

# Melodia e Harmonia

Marcelo Cicconet

# Melodia

- ... uma sucessão de tons composta de modo, ritmo e notas musicais arranjadas de modo a alcançar forma musical, sendo percebida como uma unidade pela mente ... (Virginia Tech Multimidia Music Dictionary)
- Parte obscura da música que permite ao humano “repeti-la” assobiando
- Sequência de nota musicais ao longo do tempo

# Harmonia

- Combinação consonante de notas tocadas simultaneamente para produzir acordes (Virginia Tech Multimidia Music Dictionary)
- Acompanhamento

# Melodia e Harmonia

- Ambas são fundamentadas no conceito de **nota musical** (pitch)
- Nota musical lembra **frequência**

# Melodia e Harmonia

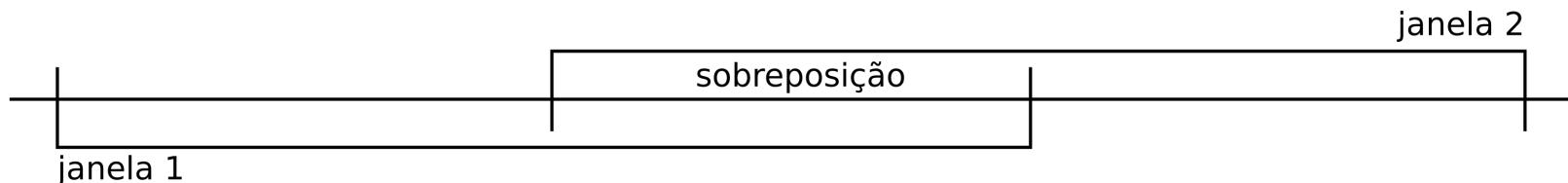
- A seguir
  - Análise
  - Composição
  - Síntese

# Análise

- Baixo nível
  - Trabalha-se sobre a onda sonora em formato digital (PCM) observando conceitos físicos como frequência e amplitude
- Alto nível
  - Trabalha-se com representações simbólicas (do tipo MIDI) levando-se em conta aspectos de teoria musical (escala, modo, acorde principal, etc)

# Análise

- Descritores harmônicos/timbrais
  - **Descriptor:** vetor representando alguma propriedade (característica física) do áudio
  - Normalmente obtidos de segmentos de tamanho constante (**janelas**) que se sobrepõe
  - Para 44100 fps o tamanho da janela varia de 512 a 4096 frames
  - Sobreposição: geralmente  $\frac{1}{2}$  do tamanho da janela



# Análise

- Taxa de eventos
  - Inversão do sinal em frames consecutivos
  - Anulação da derivada do sinal

# Análise

- Transformada de Fourier discreta (DFT)
  - Devolve a amplitude complexa com que as frequências entre zero e a metade da taxa de amostragem estão presentes no sinal

$$x = (x_0, \dots, x_{N-1})$$

$$\hat{x} = (\hat{x}_0, \dots, \hat{x}_{N-1})$$

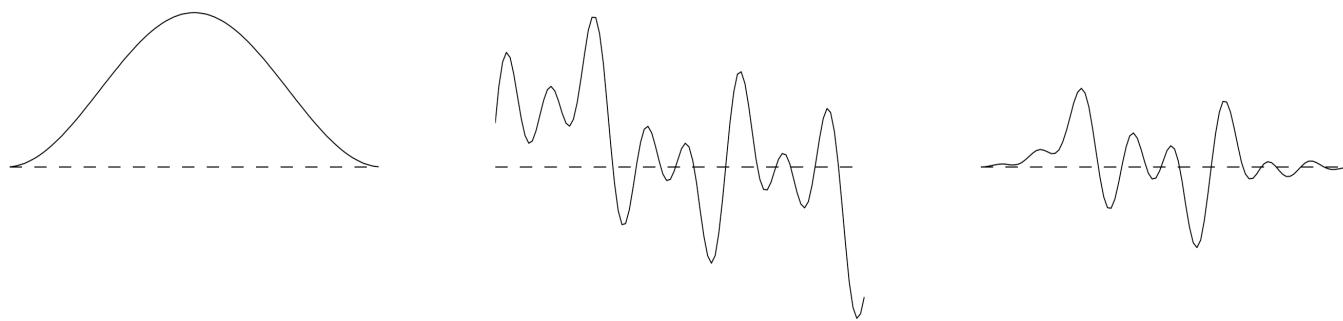
$$\hat{x}_k = \sum_{n=0}^{N-1} x_n e^{-2\pi i k n / N}$$

# Análise

- Transformada de Fourier discreta (DFT)
  - Se a taxa de amostragem é de 44100 fps e o tamanho da janela é de 1024 frames, a DFT devolve 512 valores de frequência igualmente espaçados entre zero e 22050
  - Um aumento do tamanho da janela não aumenta a gama de frequências avaliadas, mas a **resolução** de avaliação das frequências
  - Estratégia comum: “zero-padding”

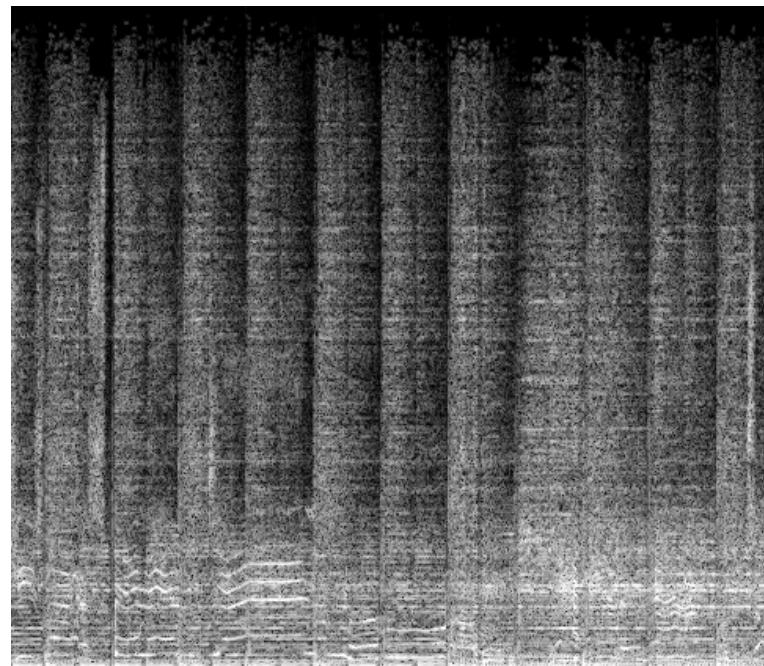
# Análise

- Transformada de Fourier discreta (DFT)
  - Janelamento (“windowing”)



# Análise

- Espectro de potências
  - Quadrados das magnitudes dos coeficientes de Fourier

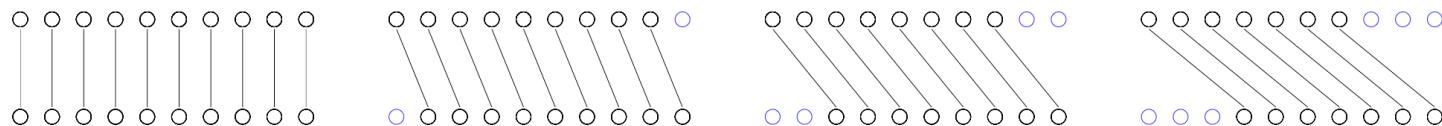


# Análise

- Amplitude (loudness)
  - Média das entradas do espectro de potências
  - Medida da energia do sinal sonoro
  - Muitas vezes representado em escala logarítmica  
(por razões perceptuais)

