

## プレイバック機能を持ったビジュアルエージェント

玉田 隆史 中村 泰明

三菱電機株式会社 中央研究所

自律的なタスク実行を目指したソフトウェア・ビジュアルエージェントにより計算機を擬人化することを目的とした研究を進めている。そのための第一歩として、画像、音声、音響、およびプログラム動作を統一的に管理・再現するハイパーシナリオを開発した。プログラムの動作は、外部からの入力（イベント）を記録し、記録したイベントをそのプログラムに送信することで再現される。即ち、対象プログラムとは独立にその動作を記録・再現することができる。シナリオインタープリタは、シナリオの解釈、アニメーション、音声による説明、操作の再現を行う複数のプロセスの起動、同期制御を行う。応用例として、プログラム自動デモシステムについて述べる。

## Visual Agent with Playback Mechanism

Takashi Tamada and Yasuaki Nakamura

Mitsubishi Electric Corp. Central Research Laboratory

We have been studying a visual agent system for autonomous processing of the office work. This paper describes some functions of our visual agent. The visual agent provides the "hyper-scenario" that manages and processes the image, audio, narration and the behavior of the software program. To store the behavior of the program, the system catches the external events given by an operator. Sending the same event to the program, the program works as if the operator did. The visual agent interprets a hyper-scenario, and makes several processes that perform the image and sound presentations, or replay the operations. As an application of the visual agent, an automatic demonstration system for the software system is also described.

## 1. はじめに

オフィスワークや日常の事務処理を代行するエージェントシステムの研究が盛んに行なわれる一方で、リアルな顔の表情や感情の表現などビジュアルなエージェントを作成するための研究も進められている[1~5]。また、複数の自律したエージェントにより、タスクを実行するという考え方に基づいたエージェント指向と呼ばれるプログラミングスタイルも提案されている[6]。ソフトウェアだけの閉じた世界では、エージェント指向による自律的なシステムの開発が可能であるかもしれない。しかし、実際のオフィスにおける人間の仕事をソフトウェアエージェントにより代行させるためには、音声による応答、会話の解釈、タスク実行のための各種判断、タスクの実行といった一連の機能が必要であり、全てを計算機処理することは難しい。ユーザとの対話性という観点からは、使い易く、親しみのあるインタフェースの実現が要求される。そのために、計算機を利用するという観点ではなく、テレビ電話で仕事を依頼するといったようなインタフェースを実現するためにも、音声による応答、顔の表情、感情変化の表現、顔を含めた身体の動きのモデル化、変形、表示方式の研究が必要である。即ち、計算機を擬人化し、人と対話する時と同様なインタフェースにより、各種タスクを依頼でき、タスクに対するレスポンスが得られるインタフェースが望ましい。

我々は、ビジュアルエージェントを計算機を擬人化するための方式として捉え、ソフトウェアの世界におけるアプリケーションプログラムの機能のデモ、教育、開発支援のビジュアルエージェントを作成するための機能を分析し、以下の機能を持つビジュアルエージェントを開発した。

### 開発機能

- (1) ソフトウェアとの対話操作の記録と再生

- (2) エージェント(デモンストレータ)表示とアニメーション
- (3) 音声・音響出力とによる説明
- (4) シナリオ作成
- (5) ハイパーカード形式の実行制御とそのインタプリタ
- (6) 複数プロセスによる実行同期制御と通信

上記のような機能を組み合わせることにより、プログラムの自動実行、デモだけでなく、プログラムの遠隔診断・教育、複数計算機利用に共同作業(CSCW)などへも応用可能であることを示す。なお、今回はX-windowシステム上でのアプリケーションプログラムのみを対象としている。

## 2. マルチメディアに対応したビジュアルエージェント

パーソナルコンピュータやワークステーションの高機能化により、音声、画像、映像情報を計算機で扱うことはそれほど難しくはない。しかし、オーディオ出力やアニメーション表示をアプリケーションプログラムと同期して実行し、かつ遠隔地の他の計算機と通信を行なうためには、処理能力の高いEWSが必要であり、かつ標準的なプラットフォームで稼働するシステムである必要がある。しかも、複数のタスクを非同期に実行する必要が生じ得る。そこで本研究では、分散環境を前提とした特定のハードウェアに依存しないユーザインタフェースプラットフォームであるX-windowシステム上にビジュアルエージェントを作成・利用する環境を構築した。

