

第 15 回 電子航法研究所研究発表会

平成 27 年 6 月 4 日

電子航法研究所の発話音声分析技術

塩見 格一

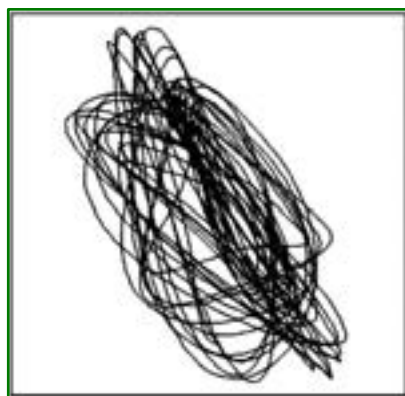
shiom@enri.go.jp

我々が音声信号を分析する目的は・・・

我々が音声信号を分析する目的は・・・

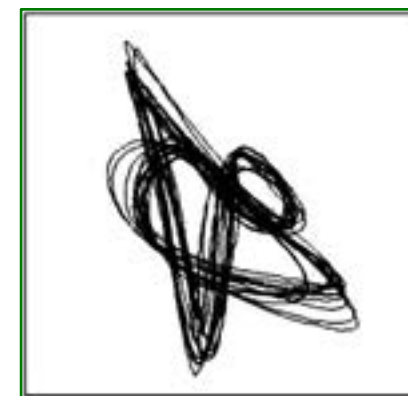
1) 1998年の発見：発話者の心身状態に依存して、音声信号から再構成されるストレンジ・アトラクタ (S.A.) が変化する。