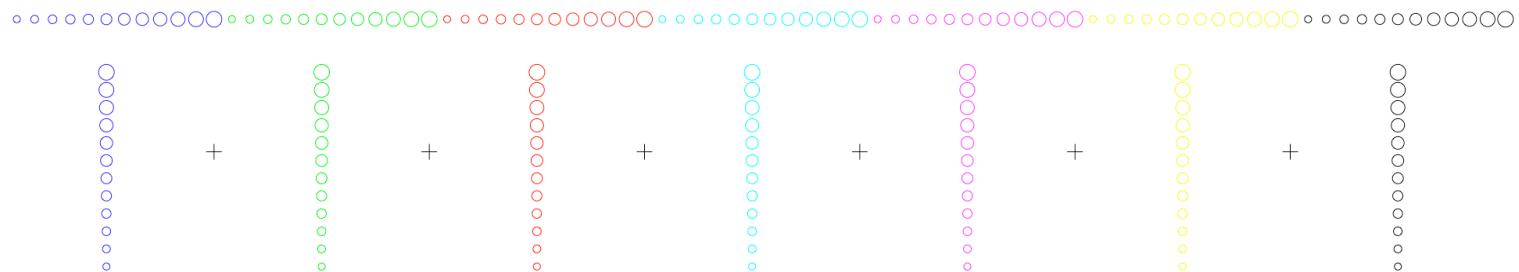
# Análise

- Auto-correlação
  - Medida de quanto um sinal é semelhante à ele mesmo transladado



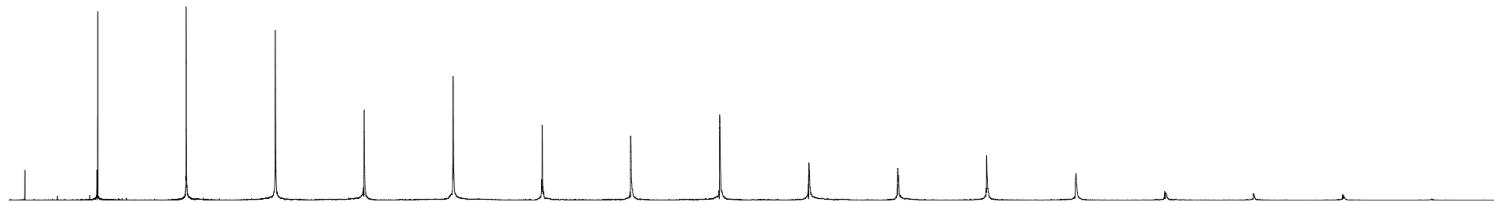
# Análise

- Vetor de croma
  - Muito usado para detecção de acordes



# Análise

- Cepstrum
  - Magnitude da transformada de Fourier inversa do logaritmo da magnitude da transformada de Fourier do sinal



# Análise

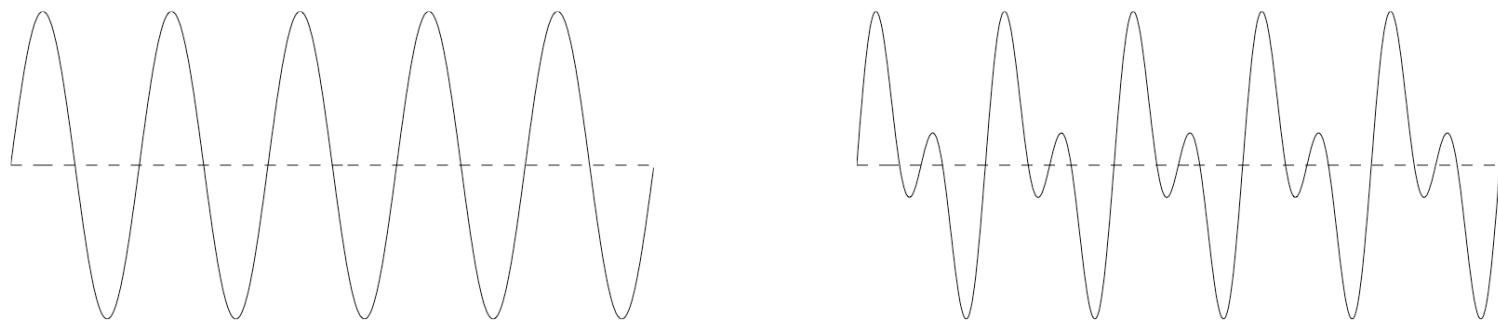
- Exemplos
  - AudioFeature
  - Audio Units

# Análise

- Detecção de frequência fundamental (F0)
  - Frequência fundamental: frequência tal que seus múltiplos explicam melhor o conteúdo do espetro do sinal
  - Aplicações: afinador eletrônico, transcrição automática de melodia, acompanhamento automático, reconhecimento de fala, transposição de notas musicais, ...
  - Os algoritmos se dividem em duas classes: os que atuam no domínio do tempo e os que atuam no domínio da frequência

# Análise

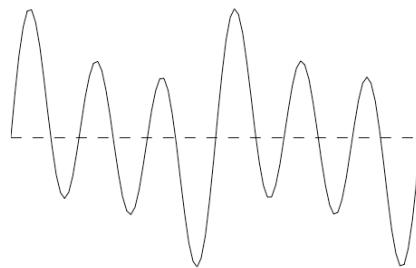
- Detecção de frequência fundamental (F0)
  - Estimação via Taxa de Eventos



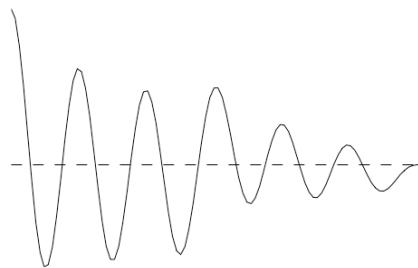
# Análise

- Detecção de frequência fundamental (F0)
  - Método de McLeod: procura-se por máximos de uma função de auto-correlação normalizada

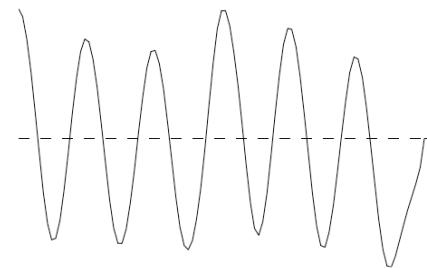
$$x_n = \sin(2\pi \cdot 2n) + \sin(2\pi \cdot 4n) + 3\sin(2\pi \cdot 6n)$$



auto-correlação



$$c_k = \frac{2 \sum_{n=0}^{N-1-k} x_n x_{n+k}}{\sum_{n=0}^{N-1-k} x_n^2 + x_{n+k}^2}$$