図1にシステムの全体構成を示す。ビジュアルエージェントは、ハイパーシナリオと呼ぶマルチメディアを制御するシナリオを解釈し、画像表示、音声出力、オペレータ操作の再現を行う機能、及び、オペレータの操作をイベントとして記録する機能を有する。以下、各機能と構

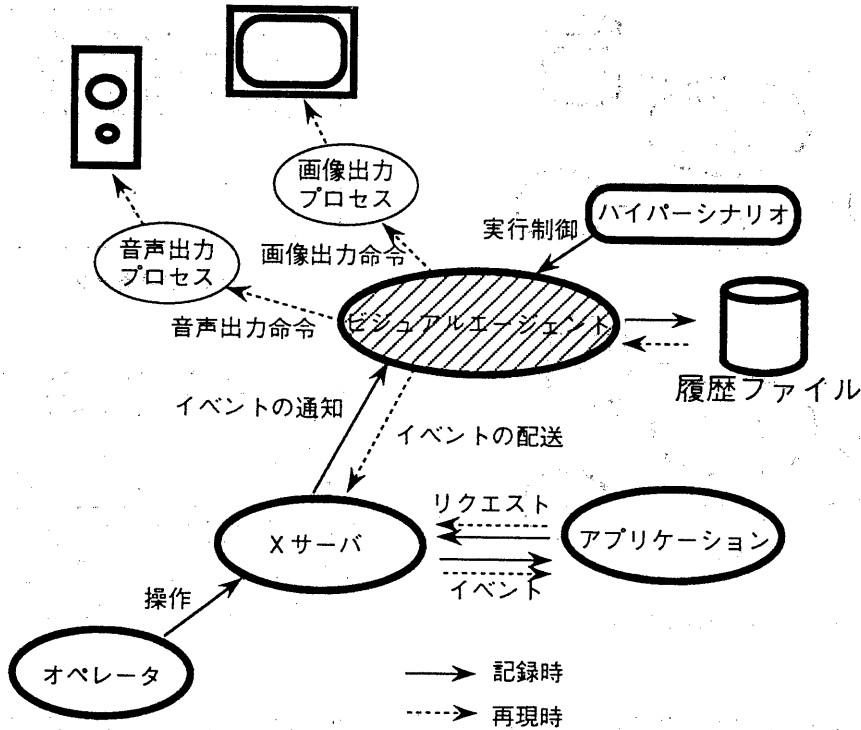


図1 ビジュアルエージェント動作時のプロセスの関係

成要素について述べる。

## 2. 1 ビジュアルエージェントの機能

ビジュアルエージェントにより計算機の擬人化を実現するためには、ビジュアルエージェントに自律的にタスクを実行する機能が備わっている必要がある。しかし、全ての処理を自動的、かつ自律的に実行することは困難である。現在の機械ロボットが、作業内容の教示・再生といったプレイバック機能だけでも多くのタスクを処理していることから、ビジュアルエージェントに人間のプログラム操作を記録・再現するプレイバック機能を装備し、自律的にタスクを実行するビジュアルエージェントの実現をめざした。

### 2. 1. 1 ナレーションとアニメーション

本ビジュアルエージェントでは、タスクの実行時に、音声（ナレーション）や動画像（アニメーション）により動作内容の説明、及び親しみのあるインタフェースの実現をめざす。音声や動画像の出力処理は、個々に独立したプロセスが実行し、動作との同期制御が可能で、ハイパーシナリオに制御内容を記述する。

### 2. 1. 2 プログラム操作記録の採取と再現

一般に、人間によるプログラムに対する操作は、マウス操作やキー入力により行なわれる。X-windowシステムでは、Xサーバがこれらの情報をイベントとして、プログラム（クライアント）に送信することによりプログラムが動作する。

