
インタラクティブドッグ

～仮想犬との3種類のインタラクション～

An Interactive Dog — Three Kinds of Interaction with a Virtual Dog —

阿部 哲也 後藤 真孝 黒田 洋介 村岡 洋一*

Summary. In this paper, we present an interactive dog that is a virtual dog on 3D computer graphics (CG) and that can interact with a user in various ways. In previous work, a user usually interacts with a CG character through only one interface. This paper proposes the following three kinds of interaction with the virtual dog through three interfaces such as a camera, a musical instrument and a computer keyboard: 1) Jumping over hurdles by directing the dog through the camera, 2) Ensemble with the dog by playing the musical instrument, 3) Hurdle-race among several users' dogs by using the keyboards. In the first interaction, the user can feel a sense of closeness with the dog by interaction using his body gestures. In the second interaction, the user can enjoy a dynamically-changing ensemble with the dog's chord play that is determined by a dog's trail. In this case, the user can indirectly change the trail by his musical play. In the third interaction, the users can enjoy game-like interaction in the hurdle-race on tracks. We implemented the system that achieved these kinds of interaction on workstations connected to the Ethernet. In our experiment, we experienced the proposed interaction with the interactive dog by combining the three interfaces.

1 はじめに

本稿で提案するインタラクティブドッグとは、映像や音などのメディアを介して人間とのインタラクションを行なうことのできる、3D Computer Graphics(CG) で表現された仮想の犬である。インタラクティブドッグでは、人間は受動的に犬の映像を眺めるだけでなく、能動的に犬とのインタラクションを楽しむことができる。本研究は、このようなインタラクティブに変化する仮想犬の実現を通じて、人間とCGキャラクタとの新たなインタラクションを考案することを目的とする。

従来の人間とCGキャラクタとのインタラクションは、ALIVE[1] や音楽演奏に反応するCGダンサー [2]、ニューロベイビー [3]、仮想生物システム [4] などで実現

* Tetsuya Abe, 早稲田大学 大学院 理工学研究科 電気工学専攻, Masataka Goto, 早稲田大学 理工学部/日本学術振興会特別研究員, Yosuke Kuroda, 早稲田大学 理工学部 情報学科, Yoichi Muraoka, 早稲田大学 理工学部

されていた。しかし ALIVE や CG ダンサーでは、一種類のインタフェース (ALIVE ではカメラによるジェスチャー入力、CG ダンサーでは音楽演奏入力) を介してしかインタラクションを行なうことができなかった。複数のインタフェースを組み合わせたインタラクションはニューロベイベーや仮想生物システムで実現されていたが、仮想犬との複数のインタフェースを用いたインタラクションは行なわれていなかった。

本稿では、複数のインタフェース (カメラによるジェスチャー入力、MIDI 楽器による演奏、計算機のキーボードによる指示) を用いた仮想犬との 3 種類のインタラクションを提案する。具体的には、1) 仮想犬によるハードルクリア、2) MIDI 楽器を用いた仮想犬との合奏、3) 複数のユーザ間でのハードルクリア競争、を提案する。これらを様々なインタフェースにより行なうことで、仮想犬とのインタラクションを多様な角度から体験できる。また、複数のインタフェースを組み合わせたインタラクションや、複数のユーザが異なるインタフェースを用いたインタラクションも可能である。

1) 仮想犬によるハードルクリアでは、ユーザはジェスチャーにより仮想犬に指示を与えることでインタラクションを行なう。実世界の犬とインタラクションする場合に一番近い感覚であるジェスチャーを用いることで、犬との一体感を得られる。2) MIDI 楽器を用いた仮想犬との合奏では、ユーザが演奏する音楽と仮想犬の演奏する和音とで合奏を行なう。仮想犬の演奏は仮想犬の位置によって変化し、ユーザは演奏やジェスチャー入力によって仮想犬に動作の指示を与えることで、ある程度仮想犬の演奏に干渉できる。これにより、ユーザは犬とインタラクティブに変化する合奏を行なえる。3) 複数のユーザ間でのハードルクリア競争では、各ユーザはそれぞれ各仮想犬の仮想の騎手となり、計算機のキーボードを手綱にみだてて指示を与えることでインタラクションを行なう。どの仮想犬が速くハードルをクリアできるかを競い合うことで、ユーザはゲーム感覚のインタラクションを楽しめる。

