

Mixagem Automática

Antonio José Homsí Goulart

Instituto de Matemática e Estatística
Universidade de São Paulo

Grupo de Pesquisa em Computação Musical - IME/USP
<http://compmus.ime.usp.br>

Ciclo de seminários DAFX
Este seminário é baseado no capítulo 13 da segunda edição do livro DAFX

25/06/2012

A mixagem

- Várias fontes
- Processa
- Combina
- Mistura sonora

AM-DAFX - Ontem e hoje

- 1975, Princípio - Ajuste automático de ganho
- Sistema para manter ganho constante independente do número de microfones. Baseado em adaptive gate.
- Hoje, equalização.

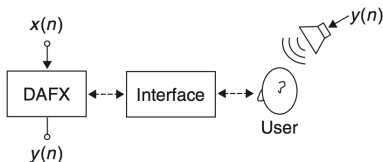
AM-DAFX - Possibilidades

- Realce de canais
- Paneamento
- Compensação de ganho / Faders
- Correção de time-offset
- Equalização, simples.

AM-DAFX - Aplicações

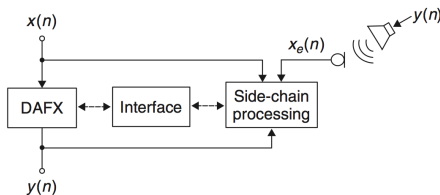
- Live Music
- Produção musical
 - ▶ Mixagem remota
 - ▶ Pré produção
 - ▶ Pós produção
- Produção audiovisual
 - ▶ Cinema, ...
 - ▶ Jogos

DAFX



- Processamento de sinal
- Usuário controla processo por GUI
- Manipulação dos parâmetros do processo para transformação desejada
- Monitoramento da saída

AM-DAFX



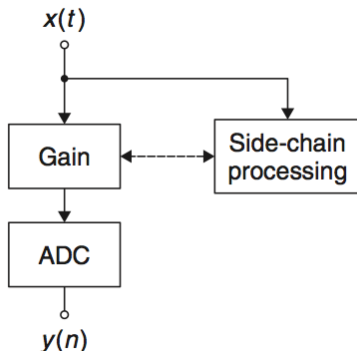
- Emular parâmetros de controle do usuário
 - Seção de processamento
 - Seção de análise (side-chain)
-
- (1) Obedecer questões técnicas de mixagem, como evitar distorção e manter faixa dinâmica adequada - Loudness War
 - (2) Simplificar as tarefas do processo mantendo um padrão similar ao de um profissional
 - (3) Permanecer livre de artefatos sonoros ($\beta \leq 1$)

Filosofia

- AM-DAFX visa tomar decisões técnicas, objetivas
- Otimização do fluxo de trabalho do engenheiro de mixagem (fica mais rápido e sem a parte tediante)
- Não realiza práticas incomuns nem decisões subjetivas
- Mas... não é esteticamente transparente

Princípio de funcionamento

Exemplo: controle automático de ganho.



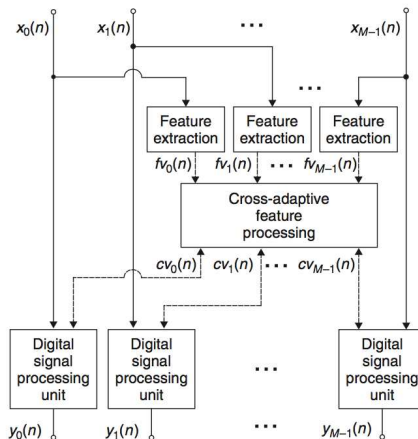
- Extração de parâmetro do sinal de entrada
- Tomada de decisão
- Libera sinal de controle para processamento

Classificação dos efeitos

- Controle direto
- Auto-adaptativo
- Adaptação externa
- Adaptação cruzada

No contexto da mixagem, as mudanças em cada faixa não são apenas devido aos seus conteúdos, mas também pela necessidade de misturá-los para obter um mix balanceado.

Arquitetura MIMO



$fv_m(n)$ Feature vector para cada faixa m

$cv_m(n)$ Control vector para cada faixa m

Extração de parâmetros

Possibilidade de extração rápida e precisa → Real-time

Quanto melhor o modelo → Melhor o resultado

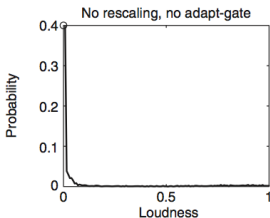
Dois tipos:

- (1) Acumulativa
 - ▶ Aumenta precisão com tempo e dados
 - ▶ Obtenção da distribuição e deriva modo
- (2) Dinâmica
 - ▶ Parâmetros instantâneos e controle instantâneo
 - ▶ Exemplo: RMS para manter vox mais a frente

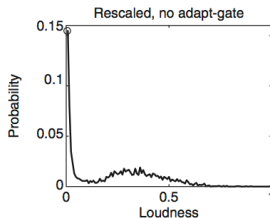
O querido Ruído

Defesa: média, validação, gating (AG adapta de acordo com ruído ambiente).