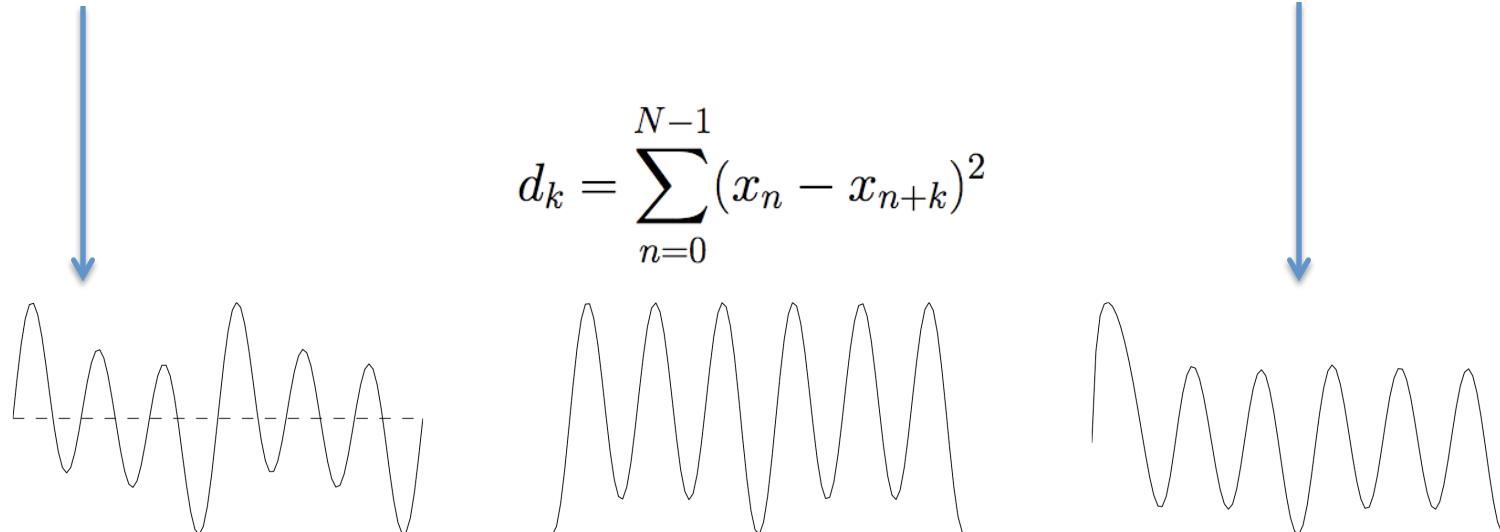


# Análise

- Detecção de frequência fundamental (F0)
  - Método YIN: procura-se por mínimos de uma **função diferença normalizada**

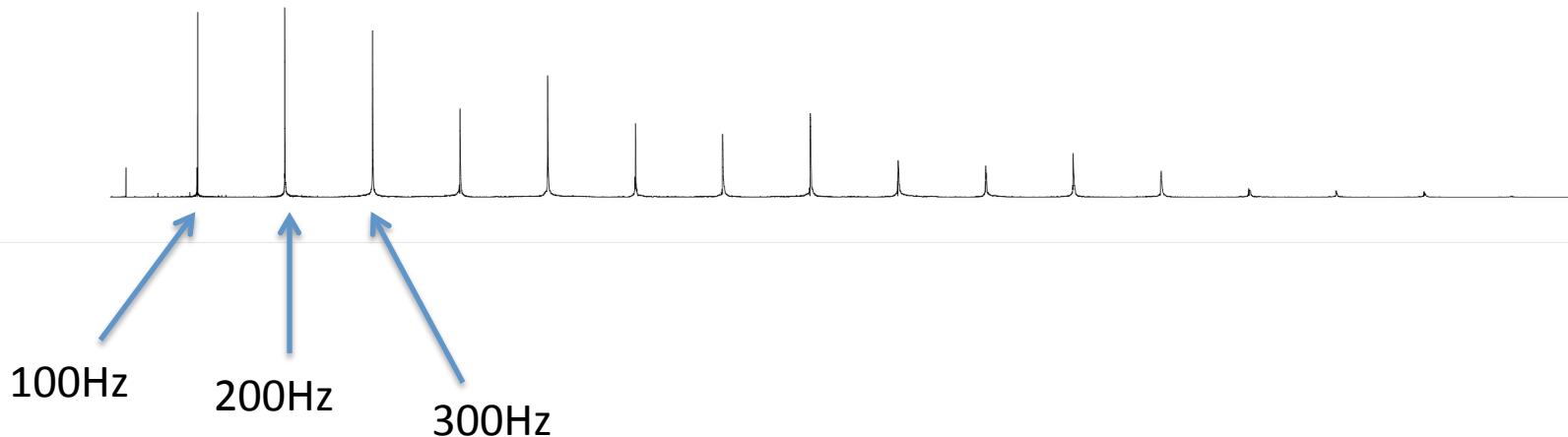
$$x_n = \sin(2\pi \cdot 2n) + \sin(2\pi \cdot 4n) + 3 \sin(2\pi \cdot 6n)$$

$$\bar{d}_k = 1_{[k=0]} + 1_{[k \neq 0]} \frac{d_k}{\frac{1}{k} \sum_{j=1}^k d_j}$$



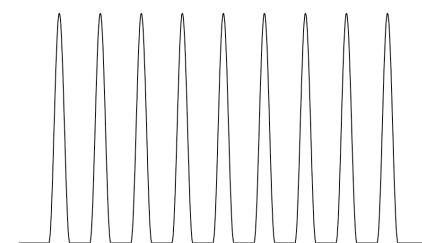
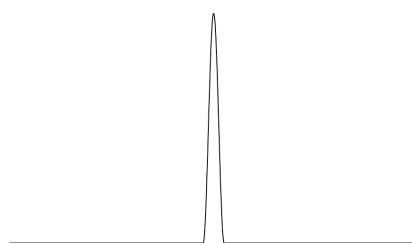
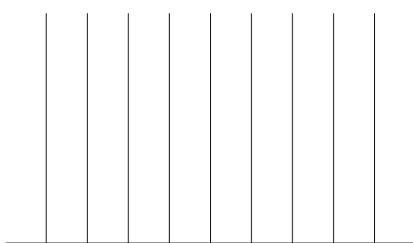
# Análise

- Detecção de frequência fundamental (F0)
  - Método HPS (Harmonic Product Spectrum): percorre-se o conjunto de frequências e verifica-se qual delas tem harmônicos cujo produto das energias é maior



# Análise

- Detecção de frequência fundamental (F0)
  - Máxima verossimilhança: cria-se um banco de dados com os chamados “espectros ideais” e, dado um espectro de cuja onda se quer saber a frequência fundamental, procura-se na base de dados o espectro mais próximo