以上のようなインタラクションを実現するシステムを実装し、それぞれのインタラクションを実験した。その結果、本稿で提案した 3 種類のインタラクションを実際に行なうことができた。さらに、複数のインタフェースを同時に組み合わせたインタラクションや、複数のユーザの協調によるインタラクションが行なえた。

2 仮想犬との 3 種類のインタラクション

仮想犬とは、計算機上に設定した仮想空間内で動作する、多関節骨格モデルを用いて CG で表現された犬のことである。仮想犬は、歩行、走行、および跳躍動作を行なう。それぞれの動作においては、移動の速度 (歩幅と動作の周期) と平面上での左右の移動方向を変えられる。跳躍動作では、自分の前に障害物があれば、そこまでの距離とその高さに応じて自律的に歩幅と跳躍の高さを調整して飛び越える。

さらに、仮想犬は自律的に動くだけでなく、インタフェースを通じてユーザの動きを検知しそれに対応して行動することができる。これにより、ユーザは仮想犬とのインタラクションを楽しむことが可能になるので、我々はこの仮想犬をインタラクティブドッグと呼ぶ。以下では、このインタラクティブドッグとの 3 種類のインタラクションを、具体的に説明する。

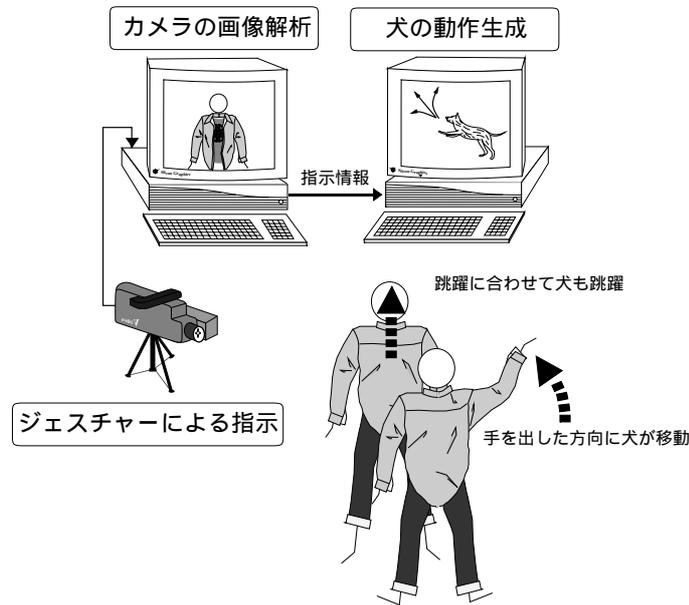


図 1. ジェスチャーによるハードルクリア

2.1 仮想犬によるハードルクリア

本インタラクションでは、仮想空間内にトラックとハードルを用意しておき、ユーザはジェスチャーによって指示を与えることで、仮想犬にハードルを飛び越えさせる(図1)。犬の目に相当するカメラでリアルタイムに取り込んだ画像を解析することで、仮想犬は、ユーザの左右の方向転換の指示や跳躍の指示を検知し、それに反応して行動する。

このインタラクションでは、犬はユーザのジェスチャーを見て移動方向を変えたり跳躍したりする。犬は、ユーザが手を水平に横に出すのを見ると、手の方向へ移動方向を変える。ユーザが跳躍するのを見ると、犬もそれに合わせて跳躍する。これにより、ユーザは犬との一体感を感じながらインタラクションできる。

