

男性両声類の女声らしさに関わ る特徴量の分析

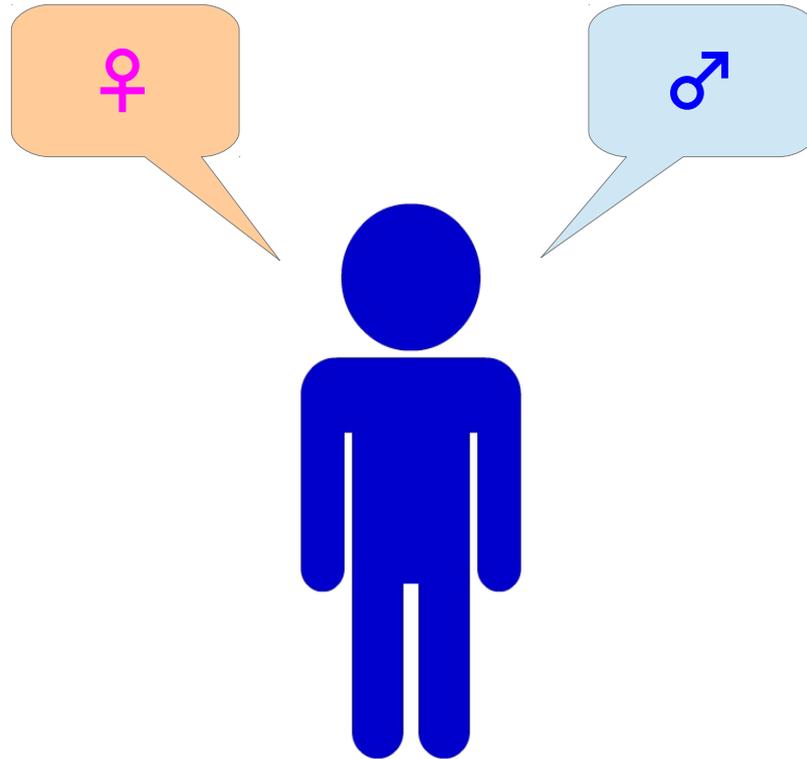
日本大学 文理学部 情報科学科
B4 長谷川 翔太

目次

- 概要
 - 研究背景
 - 研究目的
 - 用意したデータについて
- 「声」の仕組み
- 音響特徴量の抽出
- 分類実験
- 考察・追加実験
- まとめ

両声類って？

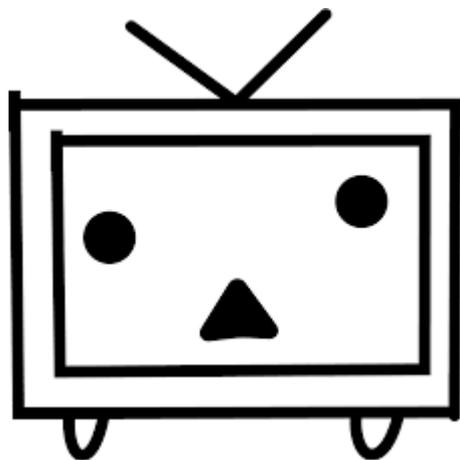
例)



定義：異性の声を使いこなす人達の呼称。

研究背景

2007年頃～



ニコニコ動画



こえ部

エンターテイメント性が強い!!

研究背景

2007年頃～

「両声類」に焦点を当てた
研究はやられていない!!

ニコ

四郎

え部

エンターテイメント性が強い!!

研究背景

関連研究

- 女性と判定された性同一性障害者(MtF)の声の基本周波数

櫻庭 京子他(2003)

両声類≠MtF

- MtF
 - 完全な女性になりたい
- 両声類
 - 男性は男性としてのアイデンティティを保持したまま**特技として**女声を出したい

研究目的

両声類が目指す目標

- 多くの人から上手いと評価されたい

→上手いと評価される人の特徴が分かると嬉しい!!

