

パノラマ画像による没入型会話環境

Panorama-Based Immersive Conversation Space

星野 准一¹ 張 磊² 中野敦²

¹博士(情報科学) 筑波大学システム情報工学研究科講師 (E-mail:jhoshino@esys.tsukuba.ac.jp)

²筑波大学システム情報工学研究科 (E-mail:{zhanglei, nakano}@esys.tsukuba.ac.jp)

本稿では、パノラマ画像による高精細な没入環境において、仮想人物と利用者が移動しながら身体的なインタラクションを行うことで、事故や災害時の知識を体験的に得られる没入型会話環境を提案する。本システムでは、パノラマ画像上に簡易に歩行領域や指示物体をアノテーションすることで、仮想人物がパノラマ環境中を移動や指示動作を行いながら事故や災害時の状況を説明できる。また、利用者からの働きかけに対して、仮想人物が身体動作を利用して答えるとともに、利用者の視野が適切に移動することができる。

キーワード：没入型会話空間、パノラマ画像、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メディアを通じて、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られる。ところ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、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ているが、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られないことも多い。

このような問題を解決する方法の一つとして、没入環境において利用者と仮想人物の会話によって、事故や災害などの安全に関する知識を体験的に得られる没入型会話システムの検討が行われている[1][7]。本稿で述べる没入型会話システムでは、事故や災害などのある特定の状況を再現して、利用者が仮想人物と身体的なインタラクションを行うことで体験的な知識を得る。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、地震によって苦労した体験や、どのような教訓が得られたかを聞くことができる。

このように教育用に仮想人物を応用した学習システムの構築は従来研究でも行われている[3]。しかし、これらのコンテンツの背景は作り込まれたものであり、災害時の風景のような複雑なCGモデルを十分に高精細に作成するためには、大きな労力とコストがかかるという問題がある。もし災害時に撮影された映像素材を簡易に利用して、疑似的に3次元環境中を歩き回るような効果を



Fig.1 パノラマ画像による没入型会話環境

生成できれば、より多様なコンテンツの制作が容易となる[5]。このため広範囲に高精細な画像情報を持つパノラマ画像の利用が考えられる。パノラマ画像とCGを組み合わせることで利用者に提示するシステムはこれまでも研究されており、高い臨場感を得られることが示されている[4][6]。

また、近年のテーマパークでは、立体視によって等身大のキャラクタを3D表示することで、テレビや映画の登場人物によるストーリーを体験できる。このように、現実感の高いストーリー体験を提供できるメディアは、今後のエンタテインメントにおいて重要な役割を果たすと考えられる[2]。

これらの背景を踏まえて本稿では、パノラマ画像による没入型対話システムを提案する(Fig.1)。本システムでは、パノラマ画像による高精細な会話空間中を、利用者が仮想人物と身体的なインタラクションを行うことで直接的な体験を可能とする。

