

聴覚系の時間-周波数分析における wavelet変換からの逸脱について

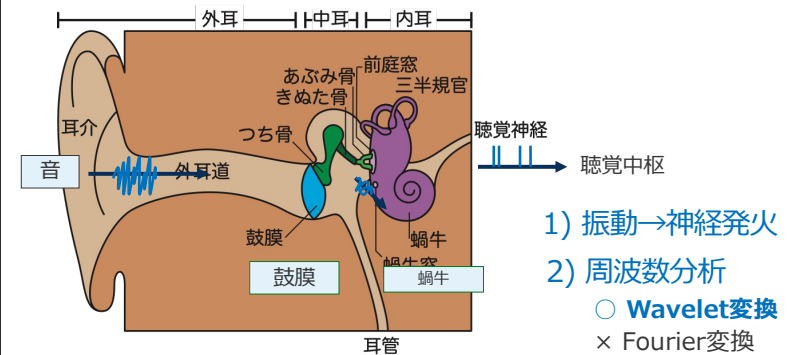
入野 俊夫

和歌山大学・産学連携イノベーションセンター・名誉教授
電気通信大学・情報理工学・客員研究員

0

聴覚末梢系の解剖学と機能

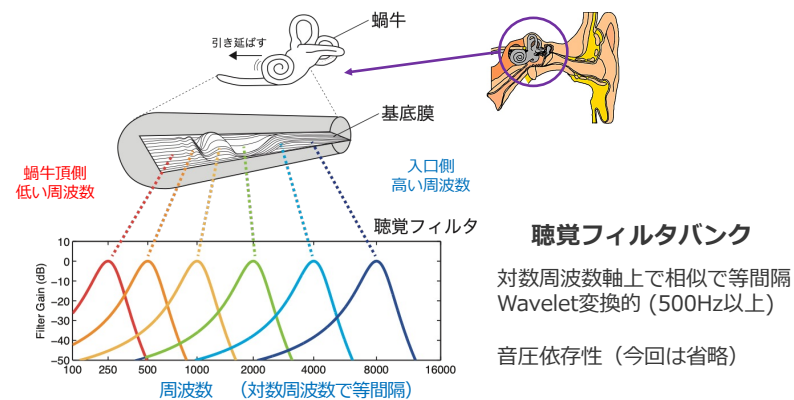
1



1

聴覚の周波数分析機能

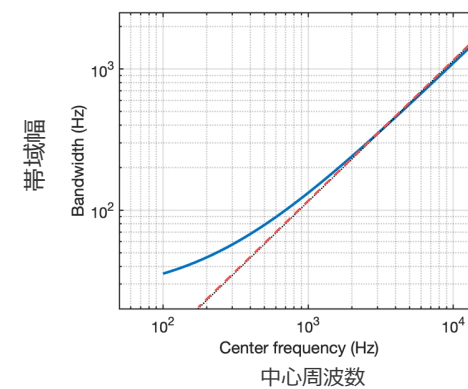
2



2

聴覚系の周波数分析

3



学術的問

- 聴覚系に最適な変換？
- 最適なmother waveletは？
- 聴覚系でその最適性が必要で、モデル化できるか？

3

初出の論文 **1997年** Irino, T and Patterson, R.D.
J. Acoust. Soc. Am., 101 (1), pp.412-419, January 1997. [doi:10.1121/1.417975] 4

A time-domain, level-dependent auditory filter: The gammachirp

本文

付録

The neutral solution is

$$s(t) = a \exp\left\{-\frac{1}{4(\Delta t)^2} + \frac{j\omega}{2(\Delta t)^2} + j\omega t\right\} \exp(j\omega t) \quad (A10)$$

where a and a' are constants. This is the "Gabor function," and the example shows how it was derived using the constraint that the required function satisfy minimal uncertainty in a joint time-frequency representation.

