

Warping - Deformação de sinais musicais em tempo e frequência

Gilmar Dias
grodias@ime.usp.br

Grupo de Computação Musical
Departamento de Ciência da Computação
Instituto de Matemática e Estatística
Universidade de São Paulo

11 de junho de 2012

Sumário

1 Introdução

Sumário

- 1 Introdução
- 2 Deformação de sinais musicais
 - Deformação em tempo
 - Deformação em frequência

Sumário

- 1 Introdução
- 2 Deformação de sinais musicais
 - Deformação em tempo
 - Deformação em frequência
- 3 Algoritmos de deformação
 - Deformação em frequência por FFT
 - Linhas de atraso dispersivas
 - Deformação baseada em Vocoder DFW

Sumário

- 1 Introdução
- 2 Deformação de sinais musicais
 - Deformação em tempo
 - Deformação em frequência
- 3 Algoritmos de deformação
 - Deformação em frequência por FFT
 - Linhas de atraso dispersivas
 - Deformação baseada em Vocoder DFW
- 4 Aplicação em música
 - Mudança de altura (pitch-shifting) de sons inarmônicos
 - Inarmonizador
 - Comb filtering + deformação e extração de sinas de excitação em sons inarmônicos
 - Vibrato, glissando, trill e flutterzunge
 - Morphing

Sumário

- 1 Introdução
- 2 Deformação de sinais musicais
 - Deformação em tempo
 - Deformação em frequência
- 3 Algoritmos de deformação
 - Deformação em frequência por FFT
 - Linhas de atraso dispersivas
 - Deformação baseada em Vocoder DFW
- 4 Aplicação em música
 - Mudança de altura (pitch-shifting) de sons inarmônicos
 - Inarmonizador
 - Comb filtering + deformação e extração de sinas de excitação em sons inarmônicos
 - Vibrato, glissando, trill e flutterzunge
 - Morphing
- 5 Conclusão

Introdução

Introdução

Warping

Introdução

Warping Deformação/empenamento de um sinal de áudio

Introdução

Warping Deformação/empenamento de um sinal de áudio

Aplicação

Introdução

Warping Deformação/empenamento de um sinal de áudio

Aplicação Pitch-shifting, inarmonizador, vibrato, morphing, etc.

Deformação de sinais musicais

- Deformação em tempo
- Deformação em frequência

Deformação em tempo

Deslocamento de amplitudes atingidas por um sinal para diferentes instantes de tempo

Deformação em tempo

Deslocamento de amplitudes atingidas por um sinal para diferentes instantes de tempo

Caracterizada por um mapa de deformação ($\theta(t)$)

Deformação em tempo

Deslocamento de amplitudes atingidas por um sinal para diferentes instantes de tempo

Caracterizada por um mapa de deformação ($\theta(t)$)

Em um sinal formado pela sobreposição de senóides, a deformação em tempo é equivalente a modular a fase de cada componente e somá-las

