

3次元動画像処理による仮想楽器システムの提案

A virtual instrument system using the 3D time-varying image processing

一ノ瀬元喜*¹ 坂東哲平*²
Genki Ichinose Teppei Bandou

*¹阿南工業高等専門学校制御情報工学科

Department of Systems and Control Engineering, Anan National College of Technology

*²徳島大学工学部

Faculty of Engineering, The University of Tokushima

Augmented reality (AR) is a physical real-world environment whose elements are mixed with (or augmented by) virtual computer-generated image. The virtual 3D objects on the real environment through devices are a new trend and are expected various applications such as surgery assistants and navigation systems. ARToolKit is a software library to create AR applications. The 3D time-varying image processing is one of the necessary technics to handle 3D objects in AR. We aim to manipulate virtual objects freely by combining AR and the technics of the 3D time-varying image processing. In this paper we developed a virtual instrument system using ARToolKit and OpenCV as a first step.

1. はじめに

現実世界に対して仮想物体を合成提示する概念は拡張現実 (Augmented Reality, 以下 AR) と呼ばれる。AR によってデバイスを通して現実の環境に仮想的な 3D 物体を提供することは、最近のトレンドであり、道案内、手術支援などといった応用も期待されている。AR をプログラミングによって実現することは最近まで難しい技術であったが、ARToolKit[Kato 99] の登場により、AR を比較的容易でリアルタイムに実現することが可能となった。ARToolKit とは AR を実現するための C 言語ライブラリである。ARToolKit を用いて仮想現実 (VR) や、テクノロジーアートなどの研究や作品が発表されている。また AR は身につけて持ち歩けるウェアラブルコンピュータとの融合も行われており、ウェアラブル AR と呼ばれるシステムでは任意の自己位置・姿勢に応じた位置依存情報を AR によって得ることができる [天目 05, 浦谷 05, 牧田 08]。

AR と人間とのインタラクションという観点からは仮想的な AR を現実の人間が自在に操作できることは重要である。AR はデバイスを通して存在し人間の目に映るので、AR によって配置された物体を操作するには画像処理による物体の認識が必要となる。OpenCV という画像処理のライブラリを使うと複雑な画像処理も容易に行うことができる。そこで本研究では AR と画像処理の技術を組み合わせて仮想物体を人間が自由に操作することを目指す。そのための第一歩として本研究では仮想的なピアノの鍵盤を生成し、人間がそれを演奏するシステムを構築した。AR の技術を用いたピアノやギター演奏支援システムなどは提案されているが [樋川 06, 元川 07, 元川 08]、本研究では作りだしたピアノオブジェクトそのものを仮想的に演奏するシステムを作成する。AR で構築された情報はこれまでは現実空間に対する付加的な側面が強かったが、本システムでは AR による仮想的なピアノオブジェクトが直接操作対象である点に特徴がある。

本研究の目的は、ARToolKit と OpenCV を用いて、仮想的な楽器を擬似的に演奏するシステムを作成することである。本研究で構築するシステムは、小規模で持ち運びできるという携帯性と仮想楽器の演奏というエンターテインメントの両面から

有益であると考え、鍵盤はヘッドマウントディスプレイ (以下 HMD) を通じて人間に認識される。画像処理によって指先を解析し、それが鍵盤と触れるのを判定し、音を発するプログラムによって本システムは構成される。本稿の構成は以下の通りである。まず第 2 章でシステム概念や必要な周辺機器、技術について概説する。第 3 章では仮想楽器演奏システムの構成について説明する。第 4 章では仮想楽器演奏の手順などを示し、第 5 章では実験結果、考察と今後の展望、最後の第 6 章で結論を述べる。

2. システムの概要

2.1 概要

本研究では ARToolKit を用いたプログラムによって仮想的な楽器を生成させ、それを OpenCV による画像認識で指の位置を検出することで演奏するシステムを構成する。本システムの模式図を図 1 に示す。本研究のシステムを構築するには様々な技術が必要である。この章ではシステムを構成するための周辺機器や周辺技術について述べていく。

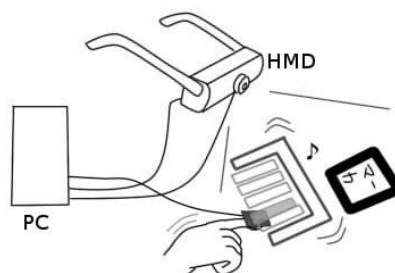


図 1: AR による仮想楽器システム

2.2 周辺機器

本研究で使用する周辺機器について順に述べる。

2.2.1 iWear VR920 (図 2)

