

# 没入型会話空間における知識共有

## Information Sharing in Immersive Conversational Space

星野 准一<sup>1</sup>

<sup>1</sup> 博士（情報科学） 筑波大学機能工学系講師 (E-mail:jhoshino@esys.tsukuba.ac.jp)

本稿では、没入環境において利用者と仮想人物が会話することによって、事故や災害などの安全に関する知識を体験的に得ることができるシステムについて述べる。本システムでは、利用者が大地震などの災害が起こった都市空間に参加して、仮想的な住人と会話することで、災害時に苦勞した体験や、どのような教訓が得られたかを聞くことができる。本稿では、没入型会話空間の概念を説明すると共に、システムを構築するために必要となる対話生成技術について述べる。

**キーワード：**没入型会話空間、バーチャルヒューマン構築基盤、会話動作生成、知識共有

### 1. はじめに

我々の日常生活では新聞やテレビなどのストーリーメディアによって、遠隔地で起きている事故や、災害時の出来事に関する情報を得ることができる。これらのストーリーメディアは社会基盤として重要な役割を果たしているが、新聞やテレビから得られる知識は、自分の生活に対して直接的に影響があると受け取られないこともあり、個人のレベルでは教訓が活かされにくい。また、大きな事故や災害の出来事は、事例集の形でアーカイブ化されたり、専門的な観点から分析して報告書にまとめることが行われるが、これらの専門知識は一般の利用者にとっては敷居が高く、自分の日常生活と直接的に関連付けが行われないことも多い。

一方で、日常生活において直接的に知識が得られる場面を考えてみると、自分が当事者として出来事に関わったり、その出来事を体験した人から話を聞くことで、様々な知識が得られている。例えば、ある大きな事故や災害が起こった場合は、その災害空間で生活する人々と会話をして、どのように大変だったのか、どのような教訓が得られたかなどを聞くことで、周囲の状況と関連付けて理解したり記憶することができる。また、単に説明を聞くだけでなく、疑問点を問い合わせたり、自分が経験した出来事を物語として他の人に説明するなど、能動的な知識の活用も行っている。

本稿では、没入環境において利用者と仮想人物が会話することによって、事故や災害などの安全に関する知識を体験的に得ることができるシステムを提案する。本システムでは、事故や災害などのある特定の状況を再現



図1 没入型会話空間の生成

して、利用者が仮想人物に能動的に働きかけることで、新しい原体験を得ることを可能にする。例えば、大地震が起こった都市の生活空間の中に参加して、その町の住人と会話することで、地震によって苦勞した体験や、どのような教訓が得られたかを聞くことができる。また、自分が当事者となって行動することで、自己の行動のコンテキストと一緒に記憶することができる。

本稿では、没入型会話空間のコンセプトを説明するとともに、システムを構築するために必要となる対話生成技術について述べる。まず、ビデオ映像から得られた会話事例を蓄積して再構成することで、表現力の高い会

話動作を生成する手法を示す。次に、仮想人物が移動しながら利用者と会話することができる移動型会話空間を生成する手法を述べる。最後に、災害地の住人から災害の様子を聞いているシーンを生成した例を示す。

## 2. 没入型会話空間の生成

図1に没入型会話空間の概念図を示す。没入型会話空間では、事故や災害が起こった町の生活空間を生成して、その町の住人と会話することによって、事故や災害時の状況や、その体験によって得られる教訓を当事者から聞くことができる。本章では、没入型会話空間の基本コンセプトと、システムの構築によって期待される効果について述べる。

### 2.1. 日常的文脈との関連付け

本システムでは従来の情報都市モデルのように抽象化されたメタファとしての都市を考えるのではなく、事故や災害時のようにある特定の状況を再現して、その状況内で行動する登場人物との会話を行う。会話における日常的文脈は、いくつかの階層でとられることができる。

- 1) 空間的文脈：会話が行われる環境での物体の配置や場所と、出来事との関係
- 2) 時間的文脈：ある出来事が起こる経緯や、歴史的な背景
- 3) 社会的文脈：自分と相手との対人関係や、集団、組織の関連

