

## 多人数チャットシステムにおける人物モデルの ジェスチャー機能

古野 文一      松並 勝      碓崎 賢一

九州工業大学 情報工学部

我々はネットワークと仮想現実システムを利用した多人数参加型のコミュニケーションシステム Virtual Community を提唱している。Virtual Community におけるチャット通信では、参加者を模した CG モデルが、必要に応じて仮想空間内の情報を参照しながら会話の内容に即したジェスチャーを行なう。そのため、仮想空間内の周囲の状況が話題に反映され、また実空間での会話のような流れや雰囲気を実現することが可能である。本稿では、仮想空間におけるコミュニケーションの言葉によらない意思伝達法として、CG 人物モデルの自動的なジェスチャーの実現方法を提案する。

## Gesture Control of 3D Human Model for Virtual Space Chat

Fumikazu FURUNO, Masaru MATSUNAMI and Ken'ichi KAKIZAKI

Department of Computer Science and Electronics,  
Faculty of Computer Science and Systems Engineering,  
Kyushu Institute of Technology, Iizuka, 820 Japan

We propose a communication system "Virtual Community" which serves the users an unexpected encounter with other users, and enables to communicate with so much presence. Using the chat system in Virtual Community, the participants of the service are represented by 3D CG human models and these models act according to the context of what they talk about. It can make the users recognize the atmosphere like an actual conversation. This paper proposes a method to make an automatic gesture of CG human model as the non-verbal interface for Virtual Community.

## 1 まえがき

パソコン通信やインターネットが大きな注目を集め、それらの利用者の増加と多様化が急速に進展しつつある。ネットワークサービスは、多くの利用者が同時に利用しており、利用者間のコミュニケーションを支援する基盤が整っている。しかしながら、現在のネットワークサービスでは、他の利用者の存在を感じることや、偶然に出会うことなどはなく、その大きな潜在能力をコミュニケーションに有効に利用できるようになっていない。

我々は新世代のネットワークサービスシステムとして、Virtual Community [松並 95] の研究と開発を進めている。Virtual Community では、コミュニケーションと情報検索機能に重点をおき、その実現のために、3次元グラフィックスを利用したサービス用の仮想空間を提供する。仮想空間内には、建物や家具だけではなく、自分や他の利用者の3次元モデルが表示され、他の利用者を含めて視覚的に情報を収集できる。各利用者は、仮想空間内で見つけ、興味を引かれた任意の利用者と会話することが可能である。本研究は、このような仮想空間内での会話に着目し、そのコミュニケーションの支援方式を提案する。

Virtual Community では、利用者間の会話は文字によるチャットで行ない、利用者の3次元モデルに会話に合わせたジェスチャーを行なわせる。これにより、会話の臨場感を増すとともに、利用者毎に異なるジェスチャーを登録することにより、仮想空間での利用者の個性を演出する。また、通常のチャットとは異なり、利用者達が同一の仮想空間を共有していることに着目し、その環境情報を利用した会話のジェスチャー方式も提案する。これらの方式により、活発あるいは興味深いジェスチャーを示しているグループを離れた場所からでも見つけることができるようになるため、話しに加わるなどの社会的な行為が自然に行なえるようになる。

## 2 仮想空間におけるコミュニケーション

ネットワークを用いた従来の通信では臨場感の再現が不十分であったため、現在では臨場感多地点

通信の研究 [岡田 95] が盛んになっている。Virtual Community では、身近な地域社会の様なコミュニケーションの場としての臨場感ある仮想空間を実現したいと考えている。このため、以下に示すような支援を積極的に行なえる方式を明らかにする必要がある。

- 仮想空間内での利用者の出会い
- 利用者の円滑な会話
- 仮想空間の環境に関連した情報検索

### 2.1 文字による会話

会話には音声を利用するのが自然であるが、一般的な利用者が安価に使用できるネットワークの伝送速度は不十分なため、ネットワークを利用した音声による会話では、伝送容量面で大きな問題を抱えている。特に、多人数による会話を実現するためには、会話データを多重伝送する必要があり、音声のようなデータ量の多い情報を利用することは困難である。このため、Virtual Community での会話は、文字を利用した文章で行なう。