← S.A. から計算されるリアプノフ指数が変化する。



健全状態のストレンジ・アトラクタ

→
changes to



消耗状態のストレンジ・アトラクタ

2) 発話音声から発話者の心身状態を評価する。

↓ 公共インフラの維持運用に係る要員の過労状態を防止する。

3) 「居眠り防止装置」等の予防安全装置を実現する。

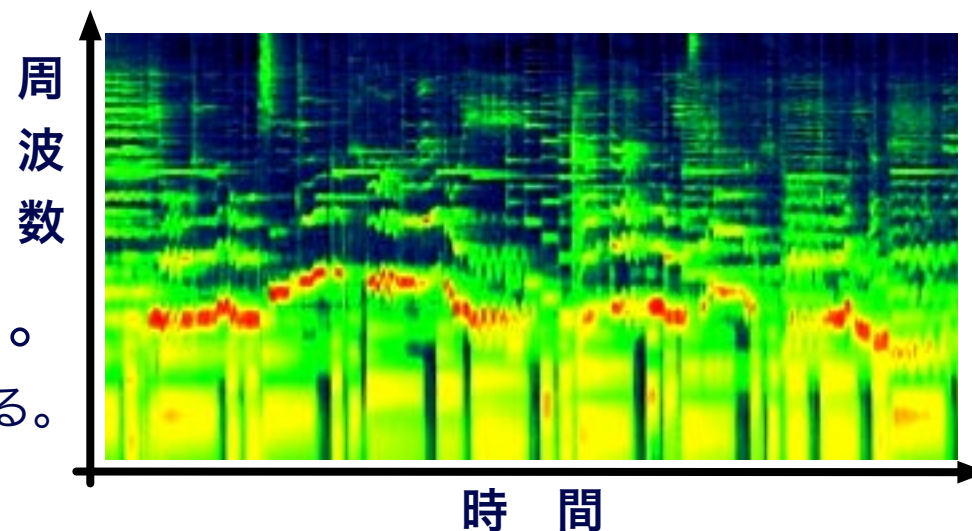
音声信号を分析する手法は幾つもあって・・・

音声信号を分析する手法は幾つもあって・・・

1) 周波数分析

→ フーリエ解析

ソノグラム上に特徴量が定義される。
色彩を強度として3次元的な特徴量が定義される。



2) カオス論的な分析

→ リアプノフ・スペクトル

アトラクタ上に特徴量を定義する。

周波数分析に比較して自由度が高く、特徴量には様々な可能性が期待される。



tu



gi



wa



ko



e



n



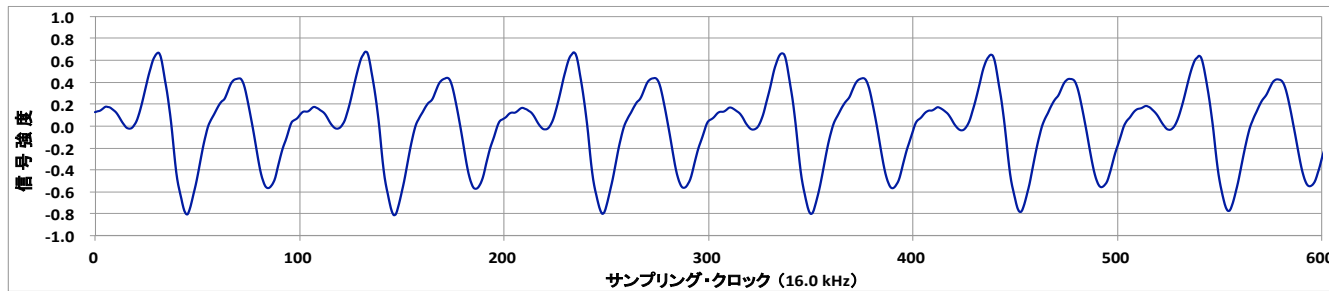
ji



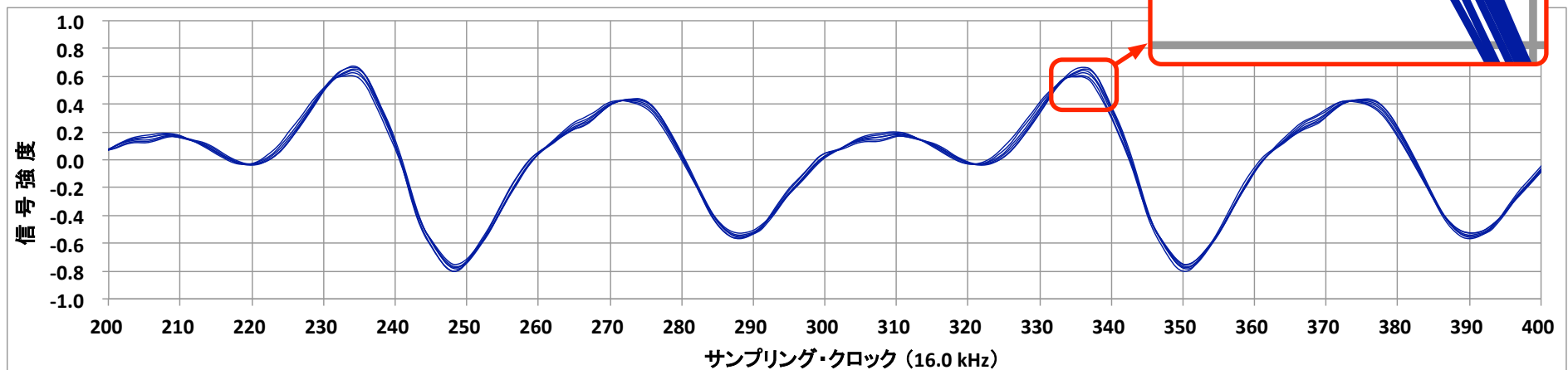
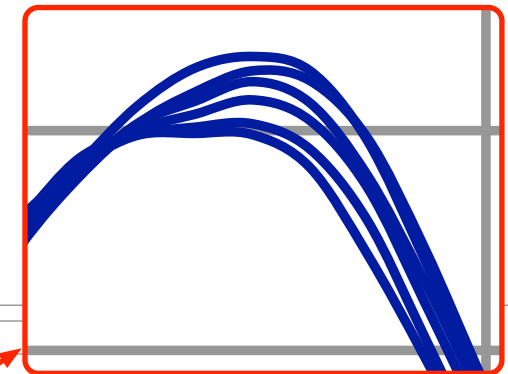
音声波形に見られる“ゆらぎ”

音声波形に見られる“ゆらぎ”

"o" 音の波形 (継続時間: 37.5ms)