# Análise

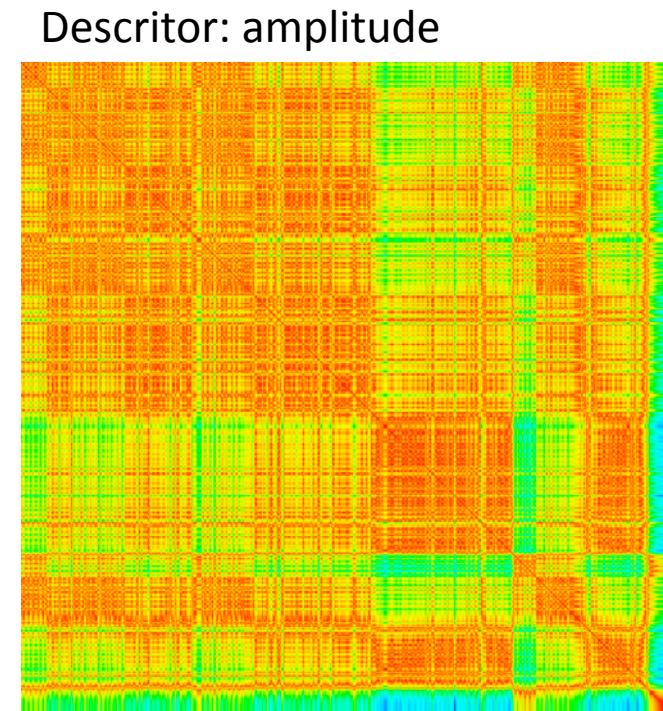
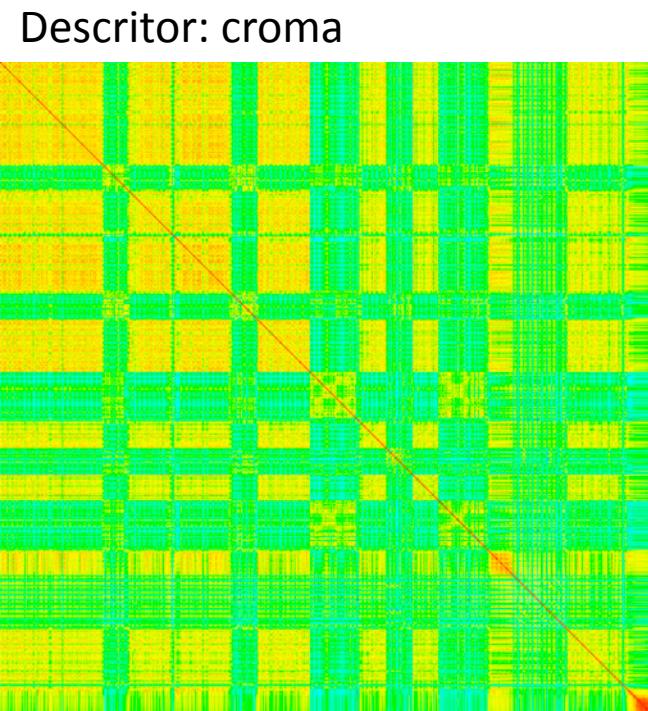
- Exemplos
  - Audio Units

# Análise

- Deteção de acordes
  - Cria-se uma coleção de vetores de croma de acordes reais, com várias amostras do mesmo acorde, as quais tendem a formar agrupamentos
  - Dado um acorde a classificar, utiliza-se algum método de Aprendizagem de Máquina ( $k$ -nn, por exemplo) para associar o acorde a um dos agrupamentos

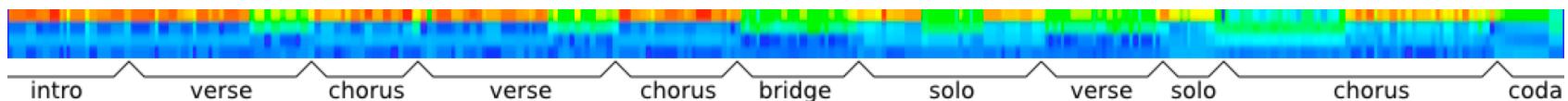
# Análise

- Análise estrutural de peças musicais
  - Matriz de auto-similaridade (SSM)



# Análise

- Análise estrutural de peças musicais
  - Matriz de auto-similaridade (SSM): Aplicações:
    - Segmentação
    - Resumo de peças musicais (thumbnails)
    - Audio fingerprint
    - Visualização

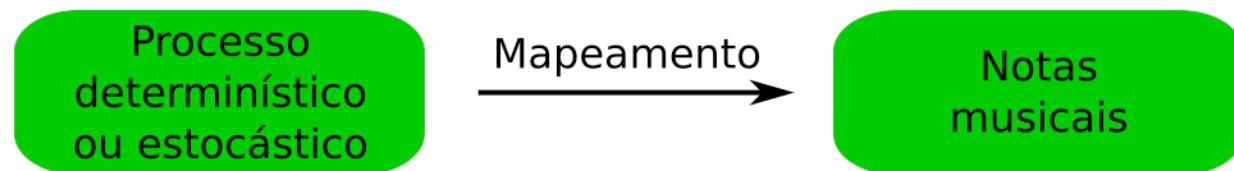


# Análise

- Exemplos
  - AudioFeature
  - EigenSound

# Composição

- Protocolo para composição automática:

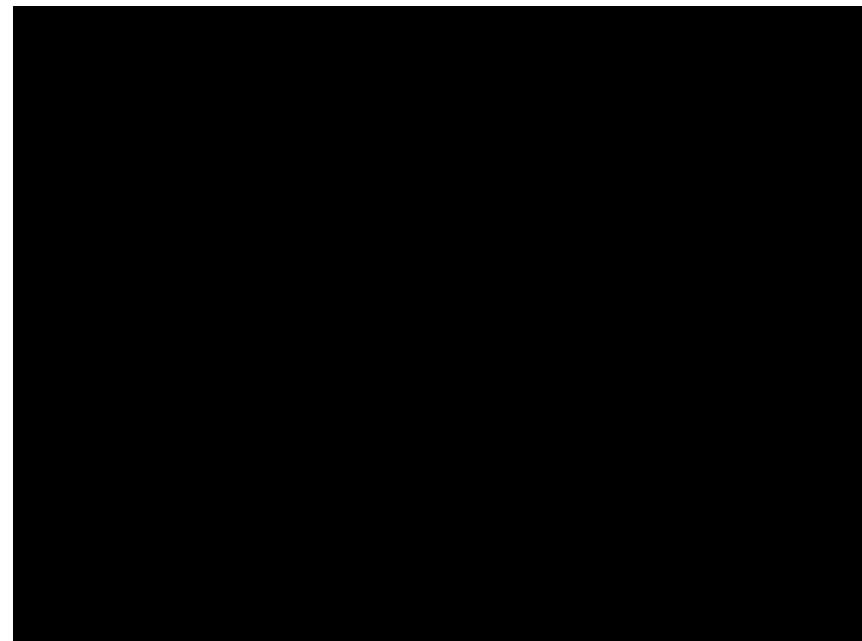


# Composição

- Algumas ferramentas utilizadas:
  - Cadeias de Markov
  - Autômatos celulares
  - Fractais
  - Redes Neurais

# Composição

- Exemplos:
  - Continuator, por François Pachet
    - <http://www.csl.sony.fr/~pachet/Continuator/>



# Composição

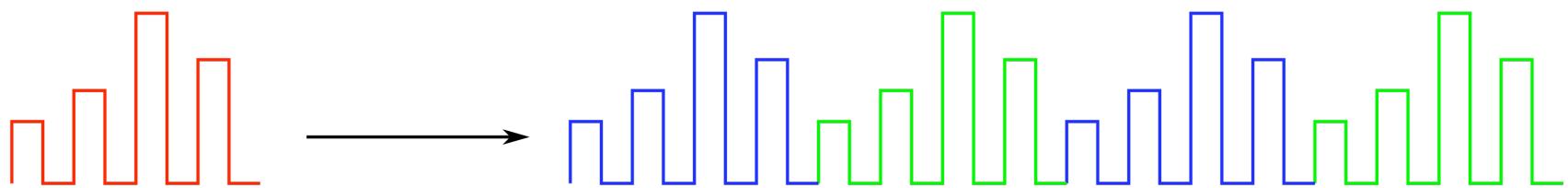
- Exemplos:
  - Music Stretching, por Tristan Jehan
    - <http://web.media.mit.edu/~tristan/>

Original (20s) 

Stretched (80s) 

# Síntese

- Por wavetable
  - Reprodução sucessiva da mesma forma de onda, armazenada numa tabela