男性両声類の女声らしさに関わる音響特徴量について調べる。

使用するデータ

両声類のジャンル

- アニメ声 
- ロリータ声 
- ナチュラル声 
- など

ニコニコ生放送ではジャンル別のコンテストも開催される

← これに焦点を当てる

ニコニコ生放送で行われたナチュラル声コンテストの音声を用意

- 男性両声類37人分
- 音声データは計534個(1~10秒程度)
- 音声はすべて会話文

主観的評価

話者ごとに各視聴者が1人1票投票できるアンケート形式

- 1) ナチュラル声に聞こえる
- 2) 他の女声に聞こえる
- 3) 男声に聞こえる

アンケート結果は百分率で表示される

アンケート結果の「1) ナチュラル声に聞こえる」の割合から音声を以下に分類

0～25%	→	very low
25～50%	→	low
50～75%	→	high
75～100%	→	very high

4段階

研究の流れ

男性両声類のナチュラル声の音声と評価

openSMILEで音声から音響特徴量を抽出



評価をラベリング

wekaで属性選択&分類実験

研究の流れ

男性両声類のナチュラル声の音声と評価

openSMILEで音声から音響特徴量を抽出

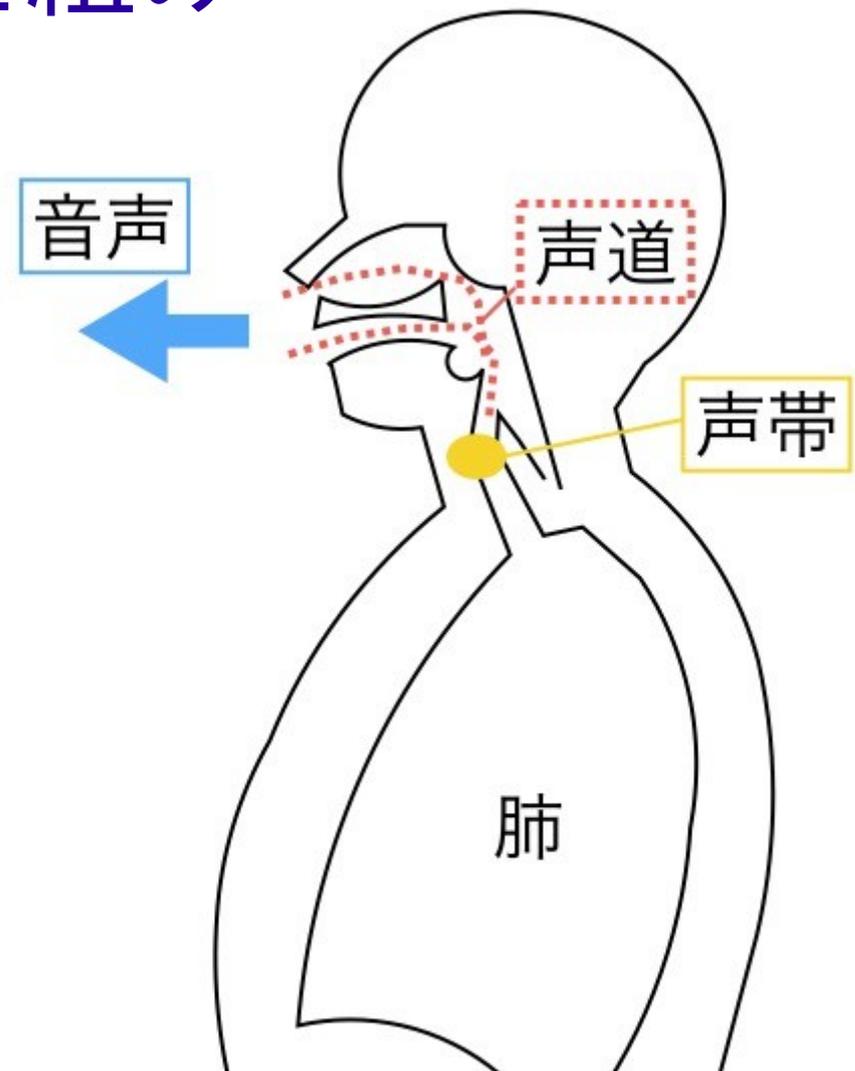
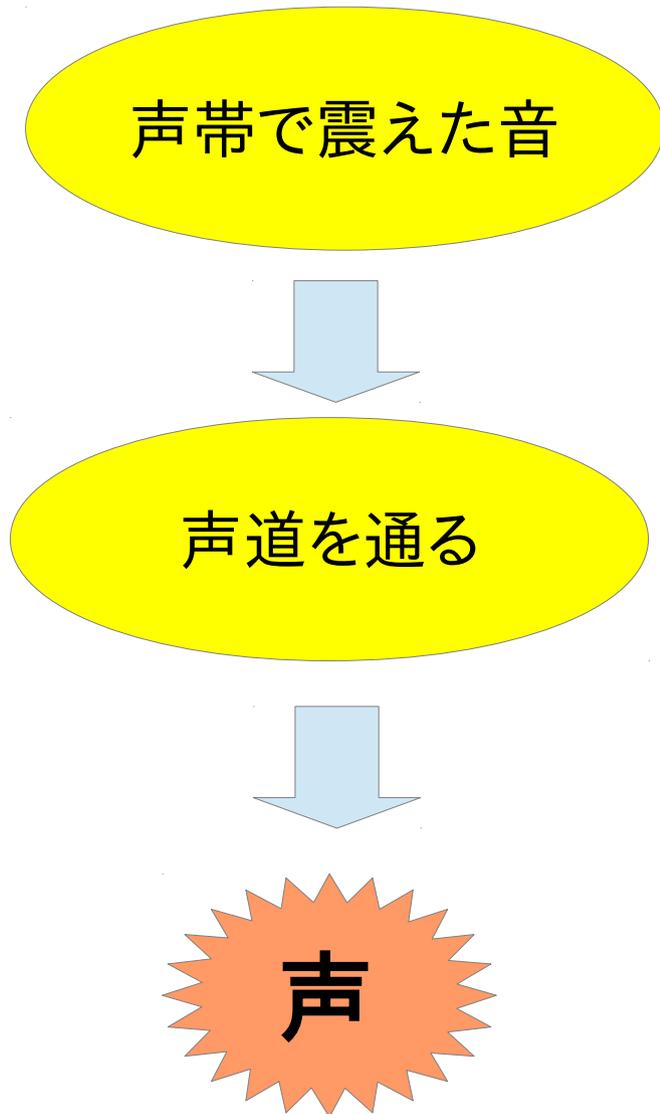