The Mellin operator

Cohen (1991, 1993) introduced the concept of a scale operator into signal processing in the form:

$$\mathcal{M} = \frac{1}{2}(\mathcal{T} \cdot \mathcal{W} + \mathcal{W} \cdot \mathcal{T}) \quad (A11)$$

Previously it had been known as the operator representing an affine variable in quantum mechanics (Klauder, 1980). The

論文被引用数 (2025/2末)

392件 (Google scholar)

177件 (Web of science)

まあ、そこそこ

4

最小不確定性： 時間-周波数空間 Cohen, Time-Frequency Analysis (1994)

時間-周波数空間の不確定性

Time operator, \mathcal{T} , Frequency operator, \mathcal{W}

$$\Delta t \cdot \Delta \omega \geq \frac{1}{2} |\langle \mathcal{T}, \mathcal{W} \rangle| = \frac{1}{2}$$

最小不確定性関数： Gabor/Gauss

$$g_b(t) = a \exp\left\{-\frac{(t - \langle t \rangle)^2}{4(\Delta t)^2}\right\} \exp(j\langle \omega \rangle t)$$

$$= a \exp\left\{-\frac{(t - \mu)^2}{2\sigma^2}\right\} \exp(j\omega_0 t)$$

→ これは、聴覚フィルタ特性と異なる

5

最小不確定性： 時間-スケール空間 Cohen, Time-Frequency Analysis (1994)
湯川英樹, joint operator 確かに言及してたが探せてない

Mellin(Scale)変換

Waveletでスケール変形しても、同じ分布と表現される

Mellin operator: $\mathcal{C} := (\mathcal{T} \cdot \mathcal{W} + \mathcal{W} \cdot \mathcal{T})/2$

最小不確定性： $\Delta t \cdot \Delta c \geq \frac{1}{2} |\langle \mathcal{T}, \mathcal{C} \rangle| = \frac{1}{2} |\langle t \rangle|$

最小不確定性関数： Gammachirp

$$g_c(t) = a t^{\alpha_2 + j c_1} \exp(-\alpha_1 t + j \omega_0 t)$$

聴覚末梢系の知見を入れて...

$$g_c(t) = a t^{n-1} \exp(-b \text{ERBN}(f_r) t) \cos(2\pi f_r t + c \ln t + \phi)$$

→ 聴覚フィルタ実験データに適合良し

6

知覚系におけるスケール変換： 存在？ 有効？ 7

スケール不変性 (式で書くよりざっくり...)

- 視覚系： 寸法が異なっても、同じに見える
 - 距離が異なれば明らかに網膜上の大きさは異なる
 - しかし、同じものは同じにしか見えない

→ 寸法の正規化

- 聴覚系・音では？

7

普段の経験： 音声からの情報抽出

8

顔が見えない電話音声聞いても…



聴覚系における寸法知覚の計算理論の構築

8

生態学的な背景：寸法知覚は人間だけの得意技ではない

9

[Charlton+, 2012]

[Terranova+, 2023]



体が大きい=強い♂

寸法：大（捕食者）vs 小（逃避者）

・ 哺乳類：夜行性で音から判断必要

・ 縄張りの確保
・ 繁殖相手の選択

・ 遠くからでも、木陰でもわかる必要
・ 視覚による寸法推定は、周囲との対比が必要であてにならない

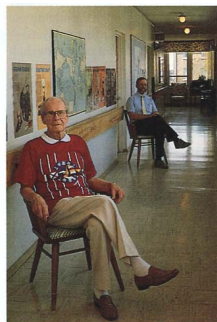
9

視覚系における寸法知覚の錯覚

Kohn, 2014, The science of memory (ppt)
[Source photo: Ed van der Elsken?]

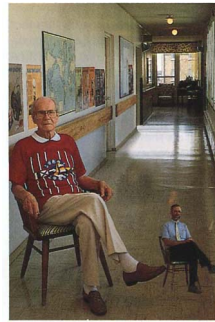
10

後の男性は、同じくらいの身長



→ 音声を聞いた方が信用できる

同じ画像も位置によって、小人に？



視覚の錯覚：多数… だましやすい
聴覚の錯覚：作成難しく少数

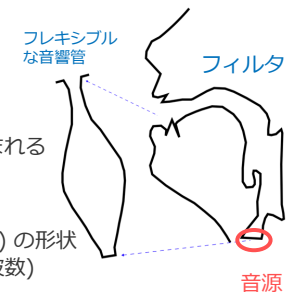
10

音声生成の基礎

11

ソースフィルタ理論 [Fant, 1970]