米国 VUZIX 社製の HMD である。PC 専用タイプの HMD であり、VGA と USB で PC に接続すれば、2.7m 先に 62 イン

連絡先: 一ノ瀬元喜, 阿南工業高等専門学校, 徳島県阿南市見能林町青木 265, 0884-23-7100, ichinose@anan-nct.ac.jp

手相当の画面で見ることができる。電源は PC から USB により供給される。また、VR920 対応 USB カメラである CamAR を取り付けることで画像を取得する。図 2 では、VR920 に CamAR を取り付けた状態となっている。



図 2: iWear VR920

2.2.2 マーカ (図 3)

ARToolKit では現実世界の映像とデジタル情報の重ね合わせの位置決め、マーカと呼ばれる黒い正方形の太枠の中に白黒またはカラーの図案があるものを使用する。またマーカを複数使う場合には、中央の模様によって、マーカの識別を行う。



図 3: ARToolKit のためのマーカ

2.2.3 入力装置 (図 4)

本システムでは、プレイステーション 2 (ソニー・コンピュータエンタテインメント) のコントローラーの基盤に導線を半田で接合し、その導線の先に押しボタンをつけ、入力を検出するようにした。指先にキャップのように付けて、指先についたボタンを押すことで入力の信号を送る仕組みである。ここで、プレイステーション 2 のコントローラーを USB 端子にコンバートしてくれる機器を使って、コンピュータに信号を伝えている。

2.3 周辺基礎技術

ここでは、本システムの構築に必要な主にソフトウェアに関わる技術について述べる。

2.3.1 ARToolKit

ARToolKit とは、AR の研究のためにワシントン大学 HIT 研究所で開発されたソフトウェアライブラリである。このライブラリを使えば、カメラで取得した現実世界の映像に対して、仮想的なオブジェクトを配置することが容易かつリアルタイムに実現できる。



図 4: 入力装置 (プレイステーションコントローラー + USB コンバータ)

2.3.2 OpenCV

OpenCV とは、Intel Open Source Computer Vision Library の略である。米国 Intel 社で開発された画像処理、画像認識用のオープンソースの C 言語ライブラリであり、Windows や Linux など複数のプラットフォームに対応している。静止画にも動画にも対応しているため、特にロボット工学やバイオメトリクスの研究によく使われている。

OpenCV を用いることによって、画像処理、構造解析、モーション解析と物体追跡、パターン認識、カメラキャリブレーションと三次元再構成、機械学習、ユーザーインターフェイスといった、コンピュータビジョンに関する高度な画像処理機能を容易に利用できるようになる。

2.3.3 OpenGL

OpenGL (Open Graphics Library) とは、Silicon Graphics 社 (SGI) が中心となって開発した 3D グラフィックスのためのプログラムインターフェイスである。

2.3.4 GLUT

GLUT とは、OpenGL に基づいて設計された C 言語のクラスライブラリのことである。OS のシステムレベルでの I/O を制御し、3D グラフィックスを描画することができる。

GLUT を用いることによって、OS が描画するウィンドウを制御したり、マウスやキーボードからの入力によって、ウィンドウを操作したりすることができる。描画されたウィンドウはジオメトリ変換による高品位な立体や球体として再現することが可能である。また GLUT は、OS を制御するプライマリレベルで機能しているため、OS に依存しない描画器である。

3. 本システムの構成

本システムでは ARToolKit により映し出された仮想的なピアノの鍵盤に対して OpenCV による画像処理で指の位置を認識し、鍵盤が押されると音を出力することによって擬似的に演奏する。全体のフローチャートを図 5 に示す。本システムの技術的な仕組みを次の節から順に述べる。

3.1 Metasequoia による仮想ピアノオブジェクトの作成

オブジェクトの作成は 3D ポリゴンモデラーである Metasequoia を用いて行う。枠 (図 6) と鍵盤 (図 7) は別々に作成し、枠の中に複数の鍵盤を配置することで、仮想ピアノを構成する。