以下、2章では仮想人物が利用者に対話するためのシステムの構成、3章ではパノラマ環境中における移動型

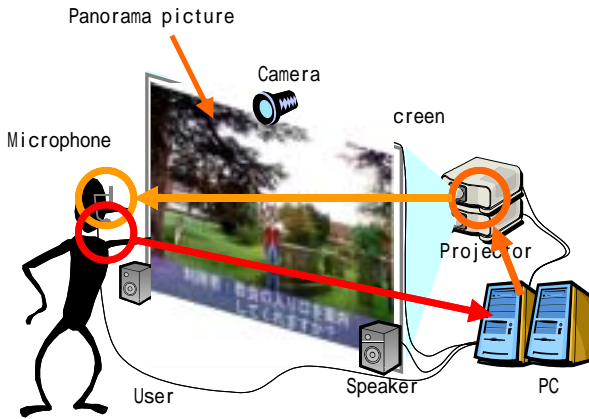


Fig.2 システムの構成

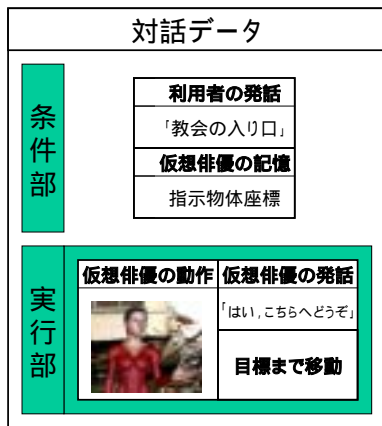


Fig.3 仮想人物の対話データ

会話の生成方法について述べる。4章では、パノラマ環境中を仮想人物が移動しながら利用者と対話する映像を示す。

2. システム構成

利用者と仮想人物がパノラマ画像をもとに直接的な体験を行うために、次のようなシステムを構成する。まず災害時の風景のような一連の場面を情報として持つパノラマ画像を背景にする。仮想人物はそのパノラマ空間を移動しながら、利用者に場面の状況を説明する。その際、仮想人物の移動に従ってカメラも移動することで、パノラマ空間中の状況を利用者は仮想人物と共に体験する効果を得る。

本システムの利用者は、仮想人物とコミュニケーションするために、ヘッドセットマイクを装着して、等身大スクリーンの前に位置し、音声対話によってスクリーンに投影された等身大の仮想人物と対話を行う (Fig.2)。また、仮想人物が利用者と対話するためには、利用者の発言に対して返答する仕組みが必要である。そのため、あらかじめ利用者の発話と説明対象の仮想人物の記憶を対応付けた条件部と説明対象の位置、その地点での指示動作や発話文といった仮想人物の行動内容を表す実行部



(a) 歩行領域の指定 (b) 領域にノードを生成

Fig.4 ノードの生成

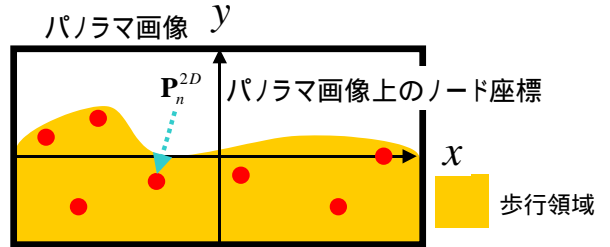


Fig.5 パノラマ画像とノード座標

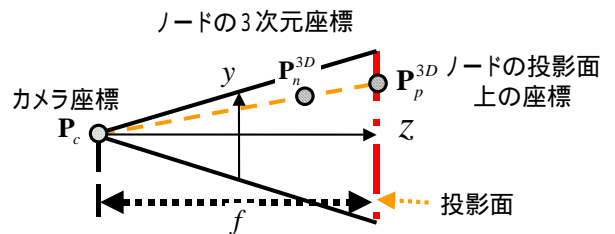


Fig.6 3次元空間のyz断面図

からなる対話データ (Fig.3) を用意する。このような対話データを多数用意し、利用者の発話文で出現したキーワードと仮想人物の記憶したキーワードとのマッチングにより、実行する対話データを決定する。これにより、対話データの実行部に記述された目標地点と行動内容をもとに、仮想人物は説明対象付近まで行き、会話動作を行いながら利用者に対象物を説明できる。

3. 移動型会話の生成

仮想人物が目標地点まで移動し、利用者に対象を説明する映像を次の手順で生成する。まず仮想人物がパノラマ空間上を自由に移動できるようにするために前処理として次のように歩行ネットワークを生成する。

1. [歩行ネットワークの生成] パノラマ画像は2次元空間であるため、その上を3次元空間に存在する仮想人物が歩行して、自然に見える映像を提示するためには、画像上にある建物などのオブジェクト上を通過しないように、仮想人物が歩行できる領域を定義する必要がある。そのため歩行できる領域を、ペイントで指定する。その後、歩行の目標地点を表すノードをペイントされた領域に生成する。それらのノードを接続したネットワーク上を仮想人物が移動することで、仮想人物がパノラマ画像上を歩行す