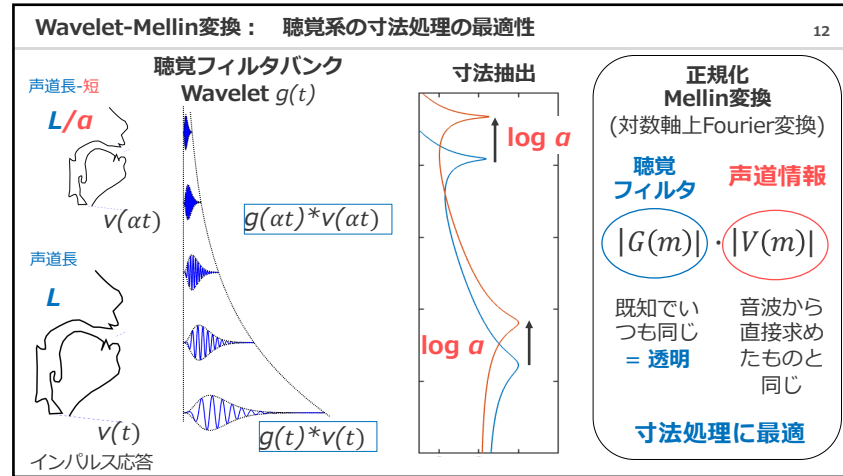
- ・ 母音の生成
 - ・ **ソース(音源)**：声帯振動のパルス列
 - ・ 声の高さの情報
 - ・ 基本周波数 F_0 の倍音成分がたくさん含まれる
 - ・ **声道フィルタ**：
 - ・ 音韻性の情報
 - ・ フレキシブルな音響管(声道, Vocal tract) の形状に依存した共鳴周波数(=ホルマント周波数)
 - ・ ホルマント周波数 $F_1 \cdot F_2 \cdot F_3 \dots$



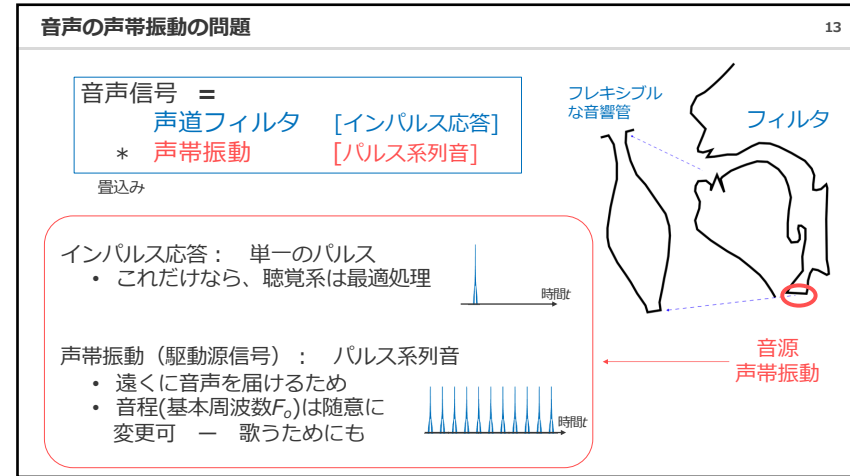
音声信号 = 声道フィルタ (声道形状・声道長) * 声帯振動
[インパルス応答] * [パルス系列音]