3.2 ARToolKit によるオブジェクトの配置

ARToolKit は、マーカを検出し、座標変換行列を作ることによって、仮想現実を与える関数を持っている。マーカを検

出するためにカメラの初期化や撮影された画像の処理を行う。その後、座標変換行列によって変換された位置に仮想的なピアノの鍵盤のオブジェクトを配置する。

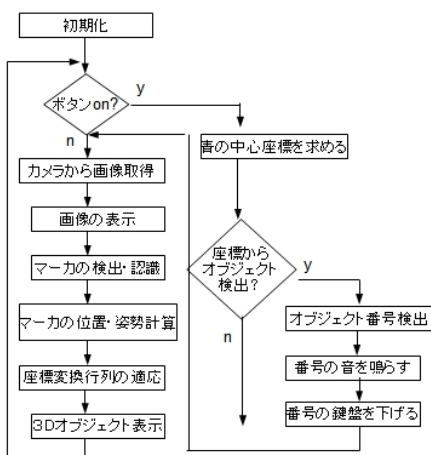


図 5: 本システム全体のフローチャート

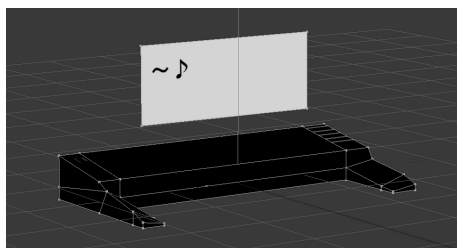


図 6: 棹オブジェクト

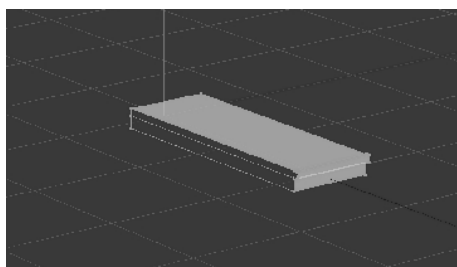


図 7: 鍵盤オブジェクト

3.3 OpenCV による指の位置計測

本システムでは、カメラによって写し出された画像より指先に青色の板を装着して、OpenCV の関数を用いて、板の中心座標を求めるようにした。ここで、直接指を用いていない理由は指先の領域が狭いことと、色が判別しにくいからである。しかし、これだけでは誤差が大きく、指先の位置を正確にとらえることが困難なので、画像を BGR 形式から HSV 形式に変換し、さらに H(色相)、S(彩度)、V(明度)の各画像に分けた。次に、それぞれの閾値により、特定の色(本システムでは青)の領域を分けて、すべての論理積を取った。そのあと、画像を拡大、縮小という処理をすることによって、ノイズによる穴を埋め精度を上げた。本システムでは、閾値を決定するために最初の一度だけ、カメラに写した指先の画像を保存し、その画像を画像編集ソフトにより、HSV の各画像に分けて、閾値による 2 値変換によって、どれくらいの値なら指先の色の部分だけの範囲になるかを測定した。この流れを図 8 に示す。

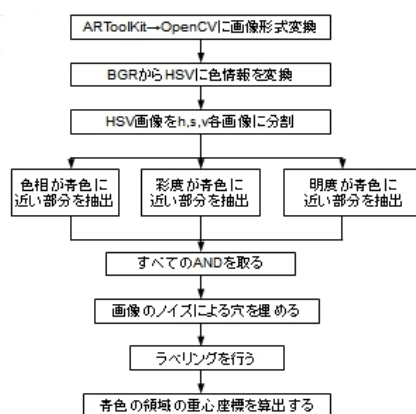


図 8: OpenCV を用いた指先の青色認識からその重心を求めるフローチャート

3.4 OpenGL による仮想オブジェクトの位置変換

ARToolKit の初期化の際に、モデルデータをロードする。そのときに、検出したいモデルデータを、検出用のオブジェクトに決定し、名前をつける。指の重心を求めても、それは 2 次元の座標であるので、そのままではどの物体の上にあるかはわからない。そのため、OpenGL にある関数を使い、その 2 次元座標上にある 3D 物体の名前と深さを配列に格納していく。そして、その深さの一番小さいもの、つまり目から一番近い物体を検出する。それが対象の鍵盤にあたる部分なので、入力ボタンからの信号があると指の重心を計算し、OpenGL の関数によって重心が対象オブジェクトの上にあった場合に、オブジェクトの座標を変更し、鍵盤の位置を下に移動し押された様に見せている。この流れを図 9 に示す。

3.5 仮想オブジェクトの移動とリンクした音の読み込みと生成

入力ボタンからの信号により、指の重心が対象オブジェクト上に存在することを判定すると、Win32 API の PlaySound 関数を用いてあらかじめ用意しておいたド～ソの音の WAV ファイルを読み込み、音を再生する。前節の鍵盤の移動はこの直後に行われる。

