

論文内容の要旨

博士論文題目 Hidden Markov Model-Based Estimation of Lip Motion from Audio Signal
(隠れマルコフモデルに基づいた音声信号からの唇の動きの推定)

氏 名 山 本 菜 里

(論文内容の要旨)

(1, 200字程度)

音声の明瞭度は、音声信号に唇画像を付け加えて提示することによって改善される。コンピュータエージェントの合成を考える場合、自然で人間らしい唇の動きを合成することが明瞭度を向上させるために重要である。この論文では、音声入力により人間らしい唇の動きを合成する方法として、隠れマルコフモデルを用いた2つの合成手法を提案する。

1 番目の提案手法は、自然な唇の動きを合成するために調音結合の効果を採用入れる。この提案手法は、HMM の Viterbi アライメントから音素系列を求めることにより文脈情報を取り込む。このとき画像パラメータの学習パターン数を削減するために口形素を用いる。口形素の利用により、少ない音声・画像同期データの場合でも文脈依存情報を取り入れることが可能になる。

2 番目の提案手法は、決定論的な音声認識手法の Viterbi アライメントを用いずに唇の動きを推定する手法である。Viterbi アライメントを用いる合成手法は、認識間違いを起こしたフレームで不正確な口形を生成する場合がある。2 番目の提案手法は、全ての HMM の状態を考慮して EM アルゴリズムを用いることにより正確な画像パラメータの推定を可能にする。提案手法を評価するために、実際の人間の動きから作成した画像パラメータと合成値間のユークリッド距離を取り、少量の音声・画像日本語データを用いた実験を行なった。

この実験結果は、1 番目の提案手法が調音結合効果を表す正確な口形を合成できることを示した。特に /h/, /g/, /k/ など後続母音に依存した口形の合成に有効であることが分かった。また、EM アルゴリズムを用いた 2 番目の提案手法は、両唇子音に相当するフレームで認識間違いにより生成された不正確な画像パラメータを改善出来ることを示した。

氏 名	山本 英里
-----	-------

(論文審査結果の要旨)

自然なヒューマンマシンインタフェースを実現する際に、自然で人間らしいコンピュータエージェントを生成することが極めて重要となる。特に、音声によるコミュニケーションを考えた場合、コンピュータからの音声合成出力は現状では品質的に未だ不十分であるが、自然で人間らしい唇の動きを加えることで、より明瞭に伝えることができる可能性が有る。本論文では、隠れマルコフモデル (HMM) に基づき、音声を入力とし入力音声に同期した自然な唇の動きを合成する方法を提案している。

本論文の成果は、次のように要約できる。

(1) 音声と唇座標パラメータの同期収録データを用い、音声と唇座標パラメータとの対応関係を HMM により自動的に学習し、合成する方法を提案している。さらに、唇の動きの調音結合に対処する方法としてビタビ復号時に遅延を入れて後続口形素を判定し、その後続口形素に依存した唇の動きを合成することで自然な唇動画像を合成する SV-HMM 法を提案している。この手法を従来提案されている VQ 写像に基づく方法、後続口形素を考慮しない HMM に基づく方法に対し、オリジナルの画像パラメータとの距離により比較し有効性を確認している。特に、後続の母音に依存した口形となる /h/, /g/, /k/ などの音素で提案法が有効であることを示している。この手法は、少量の音声・唇座標パラメータに基づき、自然な唇動画像を生成する手法として極めて効率的で高く評価できる。

(2) ビタビ復号に基づく唇画像合成法では、入力音声と音声の音素 HMM モデルのビタビアライメントが決定的であるため、アライメントの誤りが合成する唇画像パラメータの誤りに繋がる。そこで、このアライメントを非決定的するためにすべての HMM の状態をその確率込みで考慮し認識誤りによる精度劣化を回避し、さらに、音声と画像の結合確率分布を用いて、入力音声から適切な唇画像パラメータを EM アルゴリズムにより想起する HMM-EM 法を提案している。また、唇動画像生成実験を行い、HMM-EM 法を用いることにより、SV-HMM 法で誤った両唇子音の音素において大きな改善がえられることを示している。この方法は、結合確率分布に基づき入力音声から画像のパラメータを推定する新しい手法であり、きわめて高く評価できる。

以上のように、本論文は音声合成出力における明瞭度の改善を目的に、入力音声に同期した唇動画像を隠れマルコフモデルに基づいて生成する方法を提案し、有効性を示した。これらの研究成果は、情報科学、音情報処理学の分野における学術的な貢献として高く評価できる。平成 10 年 12 月 24 日開催の公聴会の結果を参考に、本博士論文の審査を行なった結果、本論文は博士(工学)の学位論文として十分に価値のあるものと判断する。