安全に関する知識を多様な日常的文脈の中で関連付けて理解することで、我々の日常生活の中で類似した状況が起こったときに適用しやすくなると考えられる。

### 2.2. 体験的な知識の獲得

利用者がある出来事を体験する方法としては、主に次の2つの形態が考えられる。

- 1) ある出来事の当事者となる場合

利用者が事故や災害が起こっている場面に遭遇して、当事者の一人として行動する。利用者の周囲では、事故や災害の状況を再現するように、様々な出来事が起こったり、仮想人物がストーリーに沿った行動を行う。この場合、利用者の心理効果は高いが、リアルタイムで状況が進んでいくため、その場の状況を総合的に理解するための時間的な余裕があるとは限らない。

- 2) ある出来事の当事者から話を聞く場合

ある出来事が起こったしばらく後に場面を設定した場合、その出来事に関する経験が整理されているので、

まとまった状態で話を聞くことができる。例えば、事故や災害のときに苦労したことや、何が問題だったのかを登場人物から聞くことができる。

この2つの参加形態を統合的に組み合わせることで、自分が体験した出来事を知識として定着させることを容易にする。我々も日常生活では、自分が体験した出来事を物語として他の人に語ることで、その出来事に関する記憶を整理したり、理解を深めることを行っている。自分が当事者としてある出来事に参加することに加えて、他の登場人物の経験を聞いたり、自分が説明するなどの能動的なプロセスによって、有機的に知識を広げることができる。

### 2.3. 暗黙的な知識の明示化

事故や災害が起こった土地では、災害時の出来事が伝聞の形で伝達されて、人々の間で共有された知識となって蓄積される。ところが、これらの知識は、原体験を共有している人々の間には伝わりやすいが、原体験を持たない人々には断片的にしか伝わらないことも多い。また、専門家が持っている知識についても、あるコミュニティの中では常識的な事柄でも、その専門外の人にとっては理解の土台となる知識や経験がないため伝わりにくいこともある。

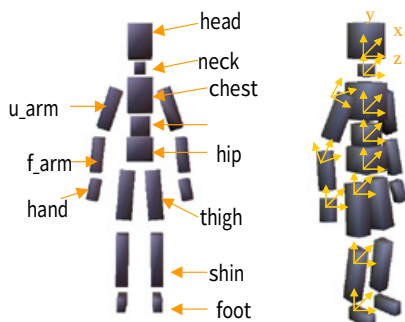
本システムでは、災害時の状況をイメージとして共有するとともに、会話空間に参加するときの「役柄」を設定することで、暗黙的な知識の共有を促進する。利用者は、町の住人から話を聞くだけでなく、町の住人の役柄になって、その出来事を知らない利用者に説明することもできる。例えば、その事故や災害を実際に体験したことがある人や、その出来事に詳しい専門家が、何も知識を持っていない人に対して説明するという役柄で参加することもできる。

### 2.4. 逐次的な知識の蓄積

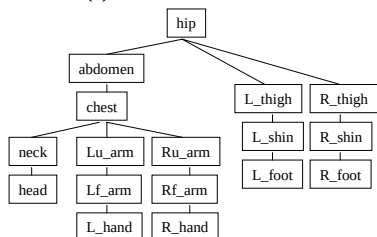
没入型会話空間では次のように幾つかの階層で知識の蓄積を行うことができる。

- 専門家や利用者が話した内容を仮想人物が記憶して、次に他の利用者から質問を受けたときに答える。
- 専門家や利用者が新しいストーリーイベントを追加することで、一連の出来事に関する知識を蓄積する。
- 世界モデルを増設することで連続的に知識空間を広げる。

