

場の雰囲気を感じて会話するシステム

2019. 9. 13

ロボティクス&デザイン工学部 システムデザイン工学科

教授 脇田 由実

1人暮らし高齢者の話し相手

家族をつなぐペットとしての役割

家族の一員としての情報検索システム

- 継続して使用されない．飽きられる
 - － 音声認識の誤りによりタスクを遂行できない
 - － ワンパターン．できる話しに限界あり．
 - － 疲れる、イライラする



“おもちゃ”以上の役割（見守る、家族をつなぐなど）
には至らないのが現状

言葉がわかる（音声認識を搭載）システムが
増加してきた



人同士の会話から、場の雰囲気や話者の状況を判断

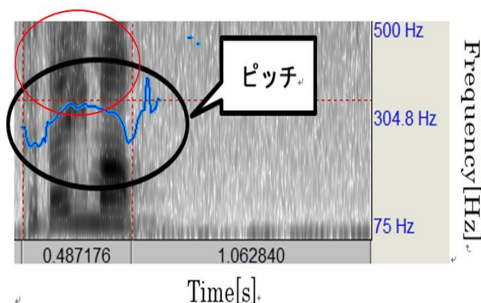
雰囲気や気持ちを理解した上で、
自然なタイミングで、話題を提供することで
会話を支援

