

## 拡張現実による演奏支援のためのギターのオンライントラッキング

元川 洋一<sup>†</sup>      齋藤 英雄<sup>†</sup>

<sup>†</sup> 慶應義塾大学理工学部情報工学科

ギターの演奏は初心者にとって非常に難しいため、ギター演奏に対する障壁を取り除き、初心者でも気軽にギター演奏を楽しめるようにするためのシステムが求められている。本稿では、拡張現実感の技術を用いて演奏支援情報をギター上に表示することによって、ギター演奏を支援するシステムを提案する。提案システムを実現するための最も重要な問題は、ギターと仮想物体との間の正確な位置合わせを行うことであり、射影行列を正確に求める必要がある。本稿では、AR Toolkit によるマーカ検出と、ギターの自然特徴の追跡を併用した射影行列算出手法を提案する。本手法ではマーカ検出が失敗した場合でも、ギターのエッジ情報を利用して射影行列を算出することで、安定性を向上させている。

### Online Tracking of Guitar for Playing Supporting System by Augmented Reality

Yoichi Motokawa<sup>†</sup>, Hideo Saito<sup>†</sup>

<sup>†</sup> Department of Information and Computer Science, Keio University

It is very difficult for the beginner to play a guitar. In this article, we propose a supporting system for guitar playing by displaying visual aid information onto the guitar using Augmented Reality. The most important problem for constructing such system is accurate registration between the visual aid information and the guitar captured with a camera, which is achieved by estimating the projection matrix of the camera correctly. We propose a method combining marker detection using AR Toolkit and tracking of natural features of the guitar to estimate projection matrix. In our method, though system fails to detect a marker, we can estimate the projection matrix using edge information of the guitar, and it makes our system stably.

## 1 はじめに

ギターは非常に人気のある楽器であり、多くの人々が趣味などでギター演奏を楽しんでいる。しかし、ギターは初心者にとって非常に演奏が難しい楽器であるため、初心者がギター演奏を気軽に楽しむことは困難である。そこで、ギター演奏に対する障壁を取り除き、初心者でも気軽に楽器演奏を楽しめるようにするためのシステムが求められる。すでに市販されているギター演奏を支援するための楽器として、「光るギター」[1]と呼ばれるものがある。この楽器は、左手で押さえる弦の部分が曲に合わせて光り、ユーザは光る弦を左手で

押さえていくだけで初心者でも容易にギターの演奏を楽しめるというものである。しかしこの楽器では、押さえる部分を光らせるということを実現するために、ネック部分の弦がスイッチになっており、一般的なギターとは全く異なった構造になってしまっている。このような構造では実際にギターを弾いているという現実感に欠けてしまうと考えられる。また、演奏支援情報の提示方法も、押さえるキーを光らせるだけとごく単純なものに限られている。ギター演奏では、特にコードを演奏する場合においては、どのような手の形を作り、どの指でどの場所の弦を押さえるか、という情報が非常に重要である。しかし、従来ではそのような情

報を組み込んで演奏者を支援することはできなかった。以上のように、楽器の構造を直接的に改造するような方法では、十分にギター初心者の要求を満たした演奏支援システムを実現することが出来なかったと言える。また、そのようなシステムを実現するためには、従来とは全く異なった新しいアプローチが必要であると考えられる。

近年、拡張現実感 (Augmented Reality) と呼ばれる技術が注目されている [2]。拡張現実感とは、仮想的な物体をあたかも現実空間に存在するかのように重ね合わせて表示する技術である。この技術を用いることにより、ただ現実空間を見ているだけでは得られないような情報を仮想物体の重ね合わせという方法で付加させることができる。この技術を応用させることで、医療や都市デザイン、ナビゲーションシステム、エンターテインメントなど、さまざまな分野での実用化が期待されている。