る映像を提示する。

2. **[3次元空間のノード座標の生成]** 1ではパノラマ画像上の2次元座標しか生成していないため、仮想人物が存在する3次元空間のノード座標を画像の2次元座標系との対応を取って生成する。

また利用者が仮想俳優と対話する際の流れは次のようになる。

3. **[説明地点の決定]** 仮想人物は利用者に対話しながら、対話データに記述された内容から説明対象の位置を決定する。その後、対象物と重ならないといった利用者に提示する映像を自然に見せるための条件を満たすノード位置を仮想人物が対象物を説明する地点とする。
4. **[視点位置の移動]** 利用者が仮想人物と共に移動し、没入型会話環境に入り込めるようにするため、視点位置を仮想人物と連動させる。
5. **[説明動作]** 対象物を指差しながら利用者に説明するために、蓄積した指差し動作パターンの中から適切な指差し動作を選択する。

本章ではこれらの項目の実装方法について述べる。

3.1. 歩行ネットワークの生成

説明時の移動プランニングを容易にするために、仮想人物が移動可能な場所をノードとした歩行ネットワークを用いる。歩行ネットワークの各ノードが背景にしているパノラマ画像上のオブジェクトと重ならないように、一つ一つノードを配置すると多量の作業時間を要してしまう。そのため、まず大まかに歩行可能な領域を指定する (Fig.4(a))。その後、領域中に一定区間毎にランダム性を持たせてノードを生成することでネットワークを構築する (Fig.4(b))。この歩行ネットワークの重みは次節で求めるノードの3次元座標間の距離とする。

3.2. 3次元空間のノード座標の生成

3.1節で生成した歩行ネットワークの各ノードは、パノラマ画像上の2次元座標を持つ。しかし、3次元空間に存在する仮想人物が2次元のパノラマ画像上を移動しているように見せるためには、パノラマ画像上の2次元座標を仮想人物の存在する3次元空間の座標に変換する必要がある。そのため本節では、2次元座標から3次元空間のノード座標を算出する方法を述べる。

本システムでは、パノラマ画像全体を利用者に提示するのではなく、その一部分を利用者に提示するため、カメラを移動すると、ノードの3次元座標は同じでも、パノラマ画像に重ねて提示した映像では、位置が異なって見えるという問題がある。そのため、カメラ座標は常に仮想人物を正面に捉え、カメラ座標に対してノード座標が定まるようにする。この条件で、カメラ座標を $P_c = (x_c, y_c, z_c)$ 、パノラマ画像上のノードの2次元座標

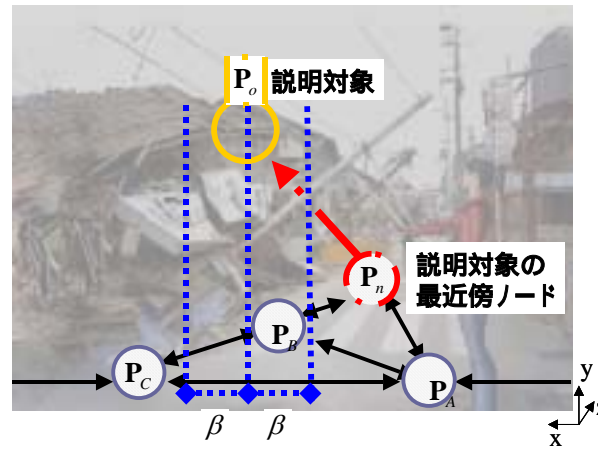


Fig.7 説明対象の最近傍ノードの探索

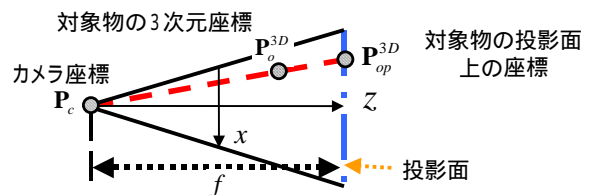


Fig.8 3次元空間のxz断面図

を $P_n^{2D} = (x_n^{2D}, y_n^{2D})$ とすると、仮想人物の存在する3次元空間の投影面座標 $P_p^{3D} = (x_p^{3D}, y_p^{3D}, z_p^{3D})$ への変換は、

