

情報科学研究科 人間情報処理領域 音情報処理分野

- ★講座HP: <http://ais.jaist.ac.jp/index-j.html>
- ★教官HP: <http://www.jaist.ac.jp/~unoki/>
- ★ライフスタイルデザイン研究センター
<http://www.jaist.ac.jp/kscenter/index-j.html>

鵜木研究室



鵜木研究室 構成員・研究体制

- 准教授：鵜木祐史
- 助教：宮内良太
- 研究補助員（2名）
- 博士後期課程（3名）
- 博士前期課程（12名）（うち東京サテライト1名）
- 修了生：15名（2006年～） 赤木研（7名）
- 共同研究者：
 - Dr. Xugang Lu (NICT)
 - Prof. Roy D. Patterson (Univ. Cambridge)
 - Prof. Dr.-Ing. habil. Rüdiger Hoffmann (TUD)
 - Dr.-Ing. Rico Petrick (Tech. Univ. Dresden)

Masashi Unoki, Ph.D

- **Master/Ph.D @ JAIST (Prof. Akagi) (1994-1999)**
- **Visiting Researcher @ ATR-HIP (Prof. Irino) (1999-2000)**



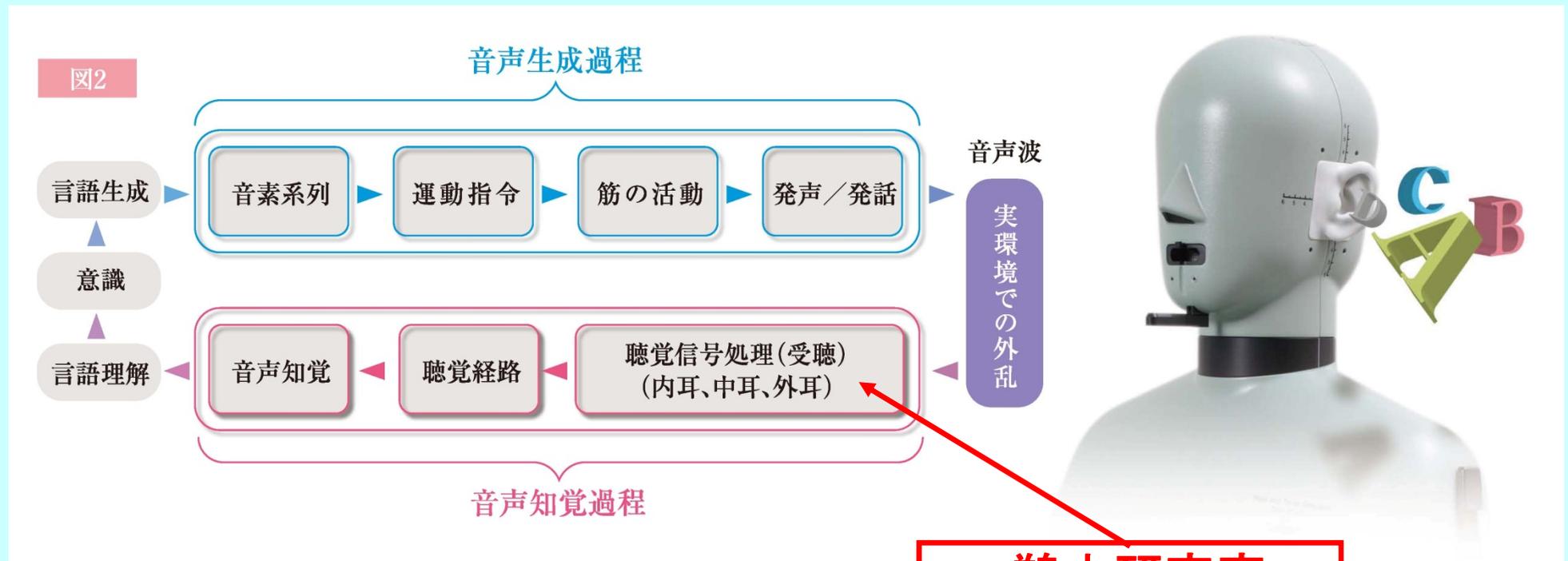
- **Visiting Researcher @ Cambridge Univ. (2000-2001)**
- **Associate@JAIST (2001-2005)**
- **Associate Prof. @JAIST (2005-)**

人間情報G／音情報処理の研究ターゲット

■ 基本路線

- **話す・聞く**は人間の営み → 人間を知り、そして、営みを記述することで、高度の音声処理システムの実現を目指す

■ 研究範囲：音声知覚，**聴覚信号処理**，音声生成



<赤木研究室と一緒に活動>

鵜木研究室

「耳」付き計算機の登場が、 音信号処理に革命を起こす!?