本研究では、拡張現実感の技術に注目して、ギターの演奏支援システムを構築することを目的とする。本システムでは、演奏支援情報をギター上に重ね合わせることで初心者のギター演奏を支援する。拡張現実感の技術を利用する利点としては、ディスプレイを通して本物のギター上に仮想物体の重畳表示を行うため、ギターの構造を変更する必要がなく、ギターを弾いているという現実感を損なわないという点が挙げられる。さらに、表示する仮想物体には制限がないために、多様な演奏支援情報を表示することができる。よって表示させる演奏支援情報を工夫することにより、ユーザはギター演奏に必要な情報を十分に得ることができると考えられる。

本システムを実現する上で最も重要な課題は、演奏支援情報をギター上に正確に位置合わせをして投影することである。拡張現実感の研究においては、現実空間と仮想物体との間の正確な位置合わせは最も重要な課題として位置づけられており、従来からさまざまな位置合わせ手法が提案されている。その例として、物理センサを用いる手法 [3]、人工マーカを配置する手法 [4]、コーナーやエッジ、平面などの自然特徴を用いる手法 [5, 6, 7, 8] などが挙げられる。また、マーカと自然特徴を併用することで位置合わせを実現する手法も数多く提案されている。Gencらは、前処理で追跡する対象物体にマーカを配置して自然特徴点の情報を学習し、その情報を追跡段階で利用することで安定な追

跡を実現した [9]。本研究では、追跡の対象物体となるギターに平面マーカを配置し、追跡段階でマーカ情報とギターの自然特徴情報の2つを利用することで、正確で安定な仮想物体の位置合わせを実現する。マーカを配置する手法はマーカが検出できる範囲内では安定で高速に処理が行えるという利点があるが、マーカの一部が隠れてしまう可能性がある場合や、マーカを誤認識しやすい環境で用いる場合には安定性が欠如してしまう。そこで本研究では、マーカとカメラ間の並進ベクトルの変化量によってマーカ検出の成功・失敗を毎フレーム判断し、マーカ検出が失敗したと判断された場合には、ギターのネック部のエッジ情報を利用することでギターの追跡を行う。

## 2 提案手法

### 2.1 システムの概要

ここでは、本研究のシステムの概要を説明する。以下の図1にシステムの概略図を示す。本システムはPCに接続されたディスプレイとWebカメラから構成されている。また、ギターはAR Toolkit マーカ [4] を取り付けたものを用意する。

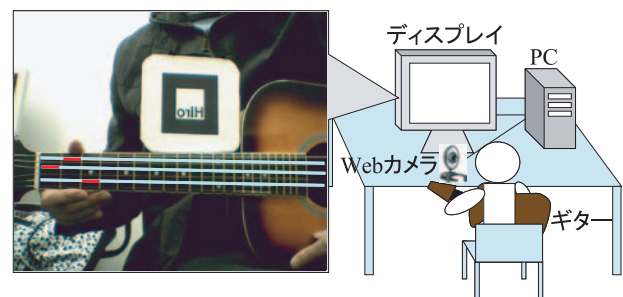


図 1: システムの概要

演奏者はマーカ付きのギターを持ち、ディスプレイの前に置かれたWebカメラの前に座る。するとディスプレイには、ちょうど鏡を見ているかのように、ギターを持った自分の姿が映し出される。そして、演奏する曲のデータが入ったテキストファイルを参照し、ディスプレイ中のギターには左手で押さえるべき弦の位置に赤い線が次々と重畳表示される。演奏者は、ディスプレイに映った自分の手と、押さえる位置が示された

ギターを照らし合わせることで、直感的に演奏することができる。

また、コードを演奏する場合には、どのような手の形を作り、どの指でどの場所の弦を押さえるのか、という情報が非常に重要である。ここで、コードにはその種類によって、図 2(a) に示すように弦を押さえる手の基本形がある程度決まっている。そこで、コードの種類別に、弦を押さえる手の形をした CG のモデルをあらかじめ作成しておき、図 2(b) のようにギター上に重畳表示させることで演奏の支援をする。ユーザは自分の手と CG の手を重ね合わせることで、複雑なコードも直感的に演奏することができる。

