

CharacterPoint: キャラクタによるプレゼンテーションシステム

山本 邦雄[†] 樽本 瑤子[†] 乃万 司[†]

プレゼンテーションなどのスライド上にユーザの代わりとなるキャラクタを配置し、ユーザが話をしながら簡単な操作で、キャラクタのアニメーションを操作できるシステム CharacterPoint を提案する。CharacterPoint では、ユーザが直接話すことで、その場の状況に合わせた説明やキャラクタのアニメーションがリアルタイムに行える。また、CharacterPoint ではユーザの肉声と操作内容を記録することで、発表内容を後で再生することができる上に、マウスによる簡単な操作でスライド単位で順序の入れ替えなどの編集が可能である。発表内容を編集することで、より適切な内容でプレゼンテーションや講義を繰り返し提供することができる。

CharacterPoint: A Presentation System with a Character

KUNIO YAMAMOTO,[†] YOKO TARUMOTO[†] and TSUKASA NOMA[†]

CharacterPoint is a presentation system with a humanoid character (avatar). CharacterPoint enables users to make presentations in realtime without external knowledge of the contents nor programs/scripts for animating the avatar. A typical user of CharacterPoint gives a talk on site, and simultaneously controls the avatar on slides. The avatar can make gestures, point specific areas, and even write/draw (circle, underline, etc) on slides by user's press of keys and mouse manipulations. The synchronized user's voice and avatar's motion are recorded on site, and they can be played back and edited (even shuffled) after the presentation. The feature greatly decrease the cost of producing on-line presentation contents, e.g. e-Learning materials.

1. はじめに

近年、プレゼンテーションや講義を、リアルタイムで、あるいはビデオで記録した上で Video-On-Demand などのシステムで配信することが多い。この場合、演者(発表者)とスライドの両方を別々の画面で伝送することが多く、受講者は両方の画面を見比べることを強いられる。また、ビデオ撮影には、演者への照明などを設定する必要があり、演者が撮影に慣れていなければ何度も撮り直しをするなど、撮影のコストも高い。

そこで、スライド上にキャラクタを配置し、実際の演者の代わりをさせることが考えられる。従来から、キャラクタにプレゼンテーションをさせるシステムは研究されており、たとえば Video Actors²⁾, PPP Persona¹⁾, Virtual Human Presenter⁵⁾, TablePresenter⁴⁾ などがある。これらは、大きく二つのタイプに大別することができる。第一のタイプは、あらかじめキャラクタに特定分野の知識や発話のテキスト生成

能力を与えておき、自動的な応答のできる、擬人化(音声)対話エージェントを目指したものである¹⁾。第二のタイプは、プレゼンテーションの内容を、何らかの形で記述させるものである⁴⁾⁵⁾。

しかし、いずれのタイプも、日常的なプレゼンテーションや講義で、演者の代役をさせるのは適当ではない。第一のタイプは一般に、一連の流れのあるプレゼンテーションをすることを目的にしたものではないし、またそれを目的とすると、準備に相当のコストがかかる。一方、第二のタイプでも、演者の発話内容をすべて記述するのは、あまり現実的ではない。

一般にプレゼンテーションにおいて演者が受講者に情報を伝達するチャネルは、発話の音声と(ジェスチャなどを含む)演者の姿の二種類である。このうち、従来のシステムで準備に手間がかかるのは前者であり、一方、実際の演者で記録のコストがかかるのは後者であることを考えると、従来のシステムのように、発話の音声も演者の姿も共にコンピュータで生成するのではなく、音声は実際の演者に任せ、演者の姿だけキャラクタに代行させるのが、最も利用コストが少ないシステムが実現できると考えられる。

