

動作も不可能であって、人間とのインタラクションを行なうことが出来るものではない。

これに対し、本研究では、実働するヒューマンインタフェースシステムに於いて、視線、表情、および身振りといった非言語メッセージを、ユーザとの通信路の確立、維持などといったメタコミュニケーションに利用している。つまり、非言語メッセージの社会的な機能に着目した利用を行なっていることが、本研究の一つの特徴である。

非言語メッセージの社会的機能は、通常人間同士の会話では、無意識的に利用され効果を発揮している。そのため、これを適切にインタフェースに組み込むことは有益であるだけでなく、そのシステムを利用するユーザにとって、認知的コスト増加をほとんど生ぜず、かつ学習不要であるという利点を持つ。

さらに、擬人化インタフェースに対しては、[Reeves&Nass96]の主張する以下の性質がより強くなり、この性質への対応能力の欠如は、そのままそのインタフェースの重大な欠点となる。

ユーザはシステムに対して社会的ルールを無意識的、自動的に適応し、かつシステムが社会的であることを期待する。

[Reeves&Nass96]

つまり、擬人化エージェントを用いたHIの社会的能力は、むしろ必須の要件である。そして、本システムは、その社会性の一部を実現したものであると言える。

実現上の観点からは、本システムは、視線検出、音声認識、および3D-CG生成と言った処理負荷の高いモジュールを複数利用して実現しているが、図6および図7に示したように必要最小限の状況だけで各モジュールを動作させているため、全体の処理負荷を小さく押えることに成功している。

なお、本技術の応用先としては、駅務/金融端末や、エンタテインメント/エデュテインメント領域などが想定できる。

4.2 知見

本システムの試用によって以下の知見を得た。

まず、本システムでは、ユーザが他の人物と(音声で)対話している途中であっても、自然にシステムへの音声入力を行なうことが出来た。つまり、ユーザが、例えば他の人物へことばで説明を行なうなど、システムへの入力以外の意図で音声を使っている状態で、本システムへの音声入力が必要になった時に、説明相手である人物の顔からエージェントCGへ視線を移すことで、特別な切替え操作無しで、音声コマンド入力によるアプリケーション操作を行なうことが出来た。また、本システムでは、ユーザが他の人物と音声で相談しながらワードプロセッサソフト(MS Word)を利用した(キーボード入力による)文書作成を行なっている状態で、ユーザがエージェントを注視し音声コマンドを発声することで、作業を中断することなく、表示形式の変更などを行なうことが出来た。

これらは、従来の音声インタフェースでは不可能であったことであるが、実際の使用状況ではよく起こり得る状況であると考えられるため、本システムの有用性を表すものであると考えられる。

また、オフィスなど、断続的な周囲雑音のある状況下では、ユーザが雑音の少ない時を見計らって、エージェントを注視することで、余分な操作無しで適切に音声入力を行なうことが出来た。また、ユーザがシステムに対して音声入力を行なっている間に、例えば他の人物がユーザに近付いてきたり、あるいは話しかけるなどしてきた場合にも、ユーザがその人物の方向を向くだけで、自動的に音声入力の受付が停止され、その人物との対話のための発話によるシステムの誤動作を回避することが出来た。さらに、その人物との会話が終了した時点で、ユーザがエージェントの方向を向き直すことで、特別な操作なしでシステム音声入力を再開することができたなど、本システムの有用性が確認された。