文字を使って会話を行なうことは、会話データの伝送量を低減するという消極的な選択のように見えるが、我々は、ネットワーク上でのコミュニケーションの可能性を広げる積極的な面に着目している。文字による会話内容は、コンピュータによる解釈と記録が容易に行なえるという特徴を持っている。Virtual Community では、これらの特徴を積極的に利用することを考えている。

### 2.2 ジェスチャーと臨場感

Virtual Community における会話は、利用者の3次元モデルが仮想空間内で対面して行なう。会話中に、各利用者のモデルが停止しているのは非常に不自然である。我々が会話を行なう際には、意識するか否かに関わらず何らかのジェスチャーを行ない、意志表示を補強している。ジェスチャーは人間の意志表示手段として注目されており、計算機に対する指示への応用 [今野 95] などが研究されている。ジェスチャーは自己表現として重要であり、Virtual Community で自然な会話を行なうためには、図 1 に示すように利用者の3次元モデルが適当なジェスチャーを示せる必要がある。

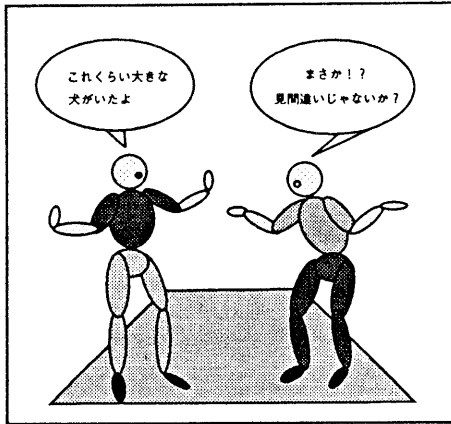


図 1: 会話に基づくジェスチャー

3次元モデルにジェスチャーを行なわせる場合、会話と並行してジェスチャーを制御する操作を行なうのは面倒であり、むしろ円滑な会話を阻害してしまうという問題がある。特に、会話が活発になればなるほど、必要なジェスチャーが頻繁になるにも関わらず、操作に割ける時間は少なくなるという矛盾がある。このような問題を解決するため、会話内容に基づいて利用者の3次元モデルを適切に動作させる方式を提案する。この方式では、会話が活発になればそれに応じて、活発なジェスチャーが行なわれるようになるという特徴がある。

会話の臨場感を向上させるためには、会話内容に対応したジェスチャーだけではなく、会話の流れを制御する人のしぐさも実現する必要がある。例えば、話の聞き手は、話題に興味がある場合には、話し手やその話題の対象を見つめることにより、話しを聞いていることや、それに関心を持っていることを表現する。また、適切な位置であいづちを打ち、円滑な会話が行なえるように会話のペース制御を自然に行なっている。直接対面していない利用者間の会話を円滑に行なえるようにするためには、このようなしぐさも表現できることが望まれる。

### 2.3 環境の認識とジェスチャー

ネットワーク上の会話（チャット）システムとしては、talk や IRC などがある。従来のチャットで

は、文字だけが情報伝達手段であり、同一環境の共有を前提として会話を行なうことはなかった。一方 Virtual Community では、利用者達が同一の3次元空間に存在し、その環境を共有しているという点で従来のチャットと大きな違いがある。

Virtual Community におけるチャットでは、共有する仮想空間内にある物事を話題の中心に据えた会話が頻繁に生じるものと考えられる。したがって、会話に基づくジェスチャーは、仮想空間の事物の指示を考慮して行なわれることが望まれる。例えば、「あその白い犬が...」という様な発言があったとすると、図2に示すように、発言者の3次元モデルが対象となる白い犬を指差すとより臨場感が増すことになる。環境情報を考慮したこのようなジェスチャーが行なえると、仮想空間への一体感と利用者間の連帯感を向上させることができ、より良いコミュニケーションを行なえるようになる。