[†] 九州工業大学
Kyushu Institute of Technology

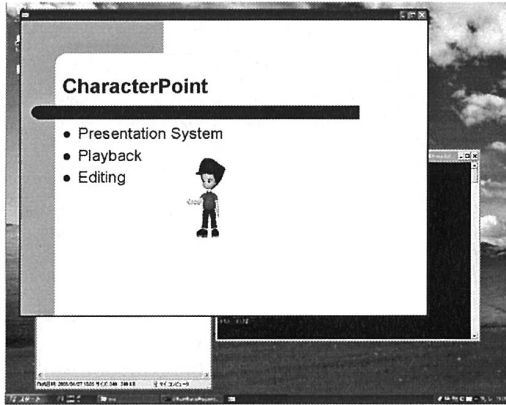


図 1 CharacterPoint の画面出力
Fig. 1 Screenshot of CharacterPoint

そこで本研究では、スライド上にユーザの代わりになるキャラクタを配置し、ユーザの肉声による説明に同期して、簡単な操作によってキャラクタのアニメーションができるプレゼンテーションシステム CharacterPoint を提案する (図 1)。CharacterPoint では、ユーザがスライドの説明をしながら、(1) マウスクリックにより、スライド上の一点を指定すると、キャラクタがその点を指示し、(2) キーを押すと、キャラクタがそのキーに対応するジェスチャをし、(3) マウスドラッグによってキャラクタがスライド上に書き込みをする。このように簡単な操作によって、ユーザはスクリプトの準備無しにその場に応じたキャラクタアニメーションによるプレゼンテーションが行える。また、一度行ったプレゼンテーションは記録することができ、その記録を再生することで、プレゼンテーションの内容を繰り返し見ることができる。さらに、プレゼンテーション内容をスライド単位で編集できるため、必要な編集をすることで、より良い内容のプレゼンテーションを提供し、繰り返し見てもらうことができる。

本システムは、プレゼンテーションを記録したり伝送したりする場合ばかりでなく、大きな会場で、演者とスライドが離れていたり、演者の姿を別のスクリーンに映さざるをえないような場合にも有用である。

2. CharacterPoint の利用

CharacterPoint の処理の概要を図 2 に示す。まず、ユーザは Microsoft PowerPoint 等でスライドを作成し、その画像をシステムに入力する。ユーザはマイクによって音声を入力しながら、マウスとキーボードのみの簡単な操作でプレゼンテーションを行える。

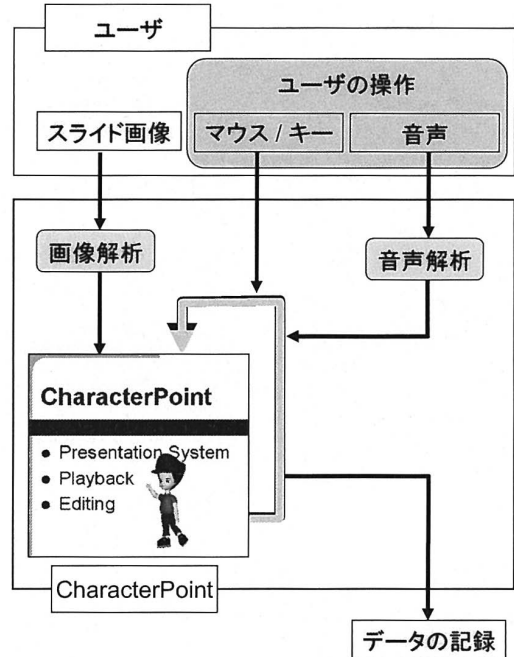


図 2 CharacterPoint の処理
Fig. 2 Processing in CharacterPoint

表 1 マウス操作の例
Table 1 Examples of mouse manipulation

操作	機能
左クリック	クリックされた文字や図を隠さないように指す
右クリック	クリックされた場所を指す
右ドラッグ	スライドに書き込む
中ドラッグ	アニメーション付線線を書き込む

2.1 マウス操作

キャラクタの移動はマウスの操作によって行う (表 1)。スライド上の一点を左クリックで指定することにより、3.1 節で述べる画像解析の結果から、クリックされた場所の文字や図をキャラクタが隠さないように自動的に移動する。これにより、ユーザはキャラクタで文字が隠れないように注意深くクリックする必要がなく、説明したい場所をクリックするだけで、邪魔にならない位置にキャラクタを配置できる。また、ユーザが、あえて文字や図の中にキャラクタを配置したい場合は、右クリックすると画像解析の結果にかかわらず、ユーザが指定した一点を指し示すようにキャラクタを配置することができる。さらに、ユーザがその場で情報を追加できるように、スライドに線や丸を書き込む機能がある。マウスの右ボタンを押したまま移動させ