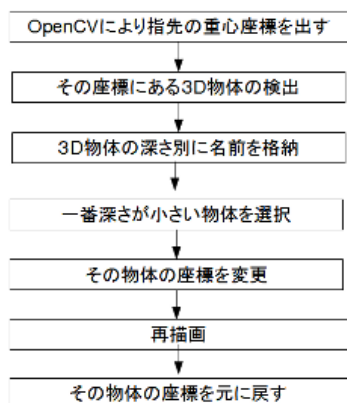


図 9: 指先座標から深さを求めて物体の位置変換までのフローチャート

4. 仮想楽器演奏の手順

前節の説明により構築した仮想楽器システムを演奏する手順は以下ようになる。

1. カメラ用の USB 接続をする。
 2. HMD 用の VGA/DVI 接続と電源用 USB を接続する。
 3. コンパイルしたプログラムを実行する。
 4. カメラの設定をする。
 5. カメラにより撮影された画像が HMD に表示される。
 6. マーカーを検出しオブジェクトを上から描画する。
 7. 青色の板がついた指を弾きたい鍵盤の位置に持っていく。
 8. 指の位置が鍵盤と重なったらボタンを押す。
 9. 押した鍵盤に対応した音がコンピュータから出力される。
- 仮想鍵盤が押されて音が鳴っている様子を図 10 に示す。

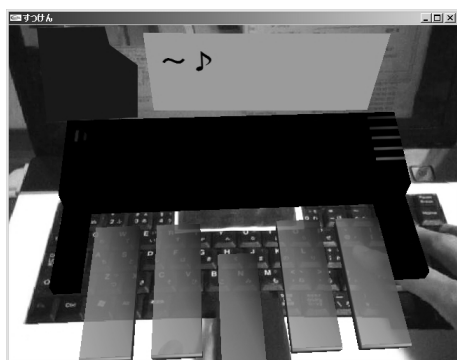


図 10: 仮想鍵盤を弾いている様子

5. 実験結果と考察

今回の実験ではド～ソまでの音を鳴るシステムを作成し、すべての音を鳴らすことができた。しかし、指先の位置の検出を青色の板の色識別で行っているため、周りの環境に左右されてしまった。中心側の精度はいいのだが、端になるほど反応しないことが多く、ドを鳴らそうとするとレが鳴ったりと誤作動も起きた。またボタンを押すことで指に圧力がかかると同時に鍵盤が落ち込み、設定された音が鳴るので演奏している感じは出ているが、音が単調な面もあった。

検出の精度が悪い点に関しては、カメラを二つ使って指先の位置を正確に検出するようにしたほうがいいだろう。また、一つの指だけでなく手を認識する HandyAR というものがあり、HandyAR を使えばもっと没入してピアノを演奏しているように感じる見込みがある。指先の感覚を伝えるのなら GravityGraber という、モーターを回転させて、指先の表面を滑らせて感覚を与えてやるとより現実感が出ると思われる。音に関しては、単調な音しかでず、押しっぱなしでも、音が途切れてしまうので、もっとサウンドに強いライブラリを使うべきである。

6. 結論

本研究では、AR と画像処理を組み合わせる仮想物体を自由に操作することを目指して、仮想的なピアノを生成し、人間がそれを演奏するシステムを構成した。本システムでは 5 キーの鍵盤を仮想的に表示し、異なる音を指先を画像認識することによって出力することが可能となった。

参考文献

- [Kato 99] H.Kato, M.Blillinghurst, K.Asano and K.Tatibana: An Augmented Reality System and its Calibration based on Marker Tracking, TVRSJ, Vol.4, No.4, pp. 607-616 (1999).
- [天目 05] 天目隆平, 神原誠之, 横矢直和: ウェアラブル拡張現実システムのための注目オブジェクトへの直感的な注釈提示手法, 日本バーチャルリアリティ学会論文誌, Vol. 10, No. 3, pp. 305-312 (2005).
- [浦谷 05] 浦谷 謙吾, 町田 貴史, 清川 清, 竹村 治雄: 拡張現実環境における奥行きを考慮した注釈提示手法の評価, 日本バーチャルリアリティ学会論文誌, Vol. 10, No. 3, pp. 343-352 (2005).
- [牧田 08] 牧田孝嗣, 神原誠之, 横矢直和: ウェアラブル AR のための移動オブジェクトの存在領域推定に基づくビューマネージメント, 日本バーチャルリアリティ学会第 13 回論文集, pp. 11-14 (2008).
- [樋川 06] 樋川直人, 大島千佳, 西本一志, 苗村昌秀: The PHANTOM of the PIANO: 自学自習を妨げないピアノ学習支援システムの提案, インタラクシオン 2006 論文集, pp. 69-70(2007).
- [元川 07] 元川洋一, 斎藤 英雄: 拡張現実表示技術を用いたギターの演奏支援システム, 映像情報メディア学会誌, Vol. 61, No. 6, pp. 789-796 (2007).
- [元川 08] 元川洋一, 斎藤 英雄: ギター演奏支援のための構造特徴追跡を利用したマーカレス AR 表示, 日本バーチャルリアリティ学会論文誌, Vol. 13, No. 2, pp. 267-278 (2008).