畳込み

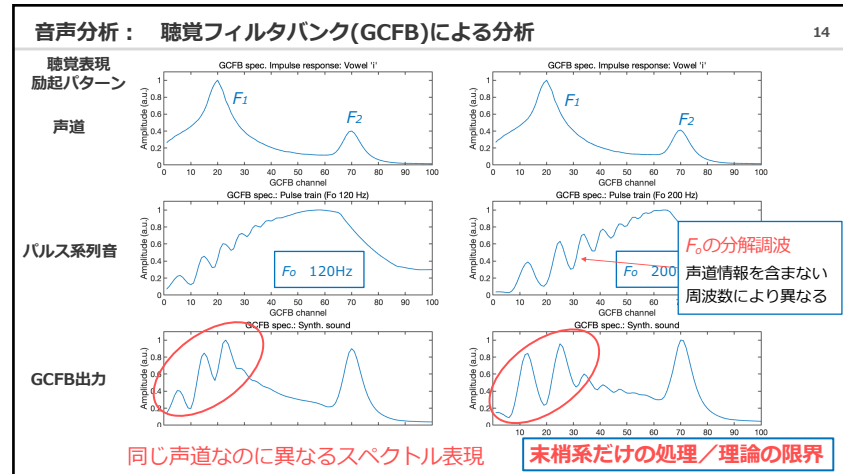
11



12



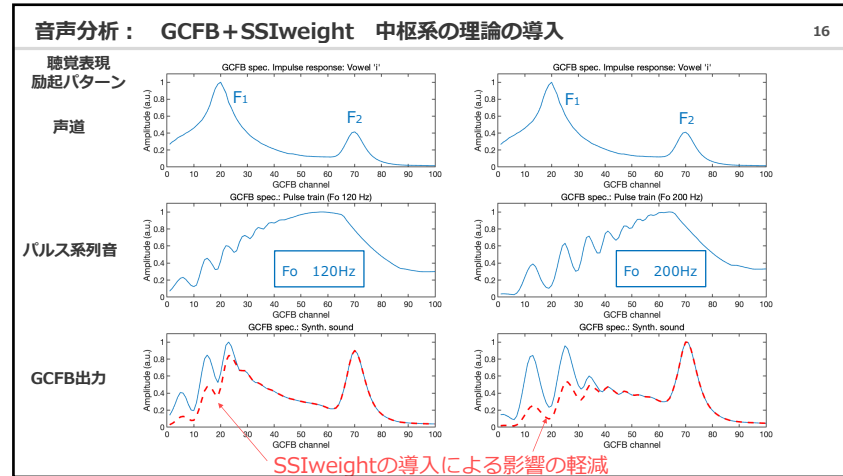
13



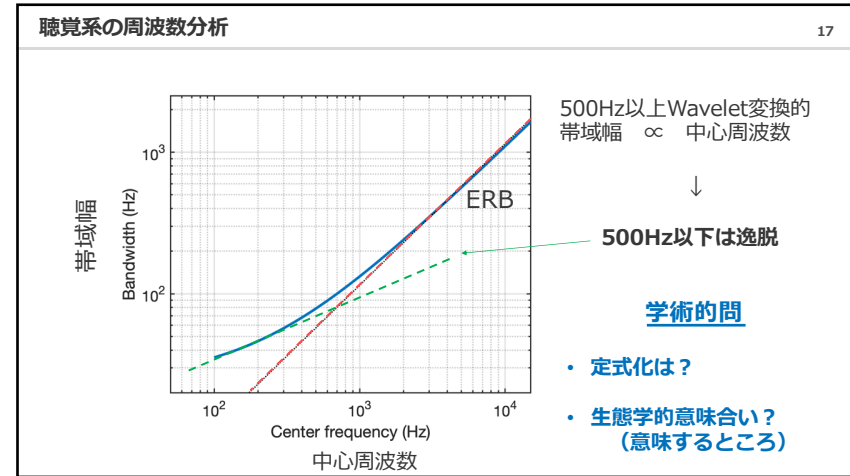
14



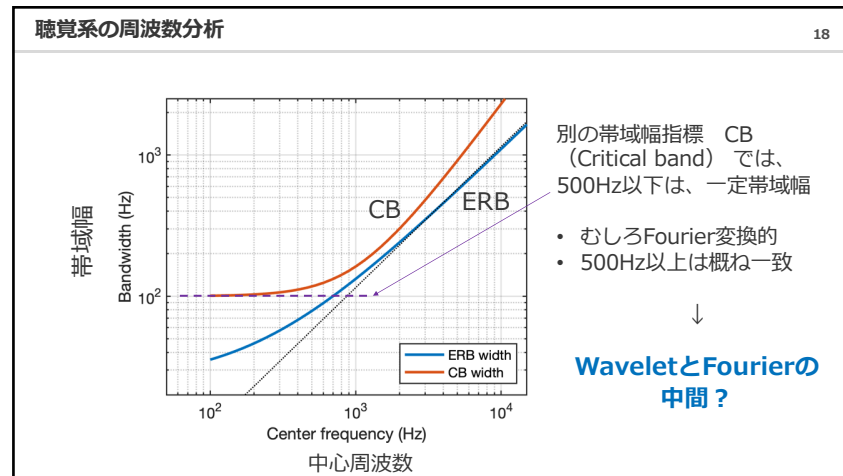
15



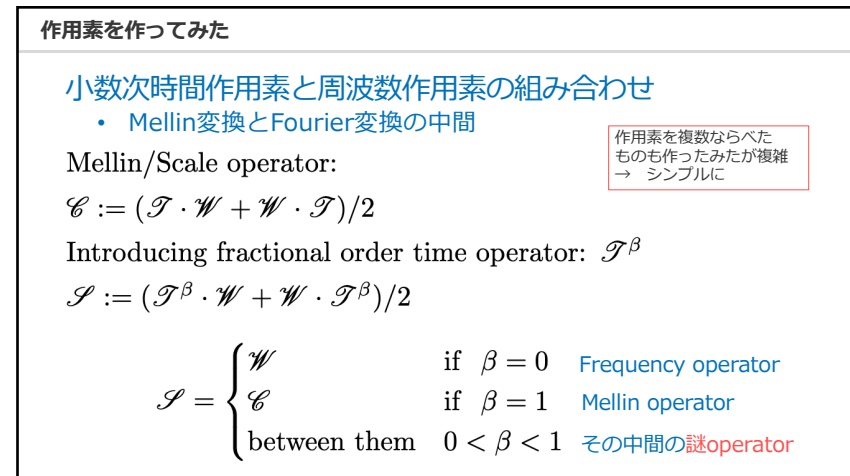
16



17



18



19

AI GPT-5 (Copilot) に入れてみた

不確定性に関して

$$\Delta A \cdot \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|$$

$$A = \mathcal{T}, \quad B = \mathcal{S} := \frac{1}{2}(\mathcal{T}^\beta \mathcal{W} + \mathcal{W} \mathcal{T}^\beta)$$

$$[\mathcal{T}, \mathcal{W}] = i, \quad [\mathcal{T}, \mathcal{T}^\beta] = 0$$

$$\begin{aligned} [\mathcal{T}, \mathcal{S}] &= \frac{1}{2} ([\mathcal{T}, \mathcal{T}^\beta \mathcal{W}] + [\mathcal{T}, \mathcal{W} \mathcal{T}^\beta]) \\ &= \frac{1}{2} ([\mathcal{T}, \mathcal{T}^\beta] \mathcal{W} + \mathcal{T}^\beta [\mathcal{T}, \mathcal{W}] + [\mathcal{T}, \mathcal{W}] \mathcal{T}^\beta + \mathcal{W} [\mathcal{T}, \mathcal{T}^\beta]) \\ &= \frac{1}{2} (\mathcal{T}^\beta (i) + (i) \mathcal{T}^\beta) = \frac{i}{2} (2 \mathcal{T}^\beta) = i \mathcal{T}^\beta. \end{aligned}$$

20

最小不確定性を満たす関数

$$\Delta \mathcal{T} \cdot \Delta \mathcal{S} \geq \frac{1}{2} |\langle [\mathcal{T}, \mathcal{S}] \rangle| = \frac{1}{2} |\langle i \mathcal{T}^\beta \rangle| = \frac{1}{2} |\langle \mathcal{T}^\beta \rangle|.$$