CharacterPoint

- Presentation System
- Playback
- Editing



図3 スライドへの書き込み
Fig. 3 Writing on slide

表2 キーボード操作の例
Table 2 Examples of keyboard manipulation

操作	機能	操作	機能
u	両手をあげる	b	お辞儀
c	腕を組む	p	指す
n	うなづく	e	気をつけ
l	左を向く	r	右を向く
↓	次のスライド	↑	前のスライド
a	怒った表情	w	驚いた表情
h	喜んだ表情	t	泣いた表情
o	表情を戻す	ESC	終了および保存

ると、キャラクタがスライドの方を向き、スライド上に書き込んでいるようなジェスチャを行う(図3)。ボタンを離れた瞬間に、キャラクタは離れた位置の真横に立ち、書き込んだ箇所を説明しているように指し示すジェスチャを行う。他にも、マウスの中ボタンをドラックすることで、アニメーション付きの波線を表示し、スライド内の文字や図を強調することもできる。

2.2 キーボード操作

プレゼンテーションを行う際は、キャラクタの移動だけでなく、ユーザの意思や感情を伝えるためのさまざまなジェスチャが必要となる場面がある。そこで、キャラクタに対し、プレゼンテーション中のジェスチャで必要となるいくつかの姿勢をあらかじめ登録しておくことで(表2)、必要な動作をキャラクタアニメーションとして表現できる。キーを押すことで即座にキャラクタがアニメーションをするため、ユーザは、その場の話の内容に応じて、キャラクタの動きや表情をつ

表3 編集時の操作

Table 3 Manipulation for editing

操作	機能
左ドラッグ	スライドの移動
右ドラッグ	画面のスクロール
d	スライドの削除

けることができる。

2.3 再利用

CharacterPoint では、プレゼンテーションで行った際のユーザの音声と操作をすべて記録し、その記録を編集し再生することで繰り返し利用することができる。スライド順の変更などはマウスドラッグにより簡単に行える(表3, 図4)。例として、図5で編集操作の流れを示す。まず、マウスの右ドラッグによって、スライド全体をスクロールさせ、ユーザが見たいスライドを自由に見ることができる(図5(a)から図5(b)への変化)。スライドの並び替えは、移動したいスライドをつかむように左ドラッグし、移動先でマウスボタンを離すと、その場所にスライドが移動できる(図5(b)から図5(c)への変化)。また、任意のスライドにマウスカーソルを合わせ、キー'd'を入力することで、不要なスライドを削除することができる(図5(c)から図5(d)への変化)。さらに、新しくスライドが必要な場合は、スライド追加モードでプレゼンテーションを行うことで、そのプレゼンテーションのスライドが既存のスライドの最後尾に追加される(図5(d)から図5(e)への変化)。

CharacterPoint ではスライド単位の並び替えや、削除と追加を繰り返し、より適したプレゼンテーションを繰り返し再生することができる。

3. CharacterPointo の実装

3.1 画像解析

システムに入力されたスライド画像は、図6に示すように、システム内で格子状に分割される。それぞれの格子に対し、格子内の画素に文字や図が含まれていないかを調べ、あらかじめそれぞれの格子に文字や図があるかないかの情報を持たせておく。ユーザが任意の部分をクリックすると、格子の情報から文字や図部分の連結領域を計算し、もっとも近い右側の余白にキャラクタを配置する。また、右側にキャラクタを配置するための十分な余白が無い場合は、そこから下側の格子を探し、余白に配置する。図6では、半透明の網掛けの部分が文字や図部分と判断された格子であり、星印(☆)が左クリックされた場所を示している。キャラクタはクリックされた場所ではなく、文字