評価をラベリング

wekaで属性選択と分類実験

分類器ごとの分類率や
分類に用いた音響特徴量などについて考察

声の仕組み



声の仕組み

声帯特性 → 声の高さ

声道特性 → 声の音色

声道特性の変化

→ 周波数の強調部分の変化

音声

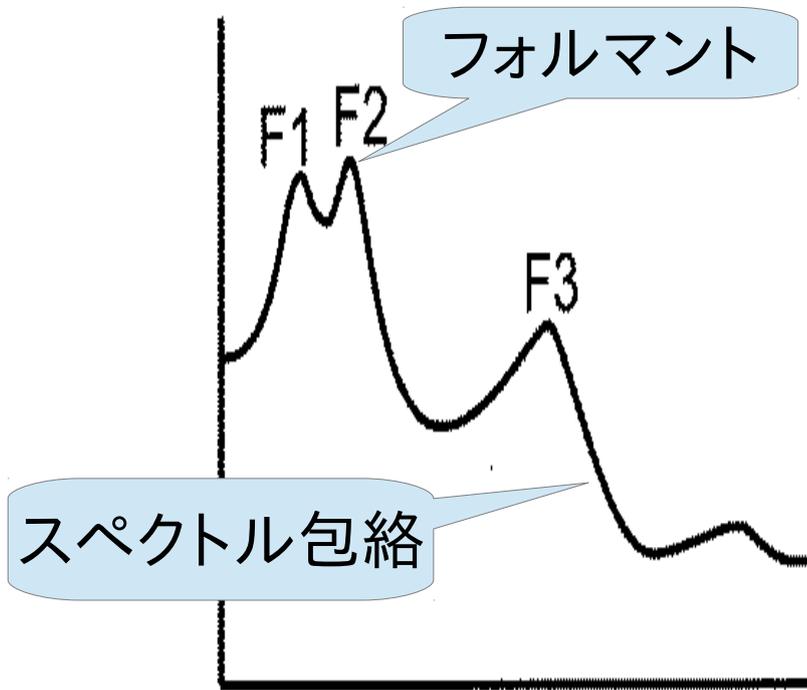
声道

声帯

肺

(a)

