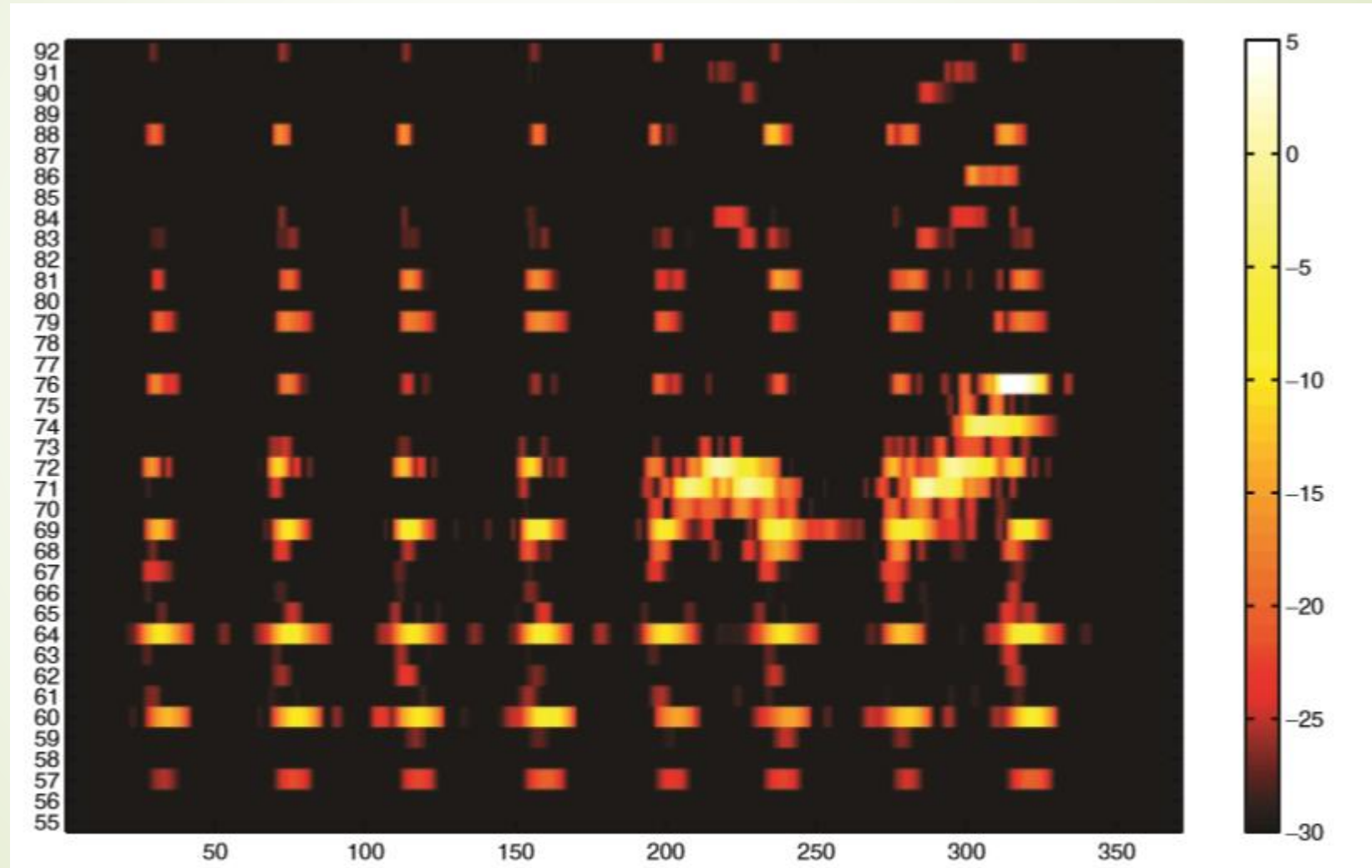


Detecção de Onsets

- Representação por lista de onsets usou entre 30 a 50 vezes menos memória nos testes realizados.
- Facilmente comparável com uma representação em midi

Ticks	Message	Ch.	MNN	Vel
60	NOTE ON	1	67	100
0	NOTE ON	2	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	2	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE ON	2	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	2	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE ON	2	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	2	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	63	100
0	NOTE ON	2	51	100
0	NOTE ON	2	39	100
240	NOTE OFF	1	63	0
0	NOTE OFF	2	51	0
0	NOTE OFF	2	39	0

Voltando ao mapa de STMSP...