話題提供後の雰囲気をセンシングし続ける

実環境にロバストな複数の特徴量から気分を推定

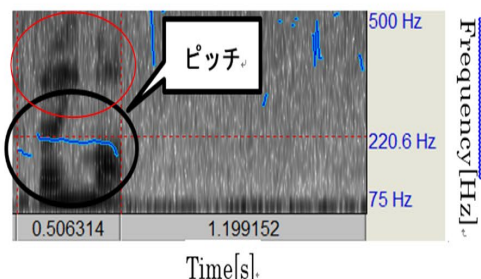
音響信号処理による 声の特徴分析

気分がいい「ただいま！」



- ・イントネーションに高低あり
発声末が上がり調子
- ・高次ホルマントも明確

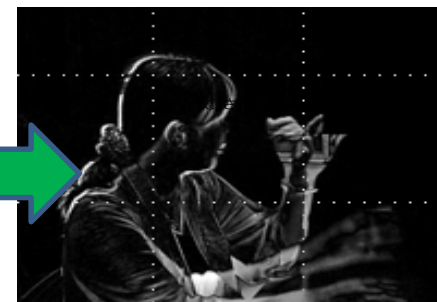
気分が悪い「ただいま..」



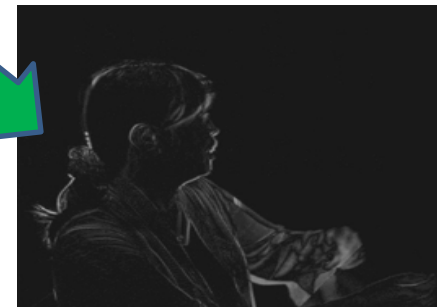
- ・イントネーションが平坦
発声末が下がり調子