波形の拡大



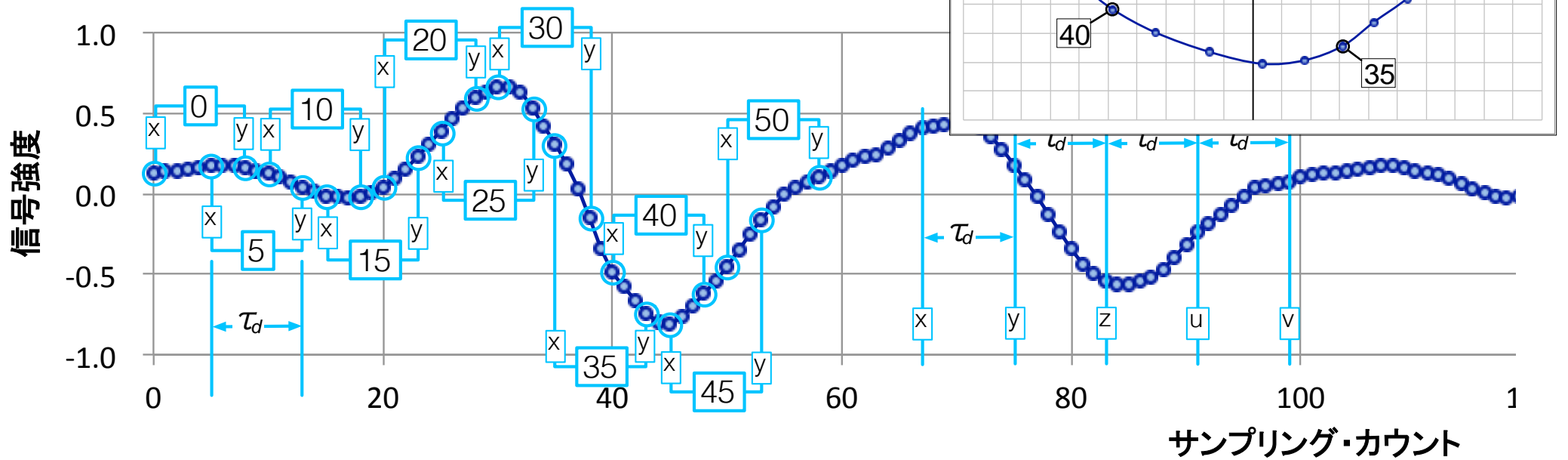
"o" 音の波形 (継続時間: 12.5ms) の重ね合わせ。

“声のゆらぎ” の定量化は・・・

“声のゆらぎ”の定量化は・・・



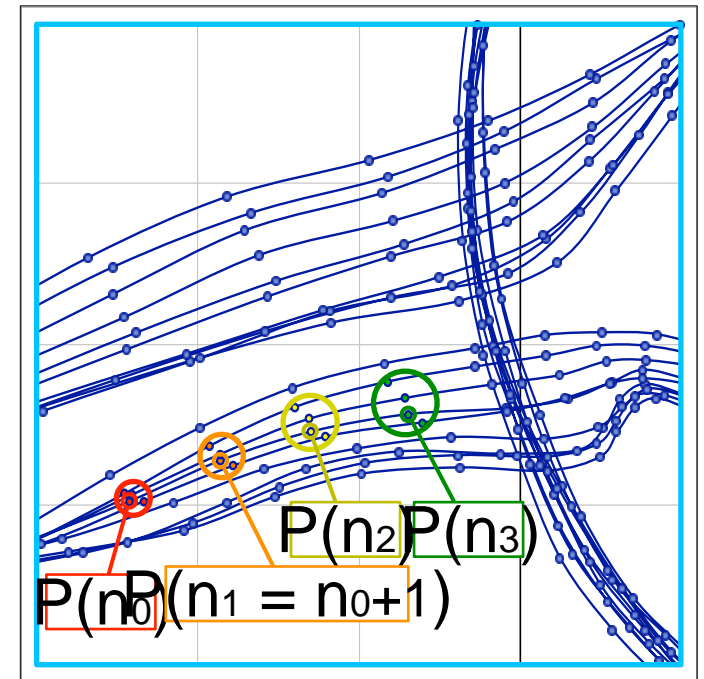
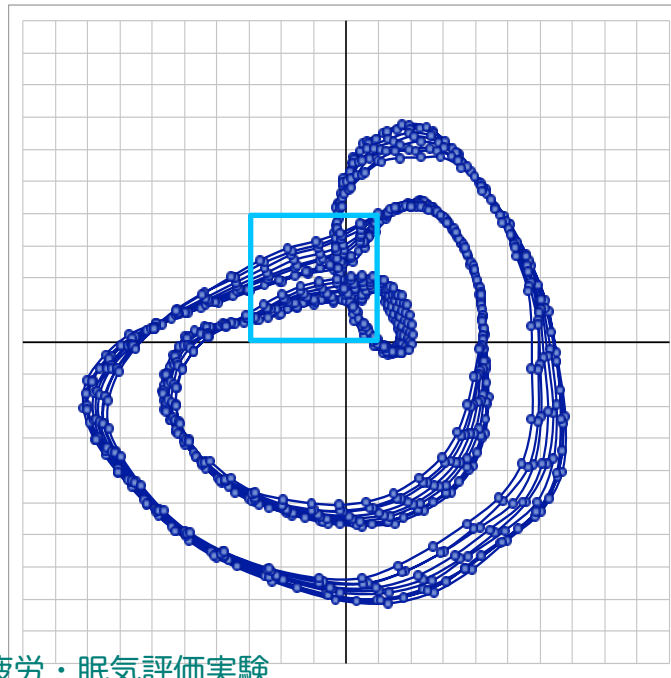
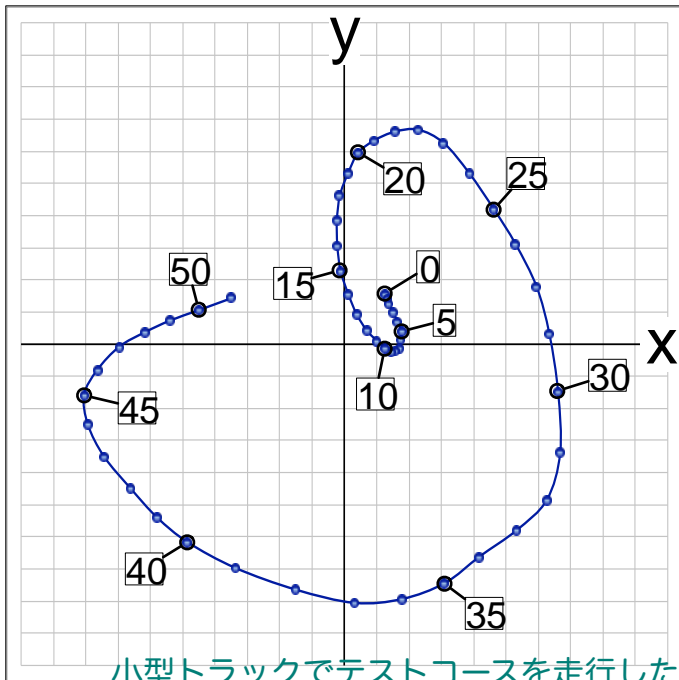
「ターケンスの埋め込み」



