

Introdução à Escuta Binaural

Thilo Koch
Grupo Computação Musical - IME - USP

18 de agosto de 2014

Introdução

Desempenho da escuta binaural

Áreas de aplicação

Modelagem binaural

Fim

O ouvir é um sentido muito importante para nós:

- na interação social,
- nos informa sobre certos aspectos do nosso ambiente até além do que nós podemos ter em vista
- é capaz de nos alertar sobre perigos (não dorme e as orelhas não fecham).

O ouvir fornece **orientação** no mundo e **consciência** situacional.

A visão técnico computacional:

- *sistema auditivo* → *computador multiuso com duas portas de entrada*
- *portas de entrada* → *duas orelhas*
 - *mesma altura em ambos os lados de um elipsoide sólido*
- *cabeça* → *elipsoide sólido como suporte de antena*
 - *6 graus de liberdade de movimento no espaço (relativas ao corpo)*
- *corpo* → *suporte*
 - *6 graus de liberdade de movimento no espaço (relativas a um ponto de referência no ambiente)*

Entrada para o sistema auditivo

Ambiente	Entrada do sistema auditivo
vibrações elásticas	→ canal auditivo (→ tímpano)
num meio de transporte	→ crânio (em ar -60dB)
(ar, liquido, sólido)	

Questões

- De que o software do computador é capaz?
- Como funciona o software e como é estruturado? (*source code*)

Monoaural → Binaural

Ouvido binaural fornece informações adicionais; codificadas nas diferenças entre os dois sinais de entrada:

ITD diferença de tempo de chegada de sinais entre as orelhas
(*interaural arrival-time differences*)

ILD diferença de amplitude de sinais entre as orelhas
(*interaural level differences*)

Usando esses dados o sistema auditivo é capaz de estimar

- localização de eventos auditivos
- mapeamento de eventos auditivos para fontes sonoras

Localização

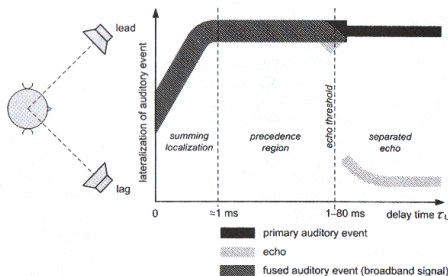
Localizar eventos auditivos no espaço:

- resolução para eventos na nossa frente com 1°
- caso monoaural: 10x pior
- melhor distinção também para elevações e distâncias
- melhor percepção da expansão espacial de um evento sonoro: evento **compacto** ou **difuso**

Reflexões

Como o sistema localiza eventos auditivos na presença de reflexões:

- **localização por somação:** ondas diretas e das reflexões determinam direção ($< 1\text{ms}$)
- **efeito de precedência:** sistema auditivo identifica um evento sonoro (fusão) as ondas diretas determinam a direção (1-80ms)
- **ecos:** vários eventos (repetições) são percebidos (limiar de eco: 1-80ms)



Reflexões

- Efeito de **dereverberação binaural** - a reverberação é reduzida na percepção
- Efeito de **decoloração** - a diminuição da coloração causada pelas atrasos pequenos (efeito filtro-pente)

Segregação do stream auditivo

Situação: campo sonoro com diferentes fontes sonoras e reflexões é escutado

- sistema auditivo analisa os sinais de entrada para formar os eventos auditivos
 - distinção entre eventos
 - identificação de fontes sonoras (mapeamento)
- pessoa é capaz de focar atenção a certas fontes ignorando outras
 - efeito festa-de-coqueteis (*cocktail-party effect*)
 - supressão de ruído aumentando a inteligibilidade

Áreas de aplicação

Além da pesquisa: modelos do sistema auditivo e suas partes são interessantes para a aplicação na tecnologia:

- digitalização e mapeamento de cenas auditivas
- análise de cenas auditivas para desenvolver representações parametrizadas no nível de sinais (síntese / codificação de cenas auditivas, audiolgia)
- análise no nível simbólico (reconhecer sirene do SAMU)
- avaliação de cenas auditivas em termos de qualidade (qualidade de MPEG-2 Audio Layer III em termos de localização)

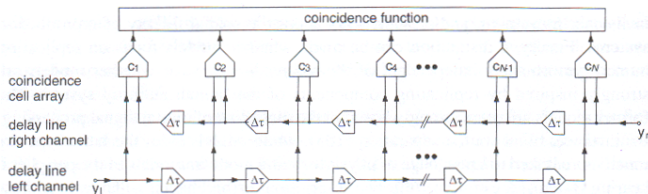
Modelagem binaural

Replicar certos aspectos comportamentais da escuta binaural com algoritmos:
modelos para localização e modelos para detecção.

