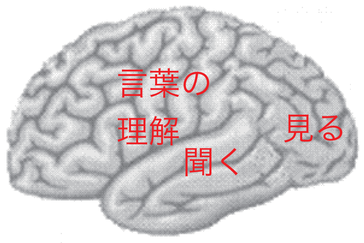
知能メディア工学講座

大西研究室

大西　昇　工藤博章　松本哲也

INFORMATION SCIENCE
NAGOYA UNIVERSITY
メディア科学専攻



概要

人間は，見る，聞く，言葉を理解するという重要な情報処理を行っています．当研究室では，人間での情報処理の仕組みの理解，人間の優れた機能に学ぶと同時にユニークなアイディアに基づく視聴覚情報処理の計算機による工学的な実現（人工知能），さらに，視聴覚機能を代替あるいは支援する福祉機器（福祉工学）の研究をしています．

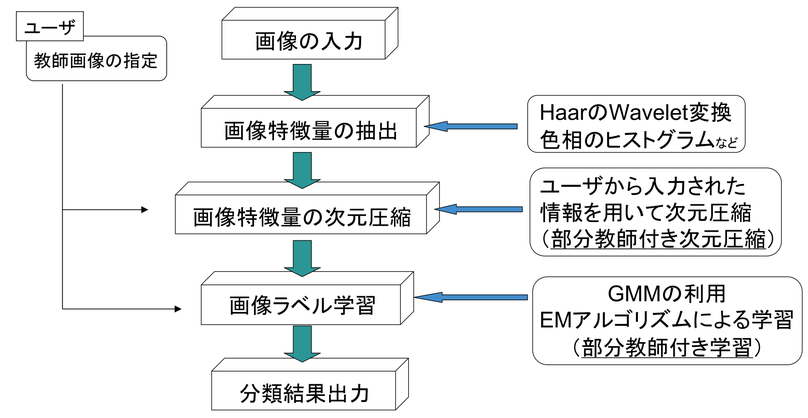
研究テーマ

- 映像評価のための視覚特性と生体信号の分析
- 人間の多様な感性に対応する柔軟な画像認識技術の実現
- 視覚障害者，聴覚障害者への感覚代行

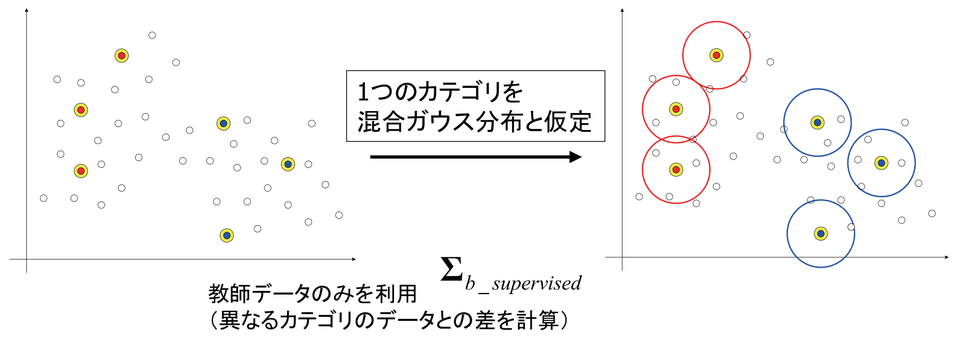
部分教師付次元圧縮による画像の半自動ラベリング

目的・システム概要

多量の画像の中から少数画像のみユーザがカテゴリを指定
与えられた情報を有効に利用し，残りを自動ラベリング



級間共分散行列の計算と射影ベクトルの計算



$$\Sigma_{b_semi-supervised} = \alpha \times \Sigma_{b_supervised} + (1-\alpha) \times \Sigma_{b_estimated}$$

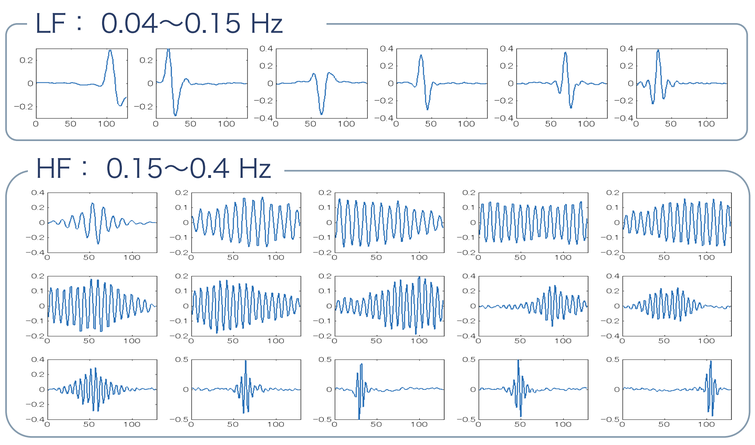
$$\Sigma_{b_semi-supervised}\mathbf{z} = \lambda \Sigma_{w_estimated}\mathbf{z}$$

$$\mathbf{s}_i = T^{\mathrm{T}}\mathbf{x}_i \qquad T^{\mathrm{T}} = (\mathbf{z}_1,\ldots,\mathbf{z}_n)^{\mathrm{T}}$$

