

言語情報を用いた談話機能推定及びロボット頭部動作生成への応用

Dialog utterance classification and nod generation for humanoid robot

劉超然¹, カルロス石井¹, 石黒浩¹
Chaoran LIU, Carlos ISHI, Hiroshi ISHIGURO
国際電気通信基礎技術研究所
¹石黒特別研究所
¹ATR/HIL

chaoran.liu@atr.jp, carlos@atr.jp, ishiguro@sys.es.osaka-u.ac.jp

Abstract

コミュニケーション中の頭部動作は話者・聴者双方に置いて、会話を円滑化する役割を果たしている。遠隔操作ロボットを介した会話では、操作者側の環境と遠隔地一致しないなどの原因で、操作者の頭部動作をロボットにマッピングするのは不十分である。本稿では、発話機能を架け橋とし、発話音声から領きを生成するモデルを提案・評価した。話者の音声発話はまず自動音声認識システムによりテキストに変換され、複数の分類器の分類結果投票によって発話機能を推定した。各発話機能クラスに置いた領きの生起確率に従い、領き動作パターンの分布から動作特徴を選出し、生成した動作コマンドを音声と合わせてロボットに送る。評価者実験では、提案手法により生成した動作をアンドロイドロボット ERICA (図 1) で再生し、自然さ・人間らしさを評価した。



Figure 1. Android robot ERICA used in this work.

1 はじめに

談話中に見られる発話に伴う頭部動作は、話者と聴者にとって重要な役割を果たしている。動作の表出は話者に心理的な影響を与える。身体動作によって、話者の内部で話が整理でき、条理的な会話に貢献する。聴者にとって、話者の動作をコミュニケーションチャンネルとして、相手の意図や感情を把握することができる。このように、頭部動作は発話内容と強く関連している。コミュニケーションロボットを経由した遠隔談話において、動作の理解や自然な動作の生成は談話を深め、対話者間の接続を強くすることができる。

遠隔操作コミュニケーションロボットに置いて、最もシンプルな動作生成は操作者の動作を計測し、ロボットにマッピングする手法である。しかし、先行研究[Tamaki 2011]で報告されたように、メディアを

介した遠隔会話に置いた相槌や肯定などの意図を表す際、動作を伴わず音声のみで伝える傾向が見られた。単純な動作マッピングは対面対話時と違った動作をロボットに再現してしまう恐れがある。更に操作者が環境に反応した不意な動作も遠隔地に伝わるべきではない。故に、本稿では操作者の発話に着目し、音声から動作を生成するモデルを提案した。

2 関連研究

発話音声の基本周波数やパワーなどの韻律特徴と話者頭部動作の相関関係は、多数の研究により報告されている[Yehia 2002] [Sargin 2006] [Munhall 2004]。例えば、英語話者の頭部 6 自由度の回転と基本周波数 F0 の相関係数は 0.39 から 0.52 の間に対し、日本語話者に置いて、この相関係数は 0.22 から 0.30 区間になる。従って、F0 と頭部動作の相関性は言語に依存しており、F0 周波数のみで頭部動作を生成するのに不十分だと言える。

しかしながら、韻律情報から頭部動作動作を生成する試みも複数あった。Busso らは HMM ベースのアルゴリズムで、韻律情報から話者の感情を推定し、頭部動作を生成するモデルを提案した[Busso 2007]。渡辺らは、音声の ON-OFF 情報を入力とした領き生成モデルを考案し、CG イーゼントやヒューマノイドロボットに応用した[Watanabe 2004]。

複数の言語において、頭部動作と発話音声の関連性が報告されている。英語において、言葉を強調するとき、上昇調を使う傾向がある[Graf 202]。スウェーデン語において、強調アクセントに伴う表情や頭部動作が非強調時より多く見られる[Beskow 2006]。日本語において、発話権の変更や相槌などを考慮した以下のような談話機能リストが提案され[Ishi 2010]、頭部動作と談話機能の関連関係が報告された。

- k (keep) : 話者が発話権を保持。ポーズないしはつきりしたピッチのリセットが伴う(明瞭な句境界)
- k2 (keep) : 発話文の中にある不鮮明な句境界(発話権の保持)
- k3 (keep) : 話者が発話末の音節を伸ばし、考えることや発話の途中であることを表現(発話権の保持)
- f (filler) : 「えっとー」「あの一」など、考え中であることを表現する感動詞