このようなジェスチャーも、指示対象を明示的に指定するなどの、会話とは別の付加的な操作が必要になると面倒であり、かえって会話を阻害する可能性がある。したがって、このようなジェスチャーを会話の内容に従って行なう方式を提案する。

### 2.4 ジェスチャーに対する反応

臨場感を向上させるためには、発話者によるジェスチャーが行なえるだけでなく、聞き手がその動作に対する自然な反応を行なえることが必要であ

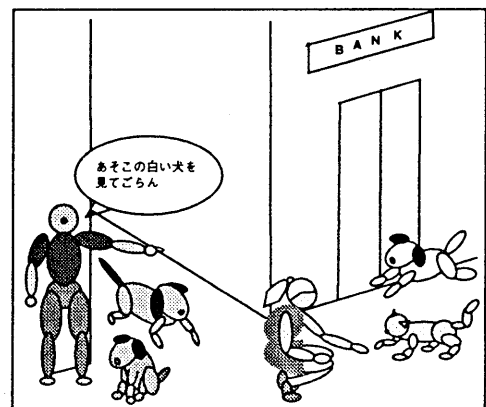


図 2: 環境を考慮したジェスチャー

る。例えば、ジェスチャーが何かを指している場合、その対象を自動的に視野に入れる機能が考えられる。このような機能により、話題の対象を自然に捉えることができ、会話をスムーズに進めることができるようになる。相手のジェスチャーに対する、このような反射的な動作を実現する方式についても提案する。

### 3 会話に基づくジェスチャー

ジェスチャーの動作は、単語に関連付けて定義し、会話文に含まれる単語に対応して実行させる。会話文にジェスチャーが定義されている単語が複数含まれる場合には、ジェスチャーが定義されている単語の出現順でその動作を行なう。

#### 3.1 会話文の特徴と基本方針

入力された文章から人物モデルの動作を決定する方式が提案 [舟渡 93] されているが、入力文は予めある程度の規則や形式が規定されており、処理に時間をかけても正確な意味解析と動作を要求するものであった。それに対して我々は不規則な文章の実時間での解析を目的としている。

会話の内容を正確に反映させたジェスチャーを行なわせるためには、会話内容を厳密に解析する必要がある。しかしながら、会話文には主語や目的語などの省略が多く、文法書などに記載されているような模範的な文は少ない。また、文は実時間で入力され推敲は行なわれないため、入力間違いが多く日本語としての文型を満たしていないものが多い。したがって、正確に解析を行なうことを試みても、その成果を期待することはできない。

そこで我々は、正確な解析を行なうことよりも適度の解析を高速に行ない、その情報をジェスチャーに利用する方式を選択する。会話内容に適合しない不適切な動作が生じる可能性があるが、人間のジェスチャーがいつも正確な意味を反映しているわけではないため、ジェスチャーに多少場違いなものがあっても良いと考えている。ジェスチャーに場違いな動きがあった場合でも、正確な内容は会話の文章を参照すれば確認することができる。また、変なジェスチャーが生じた場合にはそれ自体が新たな話題となり、会話を促進する効果を持つ

ものとして積極的に評価することも可能である。

#### 3.2 会話文の解析

会話内容を解析するためには、入力された会話文を文節毎に区切る必要がある。我々は実時間で入力される文を処理対象とするため、文の入力に利用される仮名漢字変換機能を積極的に利用することにより、文節の区切りを容易に行なうことができる。例えば、仮名漢字変換に Wnn のライブラリを利用することにより、文節単位に区切られた変換結果を入手することができる。仮名漢字変換を利用した文節解析では、誤って区切られた文節は利用者によって正しく調整されるため、簡便だけでなく正確であるという特徴がある。