インタフェースとしては、カメラによるジェスチャー入力他に計算機のキーボードによる指示を用いることもできる。この場合、仮想犬との一体感はあまり得られないが、より正確な指示を与えることができる。

2.2 MIDI 楽器を用いた仮想犬との合奏

本インタラクションでは、仮想空間内に3色のタイルを敷き詰めた床を用意しておき、仮想犬は歩きながらこのタイルを踏むことで演奏する。3色のタイルには、それぞれブルースの基本となる3つの和音が割り当てられており、踏まれたタイルの色に対応した和音が出力される。常にドラムスのパターンが一定のテンポで流れており、犬が上記のように定めた和音が小節の先頭で演奏される。つまり、仮想犬は移動することで、動的に和音進行を作り出す。

一方、ユーザはMIDI 楽器を演奏することにより、仮想犬と合奏しながら、仮想犬の動作を変えることができる(図2)。その際、ユーザには仮想犬の周辺のタイルの

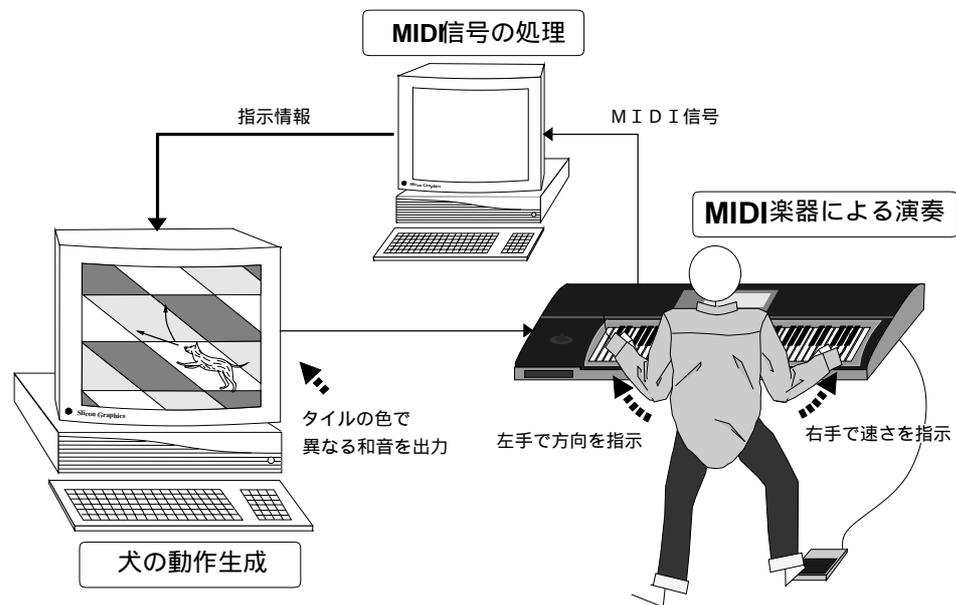


図 2. MIDI 楽器を用いた仮想犬との合奏

色が見えるので、演奏させたい和音のタイルに仮想犬が向かうように、ユーザは演奏を変化させる。こうして仮想犬が位置を変えると、その和音演奏も変わり、変化した演奏に合わせてユーザがさらに演奏を変える、というようなインタラクションが起こる。

仮想犬はユーザの演奏を聞いて、移動速度や移動方向を変える。具体的には、高域の旋律を聞いた場合と低域の旋律を聞いた場合で、以下のような異なる反応をする¹。高域の旋律（一般にはメロディー）を聞くと、仮想犬は移動速度を変える。この旋律の音を多く聞くほど、仮想犬は足早に移動し、高域でもさらに高い方の音を聞くほど、歩幅を広げて移動する。つまり、ユーザの盛り上がった演奏を聞くと、仮想犬は元気のよい動きをする。一方、低域の旋律（一般にはコードやベース）を聞くと、その音の高さに応じて犬は移動方向を変える。犬は低域をさらに三分割（低・中・高）して聞き分けることができ、それぞれに応じて左・直進・右に方向を変える。