対数振幅スペクトル

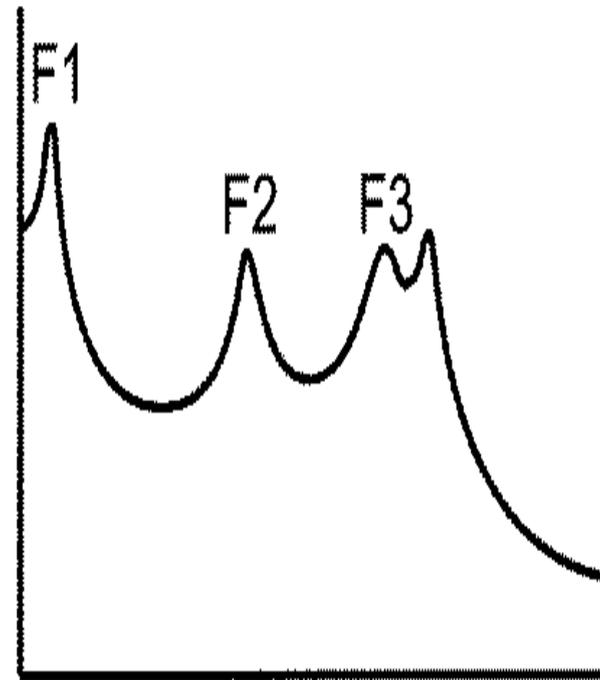


あ

周波数(Hz)

(b)

対数振幅スペクトル

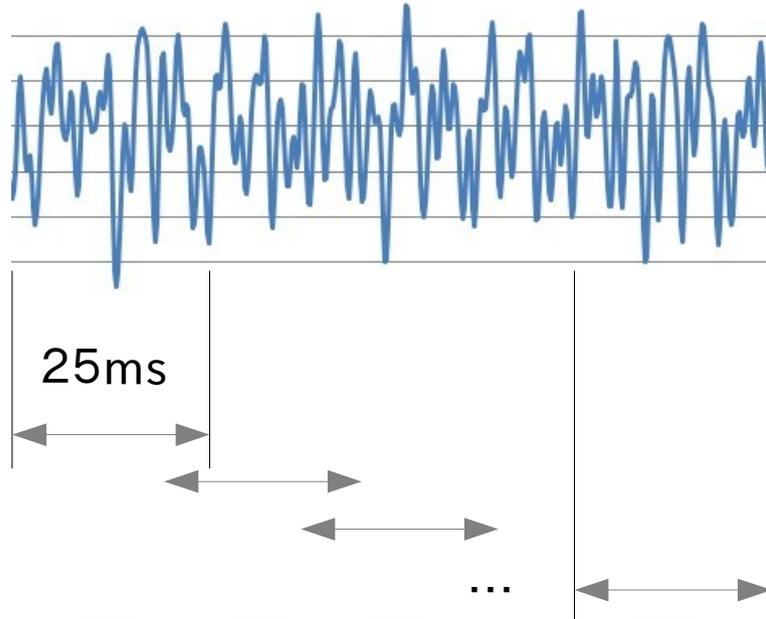


い

周波数(Hz)

音響特徴量

openSMILE



+ ラベル

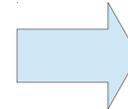
- ・very low
- ・low
- ・high
- ・very high

19種類

声の高さ



...



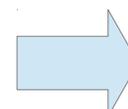
最大 最小 平均 ...



口の形に関する特徴量



...



計 52

988個の音響特徴量を得た

分類

機械学習統合環境「Weka」

- データの前処理
- 多種多様な学習アルゴリズム
- 評価

属性選択

特徴集合のうち意味のある部分集合だけを選択する手法

結果

分類器	分類率	適合率	再現率
BayesNet(18)	0.618	0.617	0.618
NaiveBayes(26)	0.742	0.741	0.742
RBFNetwork(15)	0.702	0.701	0.702
IBk(30)	0.880	0.882	0.880
J48(14)	0.642	0.644	0.642

結果

分類器	分類率	適合率	再現率
BayesNet(18)	0.618	0.617	0.618
NaiveBayes(26)	0.570	0.570	0.570
RBFNetwork(15)	0.702	0.702	0.702
IBk(30)	0.880	0.882	0.880
J48(14)	0.642	0.644	0.642

属性選択で選択された
音響特徴量の数

結果

分類器	分類率	適合率	再現率
BayesNet(18)	0.618	0.617	0.618
Naïve Bayes	0.742	0.741	0.742
Decision Tree	0.702	0.701	0.702
IBk(50)	0.880	0.882	0.880
J48(14)	0.642	0.644	0.642