雑音や残響がある環境でも、人間は聞きたい音を聞くことができる。
「もしも計算機上に、聴覚と同じ機能を持つ信号処理システムを構築できたら、さまざまな音信号処理に応用できるのでは？」
という考えの下で研究を進めているのが、鷗木准教授の研究室だ。

耳に相当する機能を 計算機につけてみたい

ICレコーダーでインタビューを録音した音声を聞き直してみると、インタビュー時には気にならなかった雑音や残響で、話し手の声が聞き取りにくいことがある。それなのになぜインタビュー時には話し手の声がクリアに聞こえたのだろうか？ それは、私たち人間の聴覚（正確には、聴覚系+脳）が、雑音や残響を排除し、狙った音だけを聞き取るという、非常に優れた機能を有しているからだ。

けれども、「私たちの聴覚は非常に優れている」ということはわかっているのに、「どのような信号処理によってそれが実現されているのか」ということは、実ははまだ解明されていないことが多々ある。自分たちの体の一部であるにも関わらず、である。

「もしもその信号処理のしくみが全て解明できれば、音声認識や補聴システム等の音声信号処理に応用できると思うんです」と鷗木准教授は語る。そして、自らの研究スタンスを図解すると「 π 型」(図1)になると言う。

「『聞く』『話す』といった音声コミュニケーションは、「ことばの鎖」(図2)の過程で行われています。情報科学研究科の音情報処理分

そこに行き着くためには、2本の柱による研究アプローチが必要です。1本目の柱は「人間の耳をまず良く知る」というアプローチです。学問的には、聴覚の各部位の機能解明を目指す『聴覚生理学』や、音の物理的変化に対する反応(知覚変化)を調べる『聴覚心理学』の研究領域です。もう1本の柱は「人間の耳を模擬して計算機上に記述する」というアプローチです。こちらは、学問的には『情報科学』あるいは『工学』の領域です。つまり、**聴覚生理学、聴覚心理学、情報科学の3つがオーバーラップする部分**が、私たちの研究領域です。JAISTの先生方や国内外の研究者に協力を仰ぎ、聴覚心理学・情報科学を本学で実施し、聴覚生理学に関しては海外の研究成果等を参考にしながら追求しています」

驚くほど快適な補聴器や 音声透かしの普及も間近

現在、鷗木研究室ではさまざまなプロジェクトが進行している。

そのひとつが、「**聴覚フィルタバンクの構築**」だ。聴覚器官では、蝸牛の基底膜で音を周波数成分に分解し、電気信号(パルス)に変換して神経に伝えている。この機能は、簡単にいうと帯域通過

耳はそもそもなぜ必要なのか？それが知りたくて研究を続けているんだと思う。

情報科学研究科 准教授

鵜木 祐史 Associate Professor Masashi Unoki

【趣味】 学生時代はラグビーをやっていました。根気よく研究を続けるには、体力や強い精神力が必要ですが、ラグビーを通じて研究に向かっていく力が身についたと思います。

【この研究を続ける理由】

聴覚はなぜ必要なのか？ どういう機能を持っているのか？ 科学を志している人間として、その一端を垣間見たいという純粋な理由で研究を続けているんだと思います。でも、生きている人間の耳を解剖するなんてことは不可能ですよ。そんな中で人間の耳の機能を解明するためには「人間の耳の各部位の機能は何か？」「人間の耳と等価なものを作るにはどうすればいいか？」ということを丹念に調べて、データを積み重ねていくしかない。遠回りに見えるけれど、実はそれが本質的でいちばんの近道なのかもしれないと思うんです。

学位

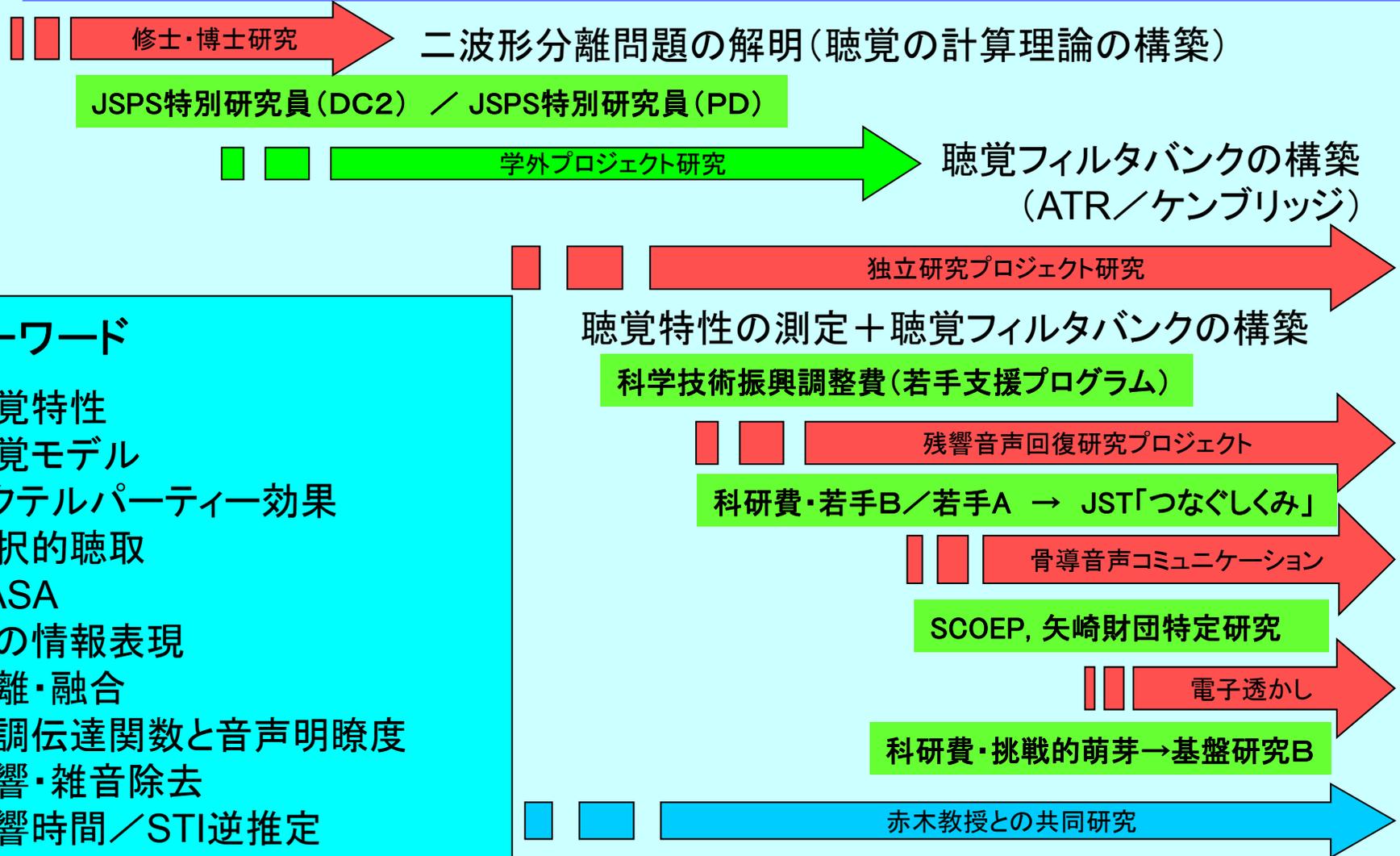
北陸先端科学技術大学院大学 情報科学修士(1996) / 北陸先端科学技術大学院大学 情報科学博士(1999)