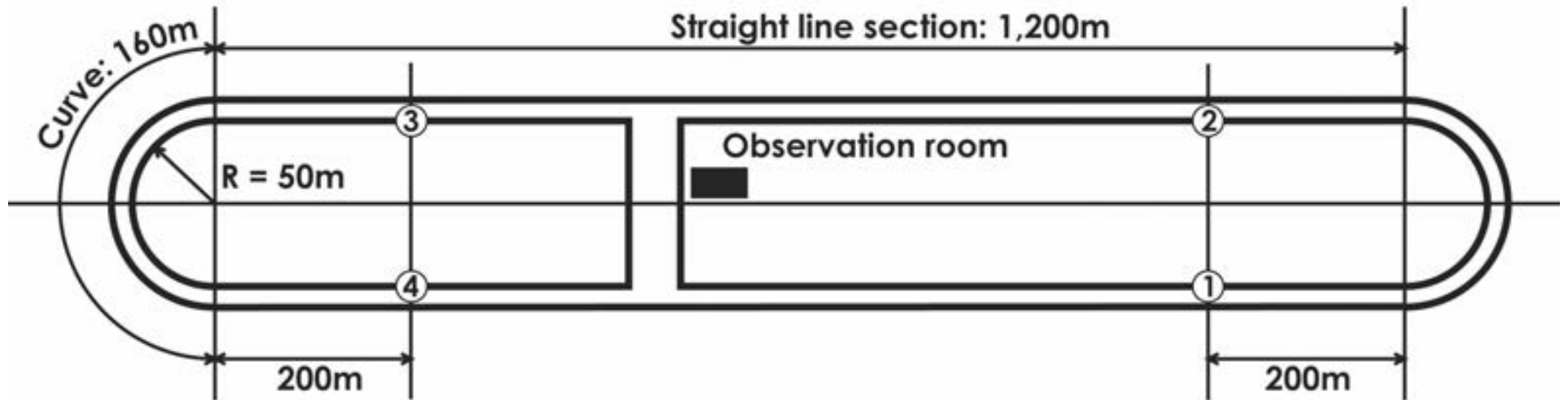
ターケンスの埋め込み定理

ターケンスの埋め込み定理

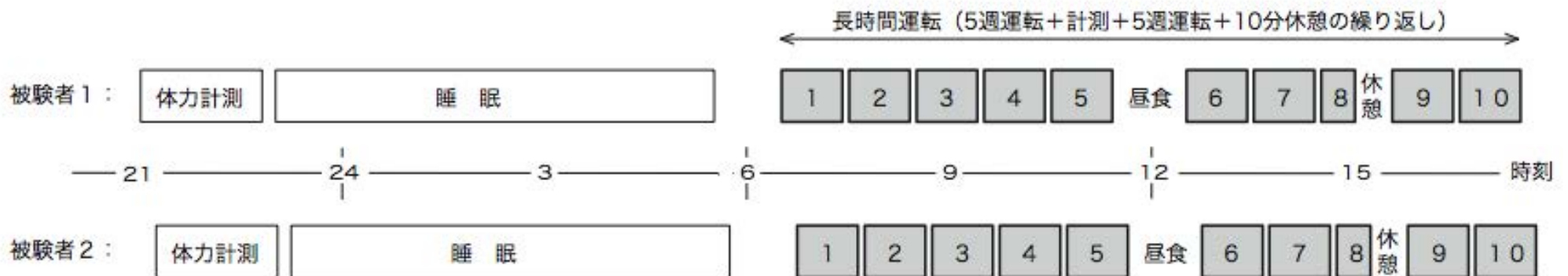
- 1) 数理モデルと時系列データとを橋渡しする。
- 2) 時間遅れ座標系への埋め込みは、力学系（数理モデル）の幾何学的構造を保存する。
- 3) ストレンジ・アトラクタが再構成されれば、時系列信号を生成させた原始機能に係る情報を取り出す事ができる。