正しく分類できた割合

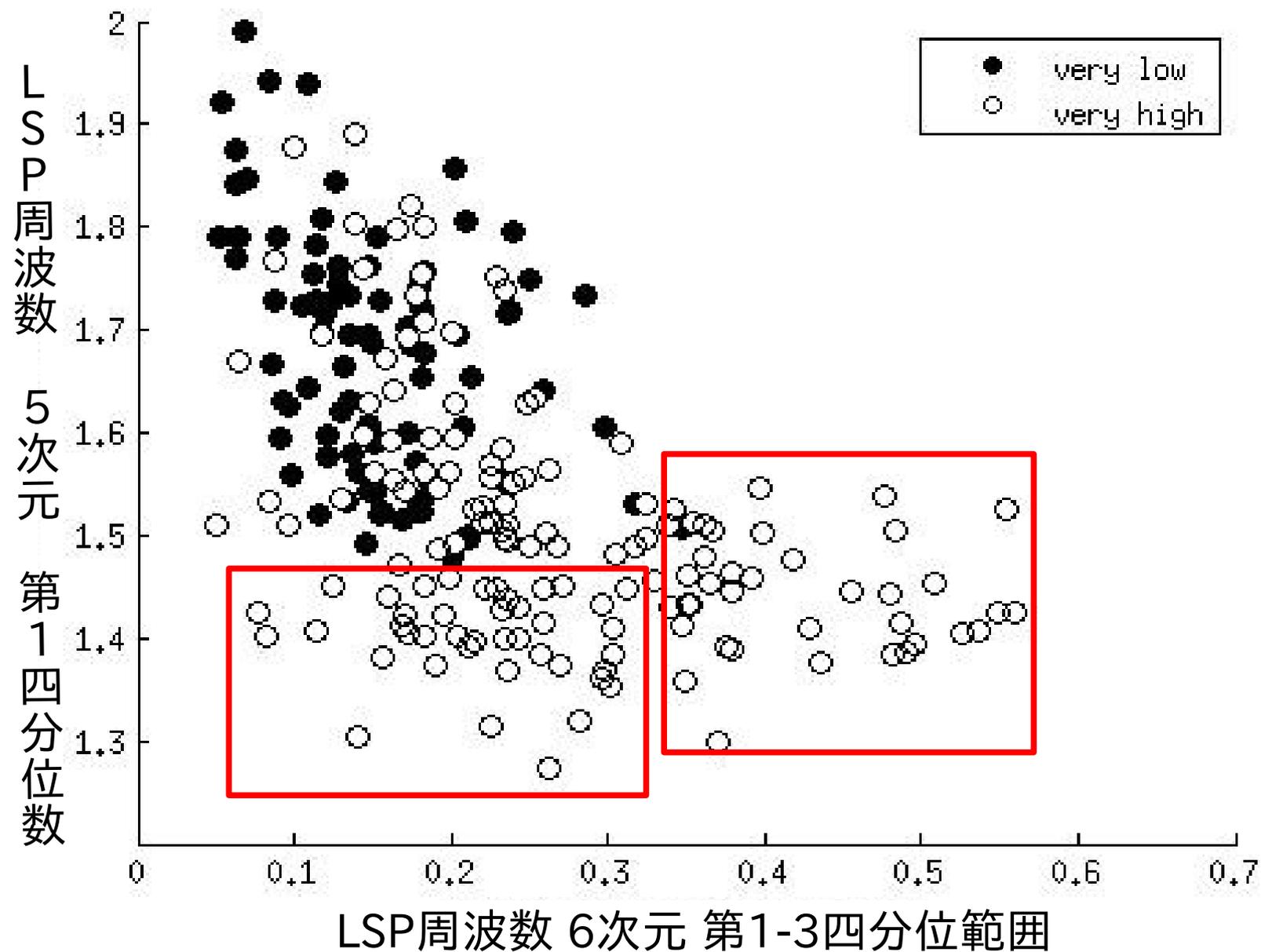
結果

分類器	分類率	適合率	再現率
BayesNet(18)	0.618	0.617	0.618
NaiveBayes(26)	0.742	0.741	0.742
RBFNetwork(15)	0.702	0.701	0.702
IBk(30)	0.880	0.882	0.880
J48(14)	0.642	0.644	0.642

