

社会的情報共有のための没入型会話環境の提案

IMMERSIVE CONVERSATIONAL ENVIRONMENT FOR SHARING INFORMATION IN A SOCIETY

中野 有紀子¹・西田 豊明²

¹博士 (情報理工学) 社会技術研究開発センター ミッション・プログラム 統括研究グループ
会話型知識プロセス研究サブグループ 研究員 (E-mail: yukiko@ristex.jst.go.jp)

²工学博士 京都大学大学院情報学研究科 (E-mail: nishida@i.kyoto-u.ac.jp)
社会技術研究開発センター ミッション・プログラム 統括研究グループ
会話型知識プロセス研究サブグループ リーダー

安心・安全な社会を作るうえで、専門家と市民との間での共通認識の構築が重要であり、そのためのコミュニケーションチャンネルの確立が求められている。本論文では、専門家と市民とのコミュニケーションを活性化することを目的とした没入型仮想会話環境 IPOC を提案する。IPOC では、背景画像を知覚的な共有情報としながら、専門家のアドバイスをより具体的に伝えるインタラクション機能と、市民からの現場情報をよりリアリティの高いものにするためのコンテンツ作成支援機能が装備されており、専門家と市民との双方向的なコミュニケーションの実現に貢献することが期待される。

キーワード： リスクコミュニケーション、没入型会話環境、会話エージェント、コンテンツ作成支援、ヒューマン・コンピュータインタラクション。

1. はじめに

安心・安全な社会を作るうえで、専門家と市民の間での共通認識の構築が重要であり、そのためのコミュニケーションチャンネルの確立が求められている。しかし、現実の状況に関する認識が双方で異なる場合等が生じ、具体的なレベルでリスクコミュニケーションを行うことは困難な状況である。本論文では、専門家と市民とが具体的な状況を共有しながら、リアリティのあるコミュニケーションを行える会話環境を提供することを目的とし、情報・メディア技術を利用した没入型仮想会話環境 IPOC を提案し、その実装について述べる。

本論文の構成は以下の通りである。次章では、よりリアリティの高いコミュニケーションを実現するために、情報・メディア技術がどのような側面を支援すべきかについて論じ、システムデザインに関する示唆をまとめる。3章では、没入型仮想会話環境 IPOC の概要を述べる。4章では、IPOC のインタラクション機能について述べ、さらに5章では、IPOC のコンテンツ作成支援機能について議論する。6章では関連研究について述べ、最後に7章で、結論とともに今後の課題について述べる。

2. リスクコミュニケーションにおけるリアリティの重要性

従来、専門的な知識を一般の人々に伝えることは、困難であると考えられてきた。もちろん、一般市民が専門

家と同等の知識を持つのは容易なことではないが、専門家が伝えたいこと、伝えるべきことは、市民の生活や環境に応じた具体的なアドバイスである。しかし、それでもなお意思疎通が難しい理由として、以下のような点が考えられる。

- (a) 実際には発生しているわけではない危険な状況を想像することは大変難しいため、「こうなったら危ない、要注意!」といった状況を専門家から市民に具体的に伝えることが難しい。
- (b) 一方、議論の現場となる状況について、市民から専門家に説明することも不可欠であるが、言葉だけでは説明しづらい現場の状況を専門家にうまく伝えることが難しい。

本研究では、これらの課題を解決するために、IPOC において、インタラクション機能とコンテンツ作成支援機能とを実装する。インタラクション機能では、画像・映像を提示するとともに、会話エージェントがユーザ(市民)の関心や興味を察知しながら、具体的な事例について説明する。また、コンテンツ作成支援機能では、一般に、その作成には専門的な技術が必要と言われる映像コンテンツを、写真と会話調の台本のみから自動的に生成するシステムを実装する。