- f2 (conjunctions) : 「じゃ」などの接続詞(文末を伸ばしていない短いフィラーとして捉えられる)
- g (give) : 当話者の発話が終了し、発話権を対話者への譲渡
- q (question) : 対話者に確認するなど応答を求める場合(発話権の譲渡)
- bc (backchannels) : 「うん」「はい」などの相槌を表現する感動詞
- su (backchannels) : 「うん」「はい」などの相槌を表現する感動詞
- dn (denial) : 「いいえ」「ううん」などの否定を表現する感動詞

この談話機能リストを用いて頭部動作を生成するモデルが提案された[Liu 2012]. 本研究は, [Liu 2012]の延長線上に位置する. 一部の談話機能を融合し, 大きく三つのクラスに分けた. 機械学習によって 談話機能クラスを推定し, 頷き動作生成モデルを提案した.

3 発話機能分類

本研究では, 発話フレーズを三つのクラスに分類する. 一つは明瞭な句境界を有するフレーズ(上述のタグリスト中の 'k', 'g', 'q' に該当する), 以降このクラスを 'kg' と表記する. 一つは相槌で, 'bc' と表記する. 最後はその他の 'o' クラス. ここで句境界を言語構造上完了した文と後続文の境界を指す. ポーズが明確なピッチリセットを伴うものが多数占めている. クラス 'kg' と 'bc' に分類されたフレーズに頷きを生成する.

相槌を独立したクラスに分ける理由として, 相槌に使われる語彙, 韻律特徴及び相槌に伴う頷きにユニークなパターンが見られるためである. また, 自動会話システムとして, ユーザの発話が相槌か割り込み発話かを識別する必要があり, ユーザ発話の種類に対応した発話を生成しなければならない. この研究では, ロボットが発話時の動作生成に注目し, ユーザ発話への対応は今後の研究課題として残す.

3.1 テキストのベクトル化

発話フレーズのテキストを分類器の入力として使うため, ベクトルに変換するプロセスが必要である. 本研究では Latent Dirichlet Allocation (LDA) [Blei 2003] トピック用いて言語情報を表現し, 分類を試みた.

LDA モデルは生成確率モデルの一種で, 階層的構造を持つベイズモデルである. データセット中の文章をハイパーパラメータで決まる有限個数のトピックの確率分布の形で表現することができる. 文章中の各単語が, ランダムに選択されたトピックによって, 一定の分布に従って生成された, という基本的な仮説を持つ教師なし学習手法である. LDA モデルの特徴として, トピックセットが Dirichlet 事前分布を有することを前提とし, 比較的に小規模なデータ

セットでも訓練データにオーバーフィットする傾向が少ない.

LDA モデルを用いたことによって, 発話フレーズをトピックの確率分布として表現することが可能になり, 特徴ベクトルの次元数は Bag-of-words (BOW) 表現時の辞書の長さからトピックの数まで減らすことができる. 本研究では, BOW ベクトルを tf-idf [Salton 1983]によって重み付けし, LDA トピックに変換した. ハイパーパラメータの設定は先行研究 [Hoffman 2010]に参考した. トピック数は 200 に設定した.

3.2 SVM分類

サポートベクトルマシン(SVM)[Shawe-Taylor 2000]を用いて発話機能の分類を行なった. SVMは広く研究されていた教師あり学習アルゴリズムで, 分類性能の優れたモデルとして知られている. SVMの学習プロセスは, 訓練データの特徴空間に, クラス間のマージンが最大になる超平面を探すプロセスになる.

発話データの中, 分類ターゲットとなる三つのクラスに属するサンプルの数が一様ではない. その他 'o' クラスに属するサンプルは全データの半分以上に占めており, 相槌 'kg' に属するサンプルは全体の 2 割に過ぎない. このようなアンバランスデータ対し, 分類誤りを最小化した結果, サンプルの少ないクラスに対する分類パフォーマンスは期待できない. 極端な場合, 少数クラスのサンプルを全て多数クラスとして分類しても, 全体の分類精度が大きく下がらない故である. この問題点を解決するため, 分類器を訓練する前に, 訓練データを cost-proportionate rejection sampling 法[Zadrozny 2003]を用いてダウンサンプリングを行ない, 三つのバランスが取れたサブセットを作った. これらのサブセットを用いて, 三つの SVM 分類器を訓練した.