Croma

$$C = \{ \dots, C2, C3, C4, C5, \dots \}$$

$$C\# = \{ \dots, C\#2, C\#3, C\#4, C\#5, \dots \}$$

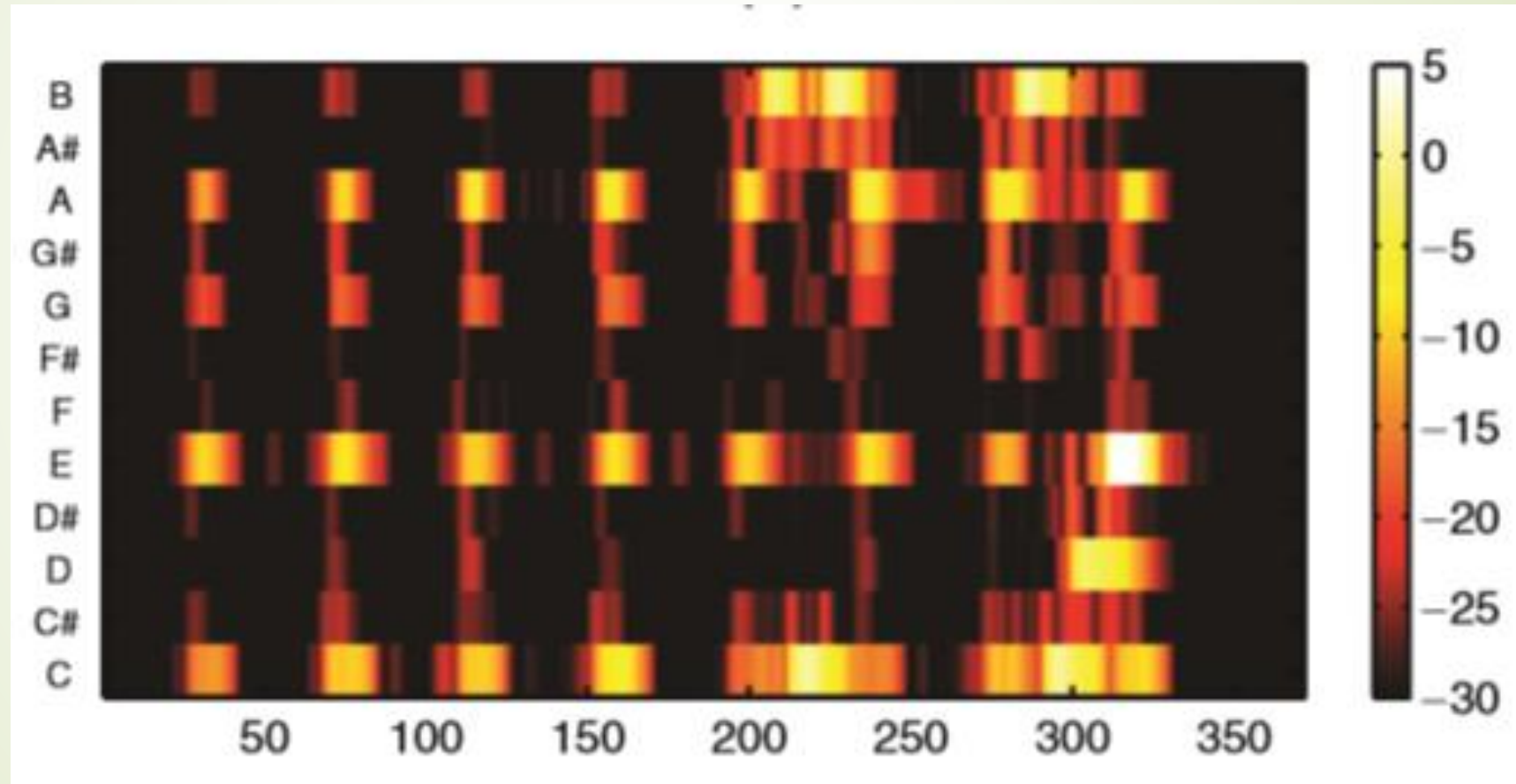
...