経歴

日本学術振興会 特別研究員(DC2)(1998) / ATR人間情報通信研究所 客員研究員(1999) / ケンブリッジ大学 CNBH客員研究員(2000-2001) / 日本学術振興会 特別研究員(PD,北陸先端科学技術大学院大学 情報科学研究科)(1999-2001) / 北陸先端科学技術大学院大学 情報科学研究科助手(2001-2005)



音環境ディバイドレス・パンフレットより抜粋



- キーワード**
- ・聴覚特性
 - ・聴覚モデル
 - ・カクテルパーティー効果
 - ・選択的聴取
 - ・CASA
 - ・音の情報表現
 - ・分離・融合
 - ・変調伝達関数と音声明瞭度
 - ・残響・雑音除去
 - ・残響時間/STI逆推定
 - ・電子音響透かし
 - ・マルチメディア情報ハイディング

歌声合成/報知音知覚/音源分離/認識
 非言語音声(SCOPE)/聴覚フィードバック

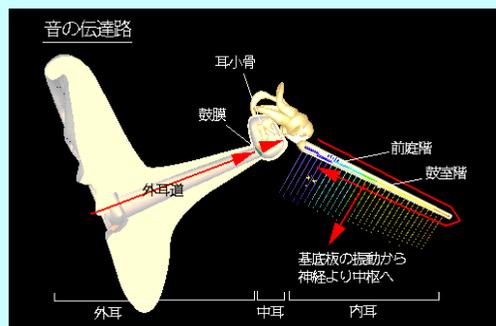
プロジェクト

聴覚的信号処理

聴覚モデル化

音（音声）
信号処理

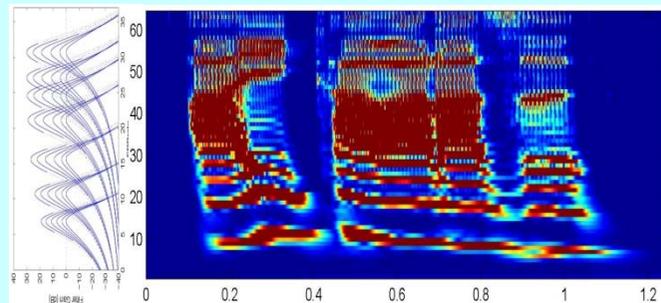
* 蝸牛モデル



* 周波数選択性



* 聴覚フィルタバンク



* 骨導音声回復



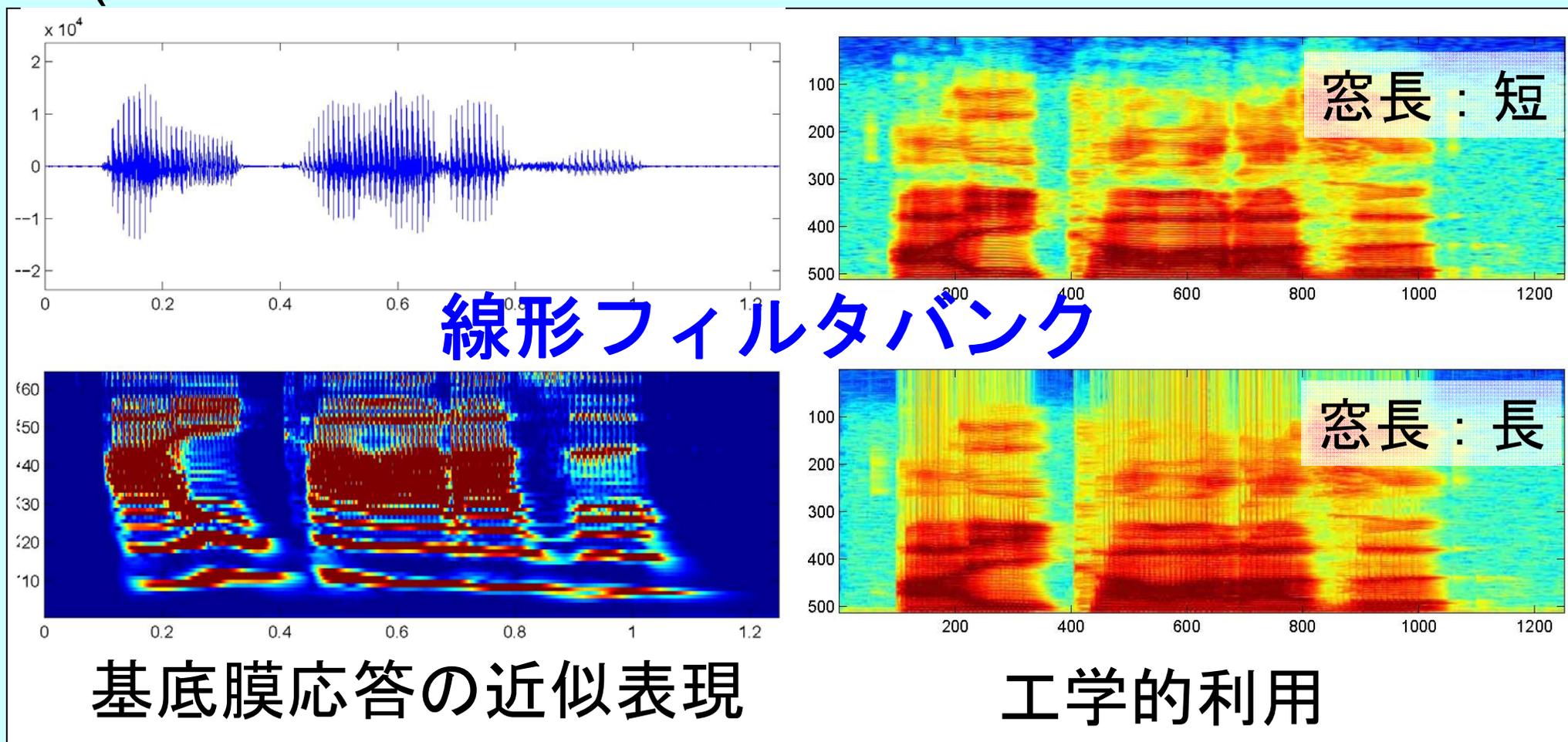
* 電子音響透かし



聴覚的信号分析技術の確立

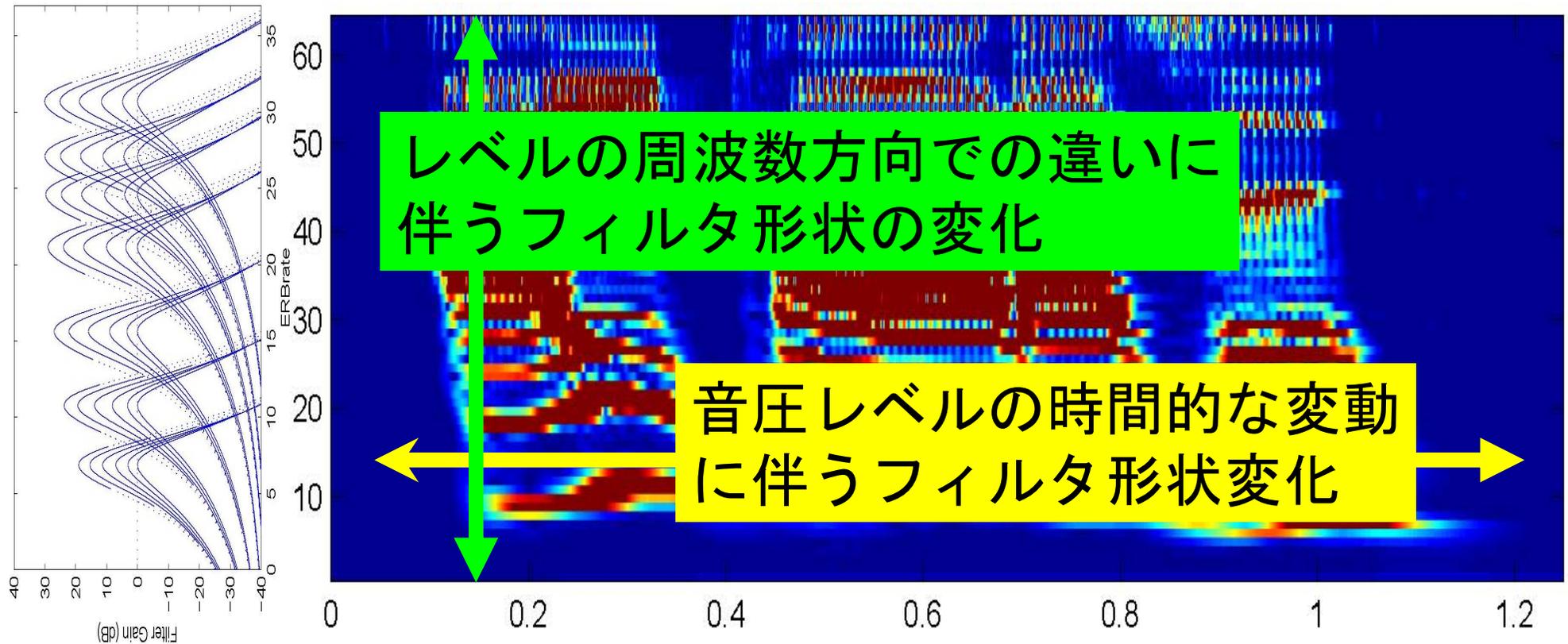
定Qフィルタバンク
(ウェーブレット変換)