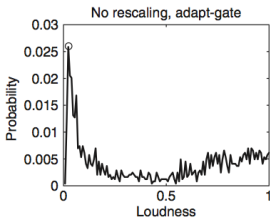
Exemplo: loudness, extração acumulativa.



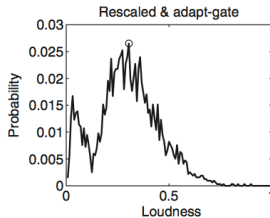
(a)



(b)



(c)



(d)

Seção de adaptação cruzada

- Tendo derivado os parâmetros dos sinais em questão, esta seção analisa a inter-dependência entre os mesmos para liberar dados de controle apropriados
- Pode-se usar regras de restrição para estreitar a inter-dependência entre os canais
- Para manter sistema estável restringe-se a contribuição geral dos ganhos a um máximo unitário

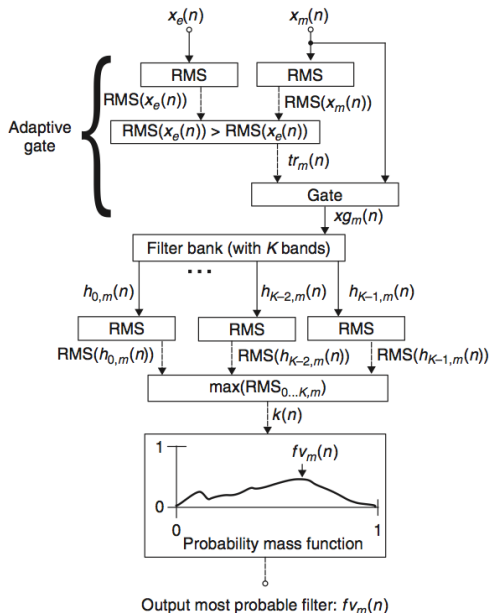
Realce de canal - Abordagem

- Aumentar ganho do canal μ a ser destacado
- Feedback e/ou distorção
- Diminui então o ganho dos outros canais
- ou melhor ainda...
- Diminui ganho dos canais espectralmente relacionados com o canal μ

Teremos um atenuador para cada canal de entrada

$$y_m(n) = cv_m(n).x_m(n)$$

Realce de canal - Extração de parâmetros



Realce de canal - Adaptação cruzada

Tendo em mão o conteúdo espectral de $x_m(n)$, precisamos mapear a relação entre o conteúdo espectral de cada canal. Para realçar o canal $x_\mu(n)$ queremos uma função com mínimo localizado em $fv_\mu(n) = k$, e que aumenta conforme distancia-se deste valor.

Parâmetros do usuário:

- μ : canal a ser realçado
- G: atenuação $\forall m | fv_m(n) = fv_\mu(n) \rightarrow$ atenuação máxima
- Q: seletividade

Parâmetros do sistema:

- M: número de fontes (canais)
- m: canal sendo processado

Realce de canal - Implementação

$$fg_m(n) = \frac{1}{Q\sqrt{2\pi}} e^{\frac{-(fr_m(n) - \mu(n))^2}{2Q^2}}$$

$$\text{freq bin} \rightarrow fr_m(n) = \left(\frac{2}{M-1}(fv_m(n) - 1) \right) - 1$$

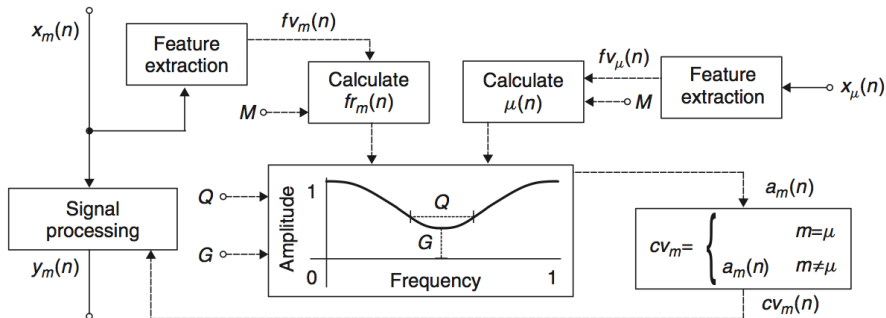
$$\text{máximo} \rightarrow \mu(n) = \left(\frac{2}{M-1}(fv_\mu(n) - 1) \right) - 1$$