$$x_p^{3D} = x_c^{3D} \tag{1}$$

$$y_p^{3D} = \frac{2C_y}{H} y_n^{2D} \tag{2}$$

$$z_p^{3D} = f \tag{3}$$

と表される。ここで f はカメラから投影面までの距離、 C_y はパノラマ画像と投影面の高さの比率、 H はパノラマ画像の高さである。

投影面座標は、仮想世界の3次元座標を2次元空間に射影するために用いる座標であるため奥行き値である z 値は等しい。そのため、この座標をそのまま3次元空間の座標として用いると、奥行きが無い映像になってしまう。パノラマ画像上に仮想人物を表示したとき、奥行き方向へ移動する映像を提示するためには、この投影面座標 P_p^{3D} を3次元空間の座標 P_n^{3D} へ変換する必要がある。カメラ座標 P_c から投影面座標 P_p^{3D} へ引いた直線の方程式は、

$$y = \frac{y_p^{3D}}{f} (z - z_c) + y_c \tag{4}$$

と表される (Fig. 6)。(4)式の z 値もしくは y 値のどちらかが決まるともう一方の値も定まる。そこで z 値を次

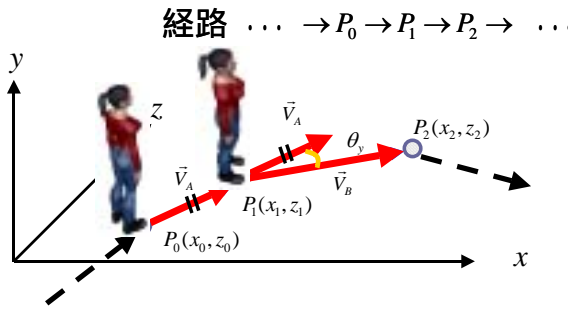


Fig.9 仮想人物の進行方向



Fig.10 仮想人物の移動に伴う背景領域の移動

式を用いてパノラマ座標上の高さ y_n^{2D} に比例した値として定める。

$$z_n^{3D} = (z_{\max}^{3D} - z_{\min}^{3D}) \frac{\left(y_n^{2D} + \frac{H}{2} \right)}{y_{\max}^{2D}} + z_{\min}^{3D} \quad (5)$$

ここで y_{\max}^{2D} はパノラマ画像上の歩行領域の上限値とする。 z_{\max}^{3D} , z_{\min}^{3D} は3次元空間で仮想人物が奥行き方向に移動できる上限値と下限値である。さらに(4)式により y 値を求める。 x 値はパノラマ画像上のノードの x_n^{2D} およびカメラ座標の x_c と同一であるため、ノードの3次元座標 P_n^{3D} が求まる。

3.3. 説明地点の決定

パノラマ画像上の対象を説明するためには、表示する画面に指示対象が入る必要がある。そのため仮想人物は対象の傍で説明を行う。対象付近のノードは、対話データにある目標地点から求める。2点間の距離を最小とする評価式(7)を用いて、

$$\left| P_o^{2D} - P_n^{2D} \right| \rightarrow \min \quad (6)$$

指示対象位置 $P_o^{2D} = (x_o^{2D}, y_o^{2D})$ の最近傍ノード座標 $P_n^{2D} = (x_n^{2D}, y_n^{2D})$ を探索する。