Representação por cromas

$$v = (C, C\#, D, D\#, E, F, F\#, G, G\#, A, A\#, B)$$

Características de croma

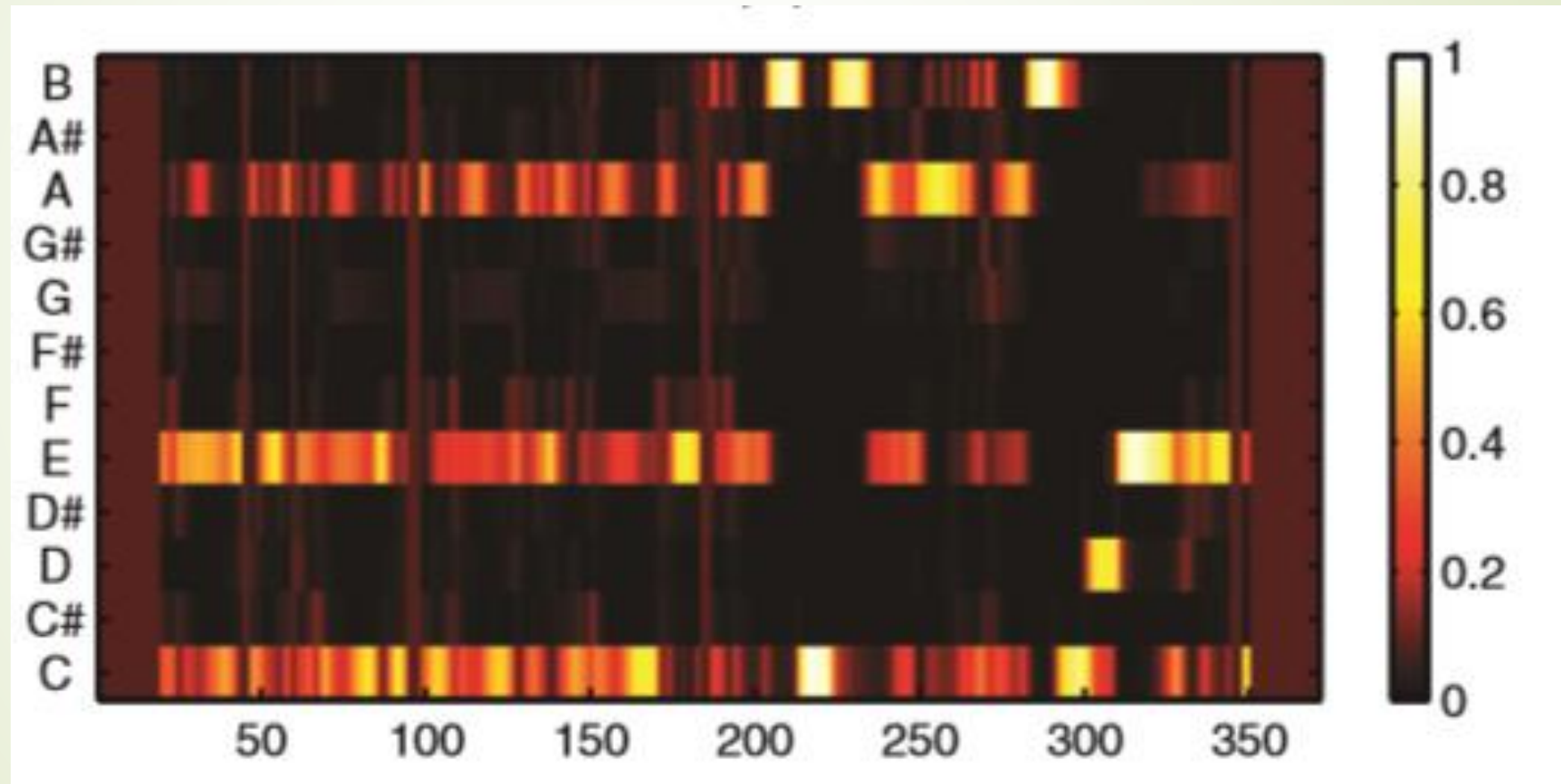




Normalização

$$\frac{v}{\sum_{i=1}^{12} |v(i)|}$$