$$a_m(n) = \left| \left(G \left(\frac{fg_m(n)}{\frac{1}{Q\sqrt{2\pi}}} \right) \right) - 1 \right|$$

$$cv_m(n) = \begin{cases} 1 & \text{se } m = \mu \\ a_m(n) & \text{se } m \neq \mu \end{cases}$$

$$mix(n) = \sum_{m=0}^{M-1} cv_m(n) \cdot x_m(n)$$

Realce de canal - Diagrama de blocos



Paneamento - Abordagem

- Separação de canais no campo estéreo
- Posicionamento pela divisão da energia do sinal entre os canais L e R
- Separar canais com espectro relacionado para posições diferentes no campo
- Não panear graves (pouco perceptível e desbalanço)
- Manter mix balanceada (separar canais correlatos simetricamente)
- Canal prioritário (importância subjetiva)
- Não exagerar no pan (prática difundida)

Extração de parâmetros similar à do realce de canal.

Paneamento - Adaptação cruzada

Tendo em mão o conteúdo espectral de $x_m(n)$, precisamos mapear a relação entre o conteúdo espectral de cada canal e determinar suas posições no campo. O objetivo é aumentar a inteligibilidade separando canais com mesmo conteúdo espectral.

Parâmetros do usuário:

- U_m : ordenação da prioridade (0: alta ... M-1: baixa)
- W: extensão do paneamento (0: pan total / 0.5: mono)

Parâmetros do sistema:

- k: banda de frequência do canal (para saber se será paneado)
- R_m : número de canais na mesma banda que x_m
- P_m : prioridade do canal (quanto menor, menos pan)

$cv_m(n)$ vai de 0(wide L) a 1(wide R), e centro em 0.5.

Paneamento - Implementação

$$R_m = \sum_{j=1}^M R_{j-1} + \begin{cases} 1 & \text{se } fv_m(n) = fv_{j-1}(n) \\ 0 & \text{se } fv_m(n) \neq fv_{j-1}(n) \end{cases}$$

$$P_m = |\{U_i : fv_i(n) = fv_m(n)\} \cap \{U_i : U_i \leq U_m\}|, \quad i = \{0 \dots M-1\}$$

$$cv_m(n) = \begin{cases} \frac{1}{2} & \text{se } R_m = 1 \\ W + [(1 - 2W) \frac{R_m - P_m - 1}{2(R_m - 1)}] & \text{se } P_m + R_m \text{ impar} \\ W + [(1 - 2W) \frac{R_m + P_m - 2}{2(R_m - 1)}] & \text{se } P_m + R_m \text{ par, } R_m \neq 1 \end{cases}$$

$$cv_m(n) = \begin{cases} \frac{1}{2} & \text{se } fv_m(n) \leq tr_{ps} \\ cv_m(n) & \text{se } fv_m(n) > tr_{ps} \end{cases}$$

$$y_L(n) = \sum_{m=0}^{M-1} \sin(cv_m(n) \cdot \frac{\pi}{2}) \cdot x_m(n) \quad y_R(n) = \sum_{m=0}^{M-1} \cos(cv_m(n) \cdot \frac{\pi}{2}) \cdot x_m(n)$$

Faders - Abordagem

- Queremos obter uma mix balanceada
- Podemos pensar em nivelar o RMS de todos os canais
- Mas não levaria em conta questões perceptuais
- RMS balanceado não remete a percepção balanceada

Solução: loudness por canal tender ao loudness médio. Assim:

- Mistura balanceada com inteligibilidade ótima entre canais
- Damos a cada canal a mesma probabilidade de ser ouvido, ou ainda...
- Damos a cada canal a mesma chance de mascarar os outros canais