本研究で扱う頷きクラス分類問題にとって, 適合性が重要な指標となっている. 先行研究[Ishi 2010]に報告されたように, 頷きがよく見られる発話機能クラス 'k', 'g' と 'bc' における頷きの出現確率はそれぞれ 49.3%, 47.1%と 84.2%である. 言い換えると, 比較的に低い再現率でも人の頷き傾向を再現するのに問題にならないのに対し, 頷くべきではない発話フレーズで, 頷きを生成すると, 不自然だと評価される可能性が高い. 従って, 少数クラスに対する適合率と多数クラスに対する再現率が望ましい分類結果となる. その為, 最終的分類結果は一般的に採用された多数投票ではなく, 合意投票の結果を用いた. 少数クラスである 'kg' と 'bc' クラスに対し, 一つの SVM が 'false' と出力すれば, 分類結果は 'false' になる ('o' クラスに属する).

4 動作分析及び頷き生成

本セクションでは, 発話に伴う動作の分析及びそ

の結果に基づいた頷き生成モデルを説明する。

4.1 動作データ

本研究で使われている談話データセットに、発話音声と話者の頭部動作が含まれている。動作の計測にはモーションキャプチャシステムを利用した。マーカーの配置は図2に示す。



Figure 2. Motion capture marker set.

計測したマーカーの三次元位置時系列データに対し、頭部位置及びピッチ、ロール、ヨー三軸回転角度への変換を施した。ラベラーにより、頷きの区間が付与されている。

発話と動作の関連関係を分析するため、人工的に書き起こしたテキスト及び発話フレーズの分類（分類器を訓練する際の正解データ）を用いた。

4.1.1 データ前処理

頭部の三軸回転データの大部分は話者の頭部回転を反映しているが、上半身を移動する際、頭部の向き付随的に変わってしまう。データの前処理ではこの部分の動きを取り除く。

上半身の移動による頭部の回転は頭部動作に比べると遅くて比較的長時間に渡るという性質を持っている。そのため短い時間で見ると、この部分の動きを線形的なトレンドと見なすことができる。各頷き区間に対し、線形回帰を施し、得られた線形成分を上半身移動と見なす。時系列回転角度からこれらの線形成分を除き、余剰成分を頭部の回転とした。

4.1.2 頷きの分析

頷きの種類を大きく二つに分けた。一つは正弦波で近似すると一周期以内のもので、‘nd’と表記する。もう一つは複数周期を含むもので、‘mnd’と表記する。データセットの中、4199の‘nd’と1482の‘mnd’が含まれている。

頷きの周期を計算するため、回転角度に対し、離散フーリエ変換を施した。クラス‘kg’と‘bc’の発話に見られた頷きの周波数と角度の平均分散を図3に示す。

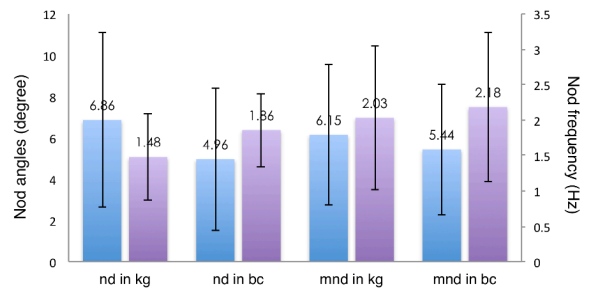


Figure 3. Averages of nod angles and frequencies. Error bars denote standard deviations

‘kg’クラスに見られた‘nd’と‘mnd’の周波数の平均値はそれぞれ1.48Hzと2.03Hzで、‘bc’クラスの平均値は1.86Hzと2.18Hzである。全体的に‘bc’クラスに見られた頷きは‘kg’クラスより速く、‘mnd’は‘nd’より速い傾向である。頷きの角度に関しては、周波数と相反した傾向が見られた。

