

非言語情報を用いたマルチモーダル対話システム

中沢 正幸† 向井 理朗† 関 進† 綿貫 啓子† 三吉 秀夫†
† R W C P マルチモーダル機能シャープ研究室 シャープ (株) 技術本部内
〒 261-8520 千葉県千葉市美浜区中瀬 1-9-2
URL <http://www.rwcp.or.jp/lab/mm-sharp/index-j.html>

Abstract In this paper, we describe a multimodal human interface system, MAICO (Multimodal Agent Interface for COmmunication), based on Dynamical Dialogue Model. This system not only integrates information of a speech processing and a gesture processing, but also controls the response timing in order to realize a smoothness and a easy interaction between a user and a computer. Our approach consists of human-human dialogue analysis, and computational modeling of dialogue.

Keywords multimodal, speech recognition, nonverbal, dynamical system, dialogue model

1 はじめに

対話における抑揚やしぐさ等の非言語情報を扱うモデルとして力学系を応用した対話モデルを提案する。さらに、本モデルをベースにした、ユーザの声の抑揚に応じて相槌を打つシステムについて紹介を行う。ユーザの非言語情報及びキーワードに基づいてCGエージェントが表情やジェスチャを交えながら、適切な間や速さ、強度で音声及び表情によってユーザに情報を伝達するプロトタイプシステムを構築した。

2 概要

我々は、抑揚やしぐさ等の非言語情報をベースにした、マルチモーダル情報の円滑なやり取りの枠組みを構築することで、人間と機械の間に円滑なインタフェースを実現することを目標としている。このために、人間と機械の間の対話モデル構築に必要な、人間同士の対話過程を記録したマルチモーダル対話データベースを構築し、解析を行っている。自然で気持のよい円滑な対話には、対話過程の解析から得られた知見を対話モデルに活かすことはもちろん、エージェント自らが自律性を持ち、マルチモーダル情報の統合処理に適した対話モデルが必要となると考えている。

人間同士の対話を解析した場合、自然で気持のよい円滑な対話には、独自のリズムやテンポがあるのでは

ないかと予想されており、我々は、その独自のリズムやテンポには相槌が重要な役割を果していると考えている。本稿で、対話における抑揚やしぐさ等の非言語情報を扱うモデルとして力学系を応用した相槌のタイミング生成手法とその対話モデルを提案する。

本システムは、ユーザの非言語情報及びキーワードに基づいてCGエージェントが表情やジェスチャを交えながら、適切な間や速さ、強度で音声及び表情によってユーザに情報を伝達するプロトタイプである。現在、モデルの定式化やパラメータは経験的に設定しているが、今後はモデルパラメータの自動学習等を行うと同時に、言語情報処理も取り入れていく。

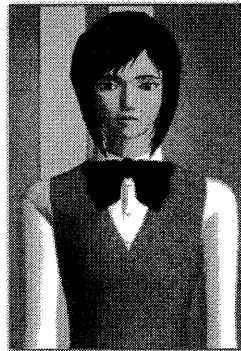


図 1: MAICO: Multimodal Agent Interface for COmmunication

“Multimodal Agent Interface for COmmunication:
- MAICO -”

Nakazawa Masayuki, Toshiro Mukai, Susumu Seki, Keiko Watanuki, Hideo Miyoshi

†Real World Computing Partnership Multimodal Functions Sharp Laboratory, Sharp Corporation