

# 高速ビジョンに基づく仮想楽器システム

山本 剛宏 (東京農工大学) 灘谷 演 (東京農工大学) 石井 抱 (東京農工大学)

## 1. はじめに

近年、1秒間に1000枚以上の画像に対して実時間処理および出力が可能な高速ビジョンの研究開発が進んでおり、人間の素早い動きにも対応可能とする実時間インタラクション技術の実現が期待されている。本発表では、このような視覚を用いた実時間インタラクションの考えとして、視覚センサ情報を音声提示情報に変換する実時間可聴化 (VA: Visual Auditorization) の概念を提案し、その有効性を示したデモンストレーションシステムとして、高速ビジョンを用いた仮想楽器システムを紹介する。

## 2. Visual Auditorization の考え

本発表で提案する VA の考えは、視覚情報の可聴化を意味し、Fig.1のように、実環境における対象の位置、さらには速度・加速度情報などを視覚センサにより観測し、その情報を基に対象物からの仮想的な音声情報に変換した上で音声提示装置に出力することで、視覚センサ情報を音声提示情報に変換する考えである。

このような VA の概念を実現する上で、既存の視覚センサではビデオ信号 (30Hz) 以上での高速化が難しく、数 kHz ~ 数十 kHz レベルでの周波数帯域を持つ音声情報に対して、十分に対応できなかった。これに対して、近年後述するような、kHz 以上のフレームレートを有しかつ実時間処理を可能とした高速ビジョンの開発が進み、音声情報の周波数帯域に対応した実時間 VA システムを実現することが可能となった。

以降では、kHz オーダーでのモーションキャプチャを実現する高速メガピクセルビジョンを説明した上で、実時間 VA システム例として、Vision-based 仮想楽器システムについて論じることとする。

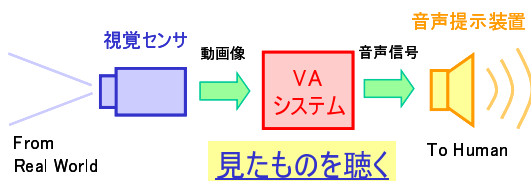


Fig.1 Visual Auditorization

## 3. 高速メガピクセルビジョン

次に、メガピクセルの高空間解像度と kHz オーダーの高速リアルタイム性を両立した高速メガピクセルモーションキャプチャシステムについて紹介する<sup>1)</sup>。

このシステムでは、Fastcom Technology 社の iMVS155 を用いている。iMVS155 には、画素単位でのランダムアクセスが可能な 1024×1024 画素を持つ

A Virtual Instrumental System using High Speed Vision  
Takahiro Yamamoto, Hiroshi Nadatani, and Idaku Ishii  
Tokyo University of Agriculture and Technology

CMOS イメージャと C 言語での汎用プログラミングが可能な DSP が搭載されたビジョンシステムであり、状況に応じて必要な画素情報のみを切り出すことにより、高空間解像度と高速リアルタイム性を両立したモーションキャプチャが可能となる。



Fig.2 iMVS155

本発表ではトラッキング対象として、標点となる白い小球などを考え、2値画像に対して高速化されたトラッキングにより、モーションキャプチャを実現する。そのトラッキングアルゴリズムの流れを Fig.3 に示す。

ウィンドウが対象から外れないためには、フレーム間の画像変化が小さい必要があるが、iMVS155 では kHz レベルでの高速フレームレートが実現されるため、トラッキングが可能となる。

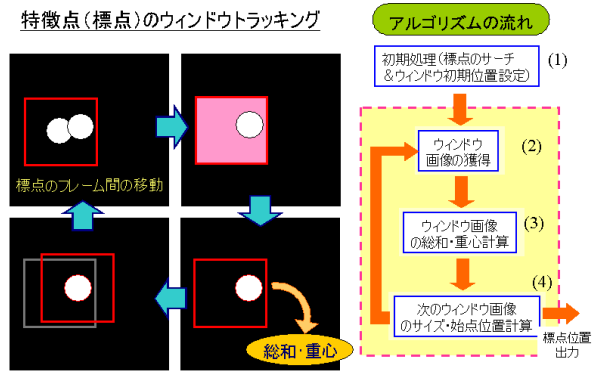


Fig.3 トラッキングアルゴリズム

対象が複数の場合は、2. ~ 4. の処理をそれぞれの対象ごとに逐次実行すればトラッキングが可能となる。

実際に 16×16 画素ウィンドウを切り出した場合、3点追跡時に 610Hz、1点追跡時に 1,830Hz のキャプチャレートが実現されており、複数個の対象に対して kHz オーダーでのモーションキャプチャが可能である。

## 4. Vision-based 仮想楽器

次に実時間 VA システムのデモンシステムの例として、Vision-based 仮想楽器について述べる。

このシステムは Fig.4 に示す以下の構成からなる。

## 高速メガピクセルビジョン iMVS155

演奏者自身あるいはマレットなどの演奏装置などの位置を、前述した高速メガピクセルビジョン iMVS155 を用いて追跡し、その位置、速度、加速度等の計算を行った上で、コンピュータへ出力する

