

K-002

## 擬人化エージェントを用いた

## ユーザ参加型マルチメディアコンテンツ流通システム

## Participatory Broadcasting System Featured with Multimedia Contents and Interface Agents

中野有紀子<sup>†</sup> 村山敏泰<sup>†</sup> 西田豊明<sup>††</sup>

Yukiko Nakano Toshihiro Murayama Toyoaki Nishida

## 1. はじめに

メーリングリストや、電子掲示板システムなど、ネットワークを利用したコミュニティを通じて膨大な情報が流通している。このようなアプリケーションを利用して、より豊富で質の高い情報を流通させ、知識として蓄積していくためには、システムデザインにおいて、コミュニティ内のコミュニケーションを活性化させるための工夫が必要である。この問題への1つのアプローチが、チャットシステムやテレビ会議システム等へのエージェント技術の適用である[1, 2]。これらのシステムでは、擬人化エージェントが仮想空間でユーザの分身として振舞い、人間同士のコミュニケーションを媒介する。一方、擬人化エージェントを用いて、人対コンピュータのコミュニケーションの円滑化を目指すシステムもある。web ページ上でのコマースエージェントや、ニュースキャスターエージェント等は最近多く見かけられる例である[3]。これらはコミュニティメンバーに情報提供を行うエージェントである。しかし、従来、このような情報提供エージェントを実装する際には、エージェントの動作やその実行のタイミングを記したスクリプトを手で作成する必要があり、コンテンツが日々更新されるようなサービスに利用することは難しかった。そこで、本稿では、擬人化エージェントの動作を自動的に決定する機能を持つ情報流通システム S-POC を提案する。S-POC は画像や映像、あるいはプレゼンテーションスライド等のマルチメディアコンテンツをテレビ放送的に流通させることを目的としたシステムであり、エージェントはユーザへの情報提示において重要な役割を担う。

## 2. 背景

S-POC システムの基本的なアイデアは、放送型コミュニティ支援システム Public Opinion Channel (POC) システム[4]に基づいている。POC は、100字程度のメッセージ、URL、画像からなる知識カードを情報発信の最小単位とする。ユーザが知識カードを POC システムに投稿すると、それがコミュニティのメンバーにも視聴可能となる。また、POC の拡張である EgoChat システム[5]では、2人の分身エ

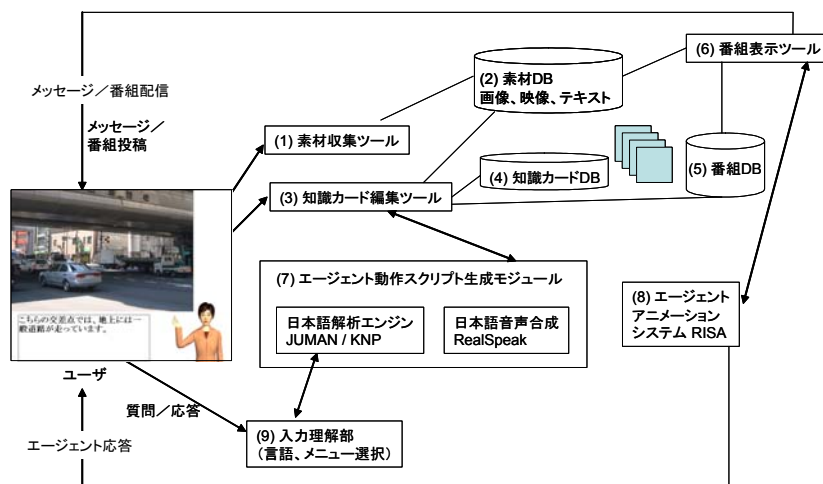


図1: S-POC システム構成図

ージェントが番組内容を会話的に演じることが可能になった。

しかし、これまでのシステムでは、映像コンテンツが扱えないなど、テレビ放送型システムとして最も重要な特徴である、動的なコンテンツの表現が不十分であった。S-POC では、特にテレビ的な表現を向上させるために、以下のような拡張を行った。(1) 映像コンテンツをストリーミングにより利用可能にした。(2) 画像コンテンツに対してズームやパンなどのカメラワークを付与することにより、画像を映像的に見せる効果を実現した。(3) エージェントのジェスチャーや顔表情を自動的に計算する機構を実装し、(4) さらに従来よりも表現力の高いアニメーションシステムを開発した。以下の章では、本システムの詳細について述べる。

