

InterSpace におけるエージェント機能に関する一考察

2W-2

宇佐美 潔忠 井上 雅之 河野 隆志 正木 茂樹

NTTヒューマンインタフェース研究所

1 はじめに

近年のインターネットの急速な普及、パソコンの高性能化・低価格化、およびインターネット上の3次元仮想空間記述言語 VRML の登場により、ユーザが家庭から手軽に3次元CGを利用した仮想空間を利用する機会が増えはじめている。

筆者らの研究グループでは、多人数参加型の上記3次元仮想空間システムのひとつとして InterSpace を提案している[1]。InterSpace は、マルチメディア通信時代の人と人との出会いや情報交換、意志疎通などの多様なインタラクションを快適に、かつ効率的に行えるようなコミュニケーション環境を提供するシステムとして開発された。

ところで、現状の InterSpace に限らず、現行の仮想空間システムサービスでは、提供する情報やサービスのほとんどは、仮想商店の商品などのサンプルを示すオブジェクトをマウスでクリックしたり、そのオブジェクトにユーザが近づくことにより、静止画情報やオンデマンド映像情報等を表示したりすることにより実現している。これらユーザの何らかの働きかけのもとで動作する受動的なサービスとは対照的に、ユーザの意図を理解し、外部からのインタラクションに柔軟に対処できるような能動的なサービスの提供が今後の仮想空間サービスをより魅力的なものにするためには必要である。

本稿では、InterSpace において、上記の能動的なサービスを提供するためにエージェント技術[2]を応用した仮想空間サービスを実現するための一環境について報告する。

A Construction of Integrated Environment for Realizing the Agents in InterSpace

Kiyotada USAMI, Masayuki INOUE, Takashi KOUNO and Shigeki MASAKI

NTT Human Interface Laboratories

2 エージェントの概要

InterSpace におけるエージェントは、その実体として、図1に示すような犬型のモデルを持つ GUI と InterScript [3]と呼ばれるスクリプト言語による内部記述により定義されている。このエージェントは、内部に音声認識・合成エンジン、および学習エンジンを持ち、「歩く」「走る」「お手」等の基本的なアニメーション動作を行う機能も兼ね備えている。今回、エージェントに犬型モデルを採用した理由は、ユーザに対して、より親しみをを持たせるためと過度の知的インタラクションを期待させないためである。

このエージェントは、ある程度の日本語を介した会話や命令の理解や発話、ユーザとのインタラクション、ユーザの手助け等が可能である。そして、ユーザは独自のエージェントを仮想空間内に任意に作成し、発信することができる。

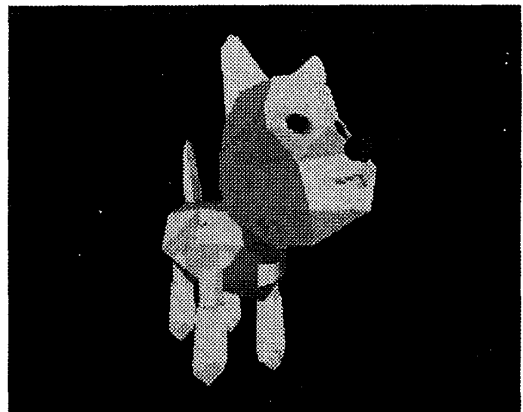


図1 犬型エージェント

3 エージェント実現環境の構成

図2に音声認識・合成技術を用いたエージェント実現環境の構成、およびエージェントの動作シーケンスを示す。エージェント・サーバは、ユーザが作成したエージェントに関するデータを管理し、各クライアント端末へのエージェントに関するデータの

入出力を行う。また、インタースペース・サーバ群 [4]に関する説明は、紙面の都合上、割愛する。なお、本研究に際し、NTT ヒューマンインタフェース研究所音声情報研究部で開発した音声認識エンジン REX および音声合成エンジン FLUET を使用した [5][6]。

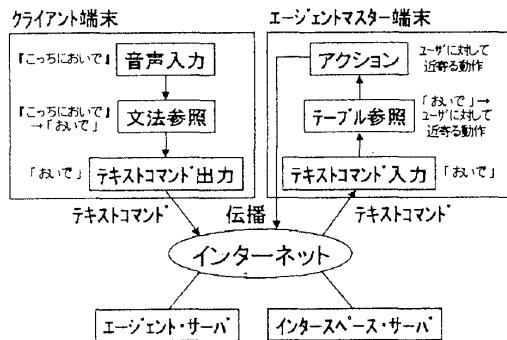


図2 エージェント実現環境の構成
およびエージェントの動作シーケンス

図2に示すように、例えば、ユーザが『こっちにおいで』と発声し音声入力を行うと、音声認識の対象であるエージェントが持つ文法により「おいで」と認識され、「おいで」というテキストコマンドがエージェントマスター端末に送信される。ここでいうエージェントマスター端末とは、共有動作を行うエージェントの動作を他のクライアント端末に伝播させる役割を受け持つ端末のことであり、エージェントの動作を伝播させる処理負担をネットワーク上に分散させる意味において、各クライアント端末に担当させることにした。エージェントマスター端末の割り当ては動的に行われ、その際に、エージェントが持つアクションテーブルを、エージェント・サーバから同時に受け取る。

テキストコマンドを受信したエージェントマスター端末は、アクションテーブルに基づいて、テキストコマンドに対するアクションを選択し、該当するアクションを開始させる。図2の場合では、テキストコマンド「おいで」に対する「ユーザに対して近寄る動作」を選択し、エージェントに対して「ユーザに対して近寄る動作」を開始させ、このアクション

を他のクライアント端末に伝播させる。

4 おわりに

本稿では、InterSpace において、従来の受動的な仮想空間サービスとは対照的な能動的なサービスを提供するためにエージェント技術を応用した仮想空間サービスを実現するための一環境について述べた。

エージェント技術と音声認識・合成技術を融合することにより、3次元仮想空間における新たな能動的サービス形態を提供することが可能となる。両技術の更なる発展に伴い、仮想空間内のバーチャル店員と対話しながら楽しくショッピングできるようなサービスが期待される。

今後は、システムの機能整備・向上を行い、システムの評価実験・考察とシステムへのフィードバックを行いながら、エージェント技術と音声認識・合成技術を融合した能動的サービスの有効性を検証する予定である。

参考文献

- [1] Sugawara, S., et al.: "InterSpace: Networked Virtual World for Visual Communication", IEICE Transactions on Information and Systems, Vol. E77-D, No. 12, pp. 1344-1349, December (1994)
- [2] 石田亨: "エージェントを考える", 人工知能学会誌, Vol. 10, No. 5, pp. 663-667 (1995)
- [3] 松浦宣彦, 菅原昌平: "共有仮想空間における動的環境制御記述言語に関する研究", 情処研報, グループウェア 19-2, pp. 7-12 (1996)
- [4] 清末悌之, 菅原昌平, 加藤洋一, 山名岳志: "インタースペースのアーキテクチャ", NTT R&D, Vol. 47, No. 4, pp. 459-464 (1998)
- [5] 野田喜昭 他: "音声認識エンジン REX の開発", 信学春季総大, D-14-9, pp. 220 (1998)
- [6] 箱田和雄 他: "波形合成法を用いたテキスト音声合成ソフトウェア (FLUET)", NTT R&D, Vol. 45, No. 10, pp. 1005-1010 (1996)