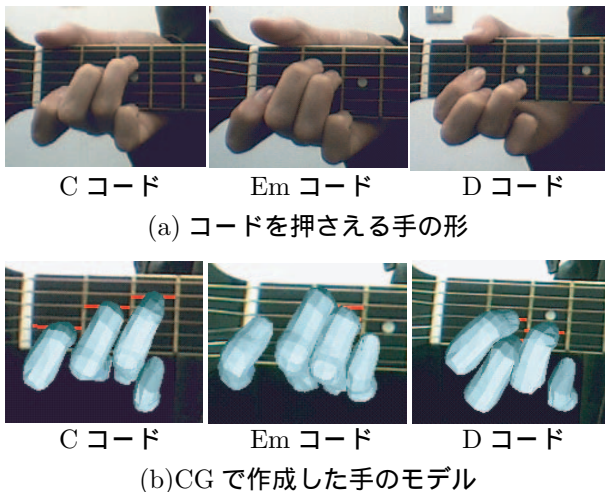


図 2: 手のモデルの重畳表示

## 2.2 手法の概要

提案システムを実現するための手法の概要について以下に記述する。演奏支援情報をギター上に正確に位置合わせをして投影するためには、ギターと Web カメラ間の回転行列  $R$  と併進ベクトル  $t$  を正確に求める必要がある。本手法では、ギターに取り付けられた AR Toolkit マーカの検出をベースにして  $R$  と  $t$  を算出する (図 3)。

次に、本手法の大まかな流れについて記述する。フローチャートを図 4 に示す。

Web カメラで撮影された入力画像から、ギターに取り付けられたマーカを AR Toolkit によって検出することで、 $R$  と  $t$  を算出する。マーカ検出が失敗した場合には、前フレームで算出された  $R$  と  $t$  を最適化することで現フレームでの  $R$  と  $t$  を算出する。求まった

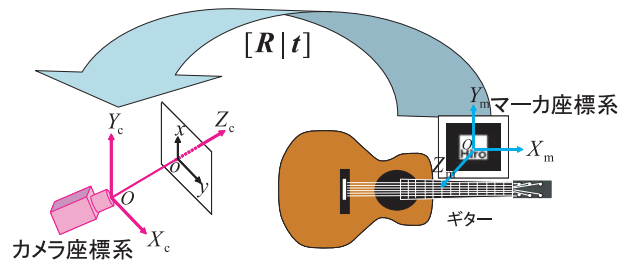


図 3: AR Toolkit を用いたカメラ位置姿勢の算出

$R$  と  $t$  とカメラの内部パラメータから射影行列を算出し、仮想物体の 3 次元座標と掛け合わせることで画像座標を算出してギター上に重畳表示する。ここで、カメラの内部パラメータと仮想物体の 3 次元座標は既知であるとする。

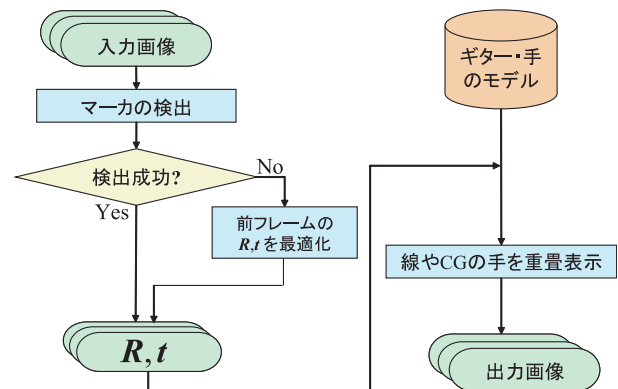


図 4: 手法の流れ

## 2.3 AR Toolkit によるマーカ検出

オンライン処理では、毎フレームで AR Toolkit によって自動的に入力画像中のマーカ検出が行われる。もしマーカ検出が成功すれば、マーカ座標系と画像平面間の射影行列を即座に得ることができる。仮想物体を投影するために必要な情報は 3 次元座標と射影行列のみであるため、マーカ検出が成功した場合には仮想物体の投影を行うことができる。

## 2.4 マーカ検出失敗の判断

もしマーカ検出が何らかの原因で失敗してしまった場合、マーカから得られる情報のみに頼ったシステムにおいてはそのフレームにおいて射影行列を求めることが出来ないため、仮想物体を正確な位置に投影する