Deformação em tempo

Para um sinal periódico,

$$\theta_0(t) = \text{rem}(\theta(t), T)$$

Deformação em tempo

Para um sinal periódico,

$$\theta_0(t) = \text{rem}(\theta(t), T)$$

Generalizando, para sinais não periódicos,

$$s_{tw}(t) = s(\theta(t))$$

Deformação em tempo

Para um sinal periódico,

$$\theta_0(t) = \text{rem}(\theta(t), T)$$

Generalizando, para sinais não periódicos,

$$s_{tw}(t) = s(\theta(t))$$

Se o mapa é invertível,

$$s_{tw}(\theta^{-1}(t)) = s(t)$$

Deformação em tempo

Para um sinal periódico,

$$\theta_0(t) = \text{rem}(\theta(t), T)$$

Generalizando, para sinais não periódicos,

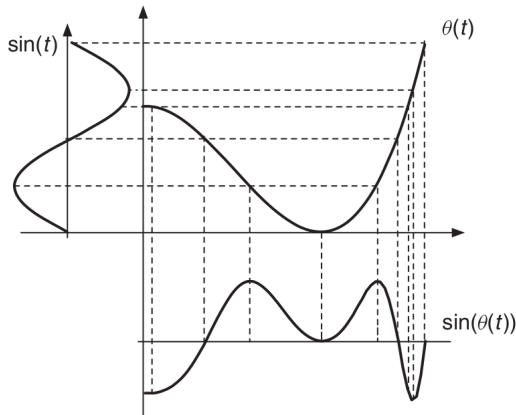
$$s_{tw}(t) = s(\theta(t))$$

Se o mapa é invertível,

$$s_{tw}(\theta^{-1}(t)) = s(t)$$

Um exemplo de aplicação musical é o *morphing* de um som em outro no domínio do tempo

Deformação em tempo



Deformação em tempo

Exemplo Pd

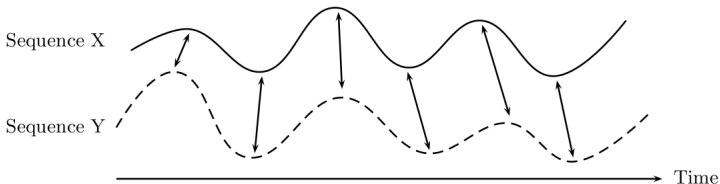
Deformação em tempo

Deformação em tempo dinâmica

- Cálculo de similaridade ("distância") entre dois sinais
- Sinais são deformados em tempo para serem alinhados (alinhamento ótimo)
- A grosso modo, é um reconhecimento de padrão
- Pode ser usado para reconhecimento de gestos
- Vídeo matching (similaridade de quadros)
- Pedal com reconhecimento de gestos para DJ's

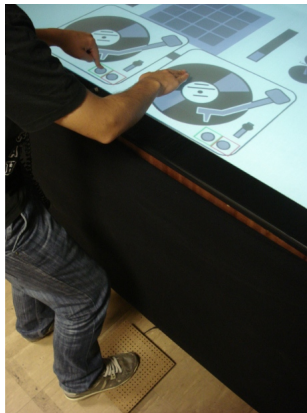
Deformação em tempo

Deformação em tempo dinâmica



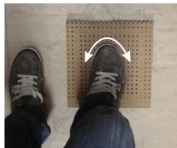
Deformação em tempo

Deformação em tempo dinâmica



Deformação em tempo

Deformação em tempo dinâmica



Deformação em frequência

Seja um sinal cuja DTFT é $S(\omega)$, formamos o sinal $s_{fw}(t)$ cuja DTFT é

Deformação em frequência

Seja um sinal cuja DTFT é $S(\omega)$, formamos o sinal $s_{fw}(t)$ cuja DTFT é

$$S_{fw}(\omega) = S(\theta(\omega))$$

Deformação em frequência

Seja um sinal cuja DTFT é $S(\omega)$, formamos o sinal $s_{fw}(t)$ cuja DTFT é

$$S_{fw}(\omega) = S(\theta(\omega))$$

Se o mapa é invertível,

$$S_{fw}(\theta^{-1}(\omega)) = S(\omega)$$

Deformação em frequência

Seja um sinal cuja DTFT é $S(\omega)$, formamos o sinal $s_{fw}(t)$ cuja DTFT é

$$S_{fw}(\omega) = S(\theta(\omega))$$

Se o mapa é invertível,

$$S_{fw}(\theta^{-1}(\omega)) = S(\omega)$$

Para um sinal real, o mapa deve ter paridade ímpar, i.e.,

$$\theta(-\omega) = -\theta(\omega)$$

Deformação em frequência

Considere um sinal periódico $s(t)$ no qual o espectro tem picos em múltiplos inteiros da frequência fundamental ω_0 . O espectro do sinal deformado terá picos nas frequências