Características normalizadas de croma

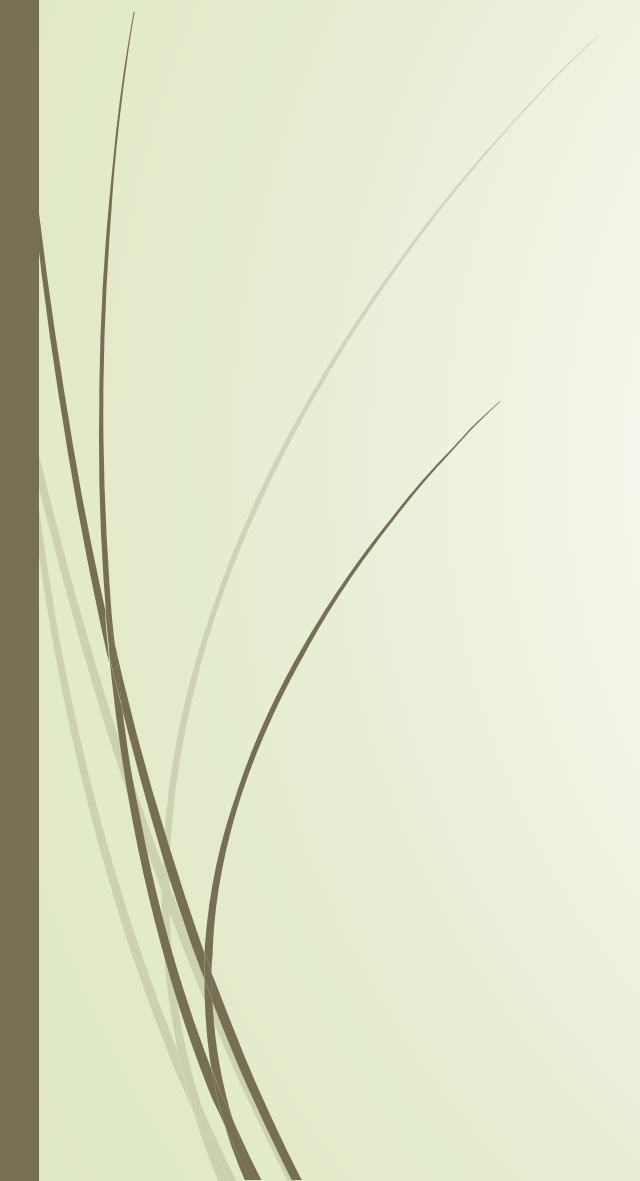




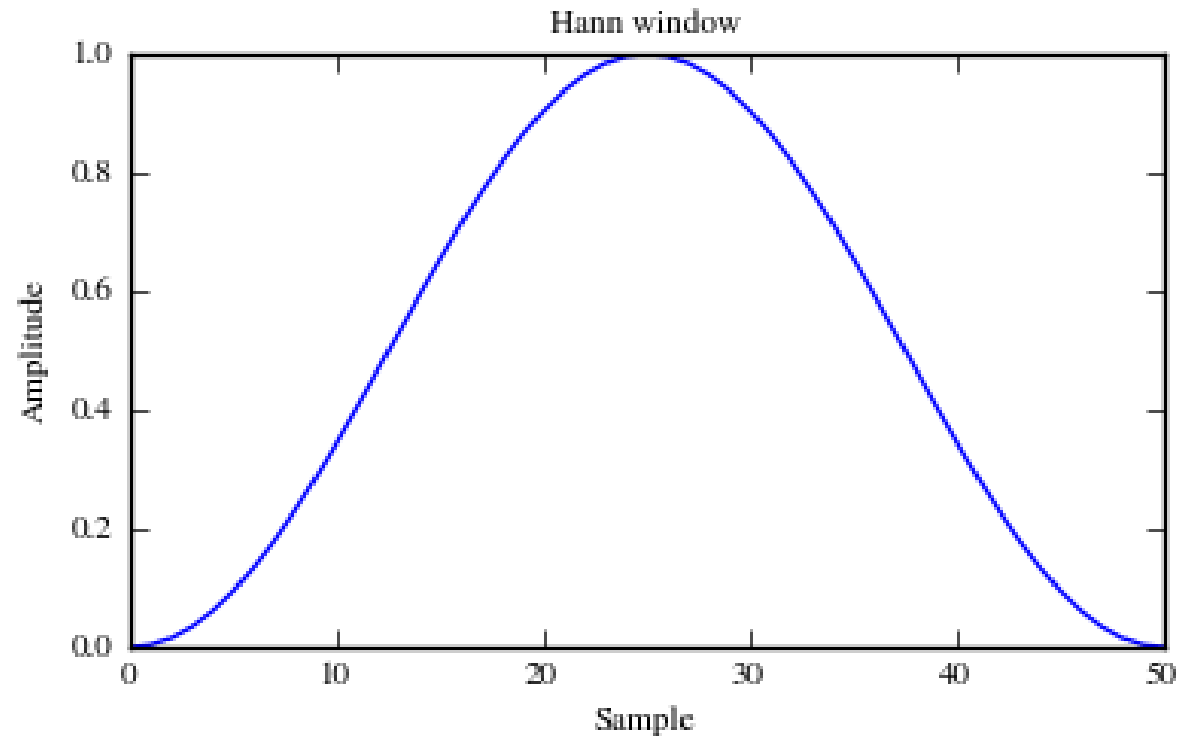
Refinando as características
em mais três passos...