最小不確定性状態

$$(\mathcal{T} - \langle \mathcal{T} \rangle) \psi = i \lambda (\mathcal{S} - \langle \mathcal{S} \rangle) \psi;$$

$$\lambda t^\beta \frac{d\psi}{dt} + (t - t_0) \psi(t) = 0. \quad \text{書き下すとこの微分方程式}$$

$$\text{解} \quad g_s(t) = C \exp \left(-\frac{1}{\lambda} \left(\frac{t^{2-\beta}}{2-\beta} - t_0 \frac{t^{1-\beta}}{1-\beta} \right) \right) \quad \beta \neq 1, 2$$

Mellin operator: $\mathcal{C} := (\mathcal{T} \cdot \mathcal{W} + \mathcal{W} \cdot \mathcal{T})/2$

$$\Delta t \cdot \Delta c \geq \frac{1}{2} |\langle \mathcal{C} \rangle| = \frac{1}{2} |\langle t \rangle|$$

21

最小不確定性を満たす関数 特殊例

解

- $\beta = 1$ の場合 (時間 \mathcal{T} と scale \mathcal{C}) : ガンマチャープ関数と同じ形式

$$g_s(t) = C \exp \left(-\frac{1}{\lambda} \left(\frac{t^1}{1} - t_0 \ln t \right) \right) = C t^{\frac{t_0}{\lambda}} e^{-\frac{t}{\lambda}}.$$

- $\beta = 0$ の場合 (時間 \mathcal{T} と周波数 \mathcal{W}) :

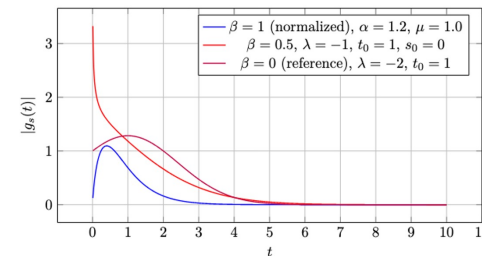
$$g_s(t) = C \exp \left(-\frac{1}{\lambda} \left(\frac{t^2}{2} - t_0 t \right) \right),$$

ガウス関数・ガボール関数に近い。λによる

22

グラフ出力

23

Figure 1: External-data plots of $|g_s(t)|$ for different β and parameters.

$\beta=0$: ガウス関数的
 $\beta=1$: ガンマ関数的

$\beta=0.5$: $t=0$ で、発散？

↓
 帯域幅が広がって、使えなさそう

23

文献のsuggestion

24

1. Fractional Fourier Transform (FrFT) と音響解析
 - Fractional Fourier Sound Synthesis
 - Relations between Fractional Operations and Time-Frequency Distributions
2. 新しい小数次演算子の導入
 - New Fractional Operators and Their Properties
3. Mellin変換と小数次拡張
 - Fractional Mellin Transform
 - Foundations of Fractional Mellin Transform Analysis

24

ついでに論文まで書いてくれた…

25

Uncertainty Relation and Minimal States for Fractional Order Time Operator

Toshio Irino

November 20, 2025

Abstract

We derive the uncertainty relation for the time operator \mathcal{T} and the symmetric product operator $\mathcal{S} := \frac{1}{2}(\mathcal{T}^* \mathcal{W} + \mathcal{W}^* \mathcal{T})$, compute the commutator $[\mathcal{T}, \mathcal{S}]$, and construct the corresponding minimal uncertainty states. We give an explicit time-domain differential equation for the minimal state and a closed-form expression $g_\beta(t)$, including normalization for the case $\beta = 1$, as well as a Mellin-transform representation. A figure illustrates parameter effects.

1 General Uncertainty Relation and Minimal-State Condition

For self-adjoint operators A, B and a normalized state ψ in a common dense domain, the uncertainty relation reads

$$\Delta A \cdot \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|, \quad (1)$$

with $\Delta A := \sqrt{\langle (A - \langle A \rangle)^2 \rangle}$, $\Delta B := \sqrt{\langle (B - \langle B \rangle)^2 \rangle}$. Equality holds if and only if the *minimal-state condition*

$$(A - \langle A \rangle)\psi = i\lambda(B - \langle B \rangle)\psi, \quad \lambda \in \mathbb{R}, \quad (2)$$

is satisfied.