Deformação em frequência

Considere um sinal periódico $s(t)$ no qual o espectro tem picos em múltiplos inteiros da frequência fundamental ω_0 . O espectro do sinal deformado terá picos nas frequências

$$\hat{\omega}_k = \theta^{-1}(k\omega_0)$$

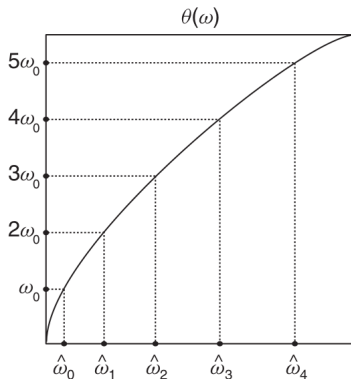
Deformação em frequência

Considere um sinal periódico $s(t)$ no qual o espectro tem picos em múltiplos inteiros da frequência fundamental ω_0 . O espectro do sinal deformado terá picos nas frequências

$$\hat{\omega}_k = \theta^{-1}(k\omega_0)$$

Se o mapa não cresce monotonicamente, o efeito obtido é similar ao obtido com um *phase vocoder*, mas com uma resolução é muito maior.

Deformação em frequência



Preservação de energia e deformação de frequência unitária

- Dilatação/contração de porções do espectro



Amplificação/atenuação de bandas

- Solução: equalizar

Preservação de energia e deformação de frequência unitária

A energia em uma banda arbitrária $[\omega_0, \omega_1]$ é

$$E_{[\omega_0, \omega_1]} = \frac{1}{2\pi} \int_{\omega_0}^{\omega_1} |S(\omega)|^2 d\omega$$

Preservação de energia e deformação de frequência unitária

A energia em uma banda arbitrária $[\omega_0, \omega_1]$ é

$$E_{[\omega_0, \omega_1]} = \frac{1}{2\pi} \int_{\omega_0}^{\omega_1} |S(\omega)|^2 d\omega$$

Substituindo ω por $\theta(\Omega)$ na integral, temos

$$E_{[\omega_0, \omega_1]} = \frac{1}{2\pi} \int_{\Omega_0=\theta^{-1}(\omega_0)}^{\Omega_1=\theta^{-1}(\omega_1)} |S(\theta(\Omega))|^2 \frac{d\theta}{d\Omega} d\Omega = \frac{1}{2\pi} \int_{\Omega_0}^{\Omega_1} |\tilde{S}_{fw}(\Omega)|^2 d\Omega$$

onde,

$$\tilde{S}_{fw}(\omega) = \sqrt{\frac{d\theta}{d\omega}} S(\theta(\omega))$$

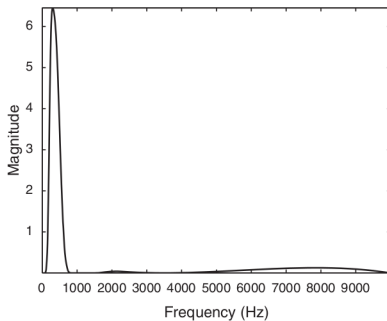
é a DTFT do sinal deformado escalado.

Preservação de energia e deformação de frequência unitária

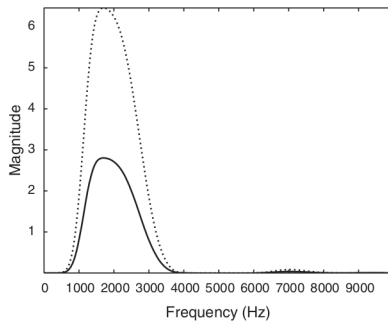
Essa equação estabelece que a energia em qualquer banda $[\omega_0, \omega_1]$ do sinal original é igual à energia do sinal deformado na banda deformada $[\theta^{-1}(\omega_0), \theta^{-1}(\omega_1)]$.

Preservação de energia e deformação de frequência unitária

Essa equação estabelece que a energia em qualquer banda $[\omega_0, \omega_1]$ do sinal original é igual à energia do sinal deformado na banda deformada $[\theta^{-1}(\omega_0), \theta^{-1}(\omega_1)]$.