小型トラックでテストコースを走行した疲労・眠気評価実験



テストコース（土木研究所 苫小牧 寒地試験道路）



タイムテーブル

小型トラックでテストコースを走行した疲労・眠気評価実験



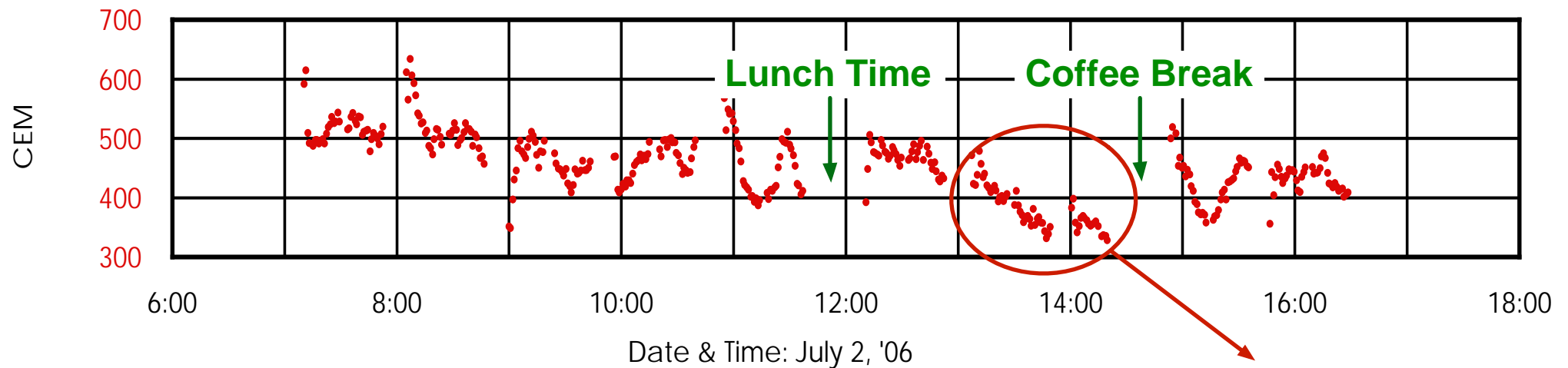
単調な運転業務による疲労・覚醒度評価実験の典型的な結果

単調な運転業務による疲労・覚醒度評価実験の典型的な結果

喚呼音声から算出された CEM 値の十分間平均値

Change of CEM values calculated from call-out voice

Experimental Result in Hokkaido, Japan

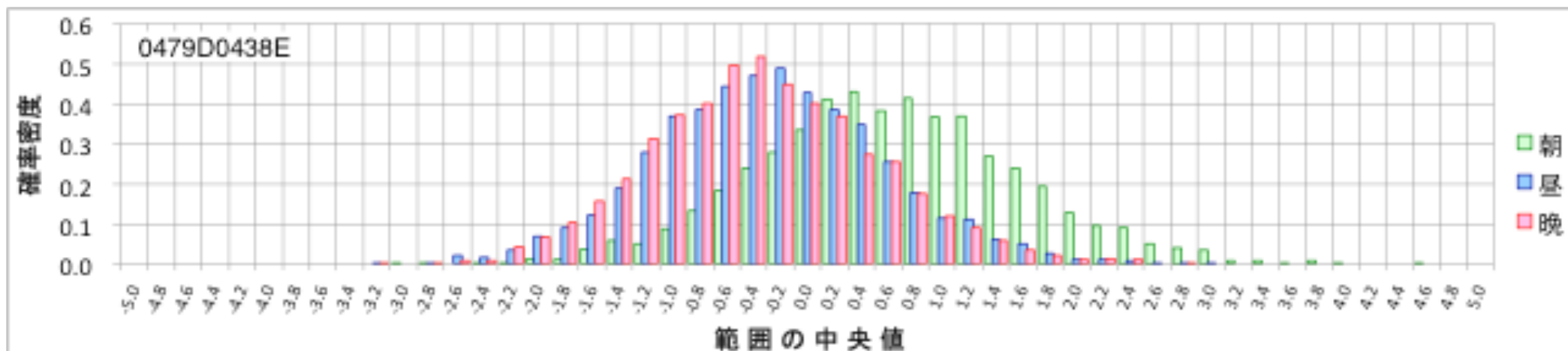


The driver must have been very sleepy.