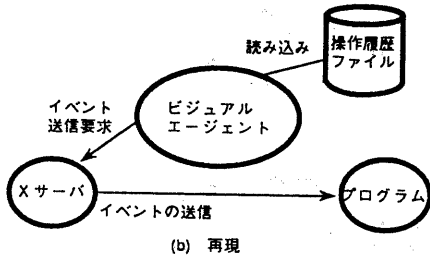
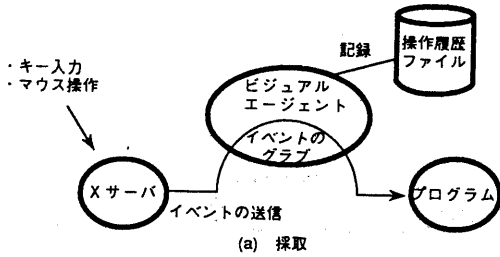


図2 プログラム操作記録の採取と再現

図1に示すように、ビジュアルエージェントは独立したバックグラウンド・プロセス（デーモン）としてシステムに常駐して、画面上で発生したイベントを監視し必要に応じ、それを履歴ファイルに記録する。また、再生時には履歴ファイルからイベントを入力し、アプリケーションプログラム（ウインド）に送信する。具体的には以下の方法でプログラム操作記録の採取と再現を行なう。

### ① プログラム操作記録の採取

プログラム操作のための人間によるマウス操作やキー入力にともなってXサーバから送信されるイベントを、ビジュアルエージェントはプログラムがイベントを受信する前にグラブ（横取り）する。そして、グラブしたイベントの内容をイベント発生時刻と共に操作履歴ファイルに記録した後、再び同じイベントをプログラムに送信する（図2-(a)）。

図3は操作履歴ファイルの一例である。操作履歴ファイルは6つの項目から構成されており（表1に操作履歴ファイルの各項目の内容を示す。）、一行が一つのイベントに対応している。

項目1	項目2	項目3	項目4	項目5	項目6
0000034	mv	0042	0815	1029	0000
0000159	br	0042	0815	1029	0003
0000281	mv	0042	0816	0005	0000
0000080	kr	0042	0816	0005	0083

図3 操作履歴ファイル

表1 操作履歴ファイルの項目内容

項目1	イベントの時間間隔(msec)
項目2	イベントの種別
項目3	イベント発生時のマウスカーソルの位置 (X座標)
項目4	イベント発生時のマウスカーソルの位置 (Y座標)
項目5	イベント発生時の修飾キー、ボタンの状態
項目6	データ

履歴保存の対象となるイベントは、ButtonPress, ButtonRelease, MotionNotify, LeaveNotify, EnterNotify, KeyPress, KeyReleaseの7つであり、Xlibの関数を用いてイベントの履歴保存を行なう。

本方式によれば、プログラムの内容を変更することなく、プログラム操作記録を時刻付きで採取することができる。

### ② プログラム操作記録の再現

上記方式で採取したプログラム操作記録の再現は、ビジュアルエージェントが操作履歴ファイルに記録されたイベント情報と時間情報に従ってイベントの送信をXサーバに要求することにより行なう（図2-(b)）。イベントを送信するウインドウはディスプレイ上での絶対座標によりウインドウを探索して決定する。このため、プログラムの改造は不要であり、オブジェクトコードだけのプログラムにもイベントの送信が可能である。また、送信するイベントの時間間隔を調節することにより、プログラム操作の再現速度の制御が可能となる。

## 2. 2 マルチメディア対応のハイパーシナリオ

### 2. 2. 1 ハイパーシナリオ

デーモンとしてシステムに常駐するビジュアルエージェントの動作制御はスクリプトファイル（ハイパーシナリオ）により行なう。ビジュアルエージェントはハイパーシナリオに記述されたビジュアルエージェントのタスク内容を読み込み、記述内容に従ってタスクを実行する。ハイパーシナリオでは、プログラム操作記録の採取・再現命令や、音声・動画像の出力命令を記述することができ、プログラム操作記録の再現に同期した音声による説明や、動画像による説明者や説明図の表示制御が可能である。表2にハイパーシナリオにおいて記述できるコマンド例を示す。

表2 ハイパーシナリオ記述コマンド

	コマンド	機能
タスク実行に関するもの	show	イメージデータ表示
	bgshow	バックグラウンド・イメージデータ出力
	wait show	バックグラウンド・イメージデータ出力完了待ち
	sound	サウンド出力
	bgsound	バックグラウンド・サウンド出力
	wait_sound	バックグラウンド・サウンド出力完了待ち
	replay	イベント再現
	bgreplay	バックグラウンド・イベント再現
	wait replay	バックグラウンド・イベント再現終了待ち
	制御命令に関するもの	card
button		ボタンの作成
go		カードの実行
gofile		実行するスクリプトファイルの変更
sleep		一時実行停止
set		変数設定
定数の定義に関するもの	setenv	環境変数の設定