全データにある頷き周波数のヒストグラム及び周波数と最大角度の分布を下図に示す。

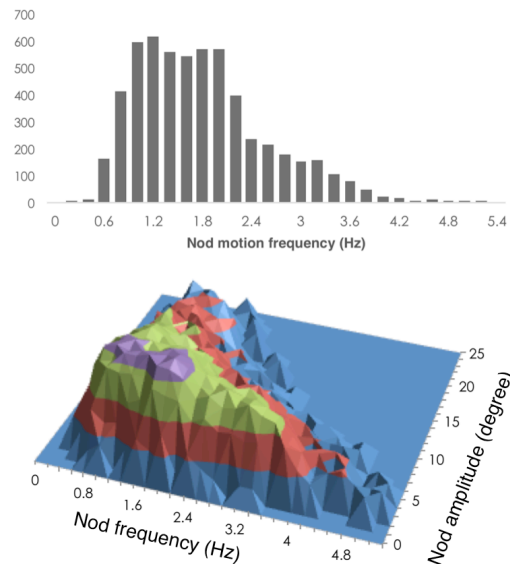


Figure 4. Histogram of nod frequencies and co-occurrence pattern between frequency and amplitude.

4.1.1 音声との関連

発話音声の持続時間、発話スピード、相対振幅、パワー変化率などの音響特徴と動作の相関関係を調べたが、大多数の特徴量に関して、強い相関関係が見られない。一つだけ動作周波数と正の相関関係が確認できたのは、‘bc’クラスにおいた音声パワー変化率である。図5に‘kg’と‘bc’両クラスの音声パワー変化率と頷き周波数の分布を示す。

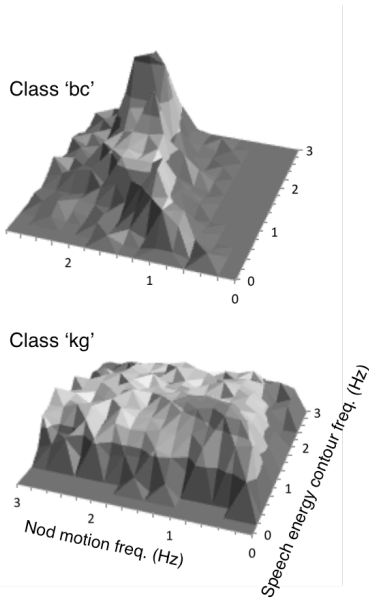


Figure 5. Co-occurrence pattern of nod motion frequency and speech energy contour frequency for classes ‘bc’ and ‘kg’.

‘kg’クラスにおいて、音声パワーの変化率と動作周波数の分布は一様分布に類似したのに対し、‘bc’クラスにおいて、特に低周波区間で強い相関が見られた。この相関は相槌時繰り返した言葉（“はいはい”や“うんうん”など）と共起した頷きだと考えられる。

4.3 頷き生成

対話・頷きデータの分析結果を利用し、発話時の頷き生成モデルを提案した。当モデルは発話音声を入力とし、頷き動作を出力する。全体の処理の流れをブロック図に示す。

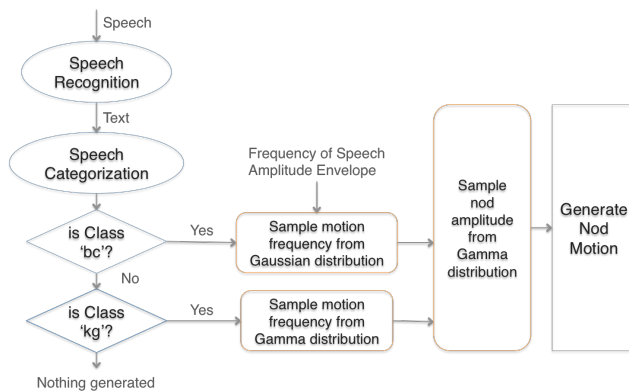


Figure 6. Block diagram of the proposed model.

まず発話音声を音声認識エンジンに送り、テキストに変換する。その後セクション 3 に説明した分類器により、発話機能クラスを推定する。分類結果はクラス ‘o’ の場合、動作を出力せず、次のフレーズに移行する。分類結果は ‘kg’ 及び ‘bc’ の場合、

頷きの動作を生成する。具体的な流れは以下に説明する。

発話機能分類の結果はクラス ‘bc’ の場合、音声パワーの変化率 f_s を利用し、頷きの周波数 f_{nd} を決める。図 5 で示した分布を正規雑音に載った線形回帰モデルで近似する。 f_{nd} は正規分布 $\mathcal{N}(\mu_{f_s}, \sigma)$ からサンプリングする。この中、 μ_{f_s} 以下のように決める。

$$\mu_{f_s} = (\mathbf{F}_s' \mathbf{F}_s)^{-1} \mathbf{F}_s' \mathbf{F}_{nd} \times f_s \quad (1)$$