ことができない．そこで，システムの安定性を向上させるためには，マーカ検出失敗時への対応が必要であると考えられる．本手法では，前フレーム情報を利用した最適化手法を使用することで現フレームでの射影行列を算出し，マーカ検出失敗時に対応する．そのためにはまず，マーカの検出の成功・失敗を適切に判断する必要がある．マーカ検出の失敗については，

- オクルージョンによりマーカの一部が隠れてしまった場合
- マーカではない部分をマーカであると誤認識してしまった場合

の2通りの場合が考えられる．1つ目の場合にはマーカの検出が失敗したことを容易に判断できるが，2つ目の場合ではマーカの検出が失敗したことを判断する特別な基準を設ける必要がある．その方法について以下に記述する．

マーカの誤検出の判断基準にはマーカとカメラ間の並進ベクトルの変化量を利用する．一般的に，処理速度が速ければ連続するフレームにおいてそれほど激しいギター的位置の変化は起こらないはずである．よって，マーカの検出が成功し続けている場合，連続するフレームにおいて並進ベクトルの変化は小さいはずである．しかし，マーカを誤検出した場合には，マーカが本来の場所とは別の場所にあるものと認識して併進ベクトルを算出してしまうため，前フレームからの並進ベクトルの変化が極端に大きくなると思われる．

そこでまず，3つの並進ベクトルの要素それぞれに対して前フレームからの変化量を算出する．あらかじめあるしきい値を設定しておき，前フレームからの変化量がそのしきい値以上になったパラメータが検出された場合に「マーカを誤検出した」と判断する．マーカ検出が失敗したと判断された場合には，以降に述べる処理によって，射影行列を算出する．

## 2.5 マーカ検出失敗時の処理

ここではマーカ検出失敗時に使用する，ギターモデルの投影とエッジ検出を利用した射影行列算出手法について記述する．処理の流れを図5に示す．

前フレームで回転行列  $R_{i-1}$  と並進ベクトル  $t_{i-1}$  が求まっていれば，現実世界に存在する自然特徴を追跡することで，マーカを利用することなく  $R$  と  $t$  を求め続けることができる．本手法では，自然特徴の一つである物体のエッジを追跡する手法 [7, 8] によって  $R_{i-1}$  と  $t_{i-1}$  を最適化して，現フレームでの  $R_i$  と  $t_i$  を求め

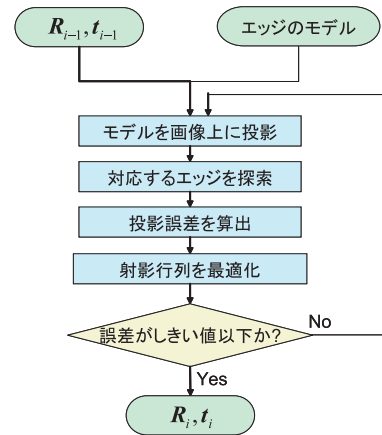


図 5: マーカ検出失敗時の処理の流れ

ることで，マーカ検出の失敗時に対応する．それぞれの処理について，以降で詳しく記述する．

### 2.5.1 エッジモデルの投影

最適化に利用するエッジとして，図6の赤線で示すようなギターのネック部の外枠を用いる．



図 6: 最適化に利用するエッジ

前フレームで算出された回転行列  $R_{i-1}$  と並進ベクトル  $t_{i-1}$  を現フレームでの推定値として用い，式 (1) を用いてエッジのモデルを画像上に投影する．ここで，エッジのモデルのマーカ座標系での3次元位置  $X_m$  はあらかじめ計測しておく．

$$\lambda x = A[R_{i-1}|t_{i-1}]X_m \quad (1)$$

$$\begin{pmatrix} \lambda & : & \text{スケールファクタ} \\ x & : & \text{画像座標} \\ A & : & \text{カメラの内部パラメータ} \end{pmatrix}$$