3. 没入型仮想会話環境 IPOC

没入型仮想会話環境 IPOC は、パノラマ写真を背景とした会話環境上に会話エージェントが存在し、背景にある建物や対象物について短いストーリーを語ることでよりインタラクティブに説明するシステムである。例えば、日本の古い町並みが背景となる場合には、エージェントはユーザに背景にある建物やそれに関連する歴史的な事柄について会話的に説明を行う。つまり、ユーザは会話エージェントとの没入的なインタラクションを通して、背景世界についての知識を得ることができる。Fig. 1 に IPOC とユーザとのインタラクションの様子を示す。以下の章では、IPOC のインタラクション機能と、コンテンツ作成支援機能について、それぞれ詳細に説明する。



Fig. 1 : IPOC とユーザとのインタラクション

4. IPOC マルチモーダルインタラクション機構

4.1. 会話における注視行動の機能

ここでは、人間同士の会話において視線がどのような役割を担っているのかをまとめ、会話エージェントのデザインの指針を得る。

Kendon⁷⁾ による挨拶行動の分析では、お互いが相手を認識し、会話を開始するまでの様々な非言語行動が報告されている。会話を始める前には、まず一定以上の距離から相手に視線を向け、小さくうなづいたり手を振ったりしてお互いを認識しあった後に、相手に近づき、距離が十分縮まったら、再度視線を合わせ、会話を開始するといった非言語的なステップが存在する。

一度会話が開始されると、それを維持する過程において、やはり注視行動は重要な役割を担う。その1つは相手に注意を向けることである。一般的には、聞き手から話し手への注視行動は、話し手から聞き手への注視行動よりも多いといわれている¹⁾。聞き手が話し手に注意を向けることは、聞き手による会話への参加意思を示し、これが話し手へのフィードバックとなる。一方、話し手は、聞き手にたびたび視線を向けることにより、聞き手の注意状況を確認する。

一方、会話中に言及される対象物を共有しながら会話

をする場合には、会話相手に対してのみならず、共有された対象物に注意を向けることは、会話遂行の目的となるタスクへの取り組みを示し¹⁷⁾、さらには発話の基盤化 (grounding) における理解の証拠として機能する¹²⁾。例えば、地図を見ながら道順を聞く場合には、聞き手が地図に視線を向け、地図の情報を共有していることを示すことは、理解の証拠となる有効な非言語的のフィードバックである。

また、ターン交代時には視線の微細なやり取りが重要な役割を担っている。従来研究では、会話が円滑に進んでいない場合には、ターン交代時のアイコンタクトの時間が長くなる¹³⁾、また、ターンをとる際、話し手は発話の開始時に聞き手から視線をそらす⁴⁾等の観察結果が得られている。

以上のように、視線は話し手・聞き手両者にとって、会話を円滑に進めるための重要な非言語情報であり、会話参加者は常に相手の視線の動きを観察しながら会話を進めていることが明らかになっている。

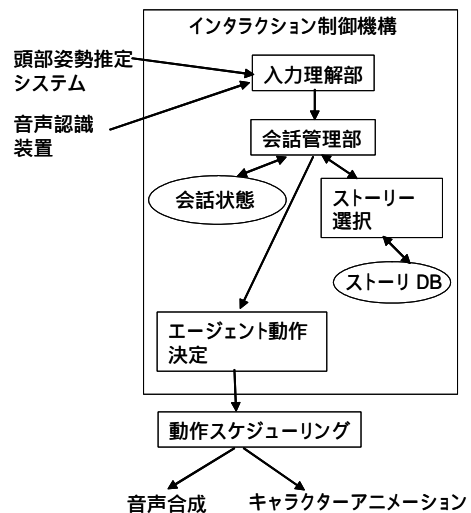


Fig. 2 : IPOC インタラクション制御機構

4.2. インタラクション機構の概要

以上の議論をシステムデザインの指針とし、IPOC のインタラクション機構では、ユーザの視線をインタラクション制御に利用した。Fig. 2 に IPOC のインタラクション制御機構の構成を示す。本機構への入力、つまり入力理解部への入力は、音声認識 Julius⁵⁾ によるユーザの言語的行動と、頭部姿勢推定システムを利用して推定されたユーザの視線方向である。会話制御部では、これらの情報にもとづき、会話の状態を更新し、更新された会話状態に基づき、エージェントによる次の行動を決定する。エージェントの言語的行動には、エージェント動作決定部により、その内容に応じて表情やジェスチャーが自動