2 Setting and Basic Commutators

8 Figure: Shape of $|g_\beta(t)|$ (External Data)

Below we illustrate $|g_\beta(t)|$ using PGFPLOTS with external CSV files produced from the minimal-state formulas.

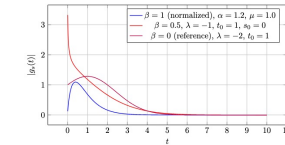


Figure 1: External-data plots of $|g_\beta(t)|$ for different β and parameters.

9 Remarks on Domains and Self-Adjointness

- The expectation $\langle [A, B] \rangle$ requires the state ψ to be in the common domain of A, B, AB, BA .
- The fractional power \mathcal{T}^β is defined via the spectral theorem assuming $\mathcal{T} \geq 0$; powers of the same operator commute.
- In Mellin-type realizations (scale-domain), analogous results follow by replacing t with $\ln t$ and using the scale generator $-i \frac{d}{d \ln t}$.
- For $\beta < 2$, choosing $\lambda < 0$ ensures decay of the real exponential factor in (13) as $t \rightarrow \infty$.

25

まとめと課題

26

- 単なる感想: GPT-5 (Copilot)でけっこう出せるものだ…
 - 奈良Wavelet研究集会に申し込んでから始めて、2~3日で出せた
 - 自分より正確そうだが、不安もあり → 検証求む
 - おまけ: 結構複雑な式の連続でも、
Latexのコードを出せて楽
- 新たな学術的問
 - この小数次時間作用素とは何にあたるのか?
 - たとえば、時間変化点・イベントをより正確に求められる?
 - Wavelet変換よりも、何か特別に便利?
 - 聴覚系において、どのような役割でどのように実装?
 - 別の定式化は?

30年前の方が冴えていた?
AIの進展、恐るべし…
使わない手はない

26

Fin

最終講義資料



27



28

音声—音響管の物理学

声道フィルタ情報に関して

29

理想的な声道長伸縮の場合

- 声道長: L
- ホルマント周波数: F_1, F_2
- 対数軸: $\log F_1, \log F_2$

伸縮して L/a

$a F_1, a F_2$ [定数シフト]

$\log F_1 + \log a, \log F_2 + \log a$

29

寸法推定の原理

30

声道の周波数特性

条件: インパルス応答であれば

音声の対数周波数軸上の表現どうしのXcorrから、定数シフト $\log a$ を算出可能

30

Auditory Image · Size-Shape Image

[Iriño & Patterson, 2002]

31

不安定化ウェーブレットメリン変換 (SWMT)