このようなシステムを実現して評価を行うためには、



(a) 各関節と座標系



(b)階層構造

図2 人物モデル

様々な基本技術が必要となる。本システムでは、利用者と仮想人物が会話をしている感覚が得られることが重要となるため、以下の章では対話生成法について重点的に説明する。

### 3. 事例に基づく会話動作の生成

空間型会話環境においては、利用者と仮想俳優の間で自然な会話を成立させることが重要となる。本手法では、会話映像から会話特徴量と動作を取得して、再構成することで、自然な会話シーンを生成する。シーンモデルを参照しながら適切な動作系列と会話を選択することで、新しい会話動作を生成する。

#### 3.1. 会話動作事例の蓄積

人物動作を計測する方法としては、一般的に磁気センサや光学式センサによるモーションキャプチャが用いられている。ところが、これらのシステムは一般的にスタジオなどの広い作業空間が必要となる。また、動作を高精度に計測するためには、高価なセンサが必要となる。手軽に入る機材で利用できるものとしては、ビデオ映像から3次元的人物動作を計測することが考えられる。従来法ではビデオ映像からの輝度値を主な情報源としているため、人物の身体部分が他の身体部分によって隠されて輝度値を取得できないとき、動作推定の結果は不安定になる。このようなオクルージョン問題を軽減する手法として、多くのカメラを用い、多視点の映像を得ることが挙げられる。しかし、多視点の映像が得られても、人物の姿勢によっては死角が発生するという問題が残る。



図3 ビデオ映像からの3次元動作の推定

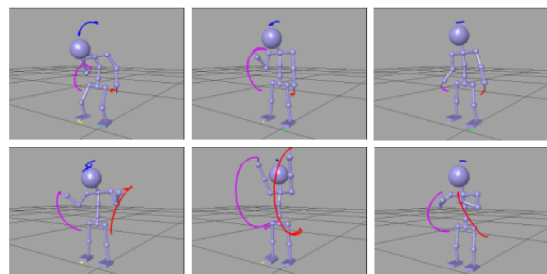


図4 会話動作事例の一部

そこで、輝度値に加えて関節駆動力を最小化することで部分的なオクルージョンに強い人物動作の推定を行う。この手法では、部分的に輝度情報が得られていない場合でも、関節駆動力を最小化する軌道を求めることができるため、より安定で正確な推定を行うことができる<sup>3)</sup>。図3に人物の3次元動作を推定しているところを示す。図4は、日常的な会話シーンのビデオ映像から典型的なジェスチャの動作セグメントを取得したところである。

#### 3.2. 非言語動作の生成

仮想人物の会話動作を生成するためには、どのような会話特徴量と動作が連動しているかを分析して再構成することが必要となる。会話動作を生成する手法としては、台詞を入力として、それに合う動作を求めることが一般的である。ところが、人間の俳優が演技をする場合でも、台詞を読んで動作を合わせるだけでは、人工的に見えるという問題がある。このように、会話動作が不自然に見える理由としては、主に次の2つが考えられる。

- 1) 会話中で参照されている物体や、話している相手に注意を向ける動作は、会話を行っているシーンの状況に依存する。これらの注意動作は、モーションキャプチャ動作を取得した状況と仮想俳優が会話している状況が異なるため、会話と連動してダイナミックに生成することが必要となる。
- 2) 会話の途中で何かを思い出したり、思考を促進するための動作は、台詞と動作が直接的に対応していないことが多い。そのため、全ての会話動作を言語的な意味を持つものとしてとらえると省略されてしまう傾向がある。