λ： 固有値
$\mathbf{z}$： 固有ベクトル
$\mathbf{s}$： 圧縮後のデータ
$\mathbf{x}$： 入力データ

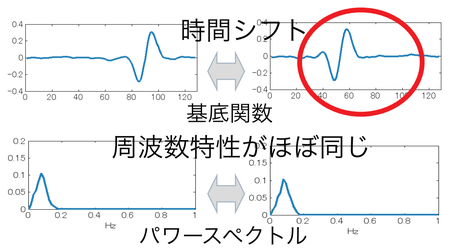
独立成分分析による心拍信号の分析

LF,HF帯域にピーク周波数をもつ基底

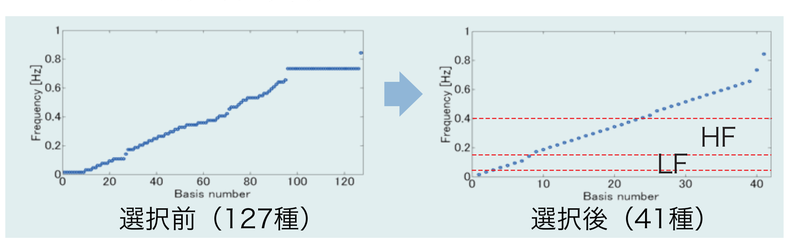


縦軸：振幅，横軸：データ点（1～128点），平均0，大きさ1

パワースペクトルによる基底の選択



時間シフトの関係にある基底は同じ属性とみなす
↓
フレーム中心にパワーが集中する基底を選択

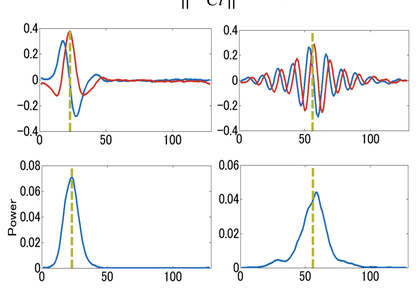


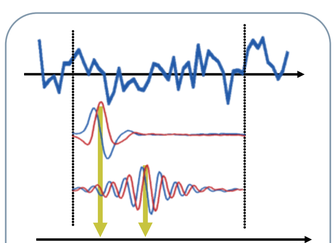
選択前（127種）　選択後（41種）

局在時間の算出

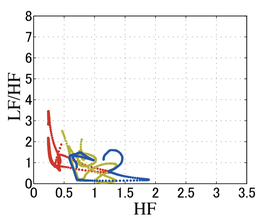
フレーム中で各基底が局在する時間はさまざま
➡基底が切り出す時間をパワーの期待値で定める

$$t_{center} = \frac{1}{\|\mathbf{a}_{Ci}\|^2}\sum_{t=1}^{64} t\,|a_{Ci}(t)|^2$$

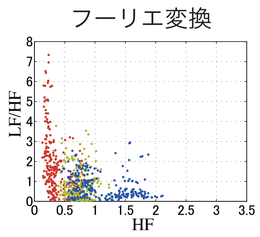




同じフレームでも基底ごとに切り出し時間を変える



フーリエ変換



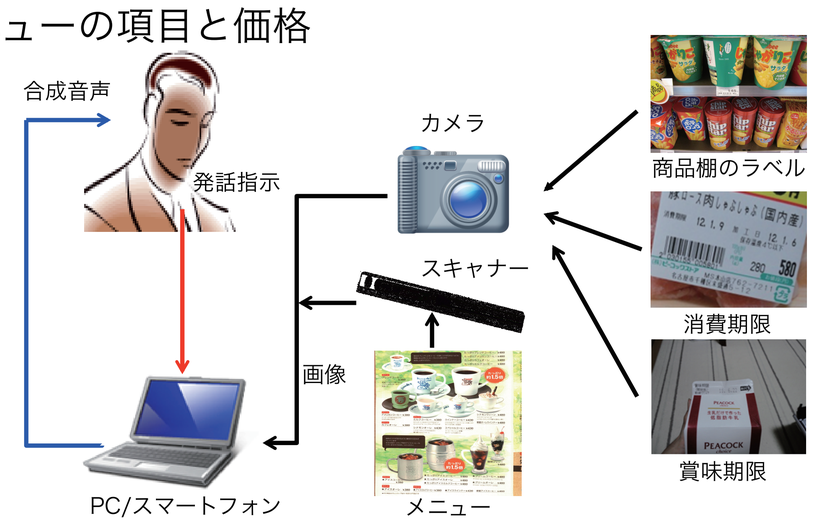
独立成分分析

視覚障害者向け商品情報取得システム

システムの概要

PC/スマートフォンにて
カメラやスキャナーで取り込まれた画像中の文字領域の抽出
OCRで文字認識後に合成音声にて読み上げ

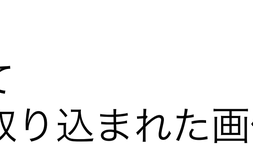
消費棚にある商品の名前と価格
食品の消費期限
メニューの項目と価格



方法

文字領域の特定の必要性

画像に直接OCRを適用すると
不十分な結果



OCRの出力

文字領域の抽出処理

1. 画像を2値化し，画素を囲む矩形領域を算出
2. 大きさとアスペクト比から文字候補領域を選択
3. 得られた矩形領域に対してSVMで
「円」，「￥」，「限」
の文字を特定

大きさの変化や回転に対する処理

矩形処理の後，大きさを正規化
SVMの学習データに回転を含めて学習

ラベルを抽出する際に，バーコード検出
ラベル内の記載配置モデルの利用

結果

ラベル認識

適合率	再現率	F値
0.83 (50/60)	0.76 (50/66)	0.78

日付認識

生鮮食品	飲料パック
0.8 (84/105)	0.4 (22/55)

メニュー認識

	完全認識		一部誤認識を許容	
重なり無し	0.83	(58/70)	0.91	(64/70)
重なり有り	0.16	(12/75)	0.55	(41/75)