前フレームと現フレームの間でギターが移動した場合には，図7のように投影されたエッジのモデルと実際

のギターのエッジとの間にはずれが生じる。これは、投影に用いた回転行列と並進ベクトルに含まれる誤差によって生じるものである。すなわち、この投影誤差が最小になるときの回転行列と並進ベクトルが、現フレームにおいて最も適切なパラメータであるといえる。そこで、以降の処理によって最適なパラメータを算出する。

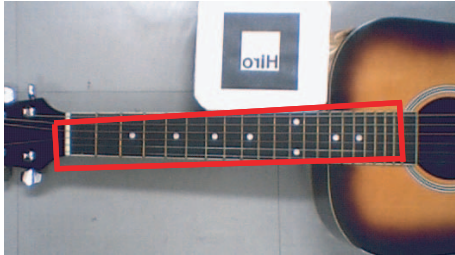


図 7: エッジの投影誤差

### 2.5.2 対応点の探索

最適なパラメータを算出するためには、投影誤差の大きさを調べる必要がある。そこで、まずモデルのエッジ上の点と画像上のギターのエッジ上の点との対応付けを行う。対応付けの方法を以下に記述する。

まず、画像上に投影されたモデルのエッジ上に一定間隔でサンプル点を設ける。次に、図 8(a) に示されるように、各サンプル点を通りモデルのエッジと直交する方向に沿って各画素の明度値  $I$  を調べていく。エッジは明度値の変化が大きい部分に存在すると考えられるので、初めに式 (2) の条件を満たした点をエッジ上の点とみなして対応付ける (図 8(b))。ここで、しきい値  $T$  は、環境によって適切な値を設定する必要がある。

$$|I_n - I_{n-1}| > T \quad (2)$$

また、一定の範囲内で対応点が検出されなかった場合には、外れ値として除外し、正常に検出された対応点のみを最適化処理に利用する。

### 2.5.3 射影行列の最適化

モデルのエッジ  $e_i$  上のサンプル点と画像上の点  $p_{i,j}$  と対応付けられたら、式 (3) によって投影誤差  $E$  を算出する。

$$E = \frac{1}{N} \sum \Delta(e_i, p_{i,j}) \quad (3)$$

(  $N$  : 対応点对の数 )

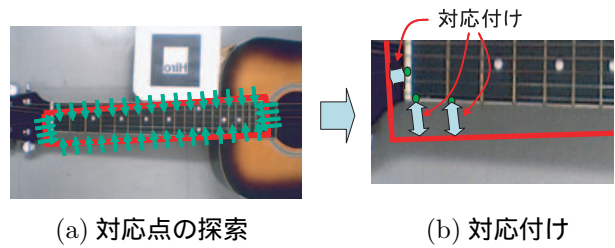


図 8: エッジの対応付け

式 (3) を評価関数として、最急降下法による射影行列の最適化を行う。

本手法において最適化させるパラメータは、回転行列  $R$  の 9 成分と並進ベクトル  $t$  の 3 成分であるが、このパラメータには定数倍の不定性があるため、12 番目のパラメータを 1 として正規化を行い、実際には 11 パラメータについて最適化を行う。最急降下法によって 11 個のパラメータを変化させていき、投影誤差が最小になるときのパラメータを最適なパラメータとして算出する。

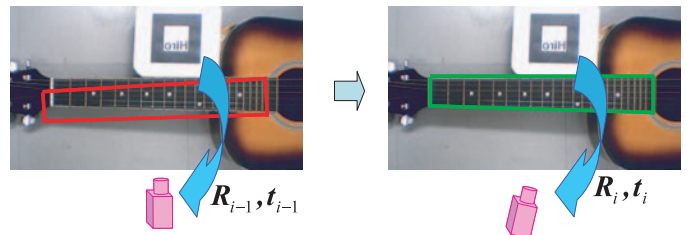


図 9:  $R$  と  $t$  の最適化

## 3 実験と検討

### 3.1 実験の概要

本研究では、本手法の有効性を示すために以下の 2 つの実験を行った。

- ギターを大きく動かした場合
- マーカ検出失敗時の処理

本実験で用いた PC のスペックは以下の通りである。

- CPU : Intel Pentium4 Processor 3.2 GHz
- メモリ : 1024 MB
- OS : Windows XP

### 3.2 ギターを大きく動かした場合

演奏中にギターが動くことは頻繁に起こりうるため、ギターが回転・並進したらそれに合わせて正確に仮想物体が投影されなければならない。そこで、今回の実験ではギターを大きく移動させてみて、仮想物体が正確に投影されているかどうかを調べた。提案手法では AR Toolkit によるマーカ検出をベースにしてギターを追跡しているが、この手法が本システムに適切であるかどうかを検討する。ギターを大きく動かした場合の結果画像列を図 10 に示す。