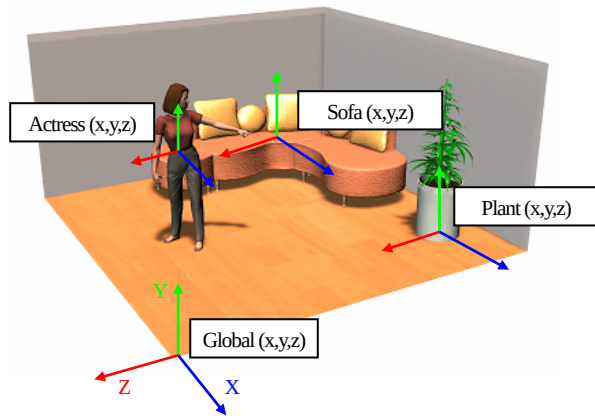


図5 仮想環境内の俳優とオブジェクト

本手法では、注意動作や思考動作のような非言語的な動作を加えるとともに、動作によって発話のタイミングを伸縮させることで、より自然な会話動作の生成を行う。

### 3.3. 会話動作の選択

本手法では、まず、連続した一区切りの会話動作を、会話動作ユニットとして蓄積する。会話動作ユニットは、動作を適用するための条件判定を行う条件部と、動作パターンを記述する動作生成部から構成される。例えば、「これ」「あれ」などの指示語に対応する指示動作や、図像的動作、メタファ的動作のように特定の言葉に対応する会話動作を蓄積する。また、会話の途中でシーン中の物体を参照するときには、その物体に対して視線を向けたり指差しをするなどの注意動作を付加する。本システムでは図5に示すように、シーン内の大まかな物体の配置と3次元位置を仮想人物の知識として持つ。これらの知識は固定されたものではなく、仮想人物の仮想センサによって動的に更新することができる。

会話動作を生成するときには、会話動作ユニットに記述された条件判定を行い、動作生成部に記述された動作を生成する。次に、注意動作や思考動作を加えて、それぞれの動作間隔に従って発話のタイミングを調節する。例えば、会話の途中である物体を参照しようとして、記憶している場所に無ければ、探索動作を起動する。また、会話の途中で大きな音がすれば、会話を中断して音がする方向を確認する動作を行う。

## 4. 会話と連動した移動動作の生成

### 4.1. 会話と移動機能

仮想人物が歩きながら会話をする場合は、会話と連動して移動を制御することが必要となる。例えば、ある建物の構造や事故現場を説明しているときには、全体的な話をするときには遠いところから指差しなどをして説明

するが、より細かい対象物を参照しながら説明するときには、その対象物に歩きながら近づく動作を行う。また、移動している途中でも説明を加えるなどの会話を行うことが必要となる。

会話の途中で仮想人物をある位置まで移動させるときに、その都度、コンテンツ制作者が環境中の物体の配置や、他の登場人物の位置関係を考えながら歩行経路を指定することは効率的ではない。特に、本システムでは仮想人物が相互に能動的に働きかけることができるため、その時々仮想人物の状態を検出して、移動を行うことが必要となる。例えば、登場人物Aが登場人物Bに近づいて話しかけるというキーイベントを指定した場合、登場人物Aは環境内の他の物体や、他の人物を避けて、登場人物Bにまで近づいて話しかけることが必要となる。

本システムでは、会話生成モジュールの条件部に、対象物までの距離や、話し相手の距離などのパラメータを含める。動作生成部に、「対象物までの距離が遠い場合は付近まで近づく」などの移動動作を記述することで、会話と連動した移動機能を実現する。

### 4.2. 知覚領域の設定

移動機能は環境から必要な情報を得るための知覚ユニットと、移動経路を求めるパスプランナから構成される。まず、知覚ユニットには、非接触で知覚する視覚、聴覚などに相当する知覚領域と、接触を行う上限を定めた不可侵領域、反発領域を設定する。知覚可能な領域はオブジェクトや、他のアクターを認識するための領域として必要となる。また不可侵領域、反発開始領域は経路の探索や動作に影響を及ぼす。

本手法では、パスプランニングのために干渉不可な領域を設定する。干渉不可な領域は、計算を簡略化するためにアクターの中心から半径  $d$  の円形状の領域とした。なお、オブジェクトに干渉を行うための領域はアクター自身に持たせるのではなく、オブジェクト自身に干渉動作実行の際に干渉可能領域（アフォーダンス距離）を設けることにより実現する。

また、ストーリー中の環境内に用意されている対象物を認識しているかどうかを表現するために認識オブジェクトリストを作成する。初期設定において、あらかじめ認識している認識オブジェクトについては、リストに状態パラメータの位置情報を設定しておく。また、実行時に新たに認識したオブジェクトについては随時追加し、認識していたオブジェクトの状態が変化したことを知覚した場合には、リストの情報を更新する。