vs. 短時間フーリエ変換
(窓関数による切取り)



マスキング特性⇒フィルタバンクの構築

- ある音（マスキュー）が妨害音（マスキュー）によって聞き取り難くなること

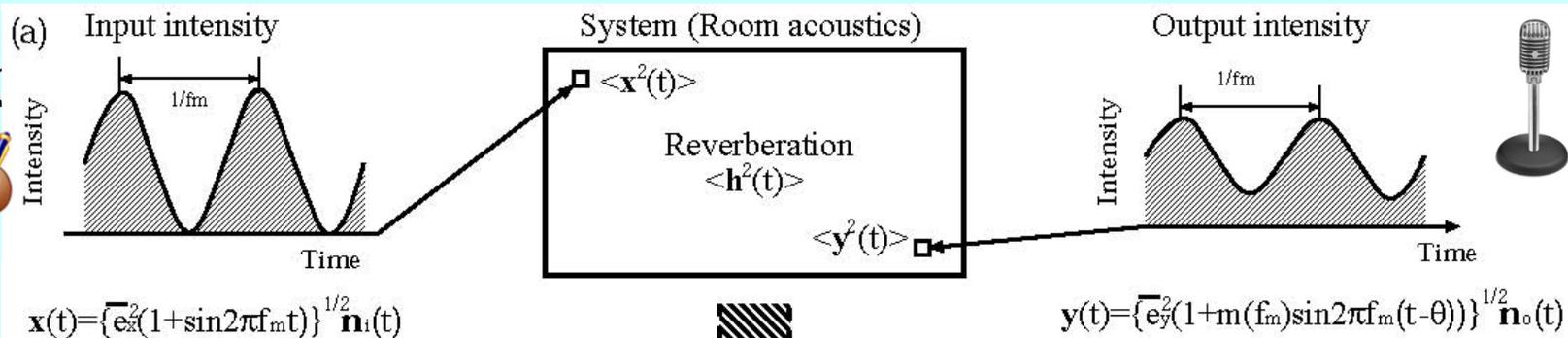


マスキング実験

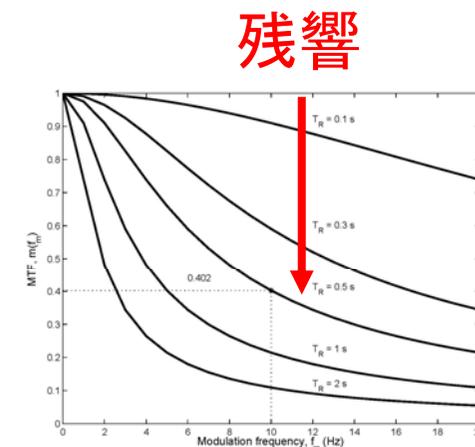
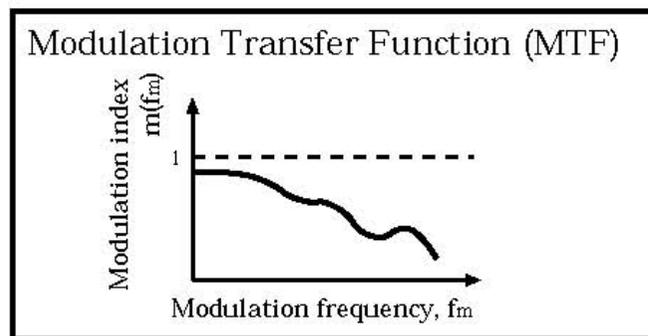
■ 実験環境の整備



変調伝達関数に基づいた信号処理



(b)