図 10: ギターを大きく動かした場合の投影結果

図 10 に示されるように、ギターが移動すればそれに合わせてモデルの手がほぼ正確な位置に投影されることがわかる。このことから、提案手法によって射影行列が精度良く求まっていることがわかる。しかし、ギターに大きく奥行きを角度をつけた場合にやや投影位置がずれてしまう箇所も見られる。これは、マーカに大きく角度がついた場合における AR Toolkit の精度上の問題が考えられる。この問題を解決するためには、エッジを利用した最適化の精度を向上させ、AR Toolkit から算出される射影行列を毎フレーム最適化する必要があるだろう。ただ、本システムを使用する上でギターに奥行きを角度を大きくつけて演奏するような状況はほとんどないと考えられるので、本システ

ムを使用する分には提案手法でほぼ十分な精度が得られていると言える。また、処理速度は約 15fps であり、システムを使用する上であまり違和感がない程度の速度が実現できている。

### 3.3 マーカ検出失敗時の処理

本手法では、マーカの検出が失敗したと判断された場合に、前フレームで算出された射影行列を最適化することで現フレームでの射影行列を算出している。最適化が有効に働いているかどうかはシステムの安定性に大きく影響する。ここでは、同じ入力画像について次の 2 つの手法による出力を比較することで、マーカ検出失敗時の判断と最適化処理が適切に行われているかどうかを検討する。

- マーカ情報のみを用いた手法
- マーカ情報とエッジ情報を併用した手法 (本手法)

ギターを上下にゆっくり移動させる入力画像列に対して、マーカ情報のみを用いた手法と本手法を比較した結果画像列を図 11,12 に示す。

図 11 において、マーカの情報のみを用いた手法では、18, 19 フレーム目においてマーカを誤検出してしまい、モデルが全く見当違いの場所に投影されていることがわかる。これに対し、マーカ検出失敗時に最適化処理を行った場合では、18 フレーム目でマーカとカメラ間の並進ベクトル  $t$  の大きな変化を検知することでマーカ検出の失敗を判断して、17 フレーム目での回転行列  $R$  と  $t$  を最適化することで 18 フレーム目での  $R$  と  $t$  を求めている。このことから、マーカ検出失敗の判断が適切に実行され、最適化処理が行われていると言える。

また、図 12 ではマーカを途中で隠した場合の結果を表している。当然、マーカ情報のみを用いた場合はマーカを検出できないために投影が失敗している。しかし、本手法の場合はマーカが隠れた後でギターを動かしても正確な位置に仮想物体が投影され続けていた。このことから、ギターをゆっくり動かす限りでは最適化が適切に実行されているといえる。

また、別の入力画像列について本手法を適用した結果を図 13 に示す。

この結果画像列はギター追跡の失敗例を示している。マーカを隠した状態において、16 フレーム目から 17



Frame17



Frame18



Frame19

(a) マーカ情報のみ (b) 本手法

図 11: マーカ情報のみを用いた場合との比較

フレーム目にかけて大きくギターが移動している．すると最適化処理が失敗してしまい，その後 20 フレーム目まで追跡が失敗している．そしてマーカが正常に検出された 21 フレーム目で，追跡が失敗した状態から回復している．これは，エッジの誤対応によるものであると考えられる．もし，エッジの対応付けをする際に，本来対応付けたいエッジ上の点以外の部分と対応付けられた場合，誤った点との誤差を減少させるように最適化が進んでしまうため，最終的な投影誤差は大きくなってしまふ．ギターにはネック部の外枠のほかに，弦やフレットなどの紛らわしいエッジが多数存在する．その紛らわしいエッジに対応点が引き付けられてしまうことが精度の低下につながる．大きくギターが移動した場合にはエッジ探索が困難になり，特に誤対応が生じやすくなってしまふのである．これを解決するためには，ギターのネック部の外枠のみを上手く抽出するようなエッジ探索手法を考へることや，弦やフレットのエッジ情報も利用した最適化手法を適用する必要があるだろう．もしエッジ情報のみの追跡