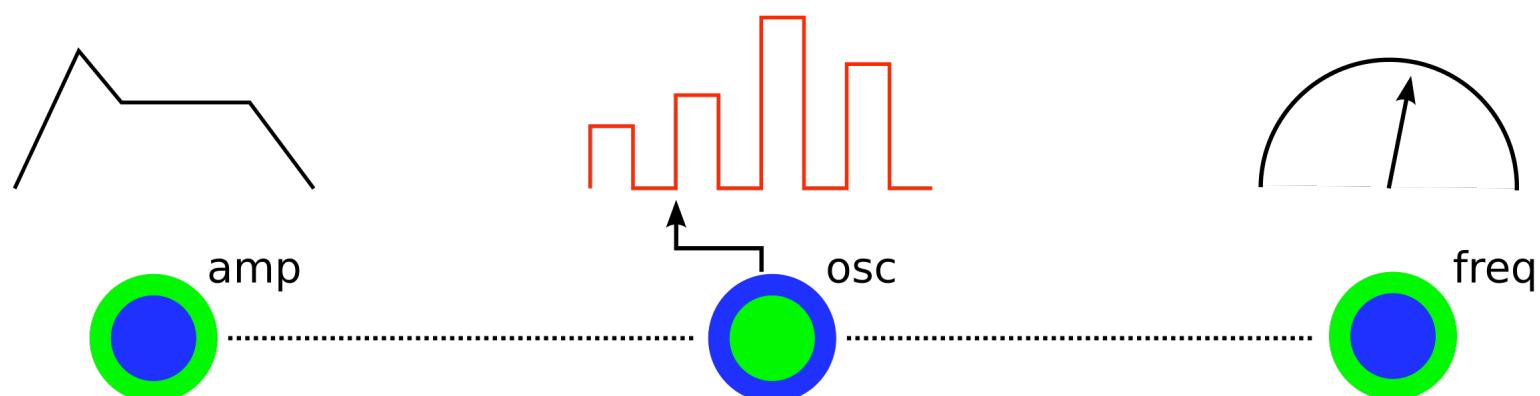


- Tais ondas podem ser sintéticas



# Síntese

- Síntese por wavetable
  - Unidade de síntese: oscilador, controlador de amplitude e controlador de frequência



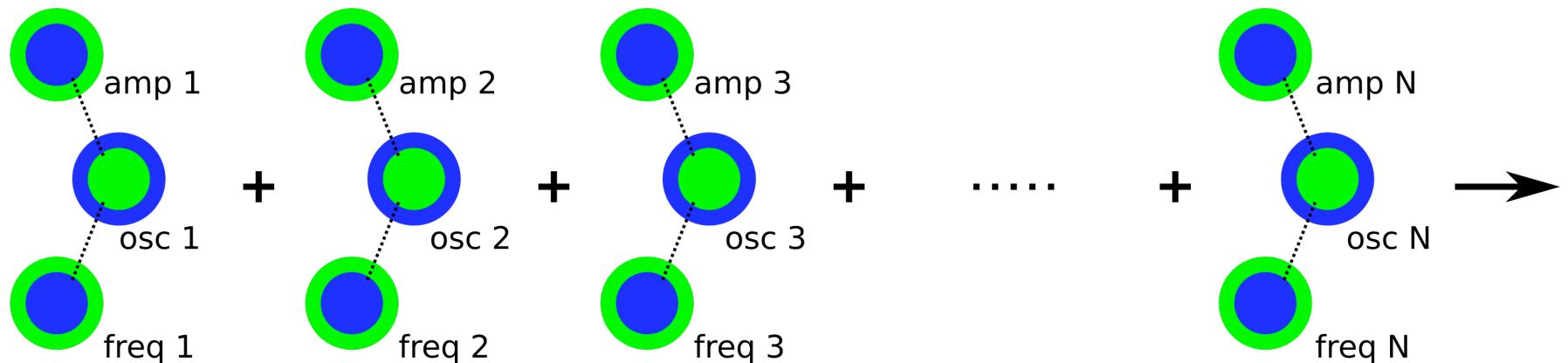
# Síntese

- Síntese por wavetable
  - Envelope de onda: (1) ataque, (2) decaimento inicial, (3) sustentação, (4) decaimento final



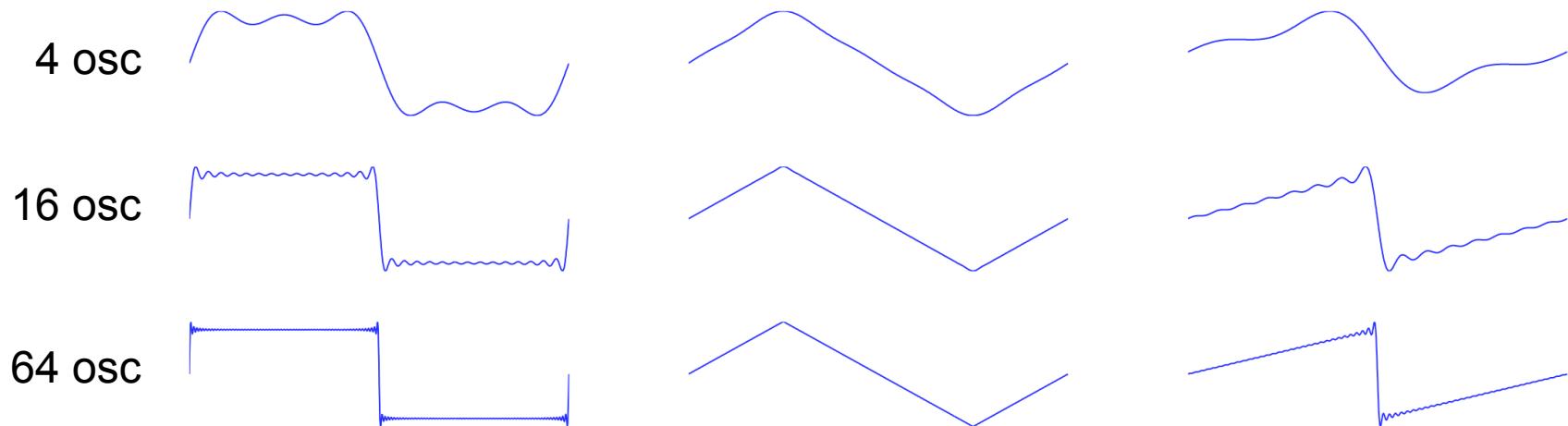
# Síntese

- Síntese aditiva
  - Soma das formas de onda geradas por um banco de fontes



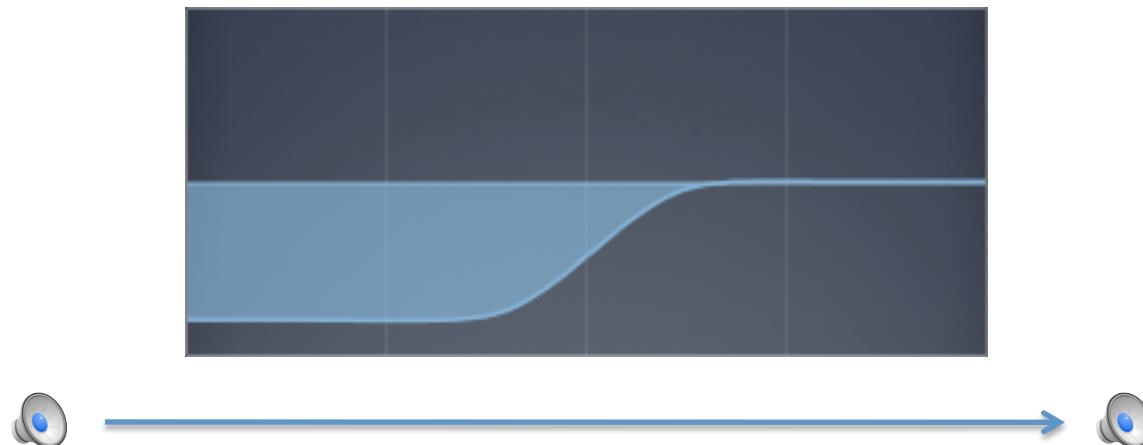
# Síntese

- Síntese aditiva
  - A aproximação fornecida pela série de Fourier pode ser vista como síntese aditiva