的に付与される。その結果、非言語情報の注釈が付与された XML 形式のデータが動作スケジューリング部に送られる。動作スケジューリング部では、日立中央研究所で開発された高品質音声合成装置 HitVoice により、合成音声が生産されると同時に、エージェントの動作と同期をとるために、各音素のタイミング情報が出力される。最後に、算出されたタイムスケジュールに従って、音声言語と同期したエージェントアニメーションが出力される。以下に、各構成素について説明する。

4.3. インタラクション機構の構成

(1) ユーザの視線認識機構

ユーザの視線推定には、頭部姿勢推定システム¹⁴⁾を利用した。このシステムでは、パーティクルフィルタにおける仮説の拡散を適応的に制御することにより、ユーザが空間中のある点を注視している場合の推定精度を高く維持すると同時に、ユーザが突発的に動作する場合にも追従性を保つことを可能にしている。このような方式により、本システムは、視線の移動に伴う頭部の微細な動きを精度よく認識できるという特長を持ち、ユーザの視線推定に有用であると判断し、採用した。また、本方式ではユーザは特別な装置を身につける必要はなく、エージェントとの自然なインタラクションを損なわないという利点もある。頭部姿勢推定システムの動作の様子を Fig. 3 に示す。



Fig. 3: 頭部姿勢推定システム動作例

このシステムを利用して、大まかなユーザの視線推定を行った。IPOC では、ユーザは 70 インチの大画面から約 1.5 メートルの離れた位置でエージェントと会話をする。そこで、大画面を 6 つの領域に分割し、ユーザがどの領域を注視しているのかを頭部姿勢推定システムから出力される頭部の位置と回転角度から推定した。各観測値には揺らぎがあるので、10 データポイントごとの平均値を毎秒 30 回算出し、その値を視線推定に用いた。また、全ての計算結果を出力するのではなく、ユーザの注視位置が変化した場合のみ、そのイベントが会話管理部に通知される。

(2) 会話管理部

会話管理部では、会話状態の更新とエージェントの行動決定を行う。ユーザの視線情報を取り入れるために、従来の言語的インタラクションのみを扱う会話管理機構を拡張した点について、以下に述べる。

(a) 視線情報の管理：会話状態は Information State⁹⁾ による会話のモデル化に準じた形式で表現されており、言語的な情報として、現在の話題 ID、発話 ID、発話の開始・終了時間、発話内容、話題の焦点となる対象等が保持されている。さらに、視線情報として、現発話中のユーザの注視行動ログと、発話中のユーザの主要な注視領域（発話区間中に最も長い時間注視していた領域）が計算され、保持されている。このように、ユーザの視線情報を発話単位で管理することにより、言語的なインタラクションに伴う視線情報を会話状態の一部として扱うことができる。

ただし、現在実装している視線認識機構では、10 データポイント（0.3 秒間のデータ）の平均値をとり、さらに発話中の主要な注視領域を算出することによって、比較的安定した注視行動を観測することに最適化したアルゴリズムを採用している。従って、非常に短時間の視点移動は観測できないという欠点も有する。例えば、ターン交代時に話し手が瞬間的に聞き手から視線をそらすといった行動は検知できない。

(b) ユーザ視線情報を利用した会話状態の更新：視線情報を会話の状態として管理することにより、会話管理部では、これを会話状態の判断に利用している。例えば、ある発話中にユーザが注視していた対象物が、その発話の焦点となる対象と合致していれば、会話管理部は、ユーザの注視行動が適切であった（話題となっている対象物に対して注意を向けていた）と判断し、その発話が正しく理解されたとみなす（少なくとも、その仮定を支持する証拠が得られたとみなすことができる）。このように、言語的な会話状態に照らして視線情報を解釈することにより、言語情報のみでの会話の状態管理よりもロバストな機構が実現する。