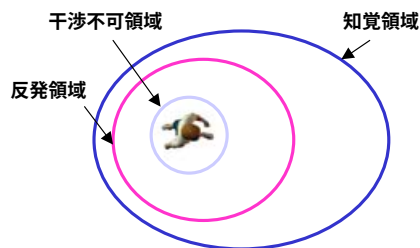


図6 知覚領域の設定

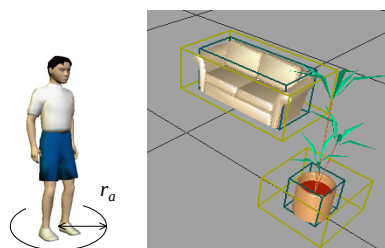


図7 非干渉領域の設定

#### 4.3. 移動経路の算出

本稿では、人が動作対象の環境を知覚した際に、通行可能な場所とそうでない場所を判別するものと考え、与えられた環境下であらかじめ多数の移動可能なパスを生成しておく手法をとることにする。移動可能なパスの生成には確率的な手法を使用することにする。

次に、移動経路の選定の手順を示す。まず初期設定として、

- (1) 3D 環境内に、ランダムに多数の移動中継点を生成する。
- (2) 生成された各点から他の点に対して、パスを生成する。

パスを生成するときは、中継点間に障害物（認識オブジェクト）があるときにはそのパスを無効にする。障害物との衝突の判定は、オブジェクト自身に簡略化した領域を設定し、アクターの不可侵領域の半径  $d$  を加えた領域で行う。以上の手順で移動可能なパスの候補である移動経路のグラフが生成される。

実際にアクターの移動するパスを選択するには以下の手順をとる。

- (3) 始点と終点を設定し、近傍の点の内で連結可能な点にパスを生成する。
- (4) ダイクストラ (Dijkstra) 法を使用して始点、終点間の最小経路を求める。グラフの重みは空間中の距離を使用する。

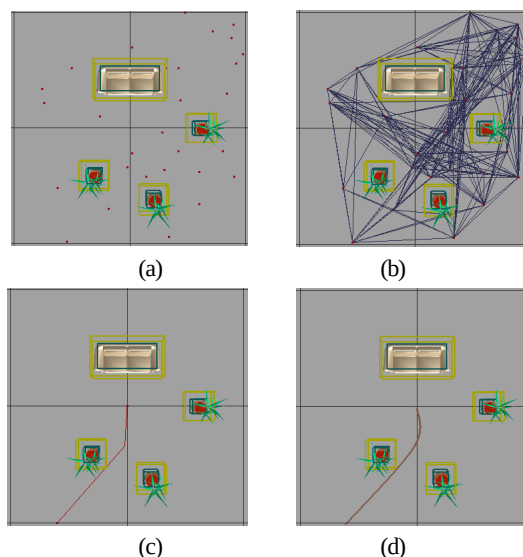


図8 移動経路の生成過程

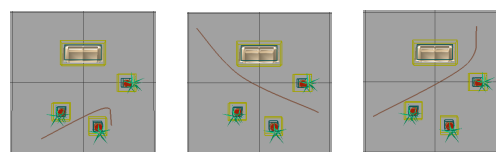


図9 移動経路の例

- (5) 求めた経路を Catmull-Rom スプライン曲線で移動動作の歩幅を考慮に入れてスムージングし、算出された曲線を移動パスとする。

即興的な動作が起きて、経路を外れた場合には、外れた地点をスタート地点として、再度経路を計算する。また、よろけたり、転ぶなどの能動的でない動作による空間内の移動情報は動作自体に含むものとする。