ユーザは音楽演奏だけでなく、2.1で述べたジェスチャーも同時に用いてインタラクションを行なえる。音楽演奏とジェスチャーを組み合わせることで、ユーザは旋律を演奏しながら、ジェスチャーで左右の方向転換などの指示を犬により正確に伝えることが可能になる。また演奏とジェスチャーを二人のユーザで別々に担当し、仮想犬と二人のユーザによる新しいインタラクションも行なえる。

2.3 複数のユーザ間でのハードルクリア競争

本インタラクションでは、仮想空間内に各ユーザ分のトラックとハードルを用意しておき、各ユーザが担当する仮想犬同士に、いかに速くハードルをクリアして走る

¹ 高域・低域は、特定の音程（スプリットポイントと呼ぶ）より高いか低いかで判定する。

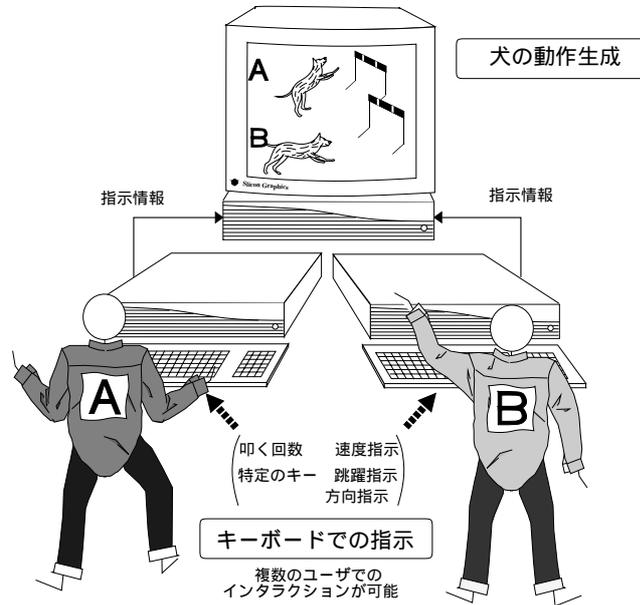


図 3. 複数のユーザ間でのハードルクリア競争

ことができるかを競わせることで、仮想犬を介したユーザ間のインタラクションを行なう(図3)。基本的にユーザは、2.1で述べたハードルクリアのインタラクションを行なうが、カメラによるジェスチャー入力を行なわない。全ユーザは一つの画面を見ながら、それぞれの計算機のキーボードで仮想犬に指示を与えて競争する。

仮想犬は、ユーザが叩いたキーの種類を検知して、移動方向を変えたり、跳躍動作を行ったりする。また、ユーザが単位時間当たりキーを叩く回数が増えるほど、犬は速度を上げて動作する。ユーザは、犬とのインタラクションに加え、他のユーザと競争することで、ゲーム感覚でインタラクションを楽しめる。

3 インタラクティブドッグの実装

インタラクティブドッグを実現するには、カメラによるジェスチャー入力とMIDI楽器による音楽演奏、計算機のキーボードによる指示という3種類のインタフェースの処理と、仮想犬の動作を生成する処理が必要である。実装に当たっては、仮想犬の動作を生成する処理と、インタラクションのためのインタフェースの処理とを分離し、異なるプロセスとした。動作を生成する処理をサーバ、インタフェースの処理をクライアントとし、その処理の間で通信を行なうサーバ・クライアント・モデルとして実装した。こうすることにより、仮想犬の動作を生成する処理とカメラ画像からユーザの動作を検出する処理という計算量が多い処理を、複数の計算機上に負荷分散することでリアルタイムに処理できる。また、インタフェースの処理を独立させておくことで、新たなインタフェースを追加するときの拡張が容易になる。