急激に覚醒度が低下している可能性が高い。

音声資源コンソーシアム提供のデータによる覚醒度の評価

毎週1回、1日3回（10時、13時、18時）ある人が50の文章を音読したデータ（AWA-LTRコーパス）を処理した結果、以下の様な分布が観測されました。



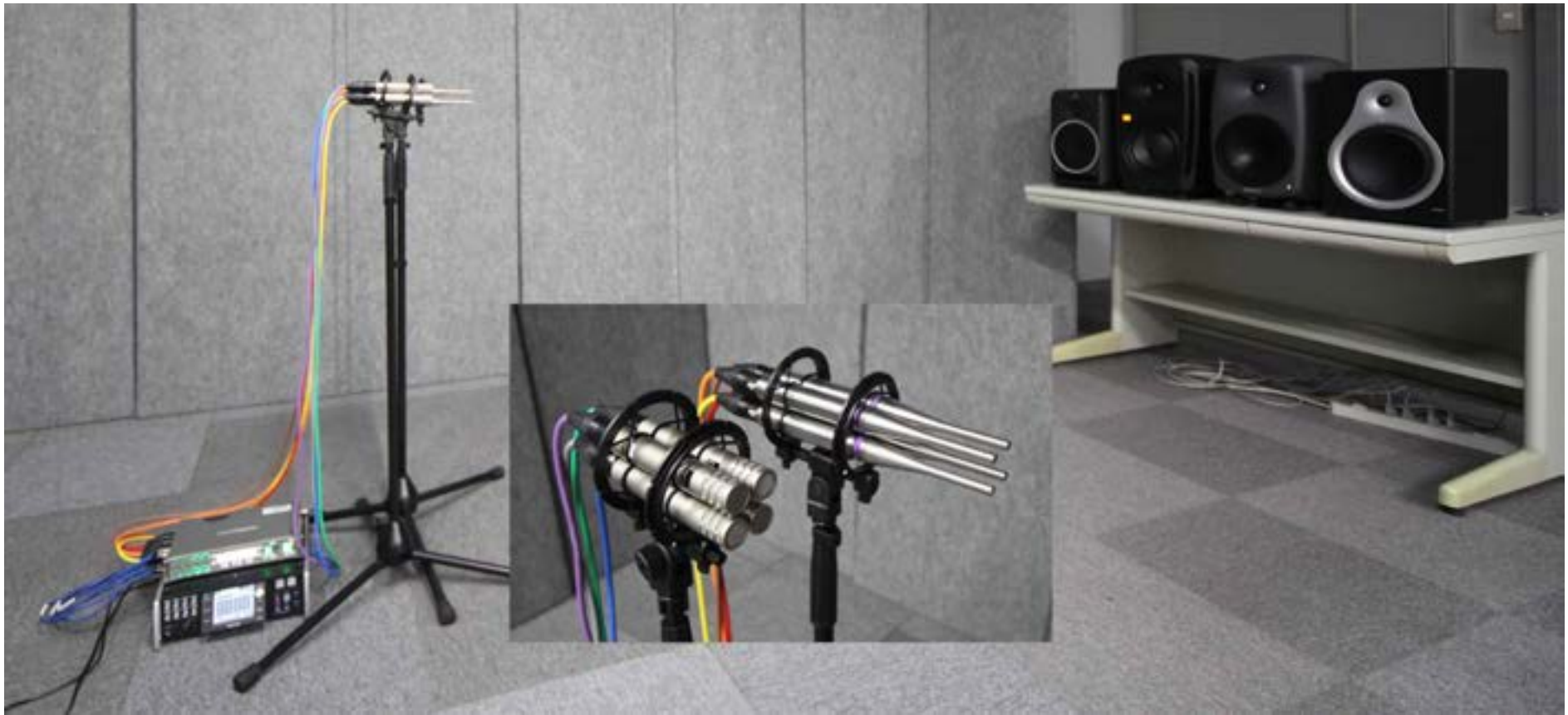
埋込み次元（D）を4～6，埋込み遅延時間（ τ ）を0.3～0.6ms程度にすれば，“平均的に”人間の覚醒度に相関すると考えられるCEM値が算出されます。

同じ音声を同時に複数のレコーダで録音した場合，周波数分析であれば，どの収録音声であっても殆ど変わらない結果が得られますが，カオス論的な手法で分析すれば，同じ型式のマイクロフォンと同じ型式のレコーダを使用した場合であっても・・・