(c) 視線情報を考慮した会話制御：会話管理部は更新した会話状態に基づき、次のエージェントの行動を決定する。ここでは、視線情報を利用して次の行動を決定するとともに、視線情報の認識誤りによる会話の失敗を防ぐための会話制御を行う。例えば、ストーリーの途中でユーザの視線が他の対象物に移行してしまった場合には、システムは、これをユーザによる非言語的な割り込みとみなし、その対象物についての話題に転換する。この時、「説明の途中ですが、こちらの建物の説明に移りますか?」といった発話を生成することにより、話題を変え

ることをユーザに確認し、視線の認識誤りによる誤った会話進行を避ける。

(3) ストーリー選択

会話管理部は、ユーザの発話や注視対象物に関連したストーリーをエージェントに語らせるために、発話内のキーワードや対象物の名称などをクエリとしてストーリー選択部に送る。ストーリー選択部では、クエリに応じたストーリーをストーリーDB から検索し、会話管理部に返す。

(4) エージェントアクション生成

エージェントによる表情、ジェスチャー、およびリップシンクはエージェント動作決定部で決定されるが、これは、CAST¹¹⁾ システムを IPOC に組み込むことにより実現されている。CAST は、日本語文を入力すると、語彙・統語情報にもとづき、エージェントの非言語行動を自動的に決定し、さらに音声合成器からタイミング情報を取得して、アニメーションのタイムスケジュールを算出するツールである。また、同時に合成音声ファイルも作成される。CAST を利用することにより、IPOC では、エージェントの発話（ストーリーを構成する各文）に応じた非言語行動を自動的に決定、出力している。

[1:S] こんにちは。
 [2:S] ここは、江戸の町を再現した場所です。
 [3:S] この町と江戸の文化についてご紹介します。
 [4:S] ではまず、こちらの駕籠屋について説明しましょう。
 [5:S] 駕籠は今で言うタクシーのようなものです。

 [10:S] 駕籠は江戸時代の最上の乗り物だったんですね。
 [11:S] 次はどこについて知りたいですか。
 (ユーザの注視対象移動)
 [12:S] この小間物屋についてお話ししましょうか。
 [13:H] はい
 [14:S] 小間物屋とは江戸時代にあった装粧品を扱う店です。
 [15:S] 江戸の初期から中期にかけて、髪結いという職業が現れ、女性の様々なヘアスタイルが登場するようになりました。

 [21:S] 女性が欲しいなと思った小物ほとんどを扱っていたのです。
 (ユーザの注視対象移動)
 [22:S] 説明の途中ですが、こちらの建物の説明に移りますか。
 [23:H] はい
 [24:S] これは、火の見やぐらです。
 [25:S] 火事が起きた時には、町火消しが火の見やぐらに登り、てっぺんに取り付けてある鐘でいち早く町内の人に火事を知らせました。

Fig. 4 : エージェントとの会話例

4.4. エージェントとの会話例

以上の機能を実装した会話エージェントとの会話例を Fig. 4 に示す。ユーザが画面上のエージェントを注視し、ユーザからの視線が、はじめてエージェントに向けられたことをシステムが検知すると、エージェントがユーザに挨拶をし、会話が開始される ([1:S])。最初の話題が終

了した後、[11:S]では、エージェントはユーザに次に何について知りたいかを尋ねると同時に、ユーザの注視点を観測している。その結果、ユーザの視線が画面下方の家屋に向けられていることを認識し、[12:S]において、それに視線を向けながら指差すことにより、ユーザとの間に共同注視を確立しつつ、話題の転換を打診している。さらに、[21:S]では、ストーリーの途中にユーザの視線が画面上方の火の見やぐらに向けられていることを感知し、これをユーザによる非言語的な割り込みとみなすことにより、[22:S]で 話題を途中で変えてもよいか訪ねている。このときにもユーザの注視対象物にエージェントが視線を向けることにより、ユーザ視線へのシステムによる Awareness を表現している。さらに、視線認識誤りの可能性を考慮し、[23:H]でユーザの言語的承諾を得た後、火の見やぐらについての話を始めている。