(a)



(b)

Algoritmos de deformação

- Deformação em frequência por FFT
- Linhas de atraso dispersivas
- Deformação em tempo curto e implementação em tempo real
- Aproximação de deformação em frequência baseada em Vocoder

Deformação em frequência por FFT

Seja a DFT

$$S\left(\frac{2\pi m}{N}\right) = \sum_{n=0}^{N-1} s(n)e^{-j\frac{2\pi nm}{N}}$$

Deformação em frequência por FFT

Seja a DFT

$$S\left(\frac{2\pi m}{N}\right) = \sum_{n=0}^{N-1} s(n)e^{-j\frac{2\pi nm}{N}}$$

Considere o mapa

$$\theta(\omega + 2k\pi) = \theta(\omega) + 2k\pi, k \in \mathbb{Z}$$

Deformação em frequência por FFT

Seja a DFT

$$S\left(\frac{2\pi m}{N}\right) = \sum_{n=0}^{N-1} s(n)e^{-j\frac{2\pi nm}{N}}$$

Considere o mapa

$$\theta(\omega + 2k\pi) = \theta(\omega) + 2k\pi, k \in \mathbb{Z}$$

De modo que a DFT de um sinal de período 2π seja

Deformação em frequência por FFT

Seja a DFT

$$S\left(\frac{2\pi m}{N}\right) = \sum_{n=0}^{N-1} s(n)e^{-j\frac{2\pi nm}{N}}$$

Considere o mapa

$$\theta(\omega + 2k\pi) = \theta(\omega) + 2k\pi, k \in \mathbb{Z}$$

De modo que a DFT de um sinal de período 2π seja

$$S_{fw}(\omega + 2k\pi) = S(\theta(\omega + 2k\pi)) = S(\theta(\omega) + 2k\pi) = S(\theta(\omega)) = S_{fw}(\omega)$$

Deformação em frequência por FFT

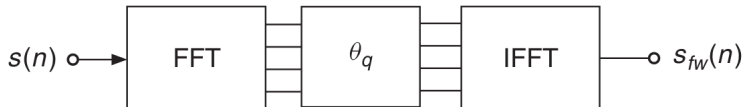
Utilizamos o mapa quantizado

$$\theta_q\left(\frac{2\pi m}{N}\right) = \frac{2\pi}{N} \text{round}\left[\theta\left(\frac{2\pi m}{N}\right)\frac{N}{2\pi}\right]$$

E calculamos a DFT aproximada do sinal deformado em frequência,

$$S_{fw}(n) \approx \frac{1}{N} \sum_{m=0}^{N-1} S\left(\theta_q\left(\frac{2\pi m}{N}\right)\right) e^{j\frac{2\pi nm}{N}}$$

Deformação em frequência por FFT



Se o mapa é uma função crescente pode-se introduzir o fator de equalização antes da IFFT.

Linhas de atraso dispersivas

Considere a DTFT de um sinal qualquer deformado em frequência e escalado,

$$\tilde{S}_{fw}(\omega) = \sqrt{\frac{d\theta}{d\omega}} S(\theta(\omega)) = \sqrt{\frac{d\theta}{d\omega}} \sum_{n=0}^{\infty} s(n) e^{-jn\theta(\omega)}$$

Linhas de atraso dispersivas

Considere a DTFT de um sinal qualquer deformado em frequência e escalado,

$$\tilde{S}_{fw}(\omega) = \sqrt{\frac{d\theta}{d\omega}} S(\theta(\omega)) = \sqrt{\frac{d\theta}{d\omega}} \sum_{n=0}^{\infty} s(n) e^{-jn\theta(\omega)}$$

Obtemos o sinal com a DTFT inversa

$$\tilde{s}_{fw}(k) = IDTFT[\tilde{S}_{fw}(\omega)](k) = \sum_{n=0}^{\infty} s(n) IDTFT\left[\sqrt{\frac{d\theta}{d\omega}} e^{-jn\theta(\omega)}\right](k)$$