結果

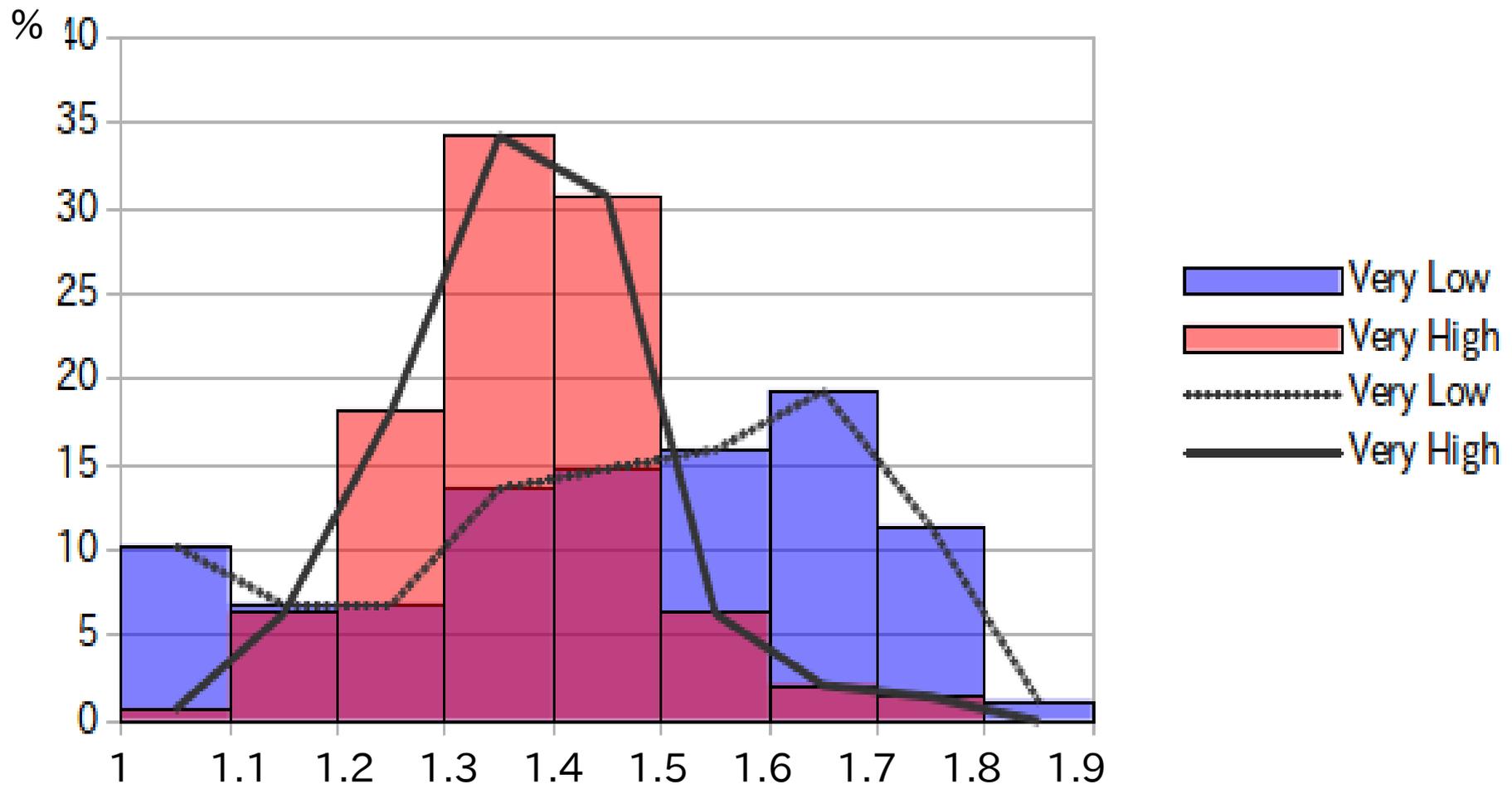
- 属性選択の結果、声道特性に関する音響特徴量が多く選ばれた。
 - MFCC
 - LSP周波数