5. IPOC コンテンツ作成支援機能

ここでは、アニメーションの会話エージェントを登場させたCG映像コンテンツの作成支援技術の開発に焦点を当てる。映像コンテンツは視聴者の注意をひきつける有効なメディアであるが、ズームやカメラ移動等のカメラワーク、ショットの切り替え、シーンの構成等、数多くの要素を含み、それらを適切に設定しなければかえってわかりにくい映像表現になってしまう⁸⁾。これは、素人が作成した映像は非常にわかりにくく、見にくいことから推し量ることができる。

この問題を解決し、一般の人でも、メッセージを伝えるのに十分なわかりやすさのCGコンテンツを手軽に作成できる環境を提供することは、市民が専門家に現場の状況を伝えるために重要である。

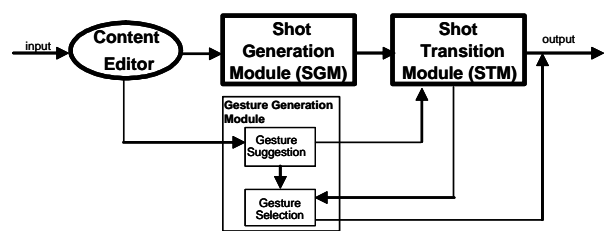


Fig. 5 : IPOC コンテンツ作成支援機構

5.1. コンテンツ作成支援機構の概要

Fig. 5 に本システムの構成を示す。コンテンツ編集画面 (Content Editor) では、利用者は、シーンの左・右・中央の写真、計3枚と、会話調の掛け合いで話を進める台本を指定するのみである。これらの情報はショット生成部 (SGM) に送られ、まず、本システムで想定した7箇所の全てのカメラからのショットが作成され、さらにショット遷移部 (STM)で、その中から適切なショットが選択される。さらに、台詞や選ばれたショットに応じて、エー

エージェントによる指差し等のジェスチャーや視線方向が決定される。最終的には、カメラとエージェントの位置、およびエージェントの動作コマンド列が出力され、それを Haptek player¹で実行することにより、作成したCGコンテンツを視聴することができる。



Fig. 6 : コンテンツ編集画面

5.2. コンテンツ編集画面 (Content Editor)

コンテンツ編集画面を Fig. 6 に示す。まず、Aの領域では、話題の焦点となる対象物がエージェントの右側にあるか左側にあるかを選択する。これは、シーンにおけるエージェントの位置の制約となる。次に、Bの領域では、背景となる写真のファイル名を指定する。Fig. 7 に示すように、写真は、焦点の対象物のミディアムショット (MS)、左側の背景 (BG1)、右側の背景 (BG2) の最低3枚が必要である。対象物をズームしたアップショット (US) は、MS から切り取るにより作成することができる。

次に、領域Cでは、MS上に、焦点となる対象物の位置をマークする。これにより、エージェントが対象物を指差す時のジェスチャーの角度が計算される。

最後に、領域Dでは、二人のエージェントの台詞を入力する。ここで、各台詞に対して、以下の2種類のタグを付与する必要がある。

- (a) 話者タグ: 本システムでは、視聴者のガイド役となるエージェント、つまり、視聴者を投影した分身の役割を担う「ガイドエージェント」と、現場の様子を説明する「説明エージェント」の2人の会話調のコンテンツを想定している。そこで、各発話に対して、その発話者がガイドエージェントである場合には G を、説明エージェントである場合には E を付与する。
- (b) 参照タグ: 当該発話が焦点となる対象物を参照し