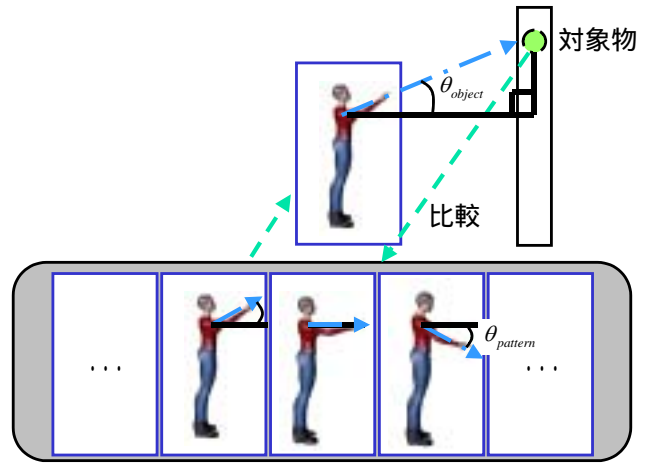


Fig.11 指差し動作の選択

仮想人物が対象物に重なると、利用者からは対象物が見えなくなり説明が分かりづらくなるため、対象物に重ならないように、以內を除く最近傍ノードを探索する条件として次式を追加する。

$$\left| x_o - x_n \right| > \beta \quad (7)$$

このノードと対応付けたカメラ座標から $P_o^{3D} = (x_o^{3D}, y_o^{3D}, z_o^{3D})$ を計算する。カメラ座標の x 値と指示対象の x 値は異なるため、ノードの場合に加えて、 x 座標方向も計算する必要がある。まず(8)式により投影面上の座標を求める。

$$x_{op}^{3D} = \frac{2C_x}{W} x_o^{2D} \quad (8)$$

$$x_o^{3D} = \frac{x_{op}^{3D}}{f} (z_o^{3D} - z_c) + x_c \quad (9)$$

次に(9)式で、対象の3次元座標を求める。

3.4. 移動経路の探索

前述したように、仮想人物はある地点で対象を説明したい場合、対象に近づく必要がある。本稿では、あらかじめ設定した歩行ネットワークを用いて説明対象に近づく。ここで述べる歩行ネットワークとは、仮想人物が静止できる地点をノードとしたネットワークを指す。

現在のノードから目的のノードへの経路はダイクストラ法を用いて求める。ダイクストラ法はネットワークのアーキに付けられた重みが最小となるように目標のノードへの最短経路を探索する手法である。

3.5. 進行方向への回転

歩行ネットワークを用いて移動シーンを生成するため、ノードを遷移する度に仮想人物は進行方向を向かな

Fig.12 パノラマ画像 1 (案内)



Fig.13 実験結果 1



Fig.14 パノラマ画像 2 (災害時)



Fig.15 実験結果 2

なければならない 仮想人物の回転角度を計算するために、ノード遷移前と後の進行方向ベクトル間の角度を、内積を用いて次のように求める (Fig.9).

$$\theta_y = \arccos(\vec{V}'_A \cdot \vec{V}'_B) \quad (10)$$

ここで \vec{V}'_A は P_0 から P_1 への正規化ベクトル, \vec{V}'_B は P_1 から P_2 への正規化ベクトルである.

3.6. 視点位置の移動

本システムでは Fig.10 のように、仮想人物の動きに合わせて、カメラも移動する. この際、利用者が仮想人物と一緒に移動している感覚を創出するために、パノラマ環境中の仮想人物の歩行速度に合わせてカメラを移動する.