\mathbf{F}_s と \mathbf{F}_{nd} は観測データ中の音声パワー変化率と頷き周波数を表す。線形回帰モデルから正規分布 $\mathcal{N}(\mu_{f_s}, \sigma)$ の期待値と分散が得られる。

又、発話機能分類の結果はクラス ‘kg’ の場合、図 5 に示したように、頷き動作の特徴と音声特徴に強い相関関係見られないため、頷きの周波数 f_{nd} を図 4 に示した全体分布からサンプリングする。この分布をガンマ分布 $\Gamma(\alpha, \beta)$ で近似する。正規分布ではなくガンマを選んだ理由は、データ分布の形及び周波数パラメータが非負である必要があるため。

頷きの動作を生成するのに、周波数と動作角度が必要であるため、上記の手法で動作周波数 f_{nd} を決めてから動作角度を抽出する。各周波数に対応した角度の期待値は図 4 で分かるように、線形回帰を使って求めることが可能である。具体的には、角度をガンマ分布 $\Gamma_b(\alpha_b, \beta_b)$ からサンプリングする。 α_b と β_b は以下のように決める。

$$\frac{\alpha}{\beta} = k_1 f_{nd} + \zeta_1 \quad (2)$$

$$\beta \propto f_{nd} \quad (3)$$

k_1 と ζ_1 は線形回帰で得られるパラメータである。

最後に、頷きの動作を上記の周波数と振幅を持つ正弦波で近似する。人の頷きは下方への回転量が上方より大きいいため、正弦波開始位相を $\pi/4$ とする。

5 評価実験

生成した頷き動作の自然さを評価するため、被験者実験を行った。先行研究で提案された正解談話機能ラベルに基づく頷き生成モデルを比較対象として用いた。

5.1 実験設定

発話の違いによるバイアスを排除するため、実際の対話ではなく、収録した発話を使って、動作生成を行った。9 の発話フレーズをランダムにデータベースから抽出した。比べやすいように、各フレーズの長さは 15 秒から 30 秒に制限した。

一つの発話に対し、四つの動作を生成した。二つは先行研究で提案された頷きタイミング制御モデルによって生成された。頷き動作は一般的に見られた

人の頷きサンプルを使用した。動作角度を図 7 に示す。その中一つは先行研究のように正解ラベルを利用し、動作の生成を行った。もう一つはセクション 3 で説明した分類結果を用いた。

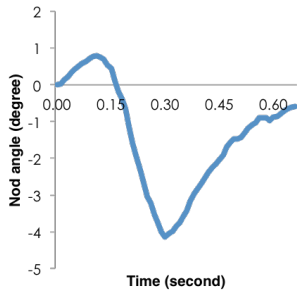


Figure 7. Fixed nod example used in previous work.

残りの二つの動作は本研究で提案した手法によって生成された。頷きの周波数、長さ、振幅は確率分布からサンプリングするので、具体的な動作は生成するたびに変わるため、同様な手法で二つの動作を作成し、検証した。正規分布からサンプリングする際、確率が低いものの、マイナスな値が得られる場合がある。そのようなサンプルはリジェクトして、再サンプリングする。また動作の周波数や角度がロボットの可動範囲を超えたようなサンプルもリジェクトする。

発話音声を変えてテキストのため、Google speech api v2 を利用した。9 の音声サンプルの認識精度は単語ベースで 64.2%であった。これらのテキストをセクション 3 で説明した 200 次元 LDA トピックに変換され、三つの分類器に送った。分類精度は 68.1%であった。分類実験時の精度より低い理由は、音声認識エラーの影響があると思われる。

生成された動作はロボットのアクチュエータコマンドに変換され、20ms 間隔でロボットに送られる。本実験ではアンドロイドロボット ERICA を使った。アクチュエータ配置は図 8 に示す。発話時の口唇動作は「」で提案された手法を用いた。

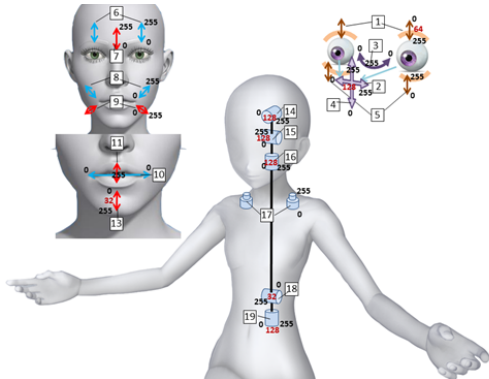


Figure 8. Actuator map for android robot ERICA.