各処理間の通信には、RMCP (Remote Media Control Protocol)²[2] を用いた。

² 基本的には文献[2]のRemote Music Control Protocolに基づいているが、複数メディアに対応す

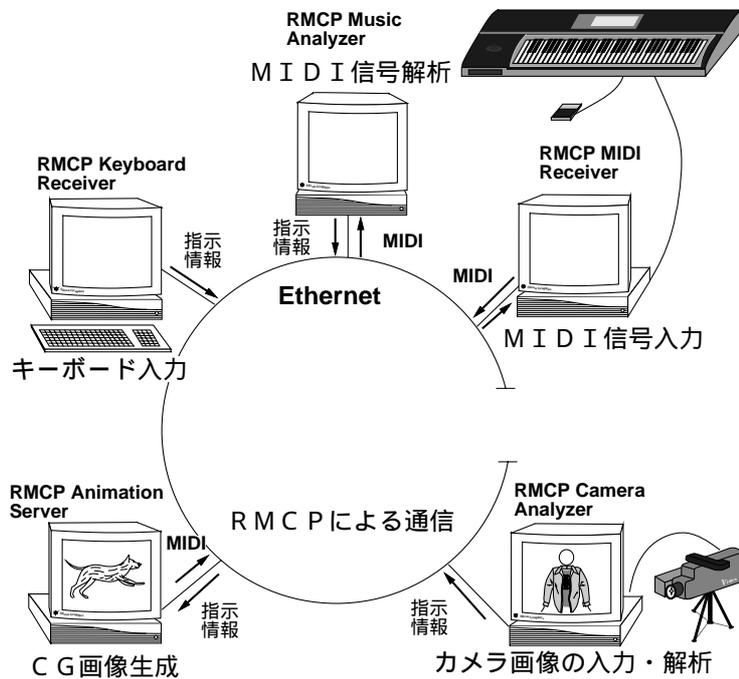


図 4. RMCP を用いた実装

インタフェースの処理と、仮想犬の動作を生成する処理は、図 4 に示すようにそれぞれ独立したプロセスとして実装した。各インタフェースの処理は、カメラの画像入力・解析を RMCP Camera Analyzer、MIDI 楽器による MIDI 信号の入力・解析を RMCP MIDI Receiver と RMCP Music Analyzer、計算機のキーボードによる入力を RMCP Keyboard Receiver として実装し、仮想犬の動作を生成する処理は RMCP Animation Server として実装した。

3.1 インタフェース用プロセス

それぞれのインタフェース用プロセスは、ユーザの入力を解析した内容を、イーサネット上に RMCP パケットとしてブロードキャストする。複数のユーザによるインタラクションを可能にするために各ユーザは異なるユーザ識別子を持っており、各クライアントはこのユーザ識別子も同時に出力する。

各プロセスの処理の概要を以下に示す。

3.1.1 RMCP Camera Analyzer

RMCP Camera Analyzer は、カメラからの画像をリアルタイムに取り込んで解析し、ユーザの左右の手の平の位置と頭の高さを出力する。前提条件としてカメラと背景は固定であるとし、ユーザはその背景の前に垂直に起立した状態でインタラクションを行なうものとする。まず最初に背景をユーザのいない状態で取り込んで用

るためにメッセージを拡張してあるので、Remote Media Control Protocol と呼ぶ。

意しておく。そして、ユーザの上半身を画像として取り込んで、以下の手順で解析する。ただし、取り込む画像の解像度は 160×120 ピクセルである。

1. 体の縦の中心線・肩の高さ・肩幅の決定

人間の上半身を正面から見た場合、肩の部分が最も横幅が広く、縦は体の中心が最も高い。そこで、各フレームで取り込んだ画像と背景の差分を取り2値化して、縦・横の各軸上でヒストグラムを作る。縦軸のヒストグラムで最大となる所を肩の高さ³、横軸で最大となる所を体の縦の中心線とする。また、肩より下の縦軸のヒストグラム中ですべての値の平均値を求め、その値の2倍を肩幅とみなす。