複数のマイクロフォンで収録した音声の比較

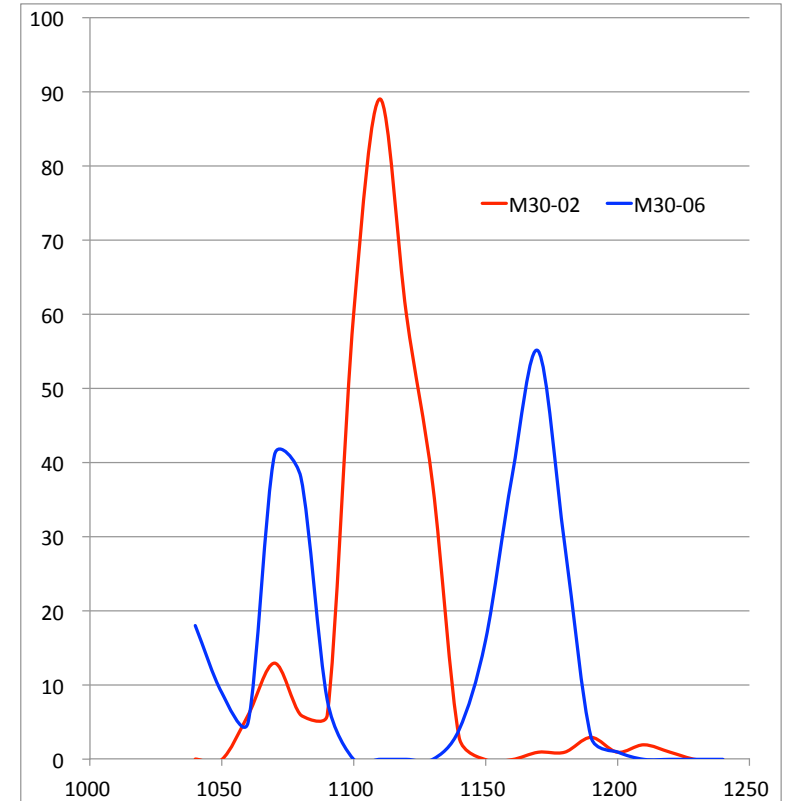
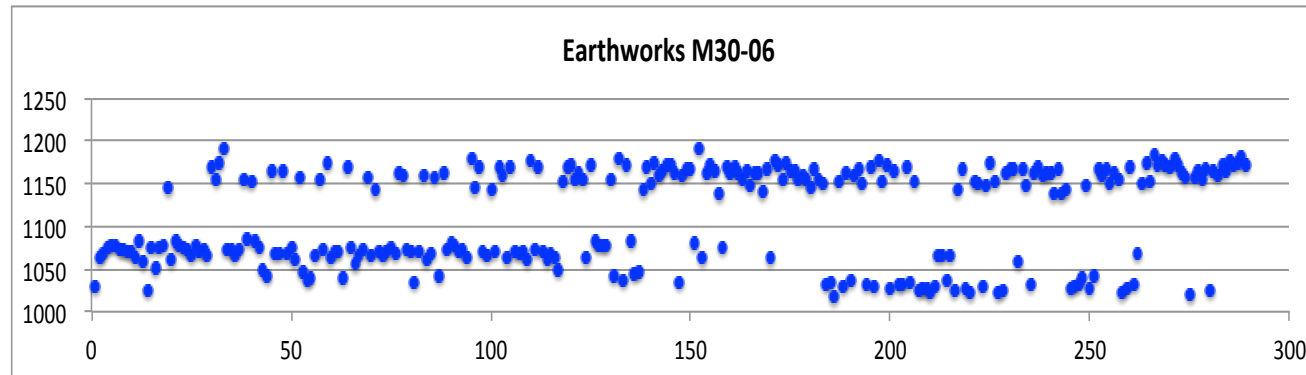
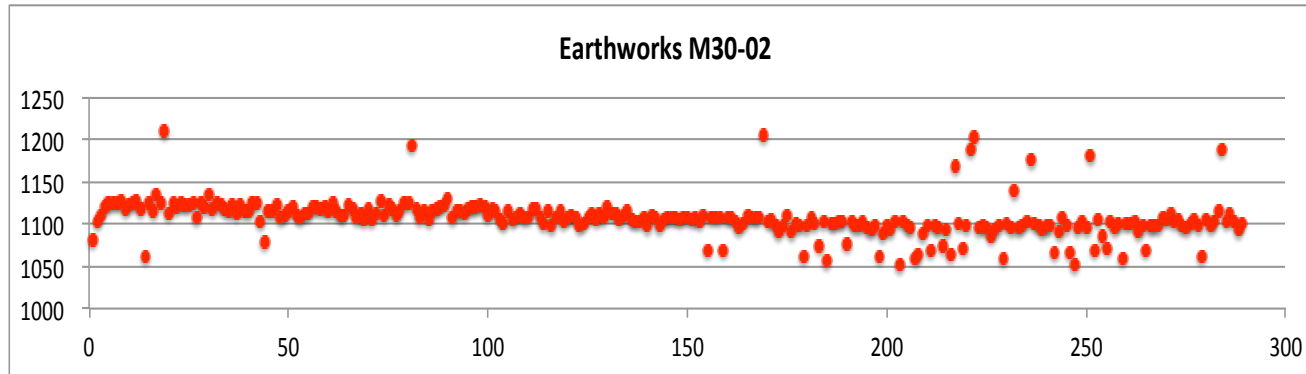
複数のマイクロフォンで収録した音声の比較

音声を再生するスピーカを決めて、特性を比較するマイクロフォンを正面に設置し、10秒程度の朗読音声を繰り返し再生し、個々の再収録音声から算出されるCEM値の平均値と標準偏差を比較する。（分布を比較する。）