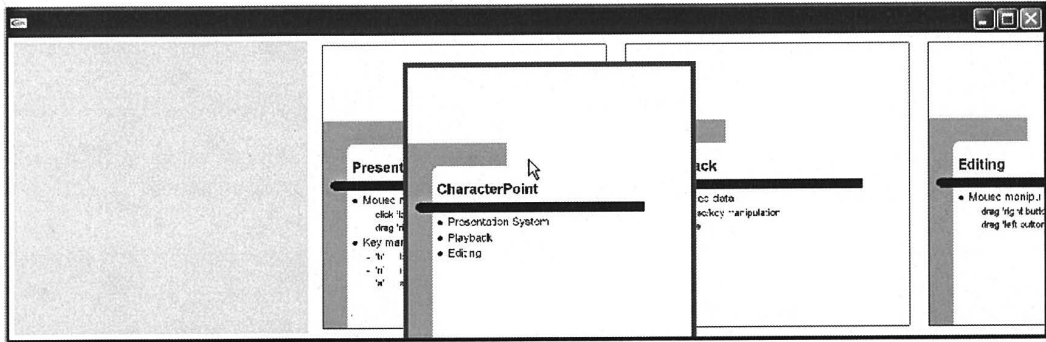


図 4 編集作業ウィンドウ
Fig.4 Screenshot of editing window

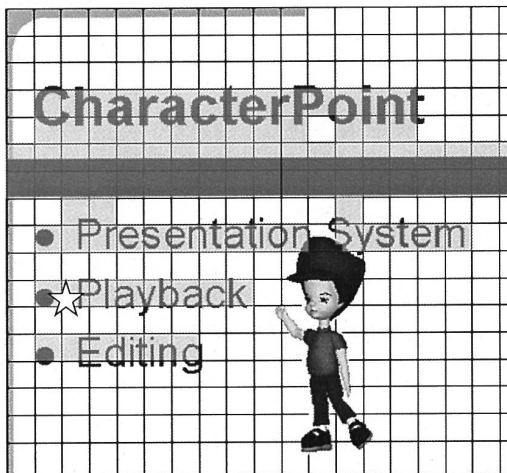


図 6 スライド画像の解析
Fig.6 Slide image analysis

や図部分でない右側の余白に配置されていることがわかる。

3.2 興味を引くジェスチャ

ユーザがスライド内容を説明しながらキャラクターのすべての動作を指定するのは困難である。しかし、入力がない場合にキャラクターがまったく動かないと、単調に感じられてしまう。そこで、「横を向く」や「手を広げる」など、プレゼンテーション中に演者が自然に行ういくつかのジェスチャは、システムがランダムで実行するようにしている。そのため、ユーザは内容を説明することに集中でき、キャラクターに意味のあるジェスチャをさせたい時だけ、キーを入力すればよい。マウスによる移動時もジャンプして移動するだけでなく、回転しながらのジャンプを含めることで、動きが

単調にならないようにしている。

また、キャラクター自体が説明しているように見せるため、ユーザの肉声に応じて、キャラクターの口を動かしている。マイクから入力された音声信号に対し、リアルタイムに周波数分析を行い、分析した周波数が人間の音声の範囲であり、かつ、ある程度の音量がある場合に、キャラクターが口を動かすようにしている。

3.3 プレゼンテーションの記録

CharacterPoint では、ユーザが行ったプレゼンテーションを再利用するために、ユーザの入力をすべて記録している。音声については、マイクからの入力を一連の音声ファイル (WAVE 形式) として保存している。また、マウスとキーの入力については、表 4 のように、入力した時間、入力に対応する音声ファイルの時間、入力キーの種類、表示するスライドの変化、および、入力時のマウスカーソルの座標を記録している。

例えば、表 4 において、スライド 1 を表示しているときに、キー 'b' でキャラクターがお辞儀をし (表 4 の ID: 1)、キー '↓' でスライド 2 が表示される (表 4 の ID: 2)。スライド 2 では、マウスの右クリックによって、キャラクターが (204, 89) に移動 (表 4 の ID: 3) した後、キー 'c' によって腕を組むジェスチャーを行う (表 4 の ID: 4) という流れになっている。ただし、ID: 0 の項目は特別な項目であり、ユーザの操作とは無関係に記録される。これは、プレゼンテーションを始めたときの最初のスライドがどのスライドであるかを記録するためのものである。