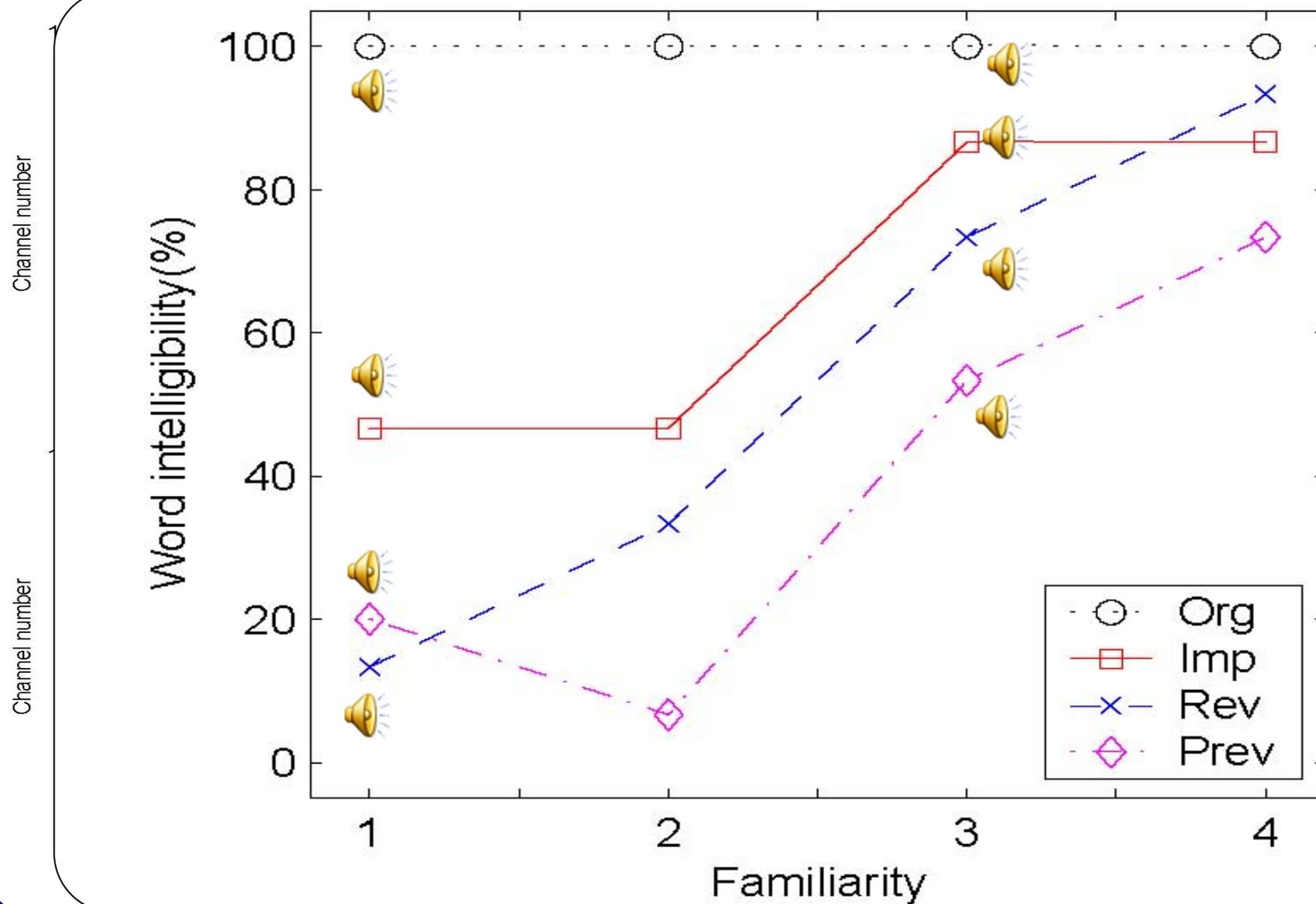


(c)

Speech Transmission Index (STI)

Prediction of speech intelligibility

残響音声回復のデモ



室内音響特性のブラインド推定



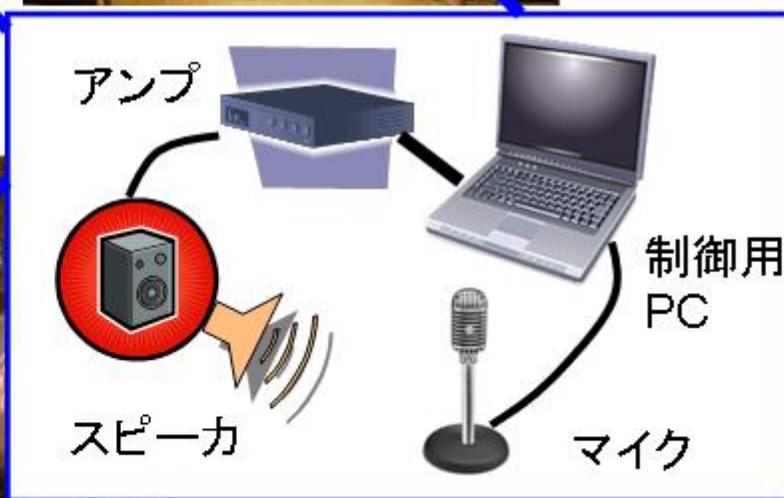
空港内



講義室



駅構内



室内インパルス応答測定装置
(大掛かりな装置)



問題点:

人を排除しなければいけない
装置を持ち込みしなければならない



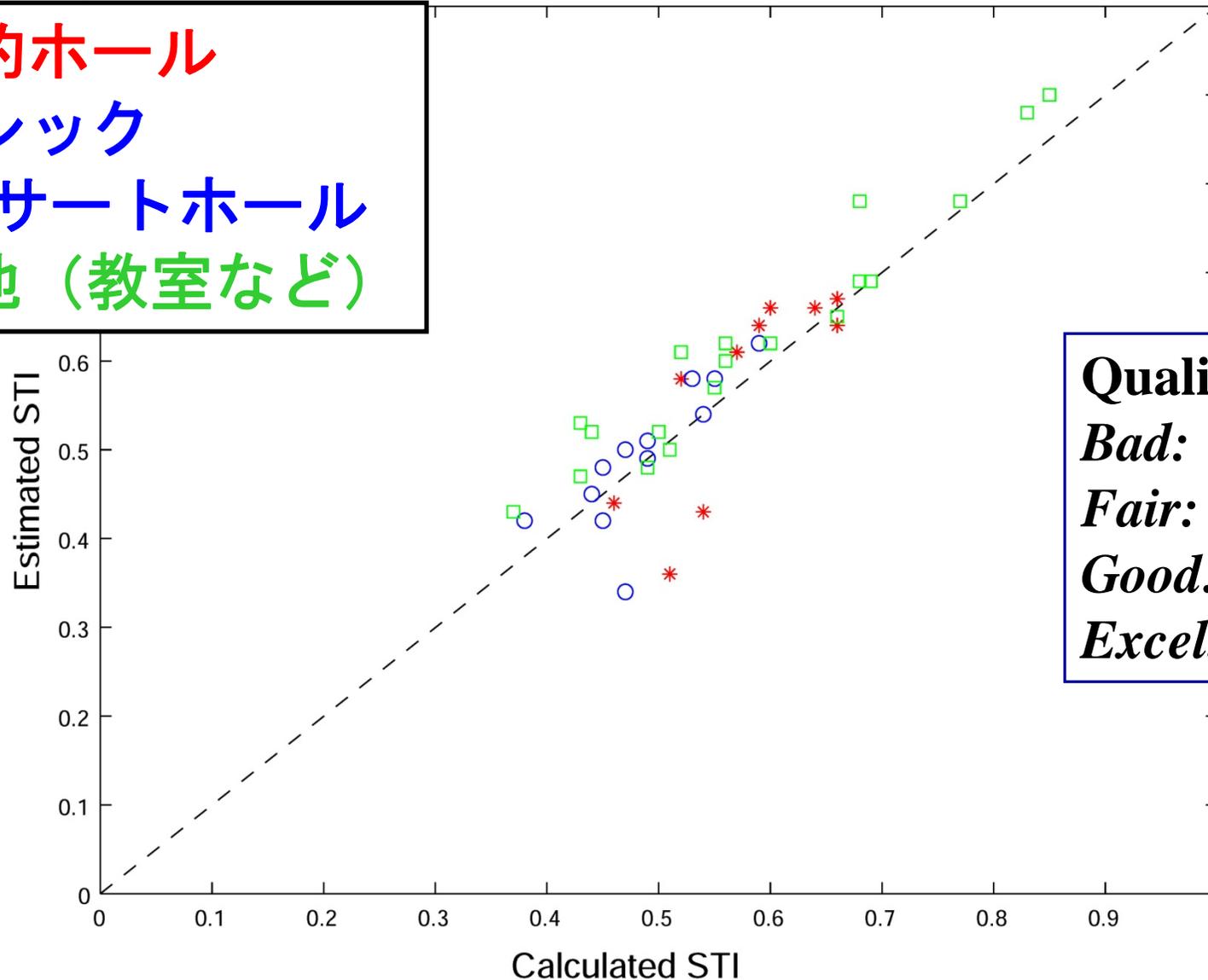
- ・実時間測定
- ・コンパクト
- ・持運び自由
- ・音声を利用



提案法

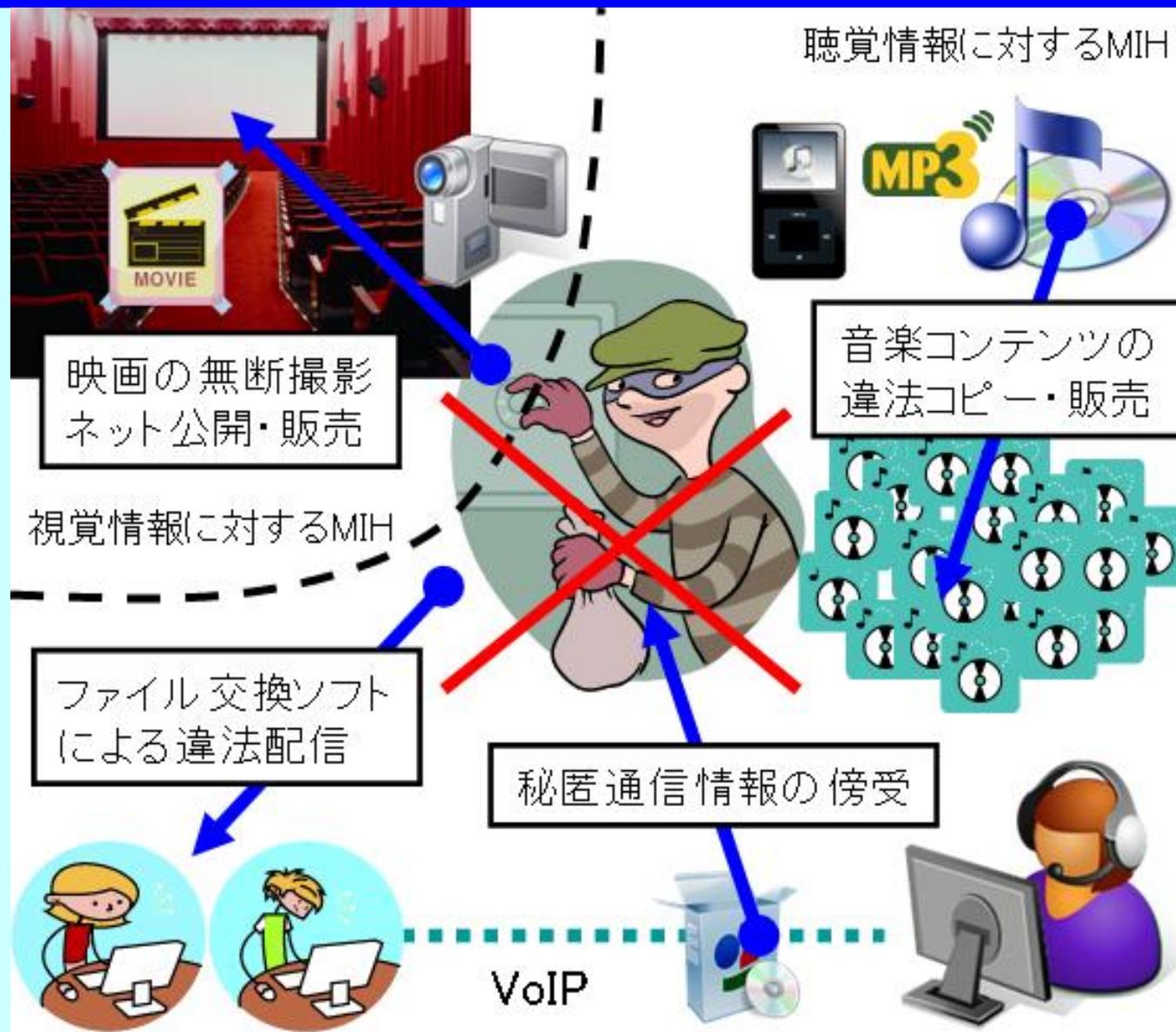
結果 (STI)

- ★多目的ホール
- クラシック
コンサートホール
- その他 (教室など)

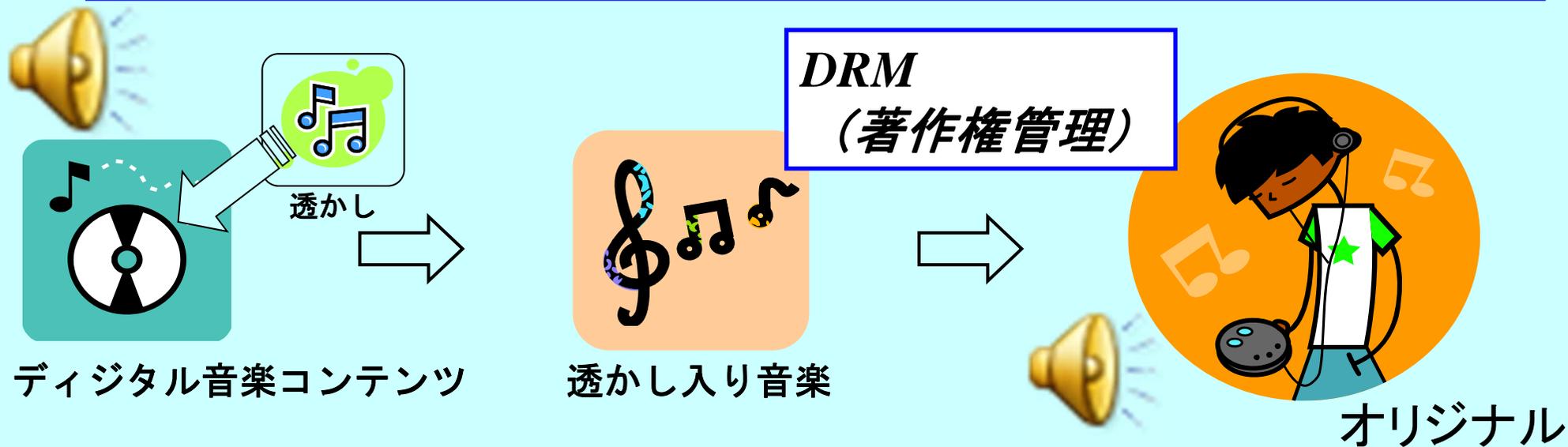


Quality SIT
Bad: 0.0-0.45
Fair: 0.45-0.6
Good: 0.6-0.75
Excel.: 0.75-1.0

マルチメディア情報ハイディング



電子音響透かし



蝸牛遅延特性 → 知覚不可能

	Echo hiding	Periodic PM	Proposed
Original			
			

音声信号への情報ハイディング

Speech database for intelligibility tests: 320 stimuli
(2 males and 2 females, 4 familiarities, 20 sets)