## 3. S-POC システム

## 3.1. マルチメディアコンテンツの作成・流通プラットフォーム

図1に S-POC のシステム構成図を示す。以下の3つのモジュールは S-POC で番組を作成し、視聴するためのユーザインタフェースを提供する。

**素材収集ツール**: カメラで収録された映像や画像を素材 DB に送信し、SPOC システム内で流通可能なフォーマットに変換して蓄積する。

**知識カード編集ツール**: これは以下の3つの機能からなる。(a) 素材 DB に蓄積されたデータに対して、文章を付加し、知識カードを作成する作業を支援する。ユーザは、素材 DB 中の映像や画像のデータを選択し、さらにどの部分を知識カードとして利用するかを選択することによって、素材の切り出しを行うことができる。また、画像データに

<sup>†</sup> 社会技術研究システム

Research Institute of Science and Technology for Society(RISTEX)

<sup>††</sup> 東京大学大学院情報理工学系研究科, UT

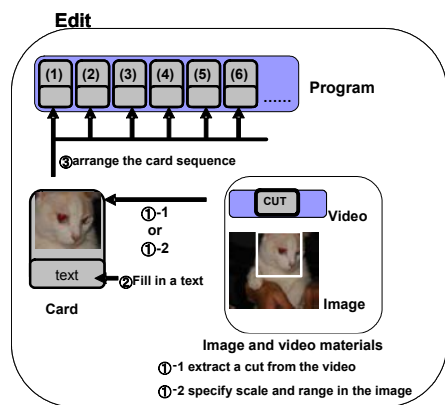


図 2 : 知識カード編集ツール

えて番組を作成する作業を支援する。ユーザが適当なカードを編集画面にドラッグして並べるとそのシーケンスが番組として定義される。(a), (b) に関して、図 2 に概略図を示す。(c) 知識カードが番組の一部として組み込まれると、そのカードの情報がエージェント動作スクリプト生成モジュールに送られる。その結果、カード中のテキストを読み上げた合成音声の WAV ファイルと、それと同期するエージェントアニメーションのタイムスケジュールが生成され、知識カード情報の一部として追加される。これによって、エージェントが各知識カード中のテキストを身振り、顔表情、リップシンクをつけて読み上げることができるようになる。

**番組表示ツール**：番組 DB に蓄積された番組情報を入力とし、素材 DB 中の映像や画像とエージェントアニメーションシステムから生成されるアニメーションとを同期を取りながら統合し、ユーザの端末へと配信する。

以上のユーザインターフェースは普及度の高い FLASH Player6.0 のみで動作する WEB アプリケーションとして実装され、バックグラウンドのサーバによって必要な処理を行っている。

### 3.2. インタフェースエージェント

以下の 2 つのモジュールでは、知識カード中のテキストの内容に対し、ジェスチャーや顔表情などの非言語情報を自動的に付与し、それをエージェントアニメーションによって表現する。

**エージェント動作スクリプト生成モジュール**：このモジュールは MIT で開発されたエージェント動作決定機構、BEAT [6] をベースにしている。テキストが入力されると、文の構造が日本語解析エンジン [7] を用いて解析され、文構造に応じて文中の強調すべき部分が決定される。次にその部分をどのような非言語情報を用いて強調するのかが決定される。例えば、新情報が導入されるときにはハンドジェスチャーを用い、眉を上げ、軽くうなづく、といったルールを記述することにより、文構造からそれに適した非言語表現を決定することができる。次にこのテキストを音声合成システムにかけ、単語と音素のタイミング情報を取得し、音声とアニメーションとを同期させるためのタイムスケジュールを計算する。また、音声合成システムから出力された WAV 形式の音声ファイルも保存される。以下にシステムから出力されるエージェントのアクションスケジュールを示す。例えば、AID[125]は、右手の手のひらを 15 フレーム