識別に有用だった特徴量の分布



識別に有用だった特徴量

LSP Frequency 4次元 第3四分位数



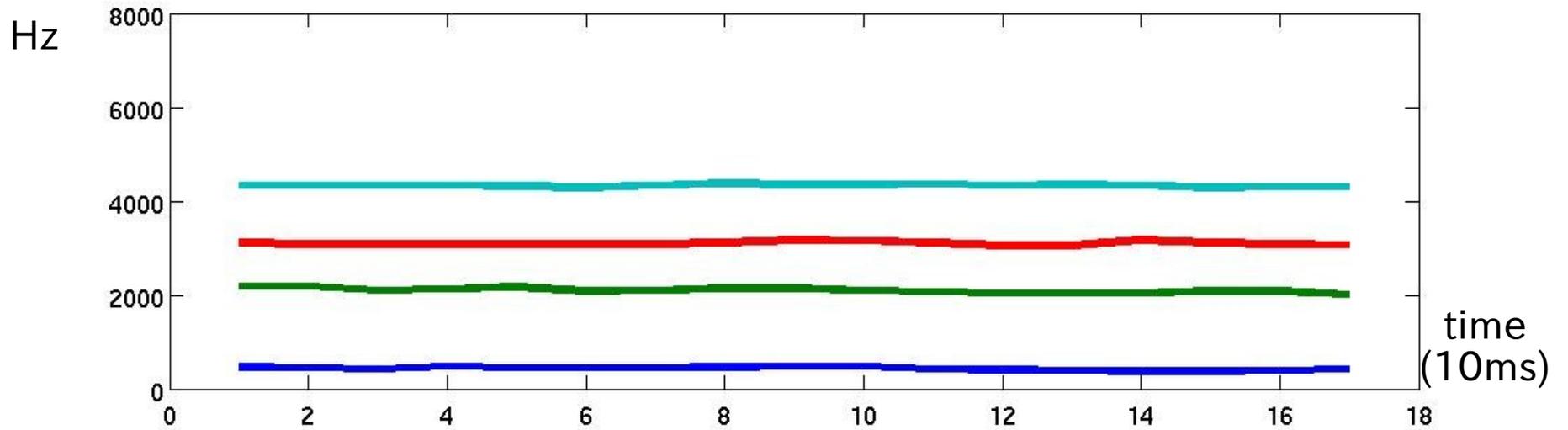
結果

- 属性選択の結果、声道特性に関する音響特徴量が多く選ばれた。
 - MFCC
 - LSP周波数
- スペクトル包絡の情報が重要である。

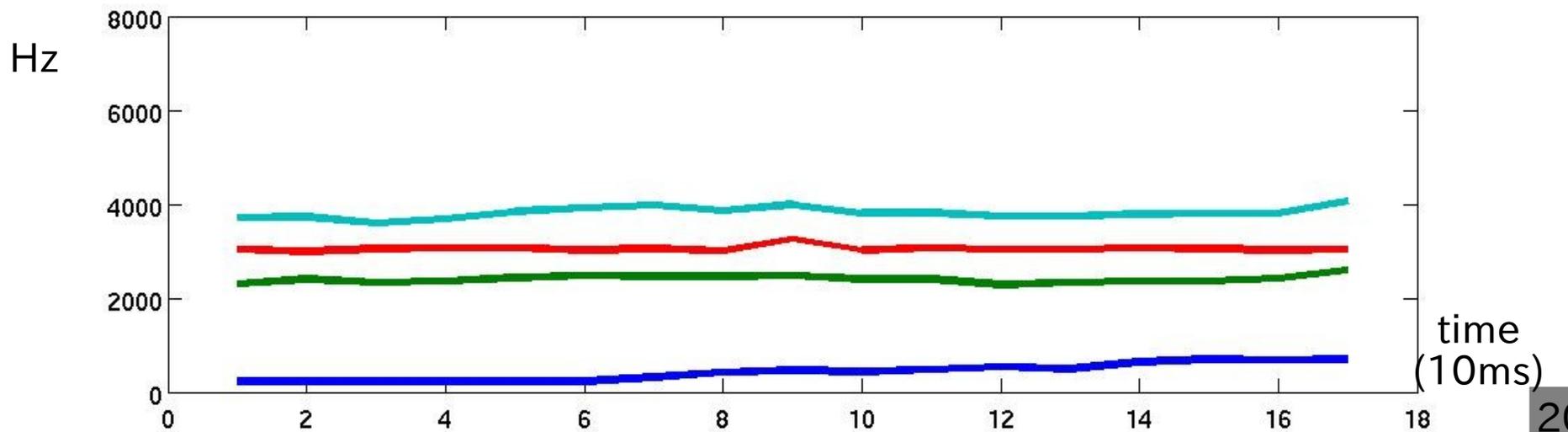
高い評価の話者と低い評価の話者の
フォルマントを分析した。

/e/のフォルマント

very high の話者



very low の話者



フォルマント周波数

- 平均

	F1(Hz)	F2(Hz)	F3(Hz)	F4(Hz)
高評価 (very high) の音声	429	2098	3100	4329
低評価 (very low) の音声	436	2414	3054	3823