ている場合には T を、それについて言及していない場合には F を付与する。

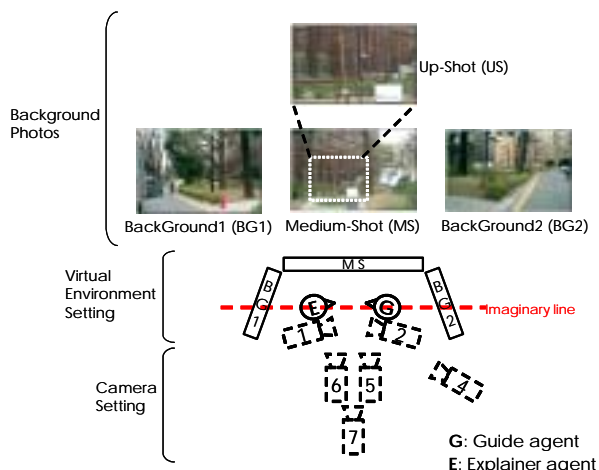


Fig. 7 : IPOC 仮想環境設定とカメラ位置

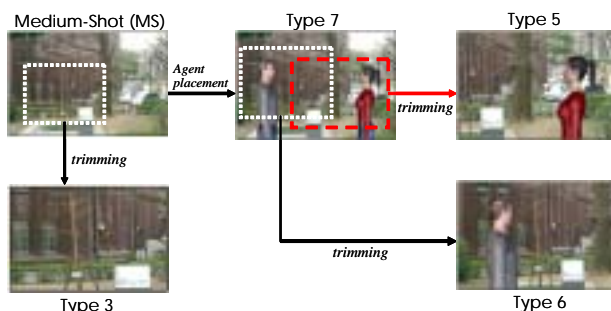
5.3. ショット生成部 (SGM)

ここでは、Fig. 7 示す仮想環境設定 (Virtual Environment Setting) に従って、以下の7種類のショットを生成する。

<ショットの種類>

- Type 1 : ガイドエージェントのみのショット
- Type 2 : 説明エージェントのみのショット
- Type 3 : 対象物のショット
- Type 4 : ガイドエージェントの肩越しから撮影した両エージェントのショット
- Type 5 : ガイドエージェントと対象物を含むショット
- Type 6 : 説明エージェントと対象物を含むショット
- Type 7 : 両エージェントと対象物を含むショット

Fig. 8 に示すように、例えば、Type7 はMSを背景とし、両エージェントをカメラ7から撮影したショットであり、Type5 はカメラ7の右側だけを、Type6 は左側だけをトリミングしたものである。また、Type3 はカメラ7から対象物をズームインしたショットである。



¹ <http://www.haptek.com/>

Fig. 8 : ショット生成例

5.4. ショット遷移部 (STM)

ショット遷移部 (STM)では,ショット生成部で作られたショットの中から,各発話のショットを選択する.

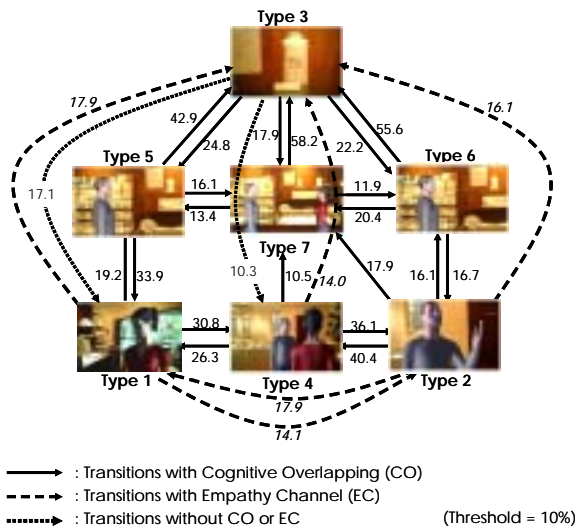


Fig. 9 : ショット遷移モデル

(1) ショット候補の決定

まず,意味的に不自然なショットを削除して,ショットの候補リストを生成する.現システムでは,(a)対象物が発話中に言及されている場合のみ,対象物を含むショットを有効とする,(b)話し手エージェントはショットに含まれていなければならない,の2点をショット候補の制約としている.

(2) ショット遷移の選択

次に,ショット候補の中から1つのショットを選択する.実際のTV番組に習うために,約30分の旅行番組3本分のデータを収集し,その中での物の紹介シーンを7種類のショットタイプを用いて分析した.その結果, Fig. 9 に示すショット遷移確率を得た.本システムでは,この遷移確率に従って,ショット候補の中から次のショットを決定している.実線,破線,点線のいずれも有効な遷移である(ネットワークの詳細については Okamoto et al.¹⁵⁾ に述べる.)例えば, Type1 が現在のノードであった場合,次のショットの候補は Type2, 3, 4, 5 であるが, 30.8%の確率で Type4 が選択され,もっとも確率の高い遷移となっている.

