

音声とハンドジェスチャーにより対話操作可能な3次元マルチメディアプレゼンテーションツール

Voice and Gesture Based 3D Multimedia Presentation Tool

福武 宏理 赤澤 由章 高野 茂 岡田 義広*

Summary. This paper proposes a 3D multimedia presentation tool that allows the user to manipulate intuitively only through the voice input and the gesture input without using a standard keyboard or a mouse device. We developed this system as a presentation tool to be used in a presentation room equipped a large screen like an exhibition room in a museum because, in such a presentation environment, it is better to use voice commands and the gesture pointing input rather than using a keyboard or a mouse device. This system was developed using *IntelligentBox*, which is a component-based 3D graphics software development system. In this paper, we introduce the proposed 3D multimedia presentation tool and describe its component-based voice and gesture input interfaces.

1 はじめに

本稿では、音声入力とハンドジェスチャー入力のみで対話操作可能な3次元マルチメディアプレゼンテーションツールを提案する。図1は提案ツールのハードコピーである。我々はこのツールを大型スクリーンを備えたプレゼンテーションルームでの使用を想定して開発した。このようなプレゼンテーション環境では、マウスやキーボードといった従来の入力インタフェースでは、講演者はPCの前になければアプリケーションの操作ができないといった制約が発生する。そのため、大画面環境におけるプレゼンテーションには図2に示すように音声やハンドジェスチャーといった入力形態によるアプリケーション操作の方が適していると考えられる。本提案ツールはインテリジェントボックス[1]と呼ばれる3次元ソフトウェア開発支援を目的としたコンポーネントウェアを用いて開発した。インテリジェントボックスは、固有の機能を持つボックスとよばれる3次元可視部品およびスロット結合とよばれるボックス同士の機能連携機構を提供し、マウスを用いた画面上の対話操作のみによる3次元ソフトウェアの開発を可能とする。インテリジェントボックスは既に様々な機能をボックスとして提供しており、それらの中には、音声入力インタフェース部品、ジェスチャー入力インタフェース部品、マルチメディアデータを扱うための部品なども含まれている。よって、インテ

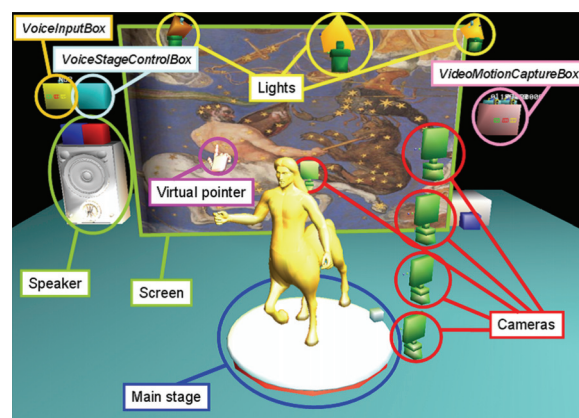


図 1. 3次元マルチメディアプレゼンテーションツール

リジェントボックスを使用する事により本提案ツールも画面上でのマウスによる対話操作のみにより開発された。我々は既に、音声操作機能を持つ3次元マルチメディアプレゼンテーションツール[2]を提案しており、今回は、ユーザのハンドジェスチャーによる操作も行えるよう機能の拡張を行った。

2 3次元マルチメディアプレゼンテーションツール

我々が提案するツールは3次元仮想空間内に実世界のステージを模倣して構築された仮想ステージである。この仮想ステージはメインステージ、カメラ、ライト、スクリーン、スピーカーとしての機能を持つソフトウェア部品群から構成される。これらのソフトウェア部品群を操作することで、ステージ上に表示された3次元マルチメディアデータのプレゼンテーションを行うという特徴を持っている。

Copyright is held by the author(s).

* Hiromichi Fukutake, 九州大学大学院 システム情報科学府 情報理学専攻, Yoshiaki Akazawa, 産業総合研究所 デジタルヒューマン研究センター, Shigeru Takano and Yoshihiro Okada, 九州大学大学院 システム情報科学研究 院 情報理学部門