- 標準偏差

	F1	F2	F3	F4
高評価 (very high) の音声	36	42	35	23
低評価 (very low) の音声	174	74	58	118

実験結果から

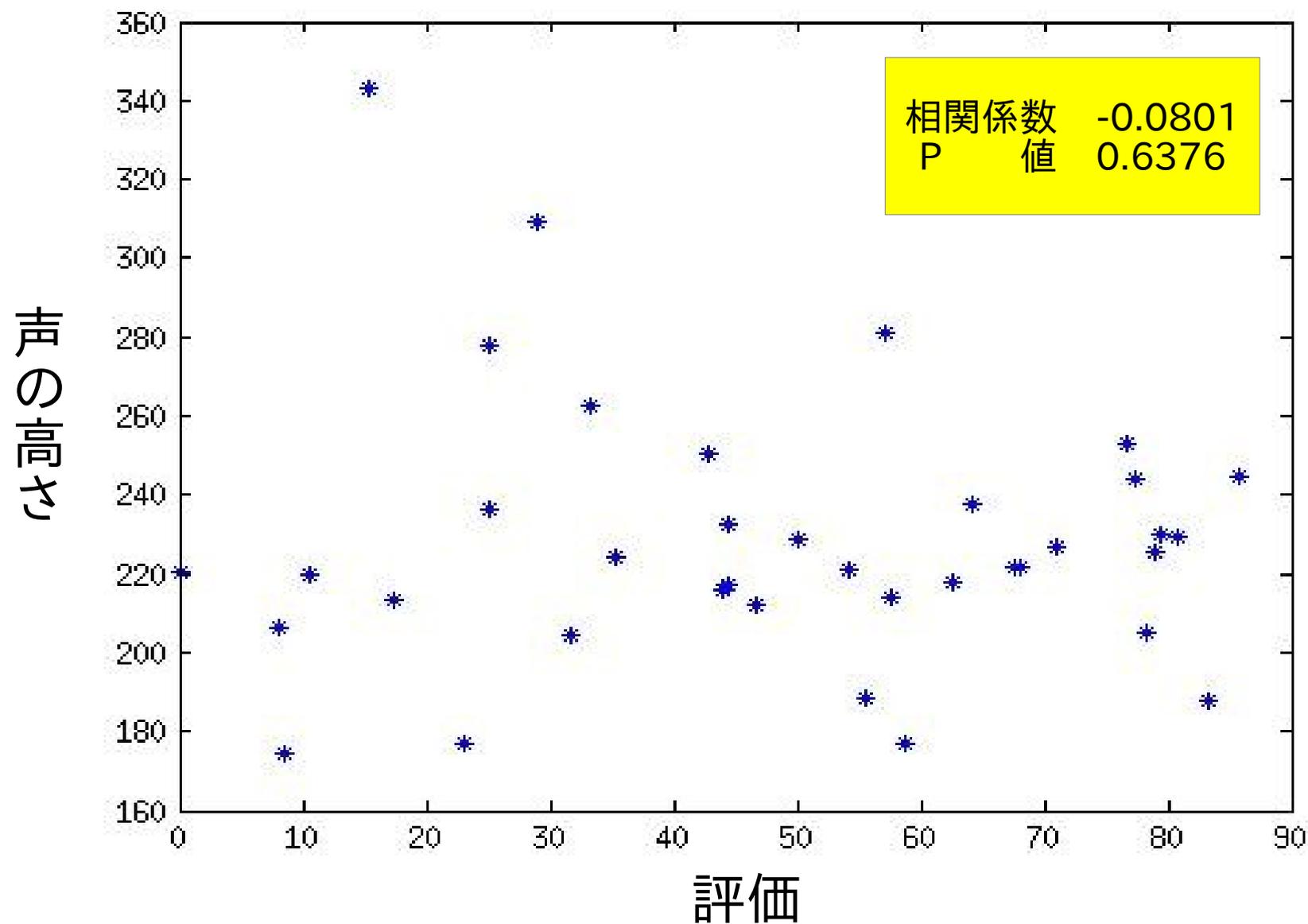
- 声道特性に関わるパラメータがナチュラル声コンテストの評価に寄与していることが分かった。
- しかし、声帯特性に関する特徴量は属性選択であまり選ばれなかった。

関連研究では声帯特性に関する記述が多かった。

- MtFの平均基本周波数 200Hz前後
- 女性の平均基本周波数 200~260Hz

→話者ごとの声の高さについても分析した。

話者ごとの平均基本周波数 (F0) の分布



まとめ

男性両声類の女声らしさについて調査した。

- 分類実験の結果、声道特性に関わる音響特徴量が多く選ばれた。
- 平均F0は高い評価を得た話者は200~240Hzに多かったが相関関係はなかった。

これからの課題

- 今回比較できなかった他の音響特徴量についての分析
- アニメ声やロリータ声に焦点を当てた分析