予想外の実験結果

予想外の実験結果

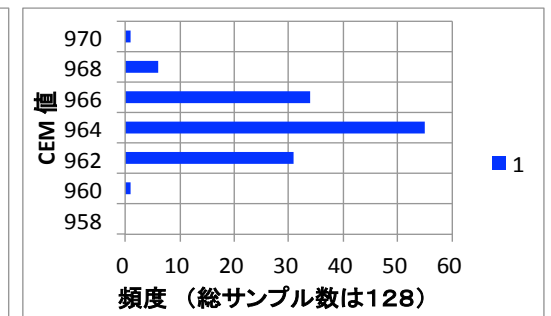
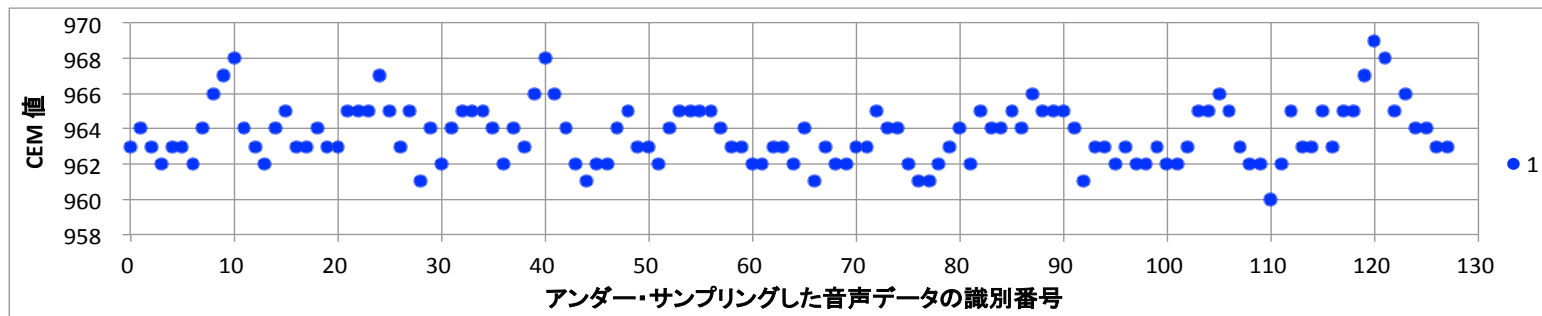


実験結果からは、10本のマイクロフォンを設置すれば、48.0kHz のサンプリング・レートとした場合、8本が赤い分布を示し、2本が青い分布を示す。

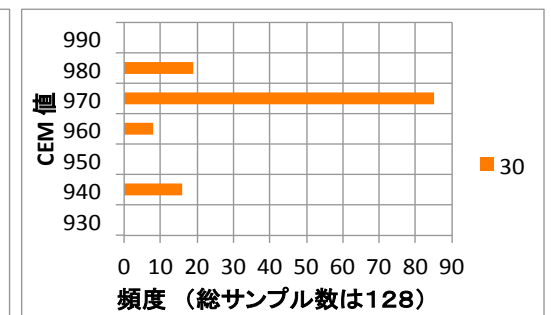
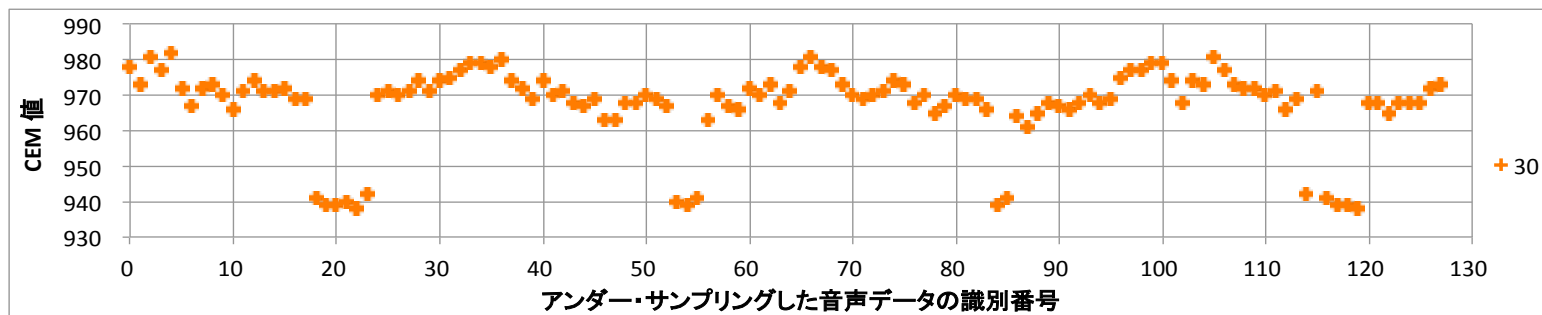
当所で収録した日本昔話朗読音声の処理結果

当所で収録した日本昔話朗読音声の処理結果

80~8.0kHzの帯域情報を確保しながら、768.0kHzのサンプリング周波数で収録した音声データから12.0kHzにアンダー・サンプリングして128種類の音声データを生成し、これらを再度48.0kHzにオーバー・サンプリングして処理すると、以下の様な結果を得る事ができます。



昔々 兎と亀の競争では 兎は麓に着く前に休んだので 亀に負けてしまいました



天女は川で水浴びをしていた時に 鋤と鍬を洗いに来た百姓に羽衣を隠されてしまいました

おわりに

おわりに

- 1) カオス論的な手法で音声进行分析すれば、主観を排除した発話者の覚醒度と高い相関を示す指数値を算出することが可能である。
- 2) 算出される指数値の分布は、カオス論的な信号処理アルゴリズムに依存し、一定の確率で双峰分布等の異常分布となる。
- 3) 音声から算出された指数値が正常に収束した値か？あるいは異常収束した値か？を区別することは、音声データをリサンプリング処理することにより可能だが、処理時間を実用的なものとするためには工夫が必要である。
- 4) カオス論的な信号処理手法を利用すれば、従来の周波数分析に比較して遙かに自由度が高い特徴量が定義可能であり、処理パラメータを調整することにより、様々な心身状態と相関する指数値が算出される可能性がある。