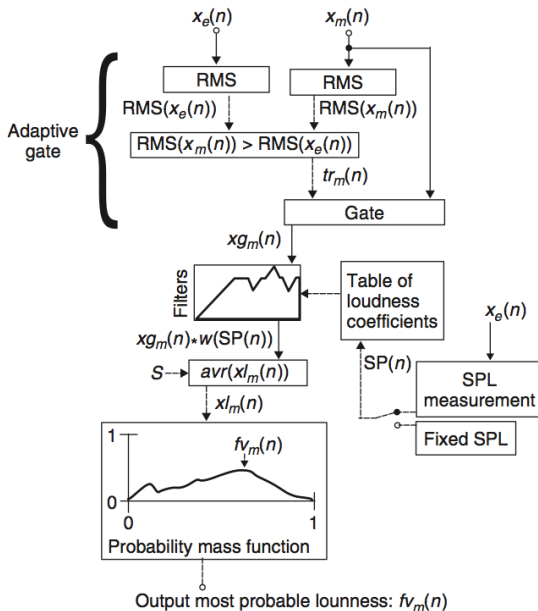
Faders - Extração de parâmetros

- Uso de adaptive gate $\rightarrow fv_m(n) = f(x_m(n), x_e(n))$
- Loudness é um atributo perceptual, precisamos de um modelo psicoacústico
- Modelos representados por w , e variam de acordo com a SPL
- Mesmo microfone que capta x_e para o gating é aproveitado aqui
- Busca-se em uma tabela os coeficientes de filtros que aproximam curvas de loudness referentes aos SPL correspondentes

$$xl_m(n) = \frac{1}{S} \sum_{i=1}^S (xg_m(n) \star w(SP(n)))_i$$

Loudness de cada canal m . Acumula-se para estimar função. Como não se sabe a faixa dinâmica do sinal, normaliza sempre que aparecer um novo valor maior que os anteriores. Acha-se então o pico da função, que é o valor de loudness mais provável da faixa m , denotado por $fv_m(n)$.

Faders - Diagrama de blocos



Faders - Adaptação cruzada

- Calcular loudness médio $I(n)$ (média aritmética dos $fv_m(n)$)
- Ganho nos canais com loudness abaixo dessa média
- Atenuação nos canais com loudness acima dessa média

Faders - Implementação

$$\frac{l(n)}{fv_m(n)} = cva_m(n)$$

$cva_m(n)$ é a variável para controle de ganho dos canais e representa na maioria dos casos um fader físico com ganho máximo limitado ($\beta \leq 1$), ou seja, precisamos normalizar $cva_m(n)$. Fazemos então com que a contribuição de ganho em todos os canais somadas tenha valor unitário.

$$cv_m(n) = \frac{cva_m(n)}{\sum_{m=0}^{M-1} cva_m(n)}$$

$$mix(n) = \sum_{m=0}^{M-1} cv_m(n) \cdot x_m(n)$$

Equalização - Abordagem

- Para uma mix com um bom balanço espectral, queremos uma mix com um bom balanço perceptual
- Equalização por canal não é feita apenas devido às propriedades individuais de cada canal, mas também porque ele será misturado com outros na mixagem
- Mesmo com o loudness balanceado algumas faixas de frequência ainda apresentam mascaramento
- Então, a idéia para a equalização é fazer uma aplicação multi-bandas por canal similar ao sistema apresentado para os faders. Assim, além de garantir loudness balanceado por canal, garante-se também loudness balanceado por canal para todas as bandas de equalização
- Em cada canal aplicaremos um equalizador gráfico (Q e f_c cte) de K bandas, denotado por $hq_{k,m}$

Equalização - Implementação

Vetor de parâmetros descrevendo cada banda k de cada canal m

$$xg_{k,m}(n) = f(x_m(n), x_e(n))$$

Modelo para obter loudness por banda - Acumular e encontrar pico

$$xl_{k,m}(n) = \frac{1}{S} \sum_{i=1}^S (xg_{k,m}(n) \star w(SP(n)))_i$$

Loudness média

$$l(n) = \sum_{m=0}^{M-1} \left(\sum_{k=0}^{K-1} fv_{k,m}(n) / K \right) / M$$

Variável de controle de ganho por banda por canal

$$cv_{k,m}(n) = \frac{l(n)}{fv_{k,m}(n)}$$

$$mix(n) = \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} cv_{k,m}(n) [hq_{k,m}(n) \star x_m(n)]$$

Conclusão

- Intenção de AM-DAFX é ajudar ou substituir algumas tarefas normalmente realizadas pelo usuário
- Não se sabe o quão autônomos os AM-DAFX se tornarão nem o quanto serão aceitos
- Aqui foram propostos sistemas que lidam apenas com questões técnicas, permitindo que o engenheiro de mixagem se concentre mais nos aspectos criativos da mix
- No futuro, modelos perceptuais mais vastos podem não só melhorar o desempenho do sistema, mas também possibilitar a tomada de decisões mais subjetivas referentes a mixagem

FIM

Brincadeira (verdadeira) entre os engenheiros de mixagem:
“Acabamos nunca terminando uma mix, mas desistindo dela.”

Dúvidas?

Grupo de pesquisa em Computação Musical - IME/USP
<http://compmus.ime.usp.br>

ag@ime.usp.br

Obrigado!