

D-14-3

感情情報を伝達する匿名音声チャットの開発

Development of Anonymous Voice Chat System Conveying Emotional Information

河津宏美¹
Hiromi Kawatsu有本泰子¹
Yoshiko Arimoto大野澄雄¹
Sumio Ohno東京工科大学大学院 バイオ・情報メディア研究科¹
Graduate School of Bionics, Computer and Media Sciences, Tokyo University of Technology

1 まえがき

オンラインシミュレーションゲームにおいて、音声を利用したチャット機能の搭載が望まれている。一方で、音声チャットが実装されたとしても、音声には個人を特定するような情報が含まれているため、不特定多数のユーザを相手にするオンラインゲームでは利用を好まないユーザもいる。このような状況を踏まえ、音声に含まれる個人を特定するような情報を削除し、匿名性を保ちながらも、音声に含まれる感情情報を損なうことなくコミュニケーション可能な音声チャットの開発を行なった。

2 システムの構成

開発した音声チャットは、ユーザの発した音声を認識することにより、言語的な情報をテキスト化するとともに、音声を手がかりとしてユーザの感情理解を行なう。そして、得られた言語的情報と感情情報をもとに自然な感情音声を合成するというものである。図1に音声チャットシステムの構成を示した。

3 感情音声の制御モデル

(1) 感情の種類

感情の種類として、Plutchikが提案した感情立体モデル[1]の基本感情である「喜び-悲しみ」「怒り-恐れ」「受容-嫌悪」「驚き-期待」の4次元8感情に加え、「平静」を取り扱った。また、基本感情に関しては、その程度まで含めた検討を行い、より細かなニュアンスの理解と合成を試みた。

(2) 感情理解のための音声資料収集とモデル化

感情理解のためのモデル構築に向けて、オンラインゲーム中の自然な対話を収録した。プレイヤーは、男性9名、女性4名で、収録に利用したオンラインゲームはラグナロクオンライン、モンスターハンターフロンティア、レッドストーンの3つである。

収録した9114発話の音声に対して、感情情報の

表1 抽出した音声の特徴量

声の高さ	男女差正規化 F_0 パターンの発話内最低値
	男女差正規化 F_0 パターンの発話内最高値
	男女差正規化 F_0 パターンの発話内平均
	F_0 パターンの発話内標準偏差
声の長さ	呼気段落の平均モーラ数
	平均発話速度 (mora/s)
声の強さ	短時間平均パワーの発話内最大値
	短時間平均パワーの発話内標準偏差
	短時間平均パワーの変動量
声質	第1次ケプストラム係数の発話内平均
	第1次ケプストラム係数の発話内標準偏差

タグ付けと、文献[2]を参考に定めた音声の特徴量(表1参照)の抽出を行ない、音声に含まれる感情を理解するためのモデルを構築した。

(3) 感情合成のための音声資料収集とモデル化

感情が表現された典型的な発話を収録するために、演劇経験者が発声した音声を収録した。収録では、10前後の発話で構成される17対話を用意し、その対話テキストに従って指定した感情を表現した音声を発声させた。話者は男性、女性、それぞれ2名ずつの計4名である。対話は男性同士、女性同士で行なった。1つの対話に対して平静の音声と、3段階の程度で感情を表現した音声を収録した。

収集した音声の特徴量を抽出し、平静の発話の音声特徴量と感情発話の音声特徴量との比較を行ない、感情を生成するためのモデルを構築した。

4 あとがき

音声に含まれる感情情報の理解と合成のためのモデル化を行ない、匿名性を有する音声チャットを開発した。

謝辞

本研究は独立行政法人情報処理推進機構(IPA)の2007年度第II期末踏ソフトウェア創造事業(末踏ユース)の支援を受け、東京大学大学院 竹内郁雄教授より開発時に指導、アドバイスを受けた。

参考文献

- [1] R. Plutchik, "Emotion - A Psychoevolutionary Synthesis", Harper, Row, 1980.
- [2] 有本 泰子, 大野 澄雄, 飯田 仁, "「怒り」の発話を対象とした話者の感情の程度推定法", 言語処理学会誌「自然言語処理」, vol. 14, no. 3, pp. 131-145, 2007.

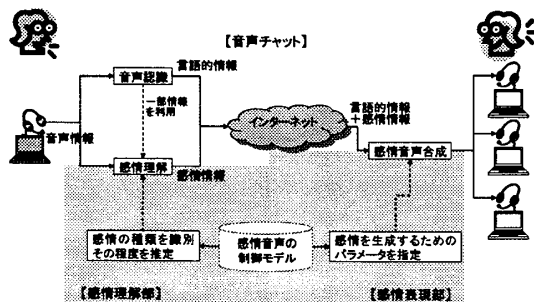


図1 開発した音声チャットシステムの構成