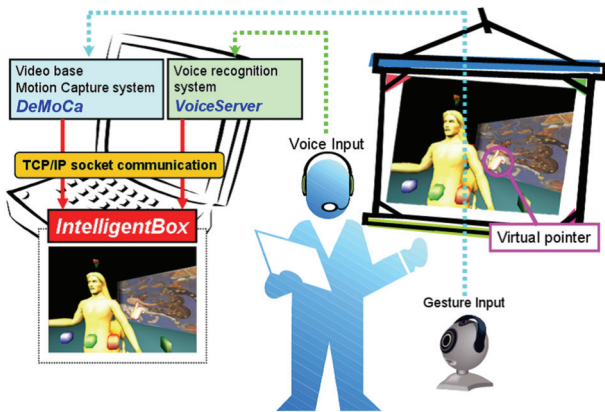


図 2. プレゼンテーション環境

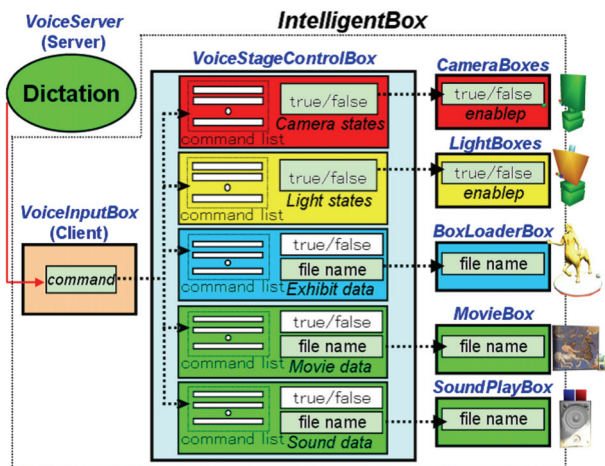


図 3. 音声入力インタフェースの構成

3 音声入力インタフェース部品

図 3 はソフトウェア部品化した音声入力インタフェースの構成である。インテリジェントボックスは音声認識機能を持っていないため、Microsoft Speech API を使用して VoiceServer とよぶサーバープログラムを開発した。VoiceInputBox は TCP/IP ソケット通信のクライアントとして機能し、VoiceServer とソケット通信を行うボックスである。VoiceServer から音声認識結果を受け取り、スロット結合により VoiceStageControlBox へ音声コマンドとして送る。VoiceStageControlBox は仮想ステージを構成するそれぞれのソフトウェア部品を操作するための値を持っている。また、音声コマンドの学習機能と学習した音声コマンドを受け取る事によりそれらの値を変化させる機能を持っている。よって、VoiceInputBox, VoiceStageControlBox, 仮想ステージを構成するソフトウェア部品群をスロット結合により合成することで、音声による提案ツールの操作が可能となる。

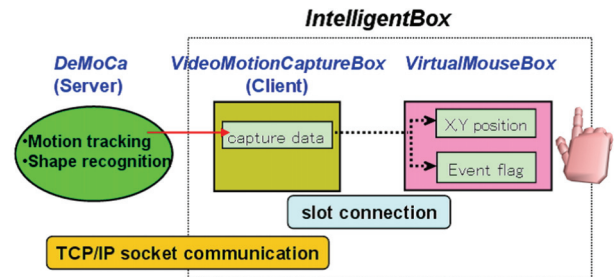


図 4. ジェスチャー入力インタフェースの構成

4 ジェスチャー入力インタフェース部品

図 4 はソフトウェア部品化したジェスチャー入力インタフェースの構成である。ハンドジェスチャーによるポインタ操作を実現するために、赤澤らによって開発された DeMoCa システム [3] を用いた。DeMoCa は 1 台のカメラのみを使用するビデオベースのモーションキャプチャシステムであり、追跡している部位の x-y 座標位置の取得と手の形状認識を行うことができる。VideoMotionCaptureBox は DeMoCa から手の x-y 座標位置および形状認識結果を受け取る機能を持つ。VirtualMouseBox はスロット結合により VideoMotionCaptureBox から手の x-y 座標位置を受け取り、インテリジェントボックスのビューポート上の対応する位置に仮想ポインタを表示する。また、形状認識結果に応じてクリック操作などのイベントを制御する機能を持っている。これらのソフトウェア部品を用いることで、ハンドジェスチャーによるポインタ操作が可能となる。

5 まとめ

音声入力とジェスチャー入力のみで操作可能な 3 次元マルチメディアプレゼンテーションツールの開発を行った。提案ツールの音声およびジェスチャー入力インタフェースはボックスとして開発しているため、インテリジェントボックスが提供する既存のアプリケーションにも容易に組み込むことができる。

参考文献

- [1] Y. Okada and Y. Tanaka, IntelligentBox: A Constructive Visual Software Development System for Interactive 3D Graphic Applications, Proc. of Computer Animation '95, IEEE CS Press, pp.114-125, 1995.
- [2] H. Fukutake, Y. Okada and K. Nijima, 3D Multimedia Presentation Tool Using Stage Metaphor and Its Voice Interface, Proc. of VSMM2005, pp. 531-540, 2005.
- [3] Y. Akazawa, Y. Okada and K. Nijima, Real-time Video Based Motion Capture System Based on Color and Edge Distributions, Proc. Of ICME 2002, Vol. II, pp. 333-336, 2002.