6. 関連研究

従来の仮想空間のエージェントとの会話システムの研究では,特定の場面についての3次元モデルを構築し,

限定された場面での会話を行う会話エージェントの研究が中心となっていた^{3,16)}.IPOCでは,非言語コミュニケーションを利用して緩やかに会話を制御することにより,コンテンツ作成のコストを削減し,ユーザが撮影した任意の場面において会話的な情報提供を行うことを可能にしている.

コンテンツ作成支援のためのカメラワーク自動生成の研究では,エージェントの発話文中の動作語に応じてカメラワークを決定する方式や¹⁰⁾,話者交代のタイミングに応じて決定する方法²⁾等が提案されている.これに対しIPOCでのコンテンツ作成支援として実現したものは,対象物に言及する会話において情報伝達に有効なカメラワークを生成することに特化した方式であり,現場の情報を伝えることを目的とする会話コンテンツの生成に適している.

7. まとめと今後の課題

本論文では,専門家と市民とのコミュニケーションを活性化することを目的とした没入型仮想会話環境 IPOC について述べた.IPOCでは,背景画像を知覚的な共有情報としながら,専門家のアドバイスをより具体的,かつ円滑に伝えるインタラクション機能と,市民からの現場情報をよりリアリティの高いものにするためのコンテンツ作成支援機能を装備している.

本論文では主にシステムの実装について述べたが,IPOCを実際の社会問題に適用する上では,まだ多くの課題が残されている.IPOCをコミュニティ内の情報流通のために利用するには,まず,システムの運用面での工夫が必要である.豊富で良質なコンテンツが必要であるが,それにはコンテンツ作成者となる市民や専門家の協力が不可欠である.例えば,地震防災のコンテンツ開発を想定した場合,IPOC番組の内容の妥当性を検証するために専門家の協力は必須であり,研究活動に多忙な専門家を動機付ける必要があるであろう.また,市民の参加を促すには,現場の状況を伝えることにより,専門家からの適切なアドバイスがもらえるという効力感が得られることが必要であろう.自分の専門知識や体験をコンテンツとして表現し,それが共有された知識となることの意義を実感させ,それによりコンテンツ制作を動機付けるような仕組みとあわせて本システムを運用していくことが必要であろう.

また,技術的な課題としては,本システムでは,専門家と市民との情報共有を目的としたが,リスクコミュニケーションにおいては,両者がさらに意見を交換しながら議論し,意思決定を行うフェーズでの支援も必要である.今後,本システムに議論支援システム⁶⁾等を統合することにより,リスクコミュニケーション全体を支援するシステムの構築を目指していかなければならない.