コマンド群は大きく、①タスク実行に関するもの、②制御命令に関するもの、③環境変数等の各種定数の定義に関するものに分類される。タスク実行に関するコマンド群においてはbg\*\* \*コマンドにより別プロセスが起動され、並列処理が可能である。また、ハイパーシナリオ内の命令はカード単位となっており、ビジュアルエージェントはカード内の命令を全て実行すると、次のカードの内容の実行を行なう。このように命令がカード単位となっているため、カードの組み合わせを変えるだけで、異なる動作制御を行なうことができる。図4、5にハイパーシナリオと画面の例を示す。

```

card1
setenv CARDSIZE 399x569
show figure1 #画像の表示命令
sound scinario1 #音声の出力命令
replay record1 #履歴ファイルの再現命令

button string "Start" position 10 10 #ボタンの生成
go cut3 #ボタンを押すと他のカード移行
end button

button string "Quit" position 10 50
go cut5
end button
end card
  
```

図4 ハイパーシナリオ記述例

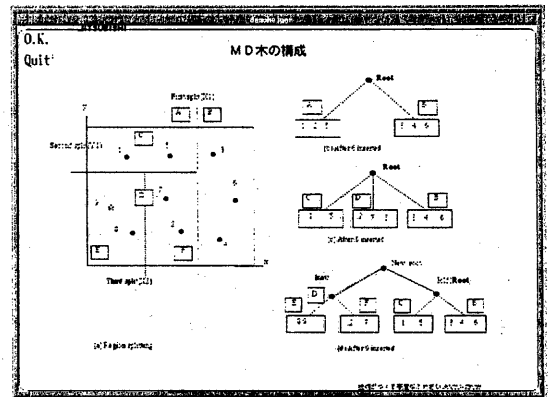


図5 画面例

### 2. 2. 2 シナリオインタプリタ

ビジュアルエージェントは図4のようなハイパーシナリオに記述されたプログラム操作記録の採取・再現、動画像表示、音声出力、ボタンの定義といった命令を一行ずつ解釈・実行していくシナリオインタプリタとしての機能を持つ。ハイパーシナリオの実行中は、ウィンドウが表示され、その中に動画像の描画やボタンの配置が行なわれる。ビジュアルエージェントはハイパーシナリオの内容に従って、操作履歴ファイル、画像データ、音声データといった各種システム内の資源を参照しながら、動画像・音声による説明に合わせてプログラムの操作を行なう。

### 2. 3 プロセス構造

ハイパーシナリオを解釈しながらの操作履歴ファイルによるプログラム操作の再現や、それと同期した音声・動画像の出力といった各種タスクを並行して行なうために、ビジュアルエージェントはプロセスを複数生成してタスクを分担する。ハイパーシナリオの解釈はビジュアルエージェント自身が行い、操作履歴ファイルによるプログラム操作の再現、動画像の出力、音声出力は、実行命令を解釈するとビジュアルエージェントが新たにプロセスを生成し、命令内容を実行させる。

操作の再現、音声・動画像の出力の同期はシグナルを利用して以下のように実現する。例えば、図6のように記述されたハイパーシナリオにおいて、プログラム操作の再現(a)とその音声による説明(b)を同期させる場合について図7を用いて説明する。ビジュアルエージェントはハイパーシナリオの第(a)行を読みとるとプロセスP1を生成して、P1にプログラム操作の再現処理を行なわせ、次に第(b)行を読みとり、プロセスP2を生成して、P2に音声出力処理を行なわせる。ビジュアルエージェントは(A)に記述された命令を実行し、第(a')行を読みとるとプロセスP1がプログラム操作の再現を終了するまで次の行の解釈・実行を待機する。そしてプロセスP1からの実行終了のシグナルを受け取ると、第(b')行を読み取り前記と同様にプロセスP2の実行終了のシグナルを受け取った後に次の行の解釈・実行を行なう。上記の方式で複数プロセスによる実行同期制御を行なうことにより、計算機の負荷の変化にともなうプログラム操作の再現時間のずれ等による影響を防ぐことができる。