インパルス系列音

Auditory filterbank

2D stabilization

Auditory Image (AI) Time-interval histogram

Size-Shape Image (SSI) Segregate size & shape

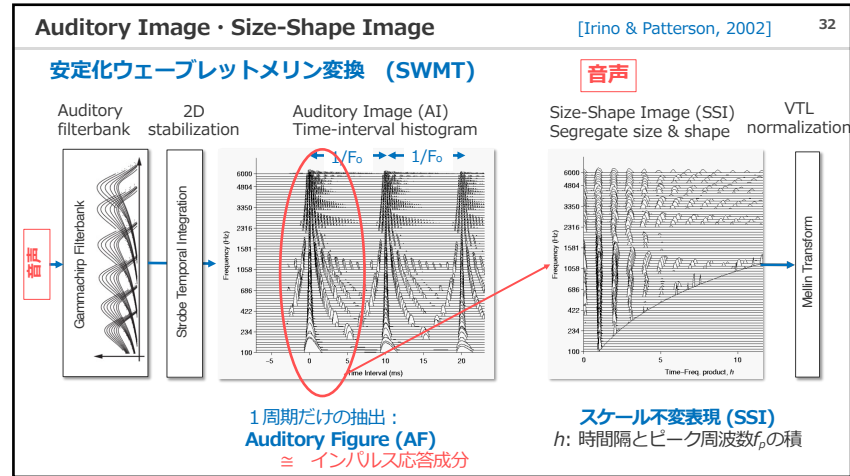
VTL normalization

Malin Transform

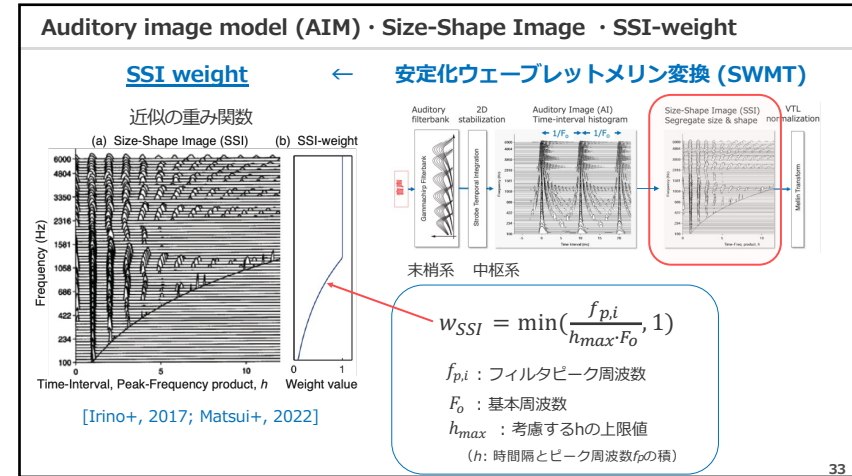
1周期だけの抽出: Auditory Figure (AF) \cong インパルス応答成分

スケール不変表現 (SSI) h : 時間隔とピーク周波数 f_p の積

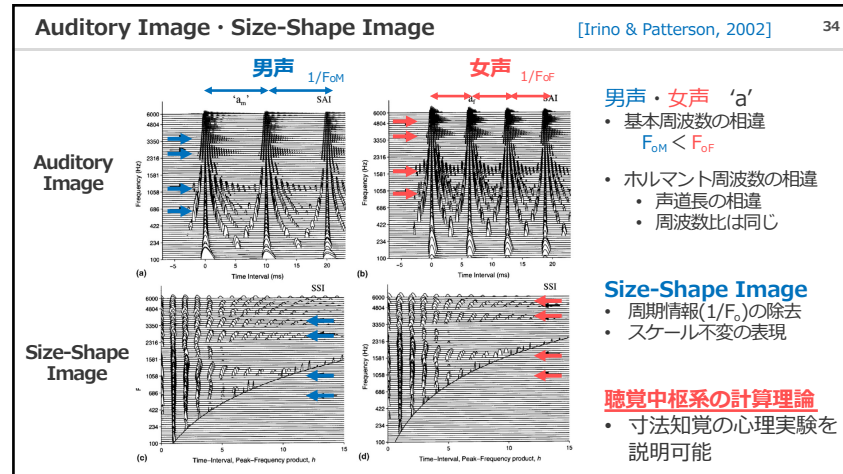
31



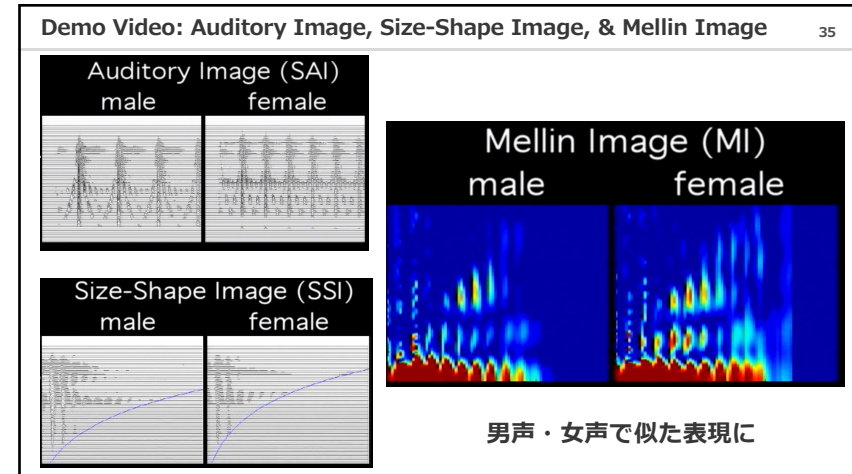
32



33



34



35