2. 頭の高さの決定

1で求めた体の縦の中心線を中心に肩幅の範囲内で縦軸のヒストグラムを作る。そして、肩の高さから上の位置で、ヒストグラムの値が最初に肩幅の $1/10$ 以下となる所を頭の高さとする。

3. 両手の平の位置の決定

体の肩より外側の部分では、手の平が最も激しく移動する。そこで、1フレーム前の画像との差分を取って2値化し、肩より外側の部分で、差分の存在した画素の座標の平均を縦・横軸について求め、それを手の平の位置とする。

以上のようにして求めた頭の高さと両手の平の位置の情報を、RMCP Camera Analyzer は、イーサネット上にブロードキャストする。各情報は十数分の一秒ごとに一回ずつ出力される。

3.1.2 RMCP MIDI Receiver, RMCP Music Analyzer

RMCP MIDI Receiver では、MIDI 楽器から MIDI メッセージを受けとり、RMCP パケットとしてブロードキャストする。RMCP Music Analyzer では、この RMCP パケット中の MIDI メッセージを受けとって、ユーザの演奏を分析する。0.1 秒ごとに、ユーザの過去の一定期間 (2 秒間) の演奏情報について分析を行ない、その結果を出力する。

出力される分析結果は、スプリットポイントを中心として、左側 (低域) で弾いた音の音程 (MIDI note number) の平均値と、右側 (高域) で弾いた音の音数と音程の平均値である。左側の音程に関しては、一定期間に弾いたすべての音に関して、弾いた強さ (MIDI velocity) による重み付き平均を求め、それを音程に正規化する。右側の音程に関しては、音程の平均値は左側と同様にして求め、音数は弾いた強さを重みとして重み付け加算したものとす。

3.1.3 RMCP Keyboard Receiver

RMCP Keyboard Receiver は、ユーザが叩いた計算機のキーボードのキーの種類と、過去の一定期間に叩いたキーの数を出力する。各キーの機能の割り当ては、ユーザが任意に変更できる。叩いたキーの種類情報はリアルタイムに出力し、一定期間に叩いたキーの数は 0.1 秒ごとに出力する。

³ このようにヒストグラムの最大値で肩の位置を検出する簡略化した実装を行なったため、ユーザはウエストが胸囲よりも細いことが要求される。

3.2 仮想犬の動作を生成するプロセス (RMCP Animation Server)

RMCP Animation Server は、各インタフェース用クライアントが出力した RMCP パケットを受けとり、それに対応した犬の動作を生成する。RMCP Animation Server は、これらのパケットを同時に受け取れるので、複数のインタフェースを同時に使ってインタラクションをすることが可能である。また、各クライアントが出力するユーザ識別子を利用して、複数のユーザが複数の仮想犬とそれぞれインタラクションを行なう画像も生成できる。

インタラクションを実現するには、仮想犬の柔軟な動作をリアルタイムに生成する必要がある。そこで、ここではそのような生成を可能にする文献 [5] の手法を用いる。一般に、多関節骨格モデルの動作をリアルタイムに生成するには、計算量の比較的少ない運動学に基づいた動作生成を行なうとよい。しかし、通常の運動学に基づいた手法では、固定した動作しか生成できない。我々が用いた文献 [5] の手法では、運動学に基づいた動作生成を行なうことでリアルタイム性を確保しつつ、通常は固定されている動作から歩幅や周期、移動方向などを変化させた柔軟な動作を生成可能にしている。これは基本的に、固定した動作の関節角度を制御パラメータで変化させることで実現する。ただし、こうして関節角度を変化させると、変化させる前に推定された身体全体の位置に矛盾が生じる。そこで、この位置のずれを各フレームごとに動的に補正することで、この矛盾を解消する。また、飛び越えるハードルまでの距離と高さに合わせて歩幅と跳ぶ高さを調整する処理には、文献 [6] の手法を用いる。