3.4 プレゼンテーションの編集

2.3 節で述べたように、ユーザは、プレゼンテーション後に、マウスを使ってスライド単位で編集を行うことができる (図 5)。例えば、図 5 における (b) から (c) の変化のように、1 番左に表示されているスラ

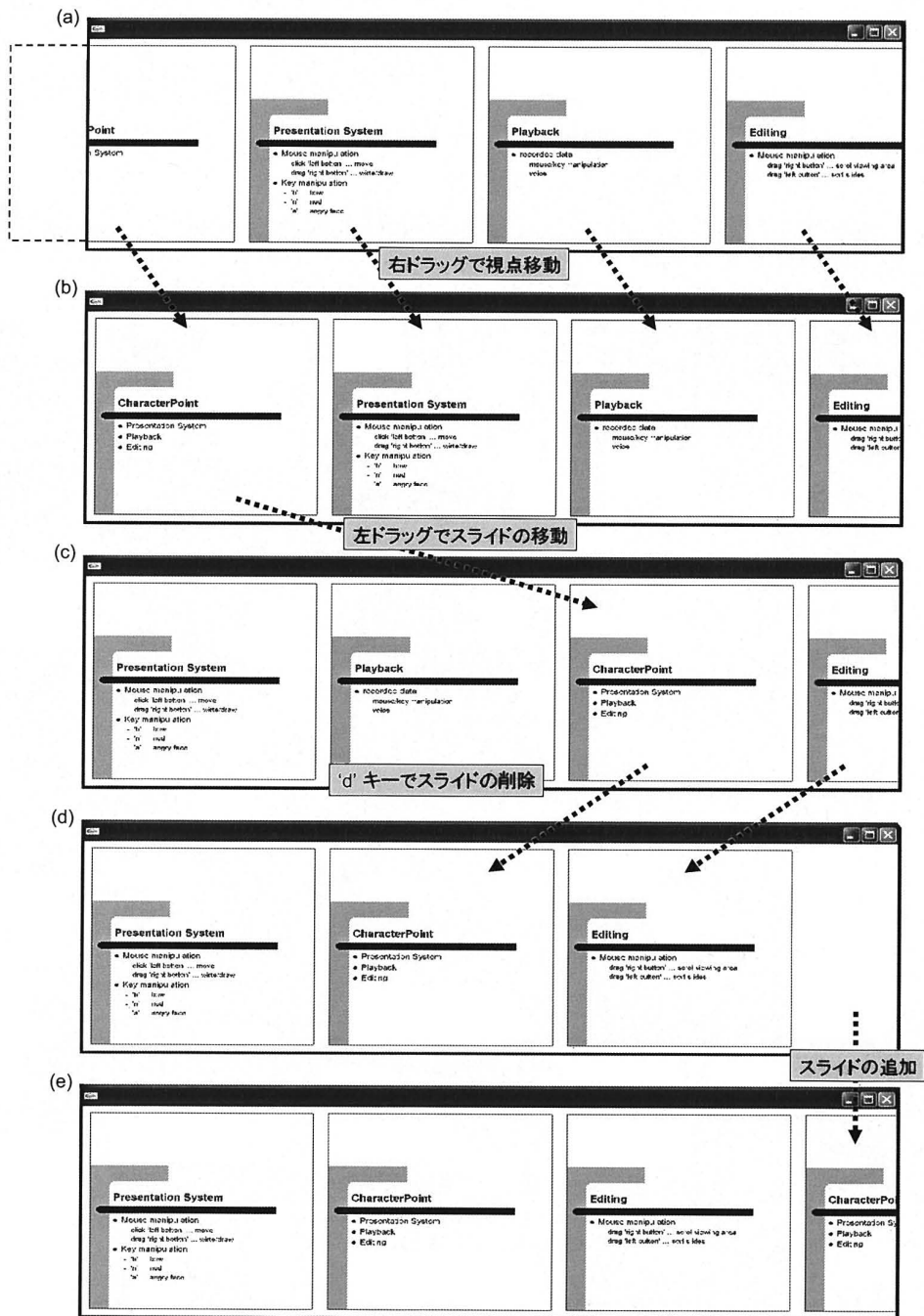


図 5 スライドの編集
Fig. 5 Editing

表 4 プレゼンテーション内容の記録データの例
Table 4 Example of recorded data

ID	0	1	2	3	4	5	6	...	n
時間	0	2468	5726	8612	10587	15221	17572	...	120932
音声時間	0	2468	5726	8612	10587	15221	17572	...	120932
キー	-	'b'	'↓'	'右ボタン'	'c'	'↓'	'↑'	...	'ESC'
状態	スライド 1	-	スライド 2	-	-	スライド 3	スライド 4	...	-
座標 (x, y)	(0, 0)	(216, 86)	(193, 102)	(204, 89)	(86, 182)	(38, 151)	(183, 82)	...	(93, 78)

表 5 編集後の記録データ
Table 5 Recorded data after editing

ID	2	3	4	5	0	1	6	...	n
時間	0	2886	4861	9495	11846	14314	17572	...	120932
音声時間	5726	8612	10587	15221	0	2468	17572	...	120932
キー	'↓'	'右ボタン'	'c'	'↓'	'↓'	'b'	'↑'	...	'ESC'
状態	スライド 2	-	-	スライド 3	スライド 1	-	スライド 4	...	-
(x, y)	(193, 102)	(204, 89)	(86, 182)	(38, 151)	(0, 0)	(216, 86)	(183, 82)	...	(93, 78)

IDを、3番目に移動したとする。この時、表4内のID: 1の項目は、スライド1上で操作された項目であるため、ID: 0の項目と一緒に移動しなければならない。このようにスライド単位での編集では、そのスライドに付随する操作も同じように移動する必要がある。さらに、移動後はスライドの順番が変わるため、操作の時間も調整する必要がある。表4では、スライド1を表示している時間は、ID: 0の項目から始まり、ID: 2の項目でスライドがスライド2に変わるまでの間なので、5726ミリ秒要することになる。そのため順番が前に来たスライド2に対応するID: 2からID: 5の項目は、5726ミリ秒分早くしなければならない。また、逆に後ろに移動したスライドに対応する項目は、ID: 2からID: 5に必要な時間の11846ミリ秒だけ追加する必要がある。移動後の記録データは、表5のようになる。音声ファイルに関しては、記録上の音声時間により、操作に対して音声ファイルのどの位置を再生すればよいか分かるので、並び替える必要はない。