文節が明らかになれば、助詞や活用を考慮して、単語の品詞を調査することができる。システムの解析部はまず文節からジェスチャーのキーワードになる単語を捜し出す。ジェスチャーのキーワードの多くは、感嘆詞、動詞、形容詞、などの状態を表す単語である。Wnn の日本語変換辞書を参照すれば、単語の品詞情報を得ることができる。キーワードが名詞であればそれにかかる形容詞、動詞であれば文中の副詞を参照して、動作の詳細を決定する。

#### 3.3 動作の大きさ

ジェスチャーによっては、動作を行なう場合でも表現したい程度により動作の大きさや速度が異なる。そのような程度の大小は、例えばキーワードにかかる「凄く」などの形容詞や副詞に依存する。程度を表す単語には、それがかかるキーワードに対応するジェスチャーがどのように変化するかを記した情報を持たせ、動作決定時に参照する。

会話文の解析結果、同じジェスチャーが繰り返し出てくる場合には、1つの大きな動作に置き換えてジェスチャーを実行する。例えば、「大きな」という単語にジェスチャーが定義されている場合には、「大きな大きな」という入力に対してその動作を2回行なうのではなく、強調された1回の動作に置き換えて実行する。

### 3.4 会話の相互作用

我々は、会話相手の発言に反応してジェスチャーを行なうことがある。このようなジェスチャーを実現することによって、より自然な会話の雰囲気表現できる。相手が発する単語に関してあらかじめ対応する動きを自分で定義しておくことで、各自の個性を演出することが可能である。ただし、過度の設定は、人の言葉で自分のモデルが過剰反応を示すことになるので注意が必要である。

### 3.5 会話の流れの制御

従来のチャットでは会話文を入力し終って確定するまで、入力操作が行なわれていることが会話相手にわからないため、同時に何人も話し始めて流れが混乱するといった問題があった。本方式では、利用者が会話文を入力し始めた場合、その情報を利用して、利用者のモデルに会話を制御するジェスチャーを行なわせる。例えば、モデルに手の平を肩の高さぐらいにあげさせ、「今から話しますよ」という注意を促す動作をさせる。このような動作により、発言権を自然な形で管理し、円滑な会話を実現することができる。

## 4 環境を対象としたジェスチャー

Virtual Community のチャットでは、会話内容に仮想空間内の物体が含まれている場合には、その対象を指さすジェスチャーを行なう。この動作を行なうためには、会話内容に仮想空間内の物体が含まれていることを検出し、その対象を特定する必要がある。

### 4.1 指示対象の検出

会話に含まれる名詞が指示対象の候補となる。会話に含まれる名詞の中には、仮想空間とは関係のないものも含まれるため、仮想空間内にあつてジェスチャーで指示すべき対象を絞り込む必要がある。

ジェスチャーによる指示対象となる名詞は、それに先立って、次に示すような位置を示す修飾が行なわれているものとする。

- 方向

右の、左の、前の、後ろの、斜め前の、右にある... など

- 距離

その、この、そのの、あそのの、むこうの、あっちの、そこにある... など

これらの修飾が行なわれている名詞を指示対象とすることにより、仮想空間内の物体が指示対象となっていることを検出することができる。

### 4.2 対象の特定

会話文である対象を指示する場合、次に示すような情報の組合せで表現される。

- 名称 (固有名、一般名)
- 特徴 (形状、大きさ、色、所有者)
- 位置 (方向、距離)

これらの情報を利用して、仮想空間内の物体のデータベースから対象を検索する。データベースには各物体の属性としてこれらの情報が記述されており、指示された複数の情報を利用して、指し示すべき物体を絞り込む。

#### 4.2.1 名称

会話内の単語に名詞が含まれている場合には、仮想空間内の物体を管理するデータベースから、その名詞が指し示す物体を検索する。データベースの検索順序は、固有名詞、一般名詞の順で行ない、これらで見つからない場合には、シソーラスを利用して、同一の物体を指す名詞を用いて検索を行なう。一般名詞として検索された対象は、多くの場合複数の候補が上げられるため、特徴や位置情報によって対象の絞り込みを行なう。