Efetuando a quantização das características de croma

$$\tau(a) := \begin{cases} 0 & \text{para } 0.00 \leq a < 0.05, \\ 1 & \text{para } 0.05 \leq a < 0.10, \\ 2 & \text{para } 0.10 \leq a < 0.20, \\ 3 & \text{para } 0.20 \leq a < 0.40, \\ 4 & \text{para } 0.40 \leq a < 1.00 \end{cases}$$


Convolução com uma janela de Hann






Seguido de uma subamostragem, e mais uma normalização (euclidiana)

$$\|v\|_2 := \left(\sum_{i=1}^{12} v(i)^2 \right)^{\frac{1}{2}}$$

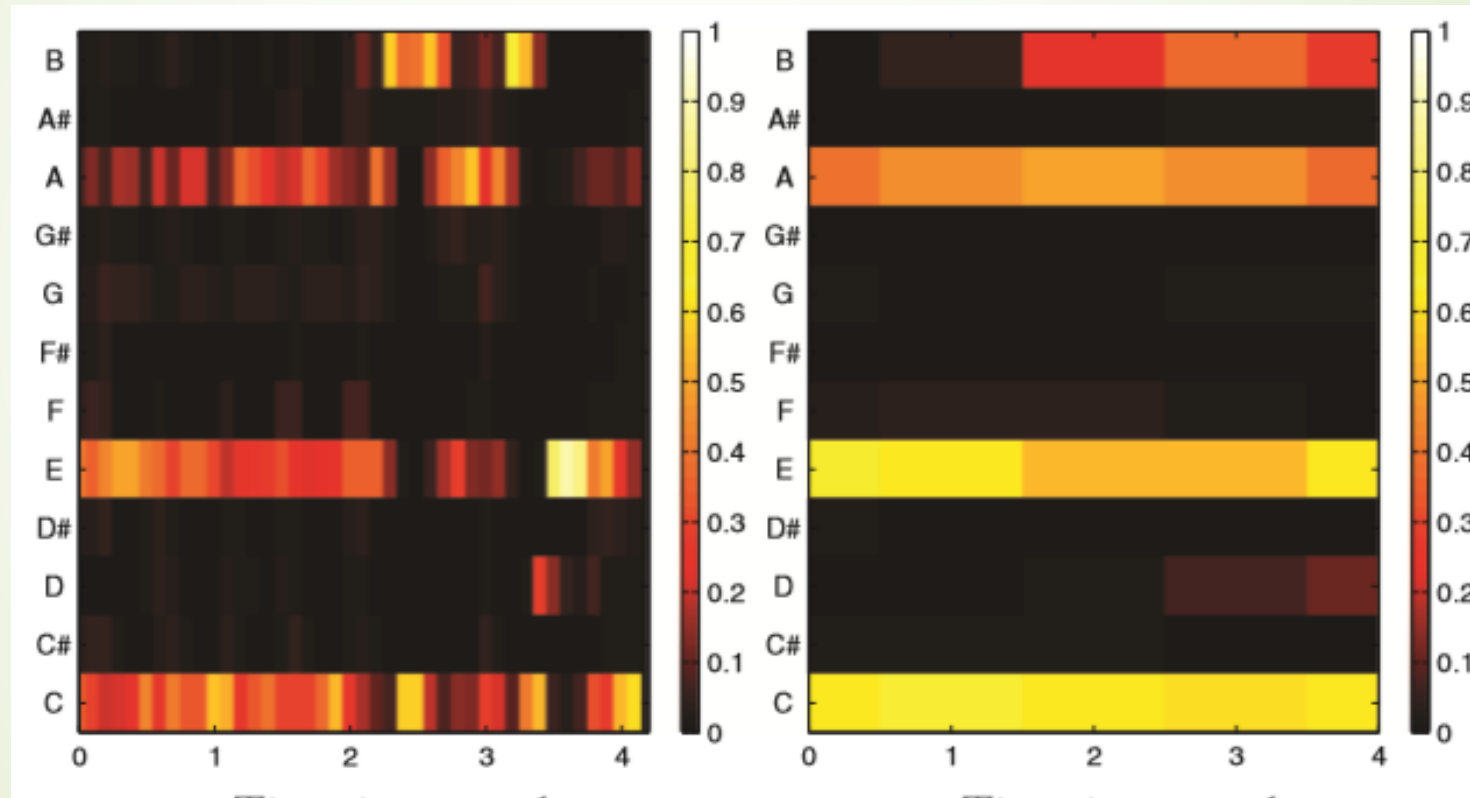
$$\mathcal{F} := \{v' = (v = (C, C\#, \dots, B)) \in [0,1]^{12} \mid \|v\|_2 = 1\}$$