対して、ズームやパンなどの簡単なカメラワークを付加することもできる。これによって画像を映像風に見せることが可能になる。

(b) 知識カード DB に蓄積された知識カードを並べ替

ム目で 45 度曲げ、21 フレーム目で元に戻すという指示である。

```
AID[110] = {VISEME type = e, StartFrame:14, EndFrame:14};
AID [120] = {VISEME type = o, StartFrame:15, EndFrame:15};
AID[125] = {Hand_R pos = 5, type = 45deg, StartFrame:15,
            EndFrame:21};
AID[130] = {VISEME type = a, StartFrame:16, EndFrame:16};
AID[140] = {VISEME type = i, StartFrame:17, EndFrame:17};
AID[150] = {EyeBrows type = up, StartFrame:20,
            EndFrame:30};
```

**エージェントアニメーションシステム RISA**：エージェント動作スクリプト生成モジュールで計算された動作スクリプトが入力されると、そのスクリプトをインストラクションセットとして、タイムスケジュール通りにエージェントアニメーションを生成する。RISA は、キャラクターの体を 12 の部品に分割し、各部品のアニメーションを組み合わせて体全体の絵を構成する手法を用いて実現されているため、顔の各パーツの動きや、腕の動きを独立に指定することができる。これによって、従来の MS エージェント等をベースとしたエージェントアニメーションに比べてアニメーションのパリエーションが格段に多くなる。また、RISA は Flash で実装されているため、S-POC システム全体を Flash をプラットフォームとしてシームレスに統合することが可能になる。

## 4. まとめと今後の課題

本稿では、テレビ番組的な情報流通を行う放送型コミュニケーション支援システム S-POC について述べた。本システムでは、メディアに関する知識や技能のない人でも、映像効果を持つコンテンツを容易に作成することができ、また、言語的メッセージを伝えるエージェントも自動的に作成される。今後は、日常生活に関するリスクや安全についての情報提供を行うサイト上のアプリケーションとして利用し、実証実験を行うことにより、システムの有用性、改善点等を明らかにしていく予定である。

### 参考文献

- [1] Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida. FreeWalk: A 3D Virtual Space for Casual Meetings. IEEE Multimedia, Vol.6, No.2, pp.20-28, 1999.
- [2] 高橋徹 武田英明 TelMeA: 非同期コミュニケーションシステムにおける Avatar-like エージェントの効果と Web ベースシステムへの実装, 電子情報通信学会論文誌 D-I, Vol.J84-D-I No.8 pp.1244-1255, 2001.
- [3] ANANOVA, <http://www.ananova.com>.
- [4] FUKUHARA, T., FUJIHARA, N., AZECHI, S., KUBOTA, H., and NISHIDA, T.: Public Opinion Channel: A network-based interactive broadcasting system for supporting a knowledge-creating community, In R.J.Howlett, N.S.Ichalkaranje, L.C.Jain, and G.Tonfoni(eds.); Internet-Based Intelligent Information Processing Systems, World Scientific Publishing, chapter 7(pp.227-268)(2003).
- [5] 久保田秀和 西田豊明 ユーザの過去の発言を利用した複数エージェントによる創造的な対話の生成, 電子情報通信学会論文誌, Vol. J84-D-I, No.8, 2001.
- [6] Cassell, J., Vilhjálmsón, H., Bickmore, T. "BEAT: the Behavior Expression Animation Toolkit." Proceedings of SIGGRAPH '01, pp. 477-486, 2001.
- [7] Sadao Kurohashi and Makoto Nagao. 1994. A Syntactic Analysis Method of Long Japanese Sentences Based on the Detection of Conjunctive Structures. Computational Linguistics, 20(4): 507-534, 1994.