(a) マーカ情報のみ (b) 本手法

図 12: マーカが隠れた場合の投影結果

が精度よく成功すればマーカをギターに取り付けなくても本システムを実現することも可能になるかもしれないが，現状の最適化精度ではマーカの使用は必須と言えるだろう．

## 4 おわりに

本研究では，拡張現実感の技術を利用してギター演奏の手順を PC のディスプレイを通してユーザに提示するギターの演奏支援システムを構築した．

本システムは従来のギター演奏システムより，以下の点で優れているといえる．

- ギター本来の特性が失われない．
- 弦を押さえる模範的な手の形を表示することができる．
- 演奏支援用の専用のギターを用意する必要がない．

また本システムを実現するために，AR Toolkit によるマーカ認識と，エッジ情報を利用した最適化処理を使い分けることによる，リアルタイムで動く安定した位置合わせ手法を提案した．また，実験により本手法



図 13: 投影の失敗例

の正確性や安定性について検証した。

今後の展望として、弦を押さえる手の3次元モデルを自動で作成することなどを考えている。これは、手のモーションキャプチャを行うことで、ギターを弾いている姿を撮影することで手の3次元モデルを自動で作成し、そのモデルを演奏支援のための仮想物体として利用するのである。これが実現できれば、例えばギターが上手い人のギター演奏している様子を撮影して手のモデルを作成し、その手の動きの様子をギター上に重畳表示することで、ユーザはギターが上手い人の手を参考にしながら演奏をする、といった使い方ができる。

他にも、エッジを利用したトラッキング精度を向上させて完全なマーカレストラッキングを実現することや、正対したディスプレイを用いるのではなく、プロジェクタやHMD、またはギターにディスプレイを取り付けることにより、通常ギターを演奏する目線において本システムを実現すること、さらには、ギターを弾いたときにその音が正しいかどうかを判定する機能

など、さまざまな応用によりシステムを発展させていきたい。

## 参考文献

- [1] <http://www.yamaha.co.jp/>
- [2] Ronald T. Azuma: "A Survey of Augmented Reality", In Presence: Teleoperators and Virtual Environments Vol6, No.4, pp.355-385 (1997)
- [3] 佐藤清秀, 穴吹まほろ, 山本裕之, 田村 秀行: "屋外装着型複合現実感のためのハイブリッド位置合わせ手法", 日本バーチャルリアリティ学会論文誌, vol.7, no.2, pp.129-137 (2002)
- [4] 加藤博一, Mark Billinghurst, 浅野浩一, 橋啓八郎: "マーカ追跡に基づく拡張現実感システムとそのキャリブレーション", 日本バーチャルリアリティ学会論文誌, Vol.4, No.4, pp.607-616 (1999)
- [5] Iryna Skrypyuk and David G. Lowe: "Scene Modelling, Recognition and Tracking with Invariant Image Features", Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2004), pp.110-119 (2004)
- [6] Michael Aron, Gilles Simon, Marie-Odile Berger: "Handling uncertain sensor data in vision-based camera tracking", Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2004), pp.58-67 (2004)
- [7] Luca Vacchetti, Vincent Lepetit, Pascal Fua: "Combining Edge and Texture Information for Real-Time Accurate 3D Camera Tracking", Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2004), pp.48-57 (2004)
- [8] Georg Klein and Tom Drummond: "Sensor Fusion and Occlusion Refinement for Tablet-based AR", Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2004), pp.38-47 (2004)
- [9] Y.Genc, S.Riedel, F.Souvannavong, C.Akinlar, N.Navab: "Marker-less Tracking for AR :A Learning-Based Approach", Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2002), pp.295-304 (2002)