Chroma Energy Normalized Statistics

*CENS*_{*d*}^{*w*}

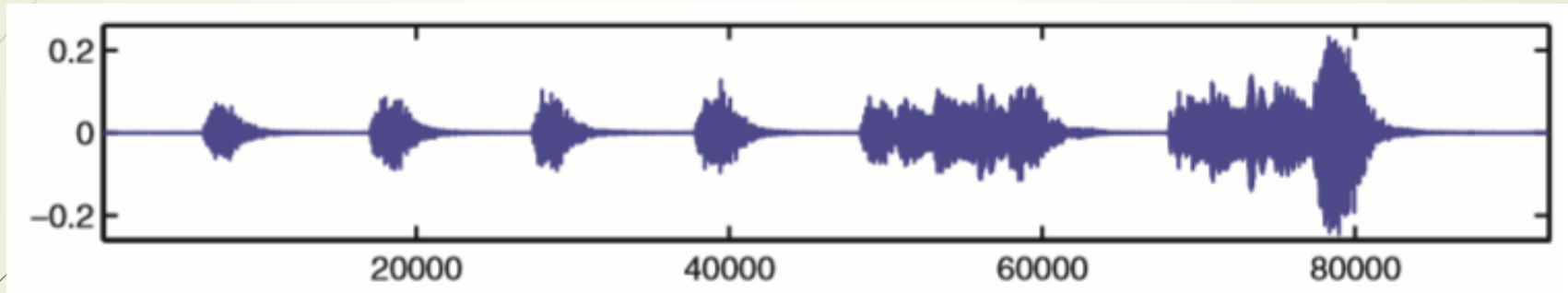

$$\mathcal{F} := \{v' = (v = (C, C\#, \dots, B)) \in [0,1]^{12} \mid \|v\|_2 = 1\}$$