## 仮想楽器モデルを有するコンピュータ

高速メガピクセルビジョンからの情報がギター、木琴など様々な仮想楽器モデルに対して、どのようなインタラクションを行ったかを計算した上で、発音、音程、音量などの MIDI 情報へ変換する。

## MIDI 音源などの音声提示装置

様々な仮想楽器モデルに対応した楽器音モデルを有する MIDI 音源を用いることにより、汎用かつ臨場感のあふれる楽器音を発生する。

本システムの大きな特徴としては、コンピュータ内でソフトウェア的に仮想楽器モデルを構築できるため、同一システム上で様々な形態の楽器が体験できる汎用性があげられる。その意味で、実在する楽器の模倣だけではなく、実在しない楽器システム、子供用の楽器など演奏者の状況に応じて設計可能なシステムであり、自由度と拡張性に優れたサウンドシステムである。

また実楽器を用いないことによる操作感の欠如が起きる場合があるが、これについては高速ビジョンで計測された速度・加速度情報などから音の強弱やタイミングなどを調節したり、さらにはソフトウェアにより演奏効果を強調するなどで対処できる。

以下に具体的な仮想楽器モデルの適用例を示す。

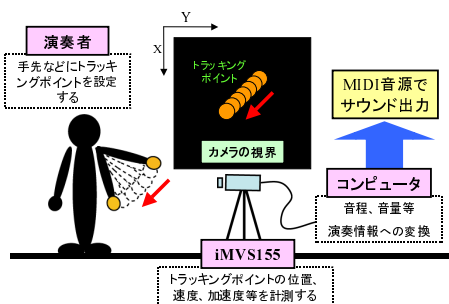


Fig.4 Vision-based 仮想楽器の概念

## 仮想楽器モデル例) パーチャルギター

パーチャルギターモデルは、人間の手の動きに合わせてギター音を鳴らす Fig.5 で示すような形態を持つ。このモデルでは、仮想的な弦を弾く右手の動きのみに着目し、音階を決める左手の動きは自動的に決定している。演奏者の右手親指の先に標点があるとし、その  $y$  座標が、 $i (= 1, \dots, 6)$  番目の仮想弦の位置  $y = y_i$  を通過した瞬間に、 $i$  番目の弦に対応した音が発音し、その大きさは通過した瞬間の速度により決定される。

## 仮想楽器モデル例) パーチャル木琴

パーチャル木琴モデルは、演奏者が動かす 2 本のマレットと空間内にある架空の木琴との位置関係により木琴の音を鳴らす Fig.6 で示すような形態を持つ。このモデルでは、架空の木琴を斜め上からカメラでみること

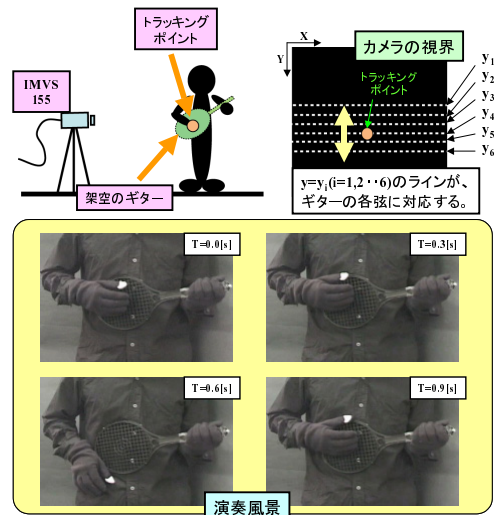


Fig.5 パーチャルギター

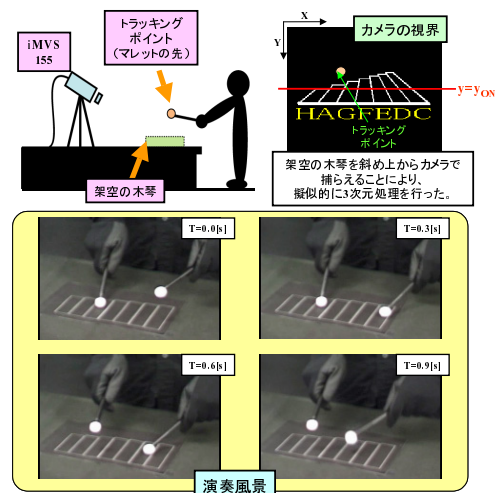


Fig.6 パーチャル木琴

で、擬似的な 3 次元処理を行い、鍵盤情報の ON/OFF を決定する。具体的には、標点を 2 本のマレットの先端に設定し、標点の  $y$  座標がある値  $y = y_{ON}$  を超えると鍵盤が叩かれた状態になる。また  $x$  座標の値から、どの鍵盤が叩かれたかを決定し、対応した音が発音される。2 本のマレットは、独立して処理されており、別々の鍵盤を同時発音させることができる。

## 5. おわりに

本発表では、VA (Visual Auditorizaion) の概念を提案し、その有効性を示したデモシステムとして、高速メガピクセルビジョンを用いた仮想楽器システムを構築した。また実際に様々な仮想楽器モデルを導入することにより、その概念の有効性を示した。

## 参考文献

- 1) 灘谷, 石井: 実時間メガピクセルモーションキャプチャシステムとそのジェスチャー認識への応用, 日本ロボット学会創立 20 周年記念学術講演会予稿集, 3A19, 2002.