3.7. 説明動作

仮想人物が対象物を指差しながら説明するためには、指差す方向を自由に変更できなければならない. あらかじめ, Fig.8 のように、上下に指差す方向の異なる動作を複数用意しておき、仮想人物と対象物のなす角度からもっとも適切な動作を次式により選択する (Fig.11).

$$|\theta_{object} - \theta_{pattern}| \rightarrow \min \quad (11)$$

4. 実験

手法の有効性を確認するために 1850[pixel] × 240[pixel] の教会周辺を撮影したパノラマ画像(Fig.12)と 2050[pixel] × 240[pixel] の災害時の状況を写したパノラマ画像(Fig.14)を用意した. このパノラマ画像に前述した手法を適用し、実際に仮想人物が利用者と対話しながらパノラマ空間中を移動し、建物の案内や災害時の体験について

話す映像を生成した。Fig.12 の実験結果では、仮想人物が利用者の「教会の入り口を案内してくれますか?」という問いかけに対して、パノラマ画像中を歩行し、利用者を入り口まで案内している。また Fig.14 では、災害時の状況を利用して、建築基準について仮想人物と利用者対話している。このように高精細な画像情報を持つパノラマ画像を背景にして、仮想人物がパノラマ画像上のオブジェクトを避けるといったあたかもパノラマ画像中の状況に存在しているような映像を提示しながら知識や教訓を提供することで直接的な体験が得られることを確認した。

5. おわりに

本稿では、パノラマ画像を用いた没入型会話環境を構築し、仮想人物が利用者に対話しながら状況を直接的に説明するシステムを提案した。仮想人物がパノラマ画像上のオブジェクトを避けて移動し、パノラマ画像を利用して直接的に状況や教訓を利用者に説明することで、直接的な体験ができることを確認した。今後はパノラマ空間に複数人の仮想人物を登場させるといった現実の状況に即した映像を利用者に提示することで、更なる会話環境への没入感の向上を計りたい。

文 献

- [1] 高沢潤, 後藤大輔, 星野准一: "対話型映画のための会話生成法", 電子情報通信学会技術報告, Vol.103, No.166, pp.29-34, 2003.
- [2] 星野准一: "ストーリー型エンタテインメント", 情報処理学会誌, vol. 44, no.8, 2003.
- [3] James C. Lester, Jennifer L. Voerman, Stuart G. Towns, Charles B. Callaway: "Coherent Gestures, Locomotion, and Speech in Life-Like Pedagogical Agents" *Applied Artificial Intelligence* 13(4-5): 383-414, 1999.
- [4] 中川雅通, 登一生, 森村淳, 佐藤宏介, 千原國宏, パノラマ実写画像と移動 CG 物体の合成, 第 42 回システム制御情報学会研究発表講演会, No.2002, pp.103-104, 1998.
- [5] 廣瀬, 渡辺, 谷川, 遠藤: "実写画像を用いた広域仮想空間構築における画像生成手法の研究", *Human Interface News and Report*, vol.13, pp.173-178, 1998.
- [6] 島村潤, 山澤一誠, 竹村治雄, 横矢直和: "全周パノラマステレオ画像と CG モデルの合成による複合現実環境の構築", 情報処理学会論文誌 : コンピュータビジョンとイメージメディア, Vol. 42, No. SIG6(CVIM2), pp. 44-53, June 2001.
- [7] 高沢潤, 星野准一: "対話型 3D 映画のための階層的動作生成法", 電子情報通信学会技術報告, Vol.103, No.515, pp.7-12, 2003.
- [1] 高沢潤, 後藤大輔, 星野准一: "対話型映画のための会話生成法", 電子情報通信学会技術報告, Vol.103, No.166,

Panorama-Based Immersive Conversation Space

Junichi Hoshino¹ Rai CHAN² Atsushi NAKANO²

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

²University of Tsukuba (E-mail {zhanglei, nakano}@esys.tsukuba.ac.jp)

We propose a system, which a virtual actor and a user move freely in the immersive environment by the panorama picture and performing by a physical interaction. In this immersive conversation environment, the knowledge at the time of the accident or a calamity can be acquired in experience. In this paper, a virtual actor performs walk and directional motion in panorama environment by annotating a walk domain or an object on a panorama picture simply. Moreover, entrainment to virtual environment is strengthened with a virtual actor not only answering using body motion, but a user's view moving appropriately to the influence from a user. This paper describes the construction technique and validity of the immersive conversation environment by using a panorama picture.

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