また、ビジュアルエージェントはX-windowシステム上に実現されており、操作対象プログラムとは別プロセスであることから、図8のような構成にしてビジュアルエージェントが端末T1, T2, ..., Tnのマウスやキーボードを操作できるような環境にすることにより、プログラムの遠隔診断・教育、C S C Wなどへも応用可能である。

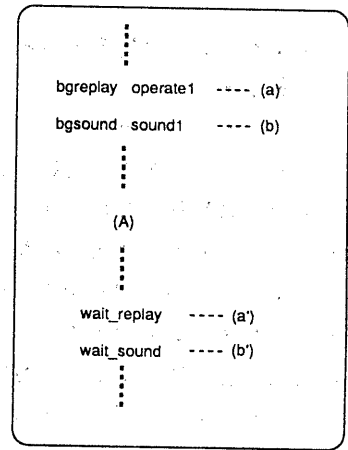


図6 ハイパーシナリオによる同期制御

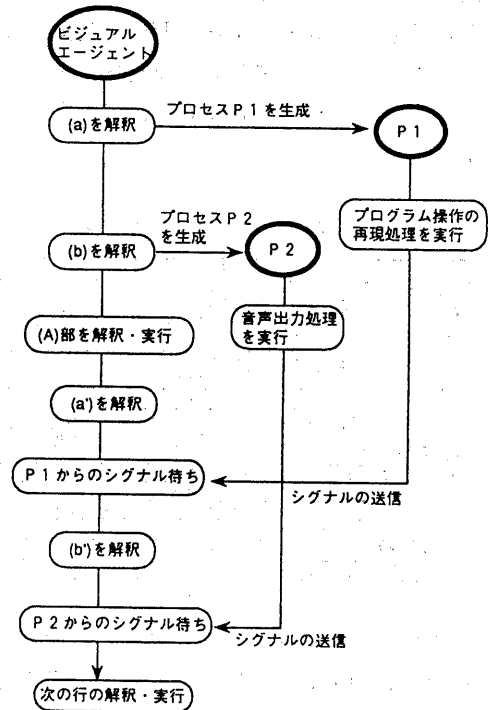


図7 同期制御のプロセス構造

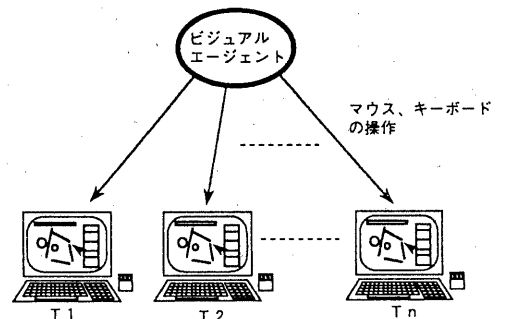


図8 分散環境におけるビジュアルエージェント



トに操作させることにより、本ビジュアルエージェントをプログラムの遠隔診断・教育やCS CWへ応用していきたい。また、映像情報が扱えるようになってきており、システムへの映像情報を取り込み、多くのEWSで多方向通信により、対話型システムへの拡張を図る予定である。

#### 参考文献

- [1]原島博:“顔画像の分析合成符号化方式の動向”, TV学技報, ICS88-17(IPA88-11), 1988
- [2]M. Ishizuka et al.:“Visual Software Agent(VSA) built on Transputer Network with Visual Interface(TN=VIT)”, Computer World 91 Osaka, 1991
- [3]M. Kaneko, A. Koike, and Y. Hatori:“Automatic synthesis of moving facial images with expression and mouth shape controlled by text”, Visual Computing, Springer-Verlag, 1992
- [4]竹内彰一:“表情豊かな顔のアニメーションを目指して”, グラフィックスとCAD論文集, 1992
- [5]玉田隆史, 中村 泰明, 阿部 茂:“プレイバック機能を持つビジュアルエージェント-Demo Companion-”, 1992年電子情報通信学会全国大会
- [6]中嶋秀之 編:“マルチエージェントと協調計算 I (日本ソフトウェア科学会MACC'91)”, 近代科学社, 1992