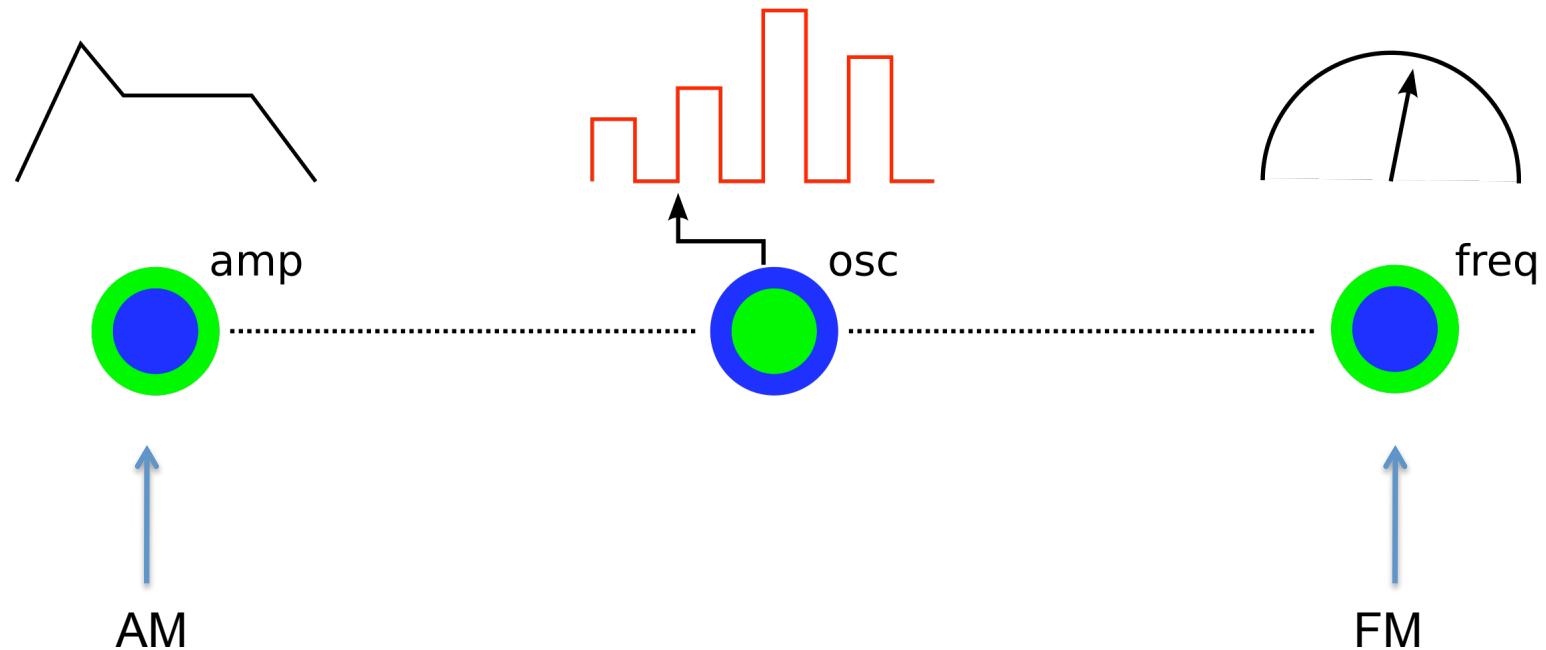
# Síntese

- Síntese substrativa
  - Confunde-se com filtragem



# Síntese

- Síntese por modulação



# Síntese

- Síntese por modulação
  - Síntese FM

$$y_n = A \sin \left( \frac{2\pi n f_c}{R} + I \sin \frac{2\pi n f_m}{R} \right)$$

Diagram illustrating the parameters of the synthesized signal equation:

- frequência base (Frequency base) points to the term  $\frac{2\pi n f_c}{R}$ .
- índice de modulação (Modulation index) points to the term  $I \sin \frac{2\pi n f_m}{R}$ .
- frequência moduladora (Modulating frequency) points to the term  $\frac{2\pi n f_m}{R}$ .
- amplitude (Amplitude) points to the coefficient  $A$ .
- taxa de amostragem (Sampling rate) points to the denominator  $R$ .

$$I = \Delta_f / f_m$$

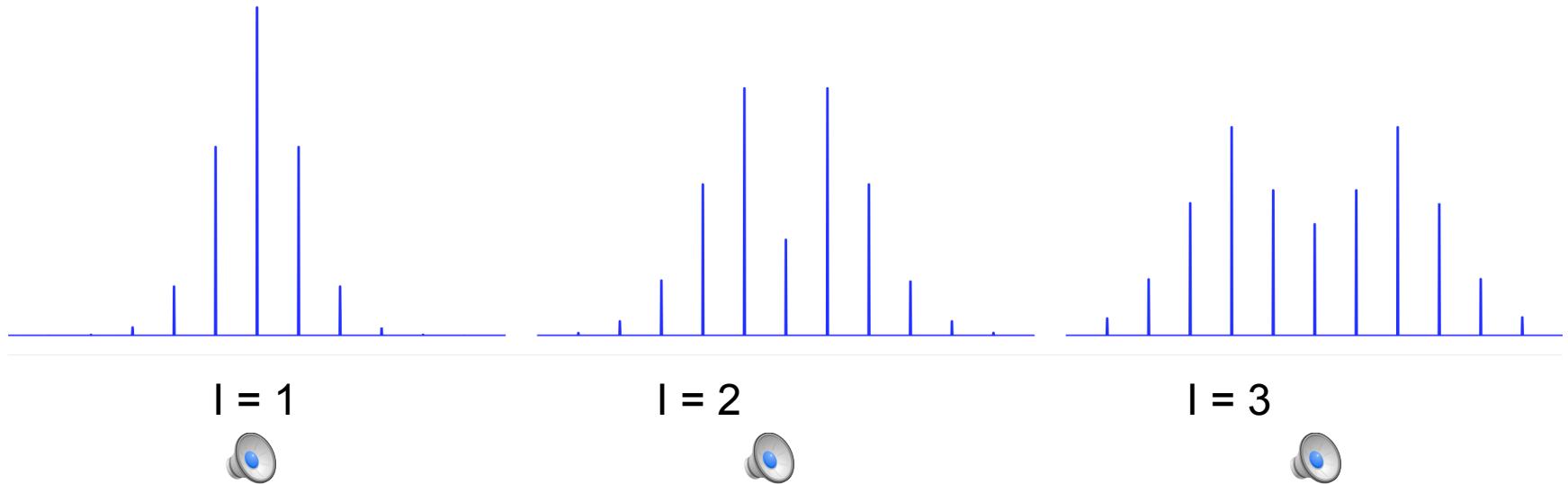
# Síntese

- Síntese por modulação
  - Síntese FM
    - Sistema não linear
    - Análise via Funções de Bessel do Primeiro Tipo
    - As frequências significativas acrescentadas ao sinal original são as que distam da frequência base múltiplos inteiros da frequência moduladora, sendo a quantidade de tais múltiplos igual a  $l+2$