Linhas de atraso dispersivas

Considere a DTFT de um sinal qualquer deformado em frequência e escalado,

$$\tilde{S}_{fw}(\omega) = \sqrt{\frac{d\theta}{d\omega}} S(\theta(\omega)) = \sqrt{\frac{d\theta}{d\omega}} \sum_{n=0}^{\infty} s(n) e^{-jn\theta(\omega)}$$

Obtemos o sinal com a DTFT inversa

$$\tilde{s}_{fw}(k) = IDTFT[\tilde{S}_{fw}(\omega)](k) = \sum_{n=0}^{\infty} s(n) IDTFT\left[\sqrt{\frac{d\theta}{d\omega}} e^{-jn\theta(\omega)}\right](k)$$

Define-se as sequências $\lambda_n(k)$,

$$\lambda_n(k) = IDTFT\left[\sqrt{\frac{d\theta}{d\omega}} e^{-jn\theta(\omega)}\right](k) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \sqrt{\frac{d\theta}{d\omega}} e^{j|k\omega - n\theta(\omega)|} d\omega$$

Linhas de atraso dispersivas

Simplificando a equação anterior, temos,

$$\tilde{s}_{fw}(k) = \sum_{n=0}^{\infty} s(n)\lambda_n(k)$$

Linhas de atraso dispersivas

Simplificando a equação anterior, temos,

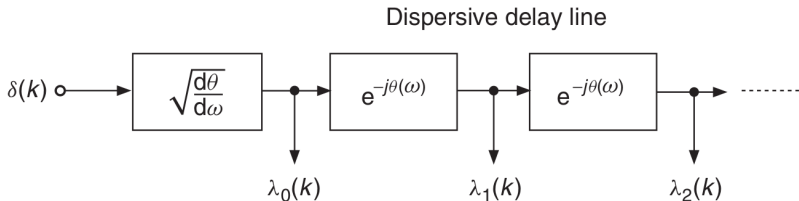
$$\tilde{s}_{fw}(k) = \sum_{n=0}^{\infty} s(n) \lambda_n(k)$$

Sabemos que

$$\Lambda_n(\omega) = DTFT[\lambda_n](\omega) = \Lambda_{n-1}(\omega) e^{-j\theta(\omega)}$$

$$\text{com } \Lambda_0(\omega) = \sqrt{\frac{d\theta}{d\omega}}$$

Linhas de atraso dispersivas



Linha de atraso em que os atrasos elementares são substituídos por filtros passa-tudo.

Linhas de atraso dispersivas

Cada filtro introduz um delay de grupo

$$\tau_G(\omega) = \frac{d\theta}{d\omega}$$

Linhas de atraso dispersivas

Cada filtro introduz um delay de grupo

$$\tau_G(\omega) = \frac{d\theta}{d\omega}$$

Inserindo um sinal genérico nessa linha, temos,

$$y(k) * \lambda_n(k) = \sum_r y(k-r)\lambda_n(r)$$

Linhas de atraso dispersivas

Cada filtro introduz um delay de grupo

$$\tau_G(\omega) = \frac{d\theta}{d\omega}$$

Inserindo um sinal genérico nessa linha, temos,

$$y(k) * \lambda_n(k) = \sum_r y(k-r)\lambda_n(r)$$

Em especial, se $k = 0$ e escolhendo um sinal de entrada $s(k) = y(-k)$,

$$\hat{y}_n(0) = \sum_r s(r)\lambda_n(r)$$

Linhas de atraso dispersivas

Podemos definir sequências transpostas

$$\lambda_r^T(n) \equiv \lambda_n(r)$$

Linhas de atraso dispersivas

Podemos definir sequências transpostas

$$\lambda_r^T(n) \equiv \lambda_n(r)$$

$$\hat{y}_n(0) = \sum_r s(r) \lambda_r^T(n)$$

Linhas de atraso dispersivas

Podemos definir sequências transpostas

$$\lambda_r^T(n) \equiv \lambda_n(r)$$

$$\hat{y}_n(0) = \sum_r s(r) \lambda_r^T(n)$$

Assim,

$$\lambda_r^T(n) = IDTFT\left[\sqrt{\frac{d\theta}{d\omega}} e^{-jn\theta(\omega)}\right](r) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \sqrt{\frac{d\theta}{d\omega}} e^{j|r\omega - n\theta(\omega)|} d\omega$$