9 の発話サンプル毎に 4 の動作タイプを生成し、ロボットで再現し、ビデオを収録した。21 人の被験者に発話毎にシャッフルしたビデオを呈示し、動作の自然さを 7 段階で評価した。‘1’は最も不自然で‘7’は最も自然を表す。

5.1 実験結果

図 9 に四つの動作タイプに対する主観評価結果を表す。左から一つ目は先行研究で提案された正解ラベルに基づいた頷き生成モデルで、‘fixed-gt’ と表記する。二つ目は同じく図 7 に示した頷き動作を利用したもので、‘fixed’ と表記した。正解ラベルの代わりに分類器の推定結果を用いた。発話機能分類時、頷きクラス (‘kg’, ‘bc’) に対して、適合率を重視したため、一部の頷きクラスの発話が頷かないクラスに分類される傾向がある。その為、‘fixed’ 動作タイプに含まれる頷きの回数は ‘fixed-gt’ より少ない。三つ目と四つ目は本研究で提案したモデルによって、生成された動作で、‘sampled1’ と ‘sampled2’ と表記する。

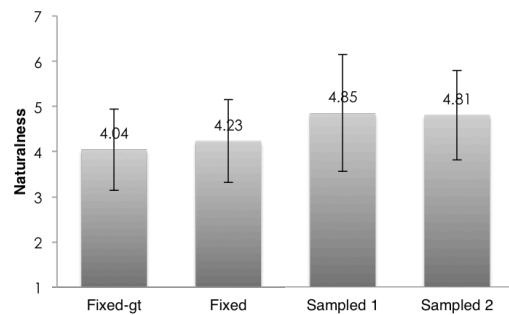


Figure 9. Perceived naturalness for each motion type.

評価結果に対し、有意水準を 0.05 に設定した被験者内一要因分散分析を施した結果、各条件間に有意差が見られた ($F(3,20) = 5.18, p = 0.003$) 。

多重比較の結果、同提案手法によって生成された ‘sampled1’ と ‘sampled2’ に有意差が見られない ($p=0.89$)。この結果は平均と分散からも容易に見られる。‘fixed-gt’ と ‘fixed’ の間有意差が見られないが、推定結果を使った ‘fixed’ は正解ラベルを用いた ‘fixed-gt’ より高い評価平均値が得られた。人の対話データに、‘kg’ と ‘bc’ クラスの発話に頷き生起確率は約 50% と 80% だったため、正解ラベルを使ってすべての頷きクラスの発話に頷きを生成すると、頷きすぎでかえて不自然になる可能性がある。最後に提案手法の評価結果はこの二つの動作より有意に高い ($v.s$ ‘Fixed-gt’: $p = 0.012$, $v.s$ ‘Fixed’: $p = 0.039$) 。

不完全な音声認識結果を用いても、本研究で提案した頷き生成モデルを使って、自然な頷き動作とタイミングを制御することが可能であること、この評価実験の結果が示した。

5 おわりに

発話に付随した頭部動作は、対話者間の感情や意図の伝達に重要な役割を担っている。又、遠隔地にいる人のアバターとしてのヒューノイドロボットにとって、自然な動作はスムーズな会話に強く影響

する。本研究では発話音声を入力とした自然な頷き動作を生成するモデルを提案した。

人の対話データから、頷きが主に発話文のクリアな句境界や相槌フレーズに生起することが分かった。この発見に基づいて、発話フレーズを句境界文・相槌・その他三つのクラスに分類した。収集した対話データからサンプリングしたサブセットを用いて、複数の SVM 分類器を訓練した。言語情報は LDA トピックの確率分布を利用してベクトライズし、SVM の特徴ベクトルとして使った。複数の SVM 出力の合意投票によって、最終的なクラス分類結果を決める。

人の頷き動作の周波数分布は、ガンマ分布で近似することが可能であり、各周波数に対応した頷きの振幅（最高点から最低点の角度）もガンマ分布に見なせることが、人の対話データから分かる。又、相槌の場合、発話音声のパワー変化周波数が低い発話において、頷き周波数とパワー変化周波数に強い相関が見られた。これらの結果より、発話音声の特徴によって、ガンマ分布若しくは正規分布から頷きのパラメータをサンプリングして、正弦波で頷きを近似する手法を提案した。