- ・高次ホルマントが不明確

画像解析による 身振り手振り分析

積極的な話し手

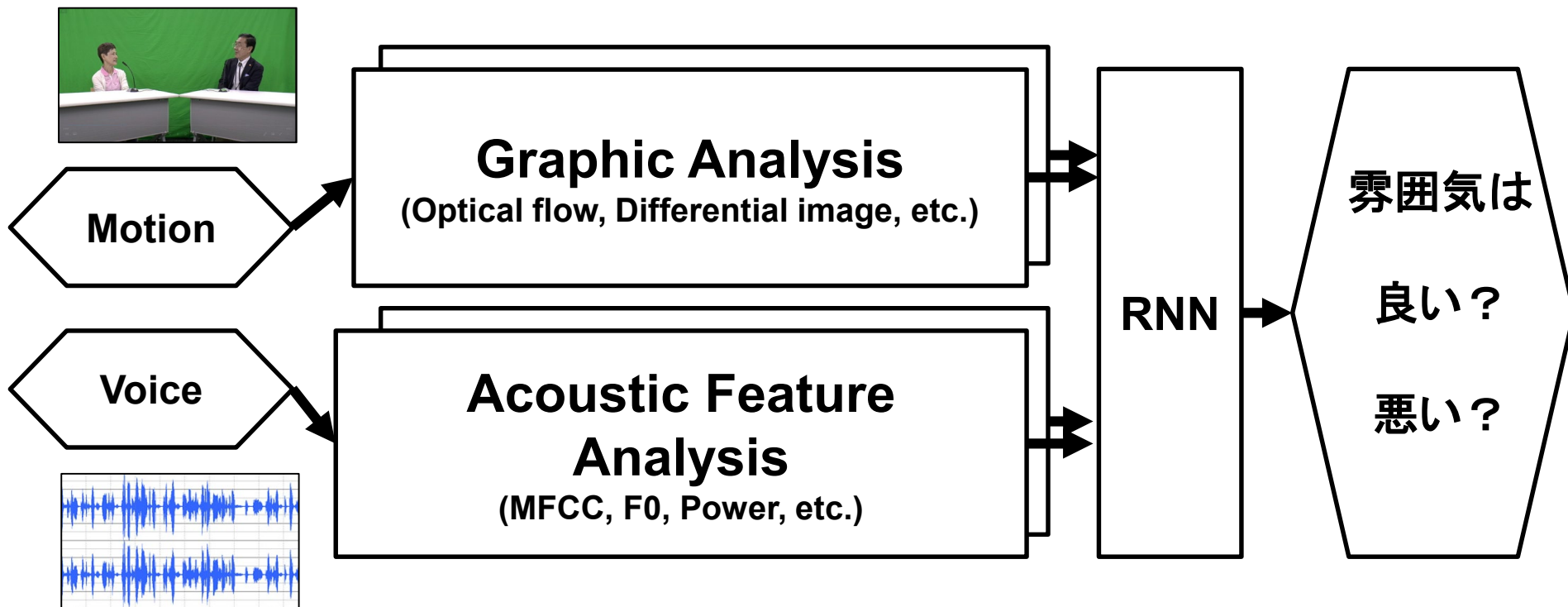


静かな聞き手



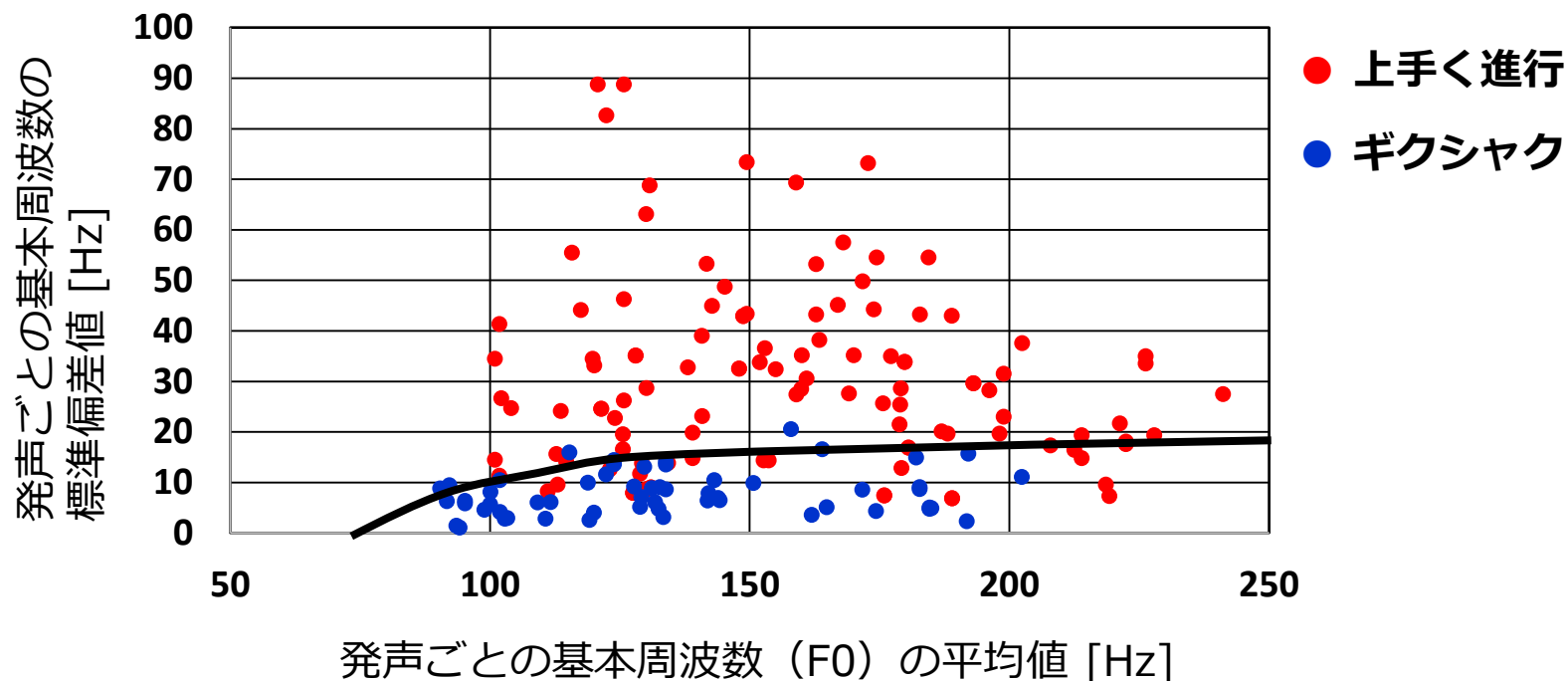
場の雰囲気 considering 話をかける





場の雰囲気推定システムのブロック図

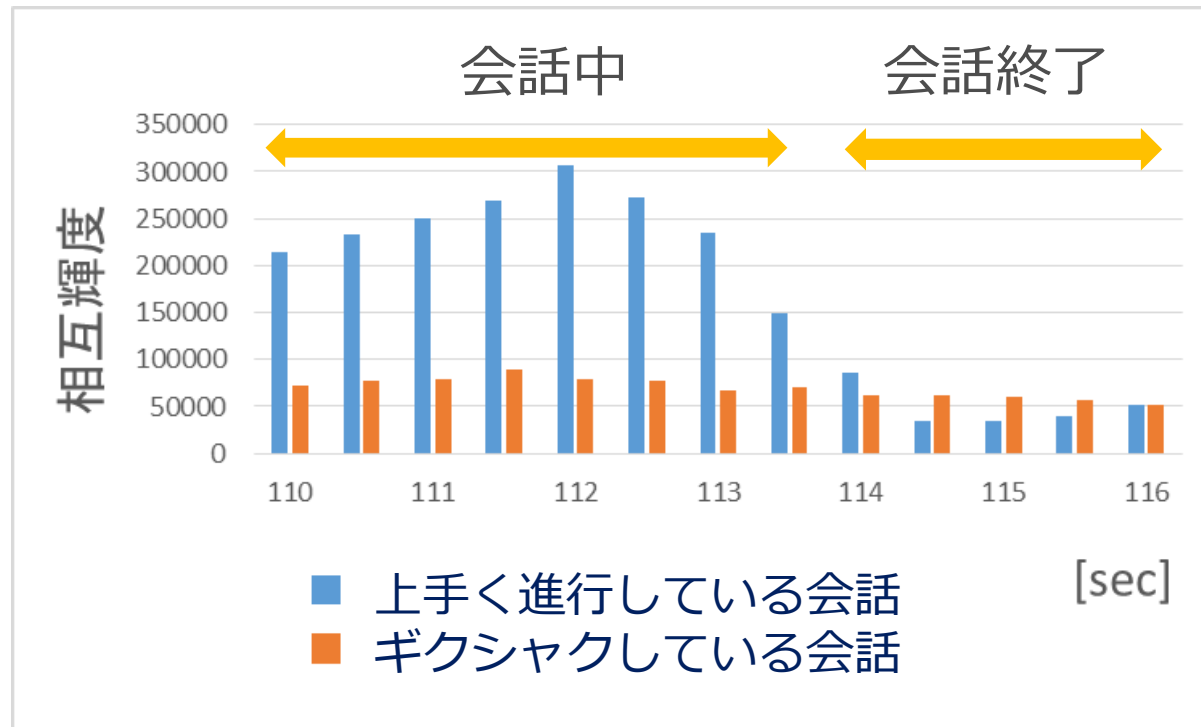
会話が上手く進んでいる時のピッチ周波数の標準偏差は、
ギクシャク時に比べて高くなる傾向（5%水準で有意差あり）