Se o mapa tem paridade ímpar,

$$\lambda_r^T(n) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \sqrt{\frac{d\theta^{-1}}{d\omega}} e^{j|n\omega - r\theta^{-1}(\omega)|} d\omega$$

Linhas de atraso dispersivas

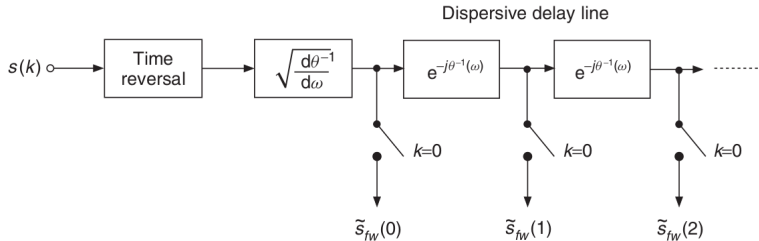
O que resulta em

$$\Lambda_r^T(\omega) = \sqrt{\frac{d\theta^{-1}}{d\omega}} e^{jr\theta^{-1}(\omega)}$$

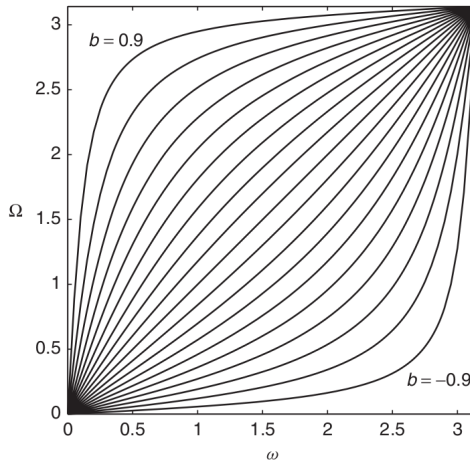
Linhas de atraso dispersivas

O que resulta em

$$\Lambda_r^T(\omega) = \sqrt{\frac{d\theta^{-1}}{d\omega}} e^{jr\theta^{-1}(\omega)}$$



Linhas de atraso dispersivas



Deformação baseada em Vocoder DFW

- Aproximação de deformação em frequência baseada em Vocoder
- Deformação em frequência com tempo variável

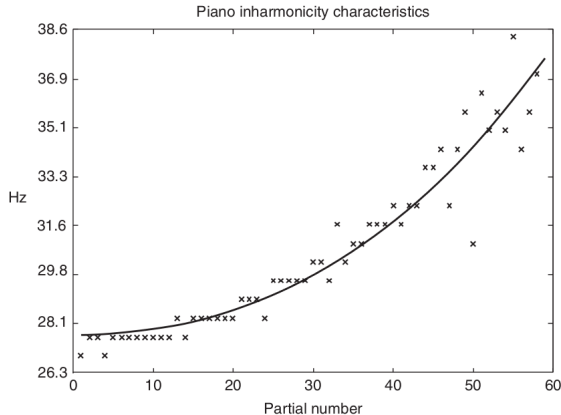
Deformação baseada em Vocoder DFW

Exemplos

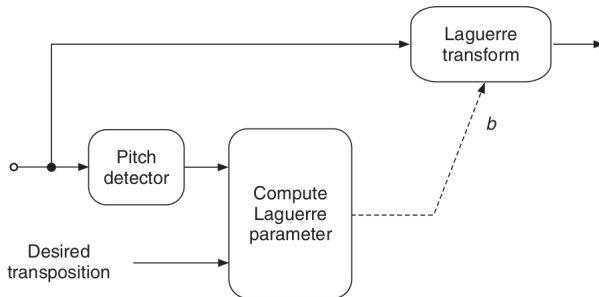
Aplicação em música

- Mudança de altura (pitch-shifting) de sons inarmônicos
- Inarmonizador
- Comb filtering + deformação e extração de sinas de excitação em sons inarmônicos
- Vibrato, glissando, trill e flatterzunge
- Morphing