$CENS_{10}^{41}$

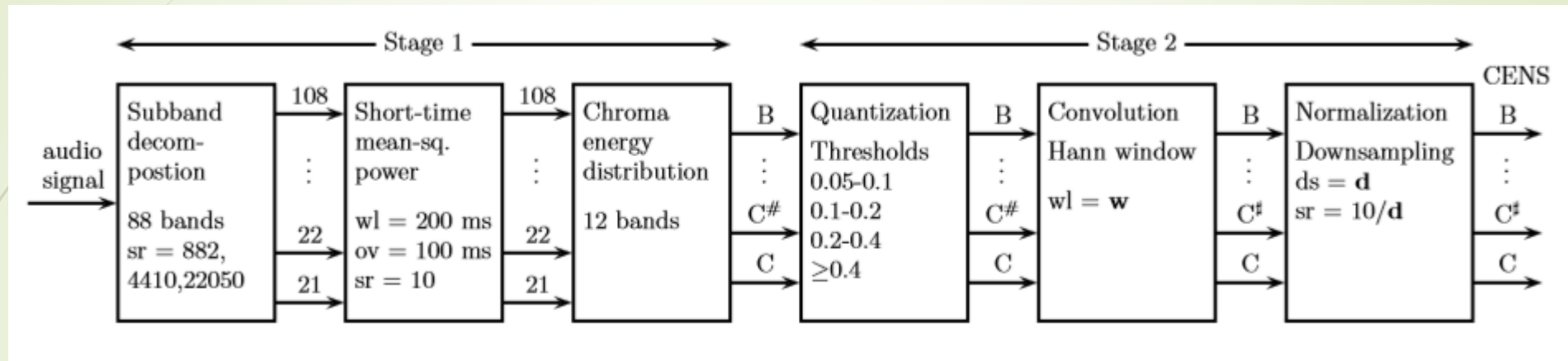




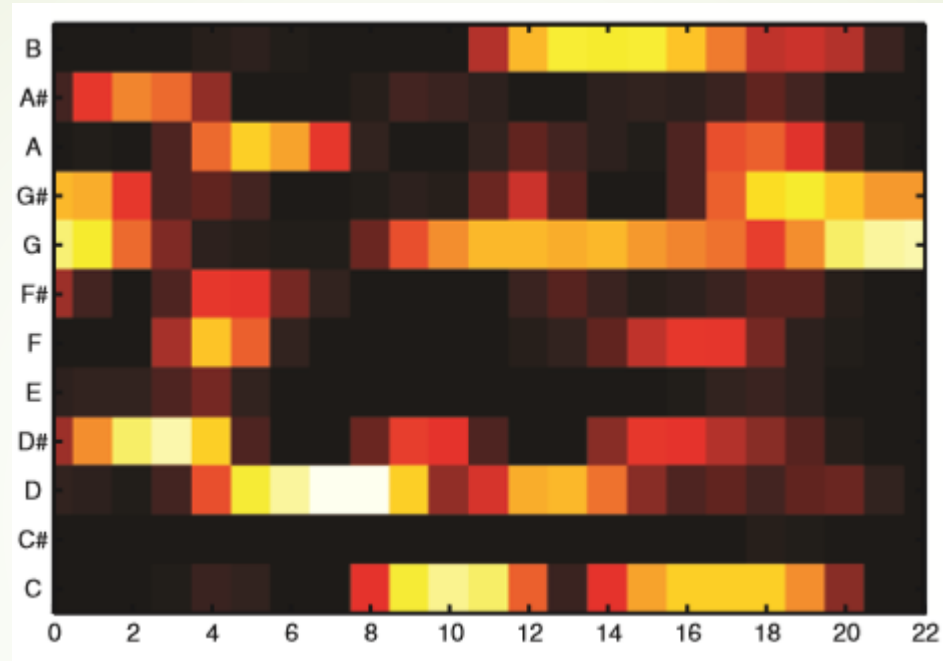
Revisando...



Um arquivo de áudio...

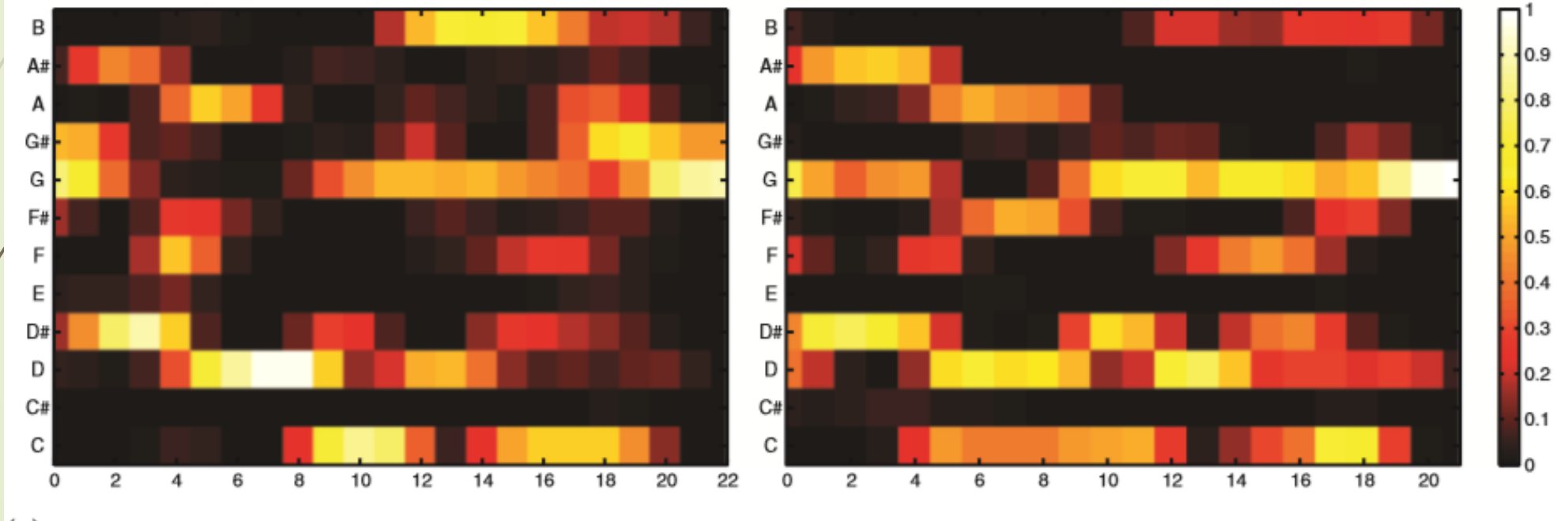


... extração de características ...



... $CENS_d^w$.