#### 4.2.2 特徴

指示対象物を特定するための特徴は、「黒いコンピュータ」の様に、対象を表す名詞に先立ち、その修飾として形容詞などで与えられる。色、大きさ、形などの物体の特徴は、物体の属性としてデータベースに記録されている。これらの特徴は、仮想空間の構築者が明示的に記録しなくても、物体の形状定義情報から取得することが可能である。

### 4.2.3 位置

あそこの、近くの、などの修飾は物体の存在範囲を示す性質を持つ。例えば、「この、その、あの」という連体詞で対象が修飾されている場合、この順に、手の届く位置、少し離れた位置、かなり離れた位置にある対象を示していると考えられる。したがって、位置情報を加味することで、後に続く名詞の検索範囲を限定することができ、その結果対象物体の特定を行なうことができるようになる。

### 4.3 対象物の確認

対象の特定を行なった結果、指示対象の候補が複数残る場合がある。この場合には、候補対象となる物体は、その優先順位に従い、自分の視野の中で丸で囲って表示される。発話者は、対象を特定したい場合には、候補の中から自分が着目している対象を選択し確認する。

対象を確認した場合や、候補が1個しか存在しなかった場合には、ジェスチャーによる指示はその対象の方を正確に向くように動作する。また、候補が複数あるにも関わらず対象を確認しなかった場合には、それらの候補が散在する中心方向を指示する。このジェスチャーにより、話題の対象を明示することができる。

### 4.4 注視点の自動的な移動

利用者の3次元モデルの注視点を会話に伴って操作することは面倒である。したがって、会話に合わせて注視点を制御する方式が必要となる。注視点の移動としては、次のようなものが考えられる。

- 複数で話している場合など、発言者に注目
- 発言者が指示した対象に注目
- 相手の手など、大きく動いた部分に注目

発話者が指示した対象は、発話者によって確認された対象であれば、聞き手の視野内でその指示対象が強調表示され、そうでなければその方向が知らされる。これにより、話題の対象を容易に認識することができるようになる。発話者や聞き手は、指示対象やその候補を選択して、その対象の名称、特性、位置などを表示させ、情報の取得や対象の確認に役立てることができる。

動きに注目する場合には、回りの人や動物の小さな動作によって注視点が安定しないことを避けるため、ある程度の大きな動作に限定する必要がある。このため、ある動きの角速度あるいは角移動量が、ある値より大きくなった場合に注目するようにする。境界値は、動作対象が小さな場合には大きく、大きな場合には小さく設定される。また、自分の前半球と後半球でも境界値が異なり、前半球の方が境界値が小さくなる。

## 5 むすび

仮想空間内での対話を支援する利用者の3次元モデルのジェスチャー方式を提案した。現在、本論文に示した基本方式に基づき、システムの設計と実現を進めている。今後の課題としては、会話文の解析の改良を考えている。

## 謝辞

本研究は、立石科学技術振興財団、電気通信普及財団の援助によっています。両財団のご厚意に感謝いたします。

## 参考文献

- [岡田 95] 岡田謙一, 松下温: “臨場感のある多地点テレビ会議システム:MAJIC”, 情報処理学会 論文誌, Vol. 36, No. 3, pp. 775-783 (1995).
- [今野 95] 今野潤, 八木正紀, 山内英明, 安村通晃: “ジェスチャー入力によるマルチモーダルインタラクティブシステムの試作”, ヒューマンインターフェース研究会資料 95-HI-60, 情報処理学会 (1995).
- [舟渡 93] 舟渡信彦, 吉川耕平, 花田恵太郎, 宮本雅之: “日本語シナリオからのアニメーションの生成”, 情報処理学会論文誌, Vol. 34, No. 6, pp. 1258-1267 (1993).
- [松並 95] 松並勝, 碓崎賢一: “Virtual Communityの提案”, グループウェア研究会資料 95-GW-13 (5), 情報処理学会 (1995).