#### 4.4. 会話開始動作

本システムでは仮想空間内をウォークスルーしながら様々な仮想俳優と会話するため、離れている仮想俳優と会話を始めることが必要となる。典型的な会話の開始制御は、1) 相手との距離を縮める、2) 注意を喚起する、3) 応答する、の3つのステップで行われる。このプロセスにはいくつかのバリエーションがある。例えば、

- 自分から相手に近づいていくと、相手が自分に気が付いてくれて、声をかけてくる。
- 近づいても相手が気が付かないので、声をかける。
- 迷っていると相手から声をかけてくれる。
- 遠くから相手に声をかけて、相互に近づいていく。

などの場合がある。これらの会話開始動作は、仮想センサからの出力を元に条件判定を行って選択する。



図 10 複数の仮想人物による会話動作の生成例



図 11 複数の仮想人物が移動しながら会話をしている動作例



図 12 利用者と仮想人物が会話している場面

## 5. 実験結果

3章および4章で述べた動作生成法を実装して、会話と連動した動作が生成できることを確認した。図10に複数の仮想人物が会話しているところを示す。会話の途中で犬が入ってきて、登場人物が注意を向けるが、確認した後で会話を継続する。図11は複数仮想人物が移動しながら会話しているところである。会話の過程で適切なタイミングで進行方向や、相手の方向に視線を向けていることが確認できる。また、図12は利用者と仮想人物が音声とジェスチャを利用して会話しているところである。利用者が問い合わせた対象物を探す動作をした後に、見つけた物体を指差ししながら説明を行っている。

図13は、利用者が震災都市に住んでいる友人に会いに行き、震災時の状況について話を聞いているところである。「道路が崩れて下敷きになった人がいる」「消防活動の支障になった」などの説明を行っている。移動機能によって説明対象の方向に近づいてから説明を行っている。

## 6. おわりに

本稿では、没入環境において利用者と仮想人物が会話することによって、事故や災害などの安全に関する知識を体験的に得ることができるシステムを提案した。また、



図 13 震災都市の住人が震災による影響について説明しているところ

システムを構成するための基本機能として、事例に基



づく会話動作の生成や、仮想空間内での移動機能を利用した会話の生成について説明した。今後の課題としては、実際に起こった事故や災害を例題としてシーンを作成して評価を行うことが挙げられる。

## 参考文献

- 1) 星野准一(2003)「ストーリー型インタラクティブシステム構築基盤」, 人工知能学会全国大会, 1B4-02
- 2) 星野准一(2003)「ストーリー型エンタテインメント」, 電子情報通信学会セミナー, TA-7-5, SS53-54
- 3) 中野敦, 星野准一(2002)「動力学モデルに基づくビデオモーションキャプチャ」, 日本バーチャルリアリティ学会論文誌 Vol.7, No.4, pp.471-480
- 4) 星野准一(2001). 「人物動作からの個性の推定と再構成」, 画像電子学会論文誌, Vol.30, No.15, pp.631-640
- 5) H. Mori, J. Hoshino(2002). Independent Component Analysis and Synthesis of Human Motion, ICASSP2002, IMDSP-P04.08
- 6) J.Hoshino, Y.Hoshino(2001). Intelligent Storyboard for Prototyping Animation”, IEEE Int. Conf. on Multimedia and Expo, ICME2001, Conference CD-ROM FAI.03
- 7) 高沢潤, 星野准一 (2002) 「対話的仮想シアターのためのオーサリング環境」, 日本バーチャルリアリティ学会第7回大会論文集, pp.187-188

---

## INFORMATION SHARING IN IMMERSIVE CONVERSATIONAL SPACE

Junichi Hoshino<sup>1</sup>

<sup>1</sup>Ph.D. (Information Science) Assistant Professor, University of Tsukuba (E-mail:jhoshino@esys.tsukuba.ac.jp)

This paper describes the information sharing method in immersive conversation space. The system generates the virtual city after the disasters such as a big earthquake and fires. Users can obtain various knowledges through conversations with habitats in the virtual city. We also describes the example-based conversational gesture generation

**Key Words:** *immersive conversational space, virtual human platform, conversational behaviors, information sharing*