提案法の有効性を検証するため、被験者実験を行った。実験結果は、提案の頷き生成モデルは、先行研究の人工的に付与した正解ラベルを使ったモデルより自然な動作を生成できることを示した。

今後の課題として、発話音声から他の動作パターンを生成するや、発話と動作の同期メカニズムを探究することが予定している。

謝辞

この研究は JST, ERATO, 石黒共生ヒューマンロボットインタラクショナルプロジェクトの一環として行われたものです。この研究の一部は JSPS KAKENHI (25220004)の助成を受けたものです。

参考文献

- [Tamaki 2011] Tamaki, H., Higashino, S., Kobayashi, M., Ihara, M. Reducing Speech Contention in Web Conferences. *Applications and the Internet*. 75-81. 2011.
- [Yehia 2002] Yehia, H.C. Kuratate, T. Vatikiotis-Bateson, E. 2002. Linking facial animation, head motion and speech acoustics. *J. of Phonetics*, Vol. 30, pp. 555-568, 2002.
- [Sargin 2006] Sargin, M.E. Aran, O. Karpov, A. Ofli, F. Yasinnik, Y. Wilson, S. Erzin, E. Yemez, Y. Tekalp. A.M. 2006. Combined Gesture-Speech Analysis and Speech Driven Gesture Synthesis. In *proc. IEEE International Conference on Multimedia*, 2006.
- [Munhall 2004] Munhall, K.G. Jones, J.A. Callan, D.E. Kuratate, T. Vatikiotis-Bateson, E. 2004. Visual prosody and speech intelligibility – Head movement improves auditory speech perception. *Psychological Science*, Vol. 15, No. 2, pp. 133-137, 2004.
- [Busso 2007] Busso, C. Deng, Z. Grimm, M. Neumann,

- U. Narayanan, S. 2007. Rigid Head Motion in Expressive Speech Animation: Analysis and Synthesis. *IEEE Trans. on Audio, Speech and Language Processing*, March 2007.
- [Watanabe 2004] Watanabe, T. Okubo, M. Nakashige, M. Danbara, R. InterActor: Speech-Driven Embodied Interactive Actor, *International Journal of Human-Computer Interaction* 17 (1) (2004) 43–60.
- [Graf 202] Graf, H.P. Cosatto, E. Strom, V. Huang, F.J. 2002. Visual prosody: Facial movements accompanying speech. In *proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition (FGR'02)*, 2002.
- [Beskow 2006] Beskow, J. Granstrom, B. House, D. 2006. Visual correlates to prominence in several expressive modes. In *proc. Interspeech 2006 – ICSLP*, pp. 1272-1275, 2006.
- [Ishi 2010] Ishi, C.T. Liu, C. Ishiguro, H. Hagita, N. 2010. Head motion during dialogue speech and nod timing control in humanoid robots. In *proc. of IEEE/RSJ Human Robot Interaction (HRI 2010)*, 293-300, 2010.
- [Liu 2012] Liu, C. Ishi, C. T. Ishiguro, H. Hagita, N. Generation of nodding, head tilting and eye gazing for human-robot dialogue interaction. In *Proc. of HRI 2012*, pp. 285-292, Boston, March, 2012.
- [Blei 2003] Blei, David M. Ng, Andrew Y. Jordan, Michael I (January 2003). Lafferty, John, ed. "Latent Dirichlet Allocation". *Journal of Machine Learning Research* 3 (4–5): pp. 993–1022. 2003.
- [Salton 1983] Salton, G. and McGill, M. editors. Introduction to Modern Information Retrieval. McGraw-Hill, 1983.
- [Hoffman 2010] Hoffman, M., Bach, F. R., Blei, D. M. Online learning for latent dirichlet allocation. *Advances in neural information processing systems*. p. 856-864. 2010.
- [Shawe-Taylor 2000] Shawe-Taylor, J., Cristianini, N. Support vector machines and other kernel-based learning methods. Cambridge University Press, 2000.
- [Zadrozny 2003] Zadrozny, B., Langford, J., Abe, N. Cost-sensitive learning by cost-proportionate example weighting. In *Proc. Of the Third IEEE International Conf. on Data Mining*. 435-442, 2003.