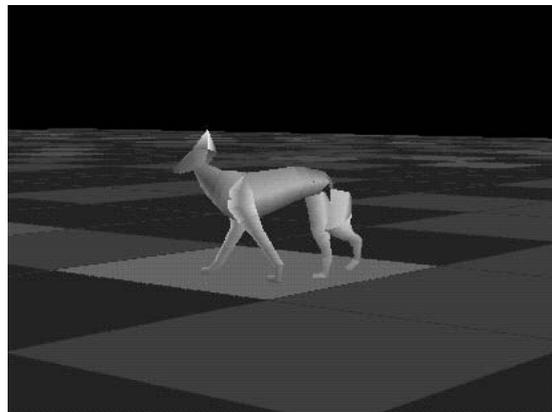
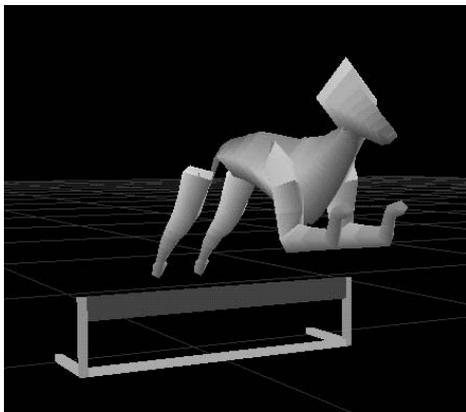


図 5. ハードルをクリアする仮想犬の画像例 図 6. 音楽演奏によるインタラクションでの画像例

3.3 実装したシステムによるインタラクションの実現

提案した 3 種類のインタラクションは、3.1、3.2 で述べた各プロセスを組み合わせることで以下のように実現する。

1. 仮想犬によるハードルクリア

RMCP Camera Analyzer と RMCP Animation Server を用いて実現する。ただし、ジェスチャーでなく計算機のキーボードで指示を与えるときには、RMCP Keyboard Receiver と RMCP Animation Server を用いてもよい。

2. MIDI 楽器を用いた仮想犬との合奏

RMCP MIDI Receiver と RMCP Music Analyzer、RMCP Animation Server を用いて実現する。さらに、RMCP Camera Analyzer を組み合わせることで、複数のインタフェースを同時に使ったインタラクションや、複数のユーザの協調によるインタラクションも行なえる。

3. 複数のユーザ間でのハードルクリア競争

RMCP Keyboard Receiver と RMCP Animation Server を用いて実現する。

上記以外にも、ユーザの要求に合わせてインタフェースを選択してインタラクション可能である。

4 実験結果

本システムを 2 台の SGI Indigo² Extreme、及び SUN Spare Station¹⁰ 上に実装した。一台の SGI を画像生成処理専用を使い、もう一台の SGI と SUN を、カメラ画像解析などのインタフェース用の処理に用いた。今回の実験で行なった 3 種類のインタラクションとその結果を以下に示す。

1. 仮想犬によるハードルクリア

インタフェースにカメラによるジェスチャー入力を用いて実験した。一秒間に十数回画像入力処理を行なっているにもかかわらず、犬の反応に若干遅れが感じられたが、体全体を使ってインタラクションをすることにより感情移入ができて、犬との一体感を得ることができた。

2. MIDI 楽器を用いた仮想犬との合奏

まず、インタフェースとして MIDI 楽器 (シンセサイザー、MIDI ギター) による音楽演奏を用いて実験した。犬が歩きながら演奏する和音の移り変わりに合わせて旋律を弾くのは、犬の動きを完全には予測できないだけに興味深かった。

次に、演奏に加えて、ジェスチャーも用いた複数のインタフェースを同時に使った実験を行なった。MIDI 楽器としてシンセサイザーの鍵盤を用いる場合には、両手で即興演奏を行なわなければならないが、片手でしか弾けない人も多い。そこで、通常左手の低域演奏で行なう左右の方向転換を、ジェスチャーによるインタラクションで行なうことで、片手でしか鍵盤を弾けないユーザでもこのインタラクションを楽しむことができた。