**Distribution of the Ave_F0 and the SD_F0
for each utterance after removing laughter parts**

上手く進行している会話では、相互輝度が高くなる傾向

- 相互輝度: 2人の輝度を掛けたもの
- $G_{xy}(t) = g_x(t) \times g_y(t)$: g_x は時間 t における参加者 x の差分画像の平均輝度

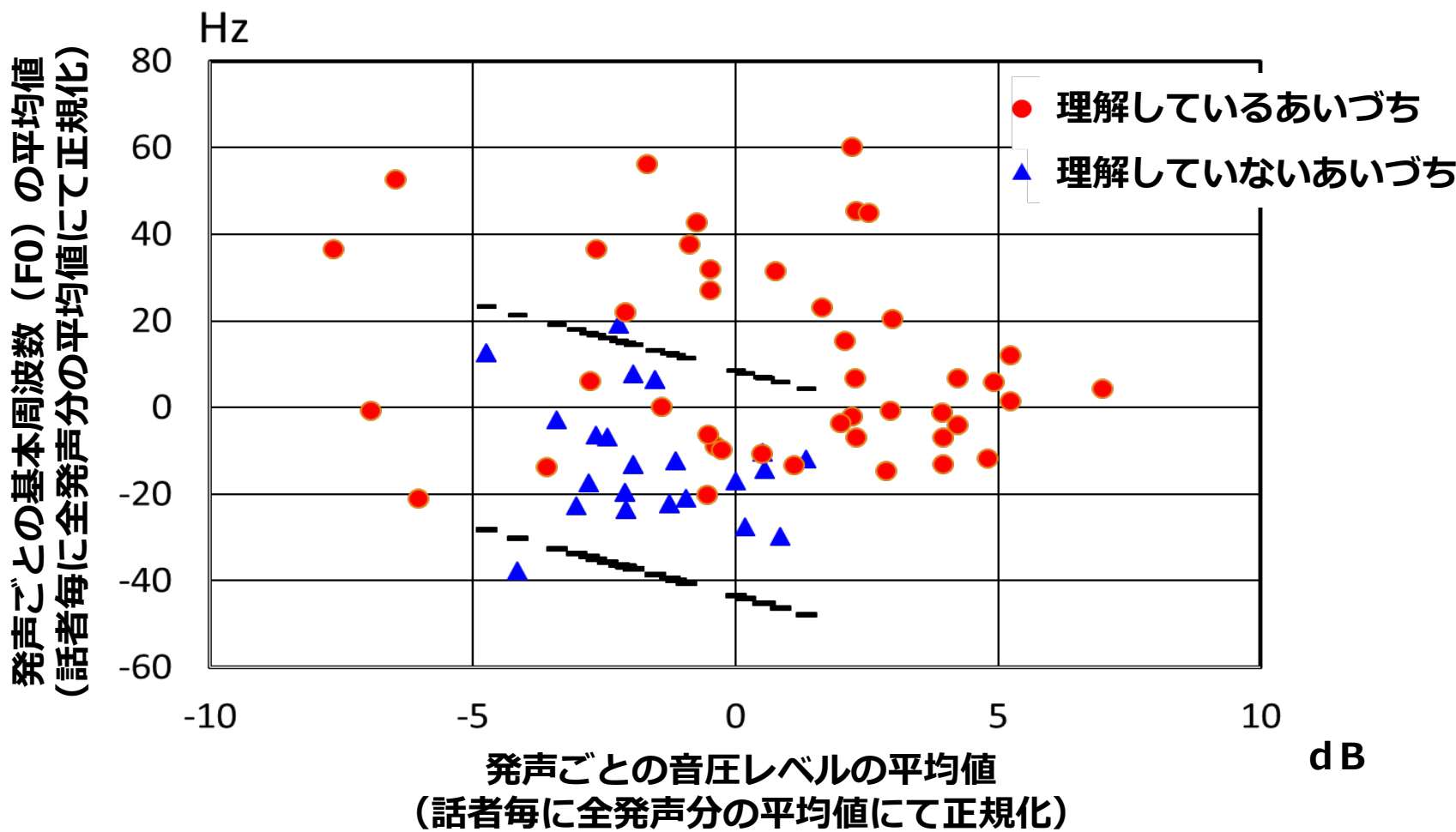


これらのパラメータは、
会話のギクシャクさのみではなく、