Mudança de altura (pitch-shifting) de sons inarmônicos



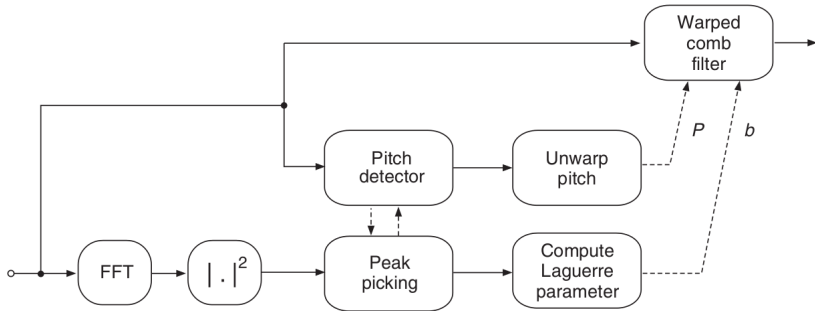
Mudança de altura (pitch-shifting) de sons inarmônicos



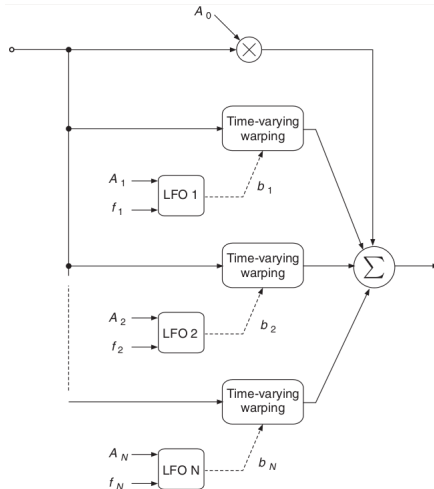
Inarmonizador

- Está entre os novos efeitos introduzidos pela deformação de frequência
- Deformação de frequência de um som harmônico (parâmetro de $\simeq 0.5$)
- O som é enriquecido por parciais inarmônicos
- A altura e a duração são alteradas, técnicas de reamostragem podem ser usadas para restaurá-las
- Envelopes de amplitudes dos parciais são alterados
- Em muitos casos, esse efeito introduz interessantes dissonâncias nos parciais
- Trompete/Sino, guitarra/piano
- Pode ser usado em síntese por modelagem física

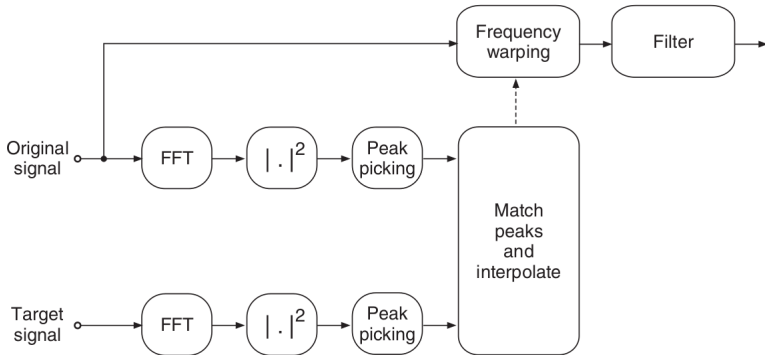
Comb filtering + deformação e extração de sinas de excitação em sons inarmônicos



Vibrato, glissando, trill e flutterzunge



Morphing



Conclusão

- Recente interesse por deformação em frequência em aplicações musicais
- Novas ferramentas para processamento de som
- Adição de expressão controlada a sons estáticos
- Mapeamento de sons harmônicos para inarmônicos
- Algoritmo eficiente por transformada de Laguerre em tempo curto torna a deformação de frequência computável em tempo real
- Encorajamento do uso de warpings por músicos e engenheiros de som

Obrigado!