Modelos de localização

Jeffress (1948)

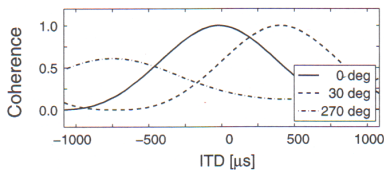
- baseado ITDs
- *delay line* para cada entrada
- sinais encontram-se numa célula de coincidência



Cherry and Sayers (1956)

- baseado ITDs
- usando IACC (*interaural cross-correlation*)
- análise por faixas de frequência (por ex.: terceira banda de oitava)

$$\psi_{y_l, r}(\tau) = \frac{\int_{t=-\infty}^{+\infty} y_l(t) \cdot y_r(t + \tau) dt}{\sqrt{\int_{t=-\infty}^{+\infty} y_l^2(t) dt \cdot \int_{t=-\infty}^{+\infty} y_r^2(t) dt}},$$



ILDs

Segunda deixa importante para localização: as ILDs

- diferenças de até 30dB de atenuação (5kHz e 60° azimuth)
- pequenas diferenças nas frequências graves (comprimento de ondas $\gg$ diâmetro da cabeça)
- em comparação: ITDs funcionam pior nas frequências agudas (comprimento de ondas $\ll$ diâmetro da cabeça)

Localização com ITDs e ILDs

HRTF (*Head Related Transfer Functions*) Resposta que caracteriza como um som chega de um certo ponto no espaço à orelha.

Método para localização

- Calcular um mapa de ITDs e ILDs usando as HRTFs (uma para cada ponto no espaço)
- Comparação do mapa com as ITDs e ILDs extraídas da cena aural

→ localizações ou ângulos de incidência

Problemas = objetos de pesquisa

- Combinação de ITDs e ILDs
- Integração sobre o tempo, sobre as frequências
- Distinção entre fontes simultâneas
- Tratamento de reflexões

Esquema de um modelo para a localização

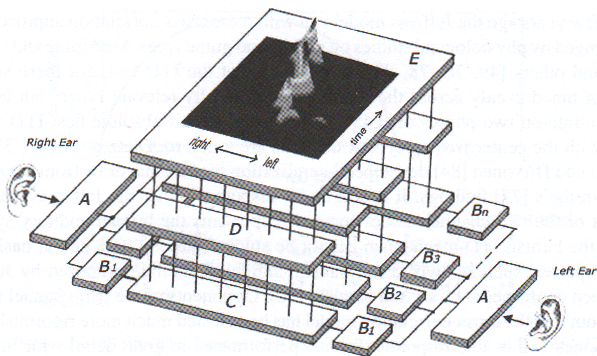


Fig. 4 Architecture of the bottom-up part of a model of binaural signal processing—from [9]. *A*... middle-ear modules, *B*_{1...n}... cochlea modules—working in spectral bands, *C*... modules to identify and assess interaural arrival-time differences, ITDs, in bands, *D*... modules to assess interaural level differences, ILDs, in bands, *E*... running binaural-activity map with the three dimensions intensity, sideward deviation and time

Modelos de detecção

Objetivo

- Detecção de sinais em cenas espaciais (até com SNR negativo)
- Melhorar o SNR / BMLD de certos sinais

BMLD (*binaural masking-level difference*) diferença do nível de ruído permitido para ainda detectar o sinal entre o caso monoaural e binaural.

Para isso os modelos consideram as **diferenças** dos sinais dos canais (ao contrário de modelos de localização que analisam **similaridades**).

Modelo EC

Durlach (1960)

Arranjo experimental

Cena espacial é dada com um sinal de interesse e um sinal mascarante (na condição de mascaramento).

Análise em 2 passos

- **E**: equalização dois canais
com SNR negativo → resulta em equalização do mascarado
- **C**: cancelamento
subtrair os canais um do outro

→ melhores (menores) limiares de detecção (SNR respectivamente BMLD)

Modelos de detecção

Desde então várias modificações e afinações desse modelo foram propostas. Entre muitas outras:

- equalizações em faixas de frequências
- conhecimento *a priori* sobre as fontes,
- usando modelos da fisiologia do ouvido (por ex.: modelos das células ciliadas)
- ...

Fontes e links

- Blauert, Jens (Ed.): **The Technology of binaural Hearing**, Berlin Heidelberg, Springer, 2013.
- **The Auditory Modeling Toolbox** <http://amtoolbox.sourceforge.net>.
- Sousa, Gustavo Henrique Montesião de: **Auralização de fontes sonoras móveis usando HRTFs**, Dissertação de mestrado, São Paulo, IME-USP, 2010.

Muito obrigado!