会話者が相手の話を理解できているかを見極める
に有効であることが分かった



明確に表面化する異変（病気など）の兆候となる、
日常のちょっとした異変（衰え、体調の悪さなど）を
本人に自覚させる、家族に知らせる、
効果を評価検討中

理解していない“あいづち”は、音響特徴量の変動が少ない



理解していない“あいづち”は、動作の同期性もない傾向

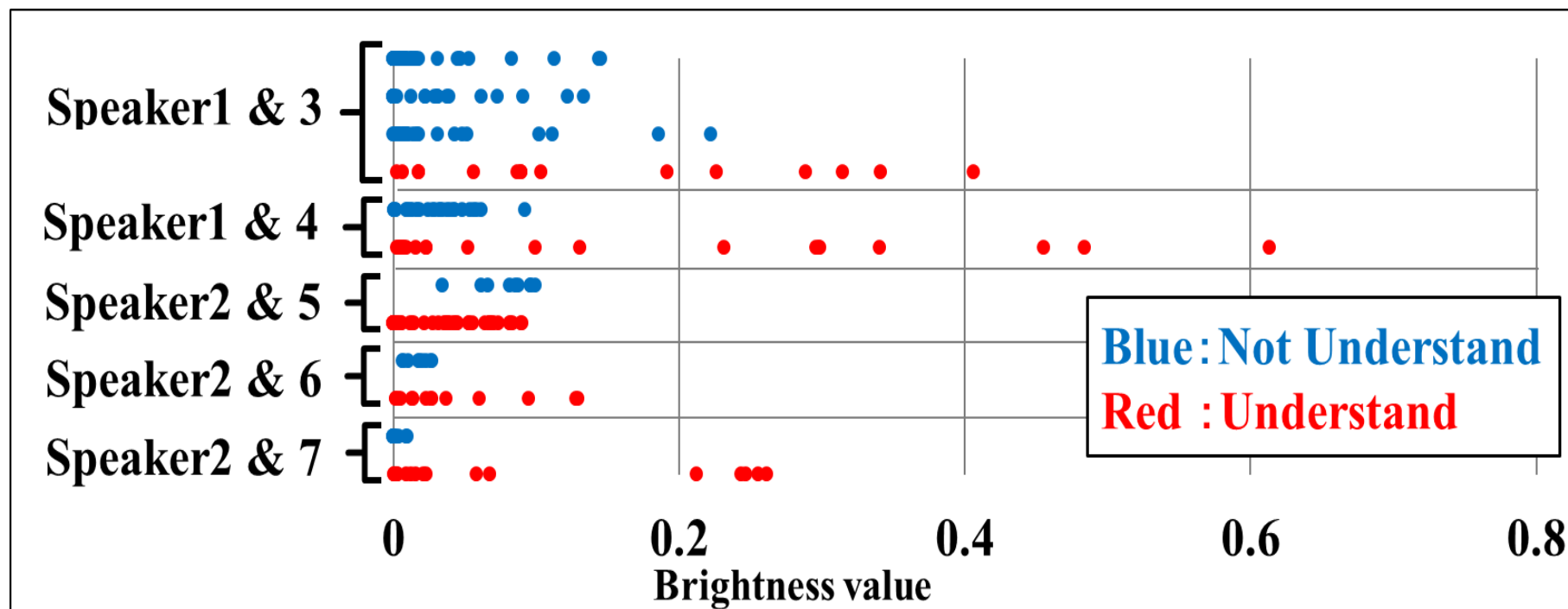


Figure 9: Brightness value of each speaker's synchronicity