さらに、音楽演奏とカメラによるジェスチャー入力を、二人のユーザがそれぞれ担当するインタラクションの実験も行なった。音楽演奏する側にとっては、他方のユーザがどのように犬の進行方向を変えるか予測できないため、次に来る和音が突然変わる場合にも対応しようとして演奏する緊迫感が味わえる。また、ジェスチャー入力をする側にとっては、自分が指示した犬の進行方向で決まる和音に対して、他方のユーザがどのように演奏を返すかを期待して楽しむことができる。

3. 複数のユーザ間でのハードルクリア競争

二人のユーザが計算機のキーボードを用いて、二匹の仮想犬で競争する実験を行なった。ここでは、ゲーム感覚でインタラクションを楽しみつつ、仮想犬を介して複数のユーザが競争するという、上記の他のインタラクションとは異なる感覚を味わうことができた。

5 おわりに

本稿では、ユーザが3種類のインタフェースを用いて3種類のインタラクションを行なうインタラクティブドッグを提案した。インタフェースとして、カメラによるジェスチャー入力、MIDI楽器での音楽演奏、計算機のキーボードによる指示の3種類を使い、1) 仮想犬によるハードルクリア、2) MIDI楽器を用いた仮想犬との合奏、3) 複数のユーザ間でのハードルクリア競争、の3種類のインタラクションを実現した。仮想犬は、ユーザのジェスチャーや演奏に応じて、移動速度や移動方向を変えたり、跳躍動作を行ったりする。これらのインタラクションを実現するシステムを実装する際に、画像生成処理とインタフェースの処理を分離したことで、負荷が分散でき、拡張性が高くなった。本システムを実装して実験した結果、それぞれのインタフェースを選択・組み合わせることで、提案した3種類のインタラクションが実現できた。

今後の課題としては、マイクやタッチセンサなどの、より多くのインタフェースを用いたインタラクションを可能にすることがあげられる。また、犬の動作に影響を与えるような動くオブジェクトを用意し、ユーザはその影響を考えながら犬に指示を与えるインタラクションを実現する予定である。さらに、現在犬のできる動作は限定されているのでこれを増やすとともに、犬以外の動物への応用も考えていく。

謝辞

本研究を行なうにあたり、御支援を頂いた早稲田大学村岡研究室の皆様には感謝します。

参考文献

- [1] Pattie Maes, Trevor Darrell, Bruce Blumberg and Alex Pentland, "The ALIVE system: Full-body Interaction with Autonomous Agents", Proceedings of the Computer Animation'95 Conference, Geneva, Switzerland, IEEE-Press, April 1995.
- [2] 後藤 真孝, 橋本裕司: 「MIDI 制御のための分散協調システム - 遠隔地間の合奏を目指して - 」, 情処研報, Vol.93, No.109, 音楽情報科学 93-MUS-4-1 (1993).
- [3] N. Tosa, H. Hashimoto, K. Sezaki, Y. Kunii, T. Yamada, K. Sabe, R. Nishino, H. Harashima, F. Harashima: "Network Neuro-Baby with robotic hand", 人工知能学会 第9回全国大会, エンターテイメントとAI, July 1995.
- [4] 藤田 卓志, 西山 聡一, 渡辺 和之, 田口 ひとみ, 福岡 俊之, 伊藤 映: 「仮想生物システム - 体験者インタフェース - 」, 情報処理学会 第46回全国大会, 3K-4, March 1993.
- [5] 阿部 哲也, 水野 裕識, 村岡 洋一: 「CG による犬の柔軟なアニメーション生成」, 情処研報, Vol.95, No.18, グラフィックスとCAD 95-CG-73 (1995).
- [6] 阿部 哲也, 水野 裕識, 村岡 洋一: 「CG 犬の柔軟な跳躍動作生成手法」, 情報処理学会 第49回全国大会, 3R-6, September 1994.