# Síntese

- Síntese por modulação
  - Síntese FM

$$f_c = 600 \quad f_m = 100 \quad A = 1$$



# Síntese

- Síntese por modulação
  - Síntese FM



WC Olo Garb (<http://www.youtube.com/watch?v=R-hML-I5RV0>)

# Tendências e Aplicações

# Localização e Identificação

- Query by humming (QbH)
  - Fase 1: criação de um banco de dados com representações de alto nível de algumas canções
  - Fase 2: dada uma canção à identificar, processá-la de modo análogo aos exemplos do banco de dados
  - Devolver o exemplo do banco de dados cuja melodia mais se aproxima à encontrada na entrada

# Localização e Identificação

- Query by humming (QbH)
  - Representações melódicas

Notas

G A B D C C

Intervalos

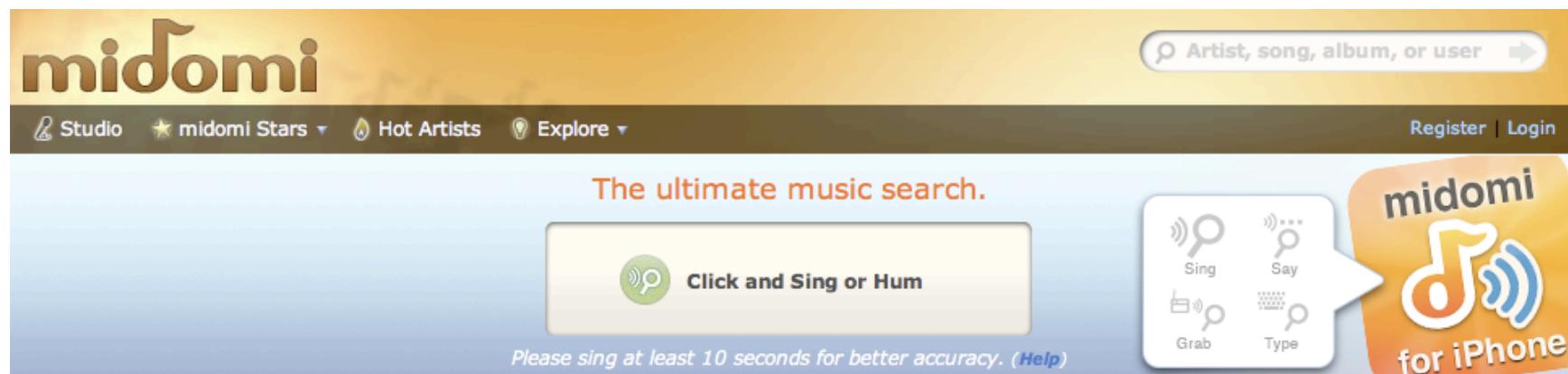
\* +2 +2 +3 -2 0

Direções

\* ↑ ↑ ↑ ↓ |

# Localização e Identificação

- Query by humming (QbH)
  - Aplicativos
    - Midomi ([www.midomi.com](http://www.midomi.com))



- Musipedia ([www.musipedia.org](http://www.musipedia.org))



# Localização e Identificação

- Monitoração de rádios
  - Técnica: “Impressão digital” do áudio
    - O arquivo é segmentado e um descritor é extraído de cada segmento
    - Um algoritmo de agrupamento é aplicado ao conjunto de descritores, sendo eleito um representante de cada grupo
    - O conjunto de representantes define a impressão digital
    - O áudio que se quer classificar passa pelo procedimento do primeiro item e os descritores obtidos são comparados às impressões digitais do banco de dados

# Recomendação por Similaridade

- iTunes Genius: recomenda canções da loja iTunes com base em uma música escolhida na biblioteca do usuário
- LastFM, Pandora: a lista de execução é composta de músicas semelhantes à que o usuário sugere

# Recomendação por Similaridade

- Abordagens para o problema de recomendação por similaridade
  - Baseada em conteúdo
    - Recomendação feita simplesmente a partir das características do produto
  - Colaborativa
    - Um produto é recomendado com base nas preferências de usuários com perfil semelhante ao do cliente
  - Métodos híbridos
    - Usam as duas abordagens, priorizando a filtragem colaborativa

# Novas Interfaces

- Princeton Laptop Orchestra (PLOrk)
  - <http://plork.cs.princeton.edu/>



# Novas Interfaces

- Princeton Laptop Orchestra (PLOrk)
  - Hardware
    - Laptop, interface firewire, amplificador, alto-falante hemisférico
  - Software
    - Max/MSP, SuperCollider, Chuck

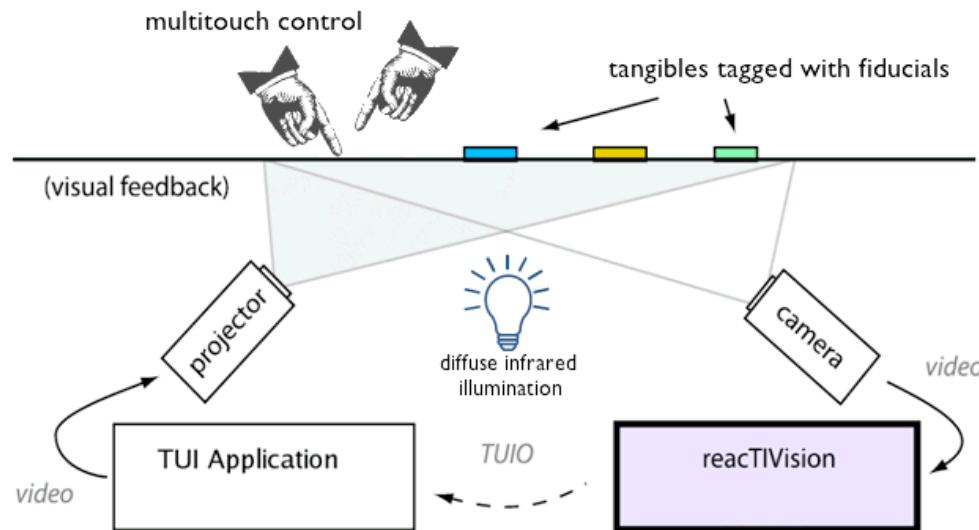
# Novas Interfaces

- Interfaces multi-touch e/ou tangíveis
  - Reactable (Universitat Pompeu Fabra)
    - <http://mtg.upf.es/reactable/>