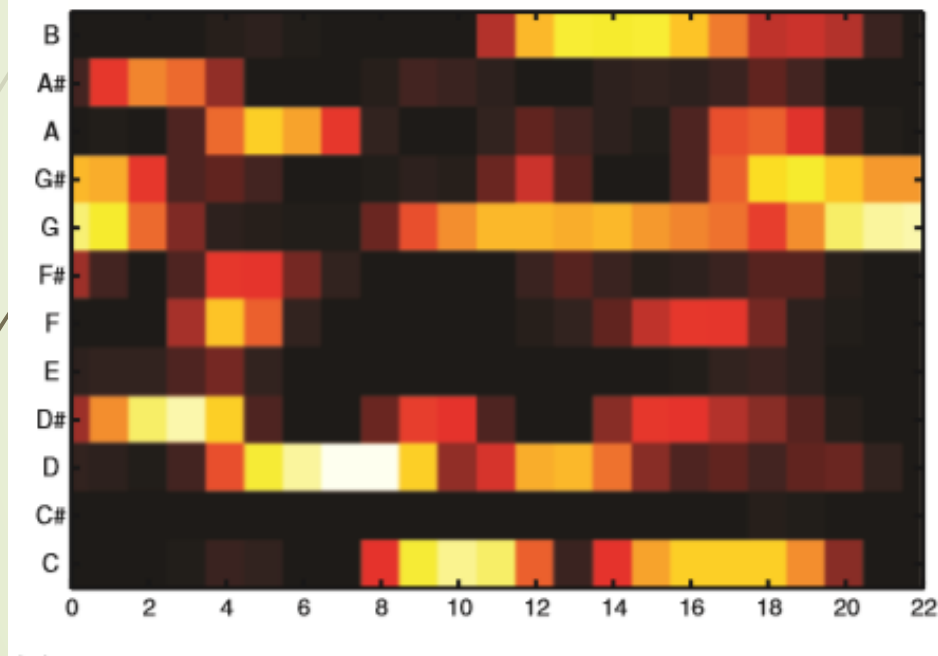
Utilizando estas características...



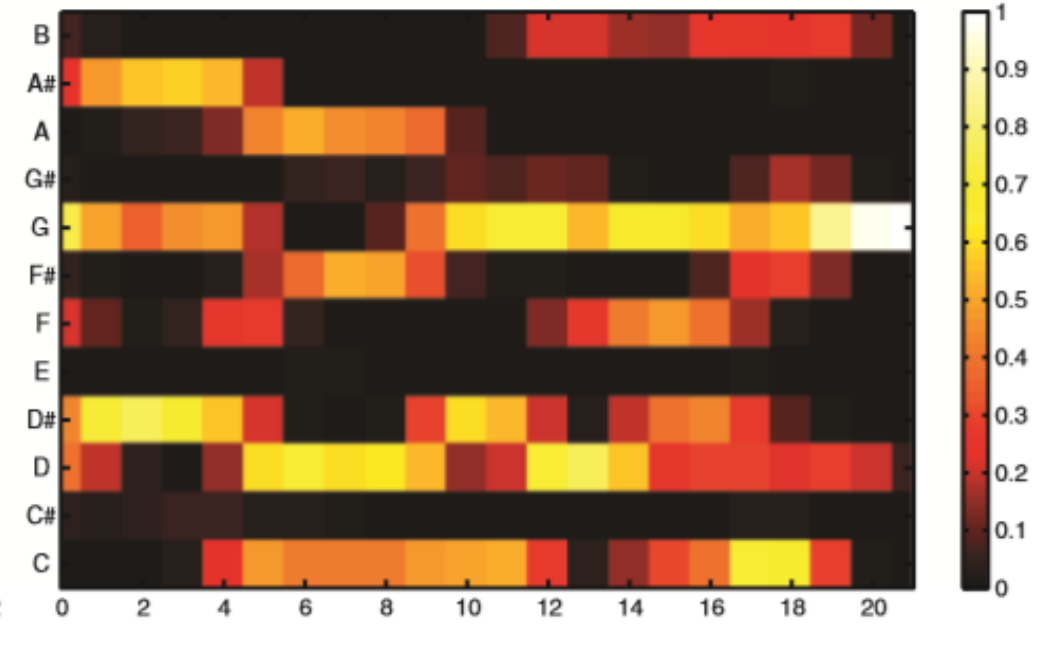
Utilizando estas características...

Quinta sinfonia de Beethoven

Bernstein (orquestral)



Gould (piano)





Perguntas?

Obrigado!

Álvaro Almeida Freire Stivi

astivi@ime.usp.br

Grupo de Computação Musical

<http://compmus.ime.usp.br>

Instituto de Matemática e Estatística – Universidade de São Paulo