また、削除する際は、対応するスライドに付随する項目をひとまとまりにして削除している。さらに、再度2章のプレゼンテーションを行うことで、それまでの記録に対して、新たな記録を追加することができる。追加されたスライドは、最後に追加されるが、編集を行うことで、任意のスライド順に変更することができる。

4. おわりに

本稿では、ユーザが発話しながら、リアルタイムでキャラクタを操作できるプレゼンテーションシステム

CharacterPointを提案した。マウスによるキャラクタの移動などの操作は簡単であるが、キーによるジェスチャ操作はキーの対応を覚える必要があり、より分かりやすい入力方法が求められる。例えば、五十嵐ら³⁾の研究のように、マウスで直感的にキャラクタを操作できるような工夫が必要となる。また、編集機能についても、機能が不足しており、今後はより細かな編集が行えるような機能を検討したい。

参考文献

- 1) Andre, E., et al: The PPP Persona: A Multi-purpose Animated Presentation Agent, *Proc. of Advanced Visual Interfaces*, pp. 254-247 (1996).
- 2) Gibbs, S., et al: Video Widgets and Video Actors, *Proc. of UIST'93*, pp. 179-185 (1993).
- 3) Igarashi, T., et al: Spatial Keyframing for Performance-driven Animation, *ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, pp. 107-115, (2005).
- 4) Kondo, K., et al: TablePresenter: A life-like agent explaining tables, *Proc. of VIIP 2005*, pp. 727-730 (2005).
- 5) Noma, T., et al: A virtual Human Presenter, *Proc. of IJCAI97 Workshop on Animated Interface Agents*, pp. 45-51 (1997).