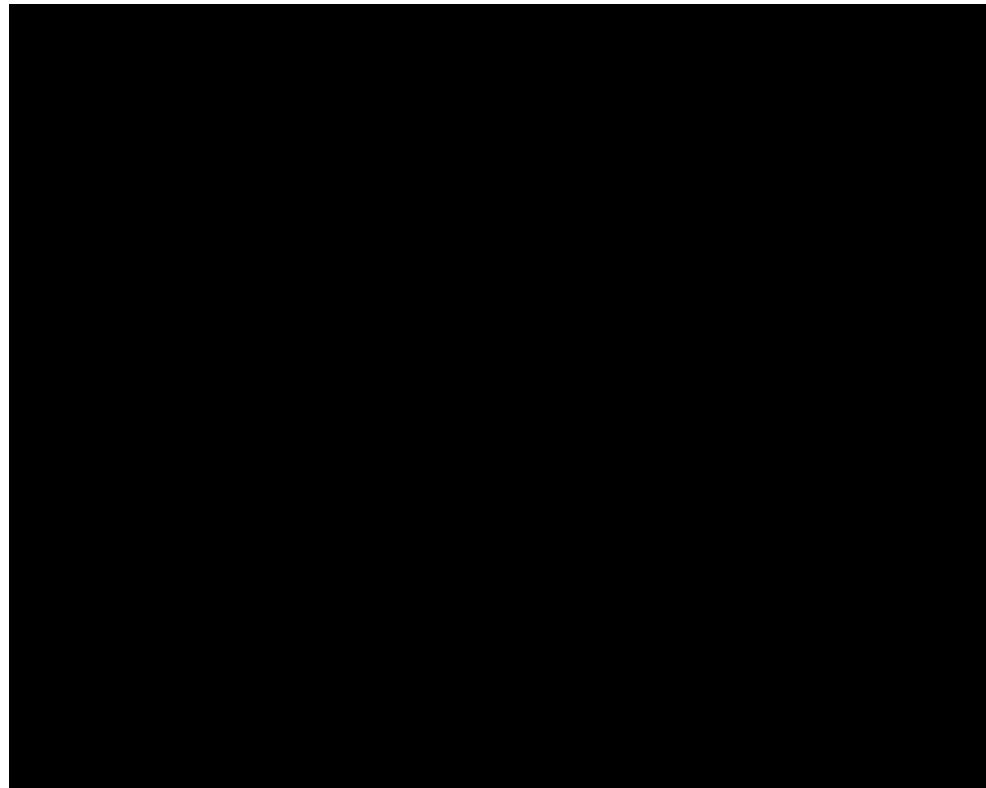
# Novas Interfaces

- Interfaces multi-touch e/ou tangíveis
  - Reactable (Universitat Pompeu Fabra)



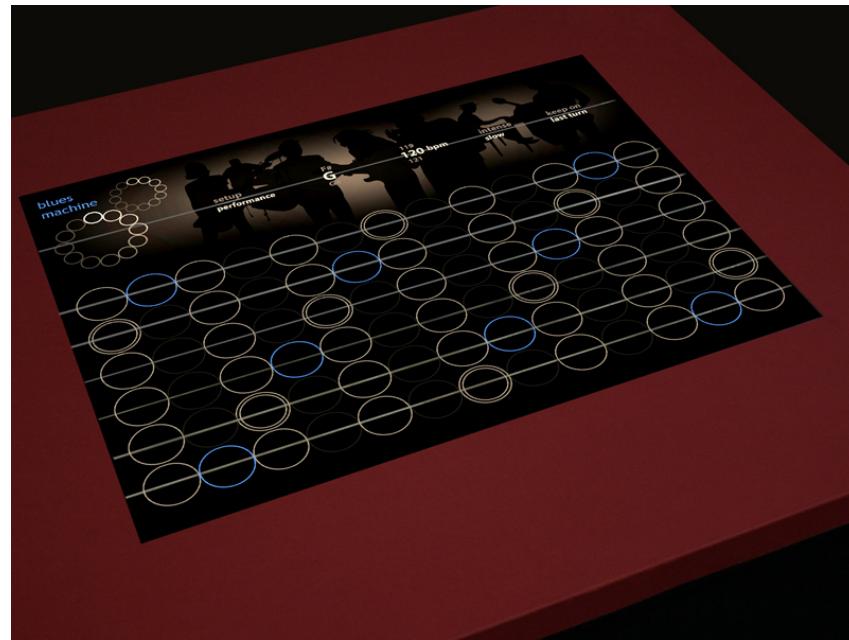
# Novas Interfaces

- Interfaces multi-touch e/ou tangíveis
  - Reactable (Universitat Pompeu Fabra)



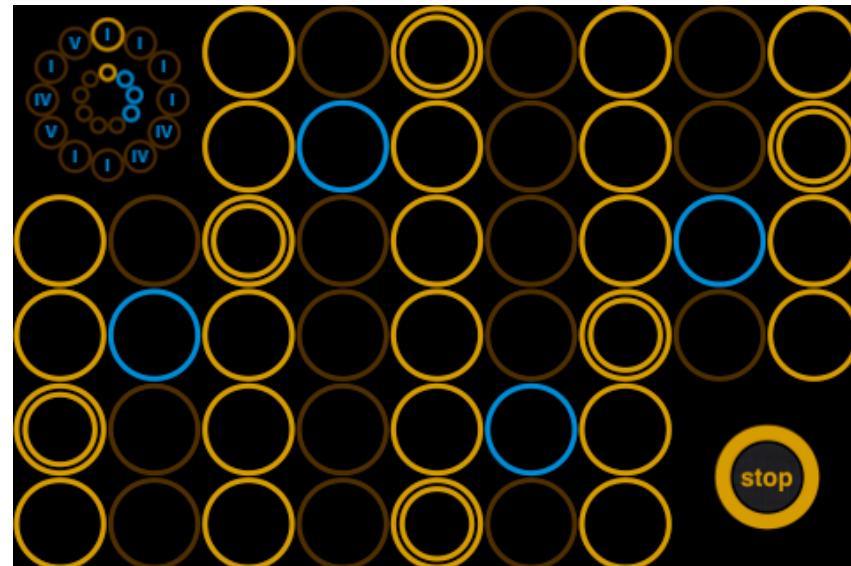
# Novas Interfaces

- Interfaces multi-touch e/ou tangíveis
  - BluesMachine (IMPA)
    - <http://www.visgraf.imp.br/bm/>



# Novas Interfaces

- Interfaces multi-touch e/ou tangíveis
  - BluesMachine (IMPA)
    - iPhone e iPod Touch (disponível na App Store)



# Referências Recomendadas

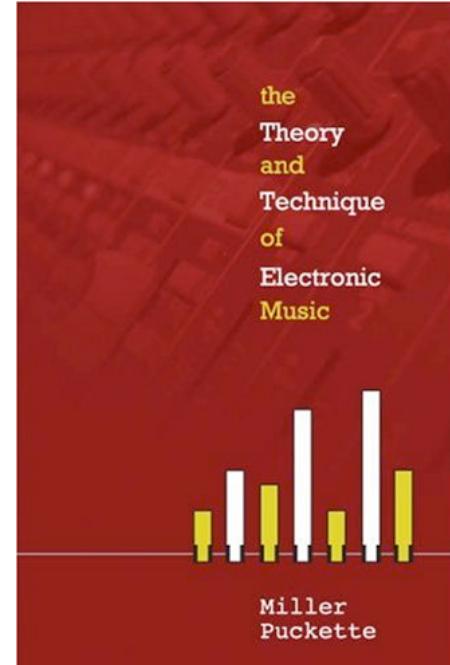
# Referências Recomendadas

- The Computer Music Tutorial (Curtis Roads)



# Referências Recomendadas

- The Theory and Technique of Electronic Music  
(Miller Puckette)
  - <http://crca.ucsd.edu/~msp/techniques/latest/book.pdf>
- PureData
  - <http://puredata.info/>



# Referências Recomendadas

- Music: A Mathematical Offering (David Benson)
  - <http://www.maths.abdn.ac.uk/~bensondj/html/music.pdf>