今後は、技術的な改善のみならず、社会への実装という観点からもシステムデザインの工夫を行っていく予定である。

参考文献

- 1) Argyle, M. and M. Cook (1976). *Gaze and Mutual Gaze*. Cambridge, Cambridge University Press.
- 2) 有安香子, 住吉英樹, 林正樹, 井上誠喜. 番組テンプレートとTVMLを用いた自動番組生成. 情報処理学会第59回全国大会, 特別セッション(1)講演論文, pp.237-241, 1999.
- 3) Cassell, Justine, Tim Bickmore, Lee Campbell, Hannes Vilhjalmsson, and Hao Yan. (2000). *Human conversation as a system framework: Designing embodied conversational agents*. In Justine Cassell, Joseph Sullivan, Scott Prevost, and Elizabeth Churchill, editors, *Embodied Conversational Agents*. MIT Press, pages 29-63.
- 4) Duncan, S. (1974). On the structure of speaker-auditor interaction during speaking turns. *Language in Society* 3: 161-180.
- 5) Julius <http://julius.sourceforge.jp/>.
- 6) 堀田昌英, 榎戸輝揚, 岩橋伸卓: 多元的議論構造の可視化手法: 社会技術としての政策議論支援, 社会技術論文集, (2003)
- 7) Kendon, A. (1967). Some functions of gaze direction in social interaction.. *Acta Psychologica* 26: 1-47.
- 8) Kraft, R. N.. The role of cutting in the evaluation and retention of film. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Vol. 12, No. 1, 155-163, 1986.
- 9) Matheson, C., M. Poesio, et al. (2000). Modelling Grounding and Discourse Obligations Using Update Rules. *1st Annual Meeting of the North American Association for Computational Linguistics (NAACL2000)*.
- 10) 宮崎誠也, 申金紅, 青木輝勝, 安田浩, 自動演出生成方式によるシナリオ入力型CG動画制作システム. 信学技報, CS2003-145, pp. 31-34, 2003.
- 11) Nakano, Y. I., M. Okamoto, et al. (2004). Converting Text into Agent Animations: Assigning Gestures to Text. *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2004), Companion Volume, Boston*.
- 12) Nakano, Y. I., G Reinstein, et al. (2003). Towards a Model of Face-to-Face Grounding. *the 41st Annual Meeting of the Association for Computational Linguistics (ACL03)*, Sapporo, Japan.
- 13) Novick, D. G, B. Hansen, et al. (1996). Coordinating turn-taking with gaze. *ICSLP-96*, Philadelphia, PA.
- 14) 岡 兼司, 佐藤 洋一, 中西 泰人, 小池 英樹, "適応的拡散制御を伴うパーティクルフィルタを用いた頭部姿勢推定システム", 電子情報通信学会論文誌 D-II, vol.J88-D-II, no.8, August 2005. (採録決定)
- 15) Okamoto, M., Okamoto, K., Nakano, Y. I., and Nishida, T. Supporting the creation of immersive cg contents with enhanced user involvement. In *Proceedings of AISB05: the Symposium on Conversational Informatics for Supporting Social Intelligence and Interaction -Situational and Environmental Information Enforcing Involvement in Conversation-*, pp. 87-96, 2005
- 16) Traum, David and Jeff Rickel. (2002). Embodied agents for multi-party dialogue in immersive virtual worlds. In *the first International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2002)*, pages 766-773.
- 17) Whittaker, S. (2003). Theories and Methods in Mediated Communication. *The Handbook of Discourse Processes*. A. Graesser, MIT Press.

謝辞

本研究を遂行するにあたり、頭部姿勢推定システムをご提供いただいた、東京大学生産技術研究所 佐藤洋一先生、岡兼司氏に深く感謝いたします。本研究で使用した頭部姿勢推定システムの一部には、オムロン株式会社のOKAO Vision 技術を利用しています。また、コンテンツ作成支援システムを共同開発いただいた東京大学情報学系研究科 岡本雅史氏、岡本和憲氏に深く感謝いたします。

IMMERSIVE CONVERSATIONAL ENVIRONMENT FOR SHARING INFORMATION IN A SOCIETY

Yukiko I. Nakano¹ and Toyoaki Nishida²

¹Ph.D. (Information Science and Technology) Research Institute of Science and Technology for Society (RISTEX)
(E-mail: yukiko @ristex.ac.jp)

²Ph.D. (Engineering), Professor, Graduate School of Informatics, Kyoto University (E-mail: nishida@i.kyoto-u.ac.jp)

In building a safe and secure society, sharing a common understanding between experts and citizens is indispensable. To accomplish this requirement, it is necessary to establish a communication channel between both of them. This paper proposes an immersive virtual conversational environment IPOC, aiming at activating communication between experts and citizens. Using IPOC interaction functions, experts can give more concrete advice to the citizens by sharing background photos as perceived information,. On the other hand, the citizens can describe the situation on their sites using IPOC contents creation support functions. These functions in IPOC are expected to contribute to bidirectional communication between experts and citizens.

Key Words: *Risk communication, Immersive conversational environment, Conversational agent, Contents creation support, Human-computer interaction.*