

Web カメラを利用したギターの練習支援システムの検討

坂牛和里[†] 植村あい子[†] 村岡眞伍[†] 甲藤二郎[†]

概要: ギターを練習する初心者はまず運指の練習が必要であり、弦を押さえて正確な音を出すことは困難である。本稿では、弦をうまく抑えられない初心者のために、画像処理を用いたギター練習支援システムを提案する。具体的には、画像処理技術を用いてギターの弦やフレット及び左手による運指情報を認識し、対応する音を出力させることで、初心者が運指の練習に集中し、音を聴きながら練習できる手法を提案する。システムはギターの初期位置を取得し、演奏中は特徴点とセンサ情報を用いてギターの位置を追跡する。また、指先を検出し押さえている弦を推定する。評価実験では、ギターを追跡するための特徴点マッチングとギターの傾き検出の精度評価及び押弦推定の精度評価を行い、それぞれの評価項目の成功率が高いことを示した。

キーワード: 音楽インタラクション, 画像処理, ギター演奏支援システム

Consideration of Guitar Playing Support System Using Web Camera

AIRI SAKAUSHI[†] AIKO UEMURA[†]
SHINGO MURAOKA[†] JIRO KATTO[†]

Abstract: Because a guitar playing requires a huge fingering training, it is difficult for beginners to play beautiful sound by pressing the correct strings. This paper suggests that a guitar playing support system using image and sensor data processing can help the beginners to improve their playing skills. To recognize strings, frets, fingering, and playing sound corresponding to fingering, our system gets an initial position of the guitar and tracks the guitar position using feature points and sensor information while playing. In addition, our system detects the finger position and estimates pressed strings. In the experiments, we evaluate a success rate of the feature point matching and guitar slope detection for guitar tracking, and pressed strings estimation. Evaluations conclude that each evaluation has a high success rate.

Keywords: Music Interaction, Image Processing, Guitar Playing Support System

1. はじめに

近年、音楽配信サイトや動画投稿サイトなどの音楽サービスの充実により、音楽はより身近なものとなってきた。動画投稿サイトでは、自身の演奏している様子を投稿する人が多い。また自身で作った曲を披露する人も増加し、音楽の配信はプロのアーティストのみに止まらず、誰でも気軽に行うことができるようになった。このような楽器演奏者の中には、独学で練習している人も多く存在する。楽器初心者の練習支援に関する研究が多数されており、MIDI演奏情報と楽譜情報からユーザの演奏を解析する研究[1]や、演奏誤りを故意に発生させることで、演奏を誤っても演奏を継続する練習を行わせる研究[2]がある。

本稿ではギターの練習に着目する。ギターの練習では初心者が弦を押さえて正確な音を出すことは困難である。そのため、演奏しながら練習するまでには時間がかかってしまう。初心者にはまず運指の練習が必要であると考えられる。元川ら[3]は、拡張現実感を用いてギターの位置を認識し、正しい位置に指のCGモデルを表示するシステムを提

案した。システム利用者はCGモデルに指を重ねることで運指の練習ができるが、自分で弦を弾いて音を出さなければならず、弦をうまく押さえられない場合は綺麗な音を確認することができない。運指の練習に専念するためには、運指に対応した綺麗な音をシステム側が提供するべきであると考えられる。

運指に対応した音を提供するためには、運指情報を認識する必要がある。Burnsら[4]はギターにカメラを取り付け、ハフ変換によりギターのカマやフレットや弦、演奏者の指先を認識している。しかしカメラを弦の付近に取り付ける必要があるため、演奏する際に障害物となる可能性がある。有賀ら[5]はKinectを使用したピアノの運指認識手法を提案した。電子ピアノの上部に取り付けたKinectから打鍵のタイミングに合わせて距離画像を取得し、手指と鍵盤の距離差分を取ることで手領域を抽出している。この手法では距離画像を取得するKinectの利用にコストがかかってしまう。

本稿では、低コストで演奏に支障を出さずに運指情報を取得し、その運指情報に対応した音を出力するギターの練習支援システムを提案する。

[†] 早稲田大学
Waseda University



図 1 USB カメラからのキャプチャ例



図 2 ギターの各部の名称

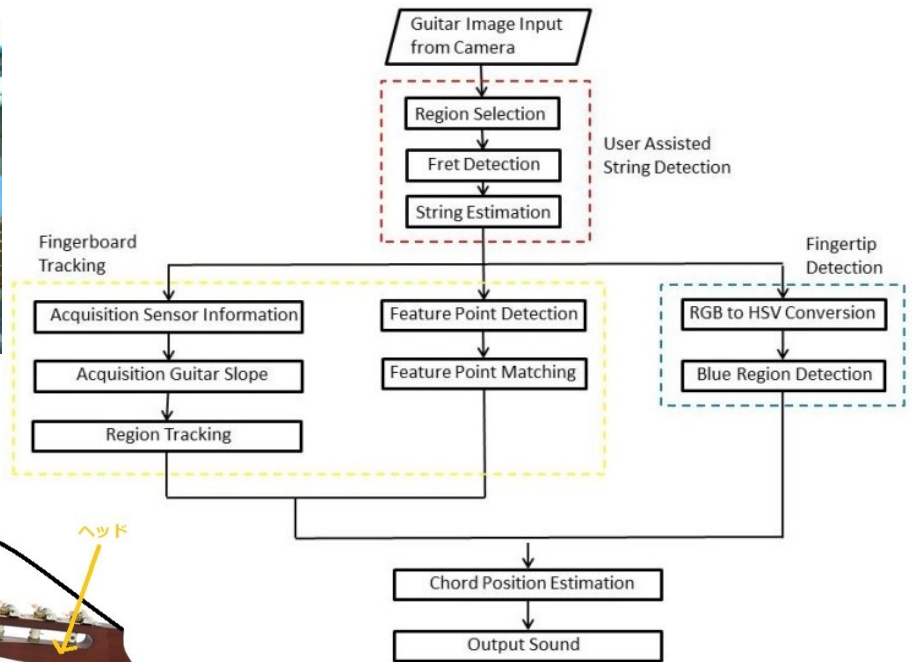


図 3 提案システムのフロー

2. 提案手法

2.1 提案手法の概要

本稿における提案システムの概要を説明する。システム利用者はパソコンの前に座り、ギターを演奏する。パソコンには USB カメラが取り付けられており、USB カメラからの映像をパソコンのディスプレイ上で確認することができる。USB カメラから得られる映像の例を図 1 に示す。また、センサ情報を利用するため、ギターにはスマートフォンが取り付けられている。

次に本稿で登場するギターの各部の名称について説明する。各部の名称を図 2 に示す。弦が張られた部分をネックという。本システムでは、指板と呼ばれるネックの表面部分を USB カメラでキャプチャする。指板にはフレットと呼ばれる金属が打ち付けられており、これにより半音階ずつ音を区別している。スマートフォンは演奏に支障が出ないように、ヘッドと呼ばれるネックの先端部分に取り付ける。また弦に関して、一番下が第 1 弦、一番上が第 6 弦である。

提案システムのフローを図 3 に示す。本システムではまず、システム利用者にギターを水平に持ってもらい、フレットの初期位置を設定し、フレットから弦の位置を推定する。初期位置が決定したら、その後はフレット位置のトラッキングと弦位置の推定を繰り返す。システム利用中、利用者には爪に青いテープを貼ってもらう。利用者の押弦中の指の位置はこの青いテープを認識することにより推定する。指の位置と弦の位置の推定結果により、利用者の押さえた弦を認識し、対応する音を出力する。この演奏中の指

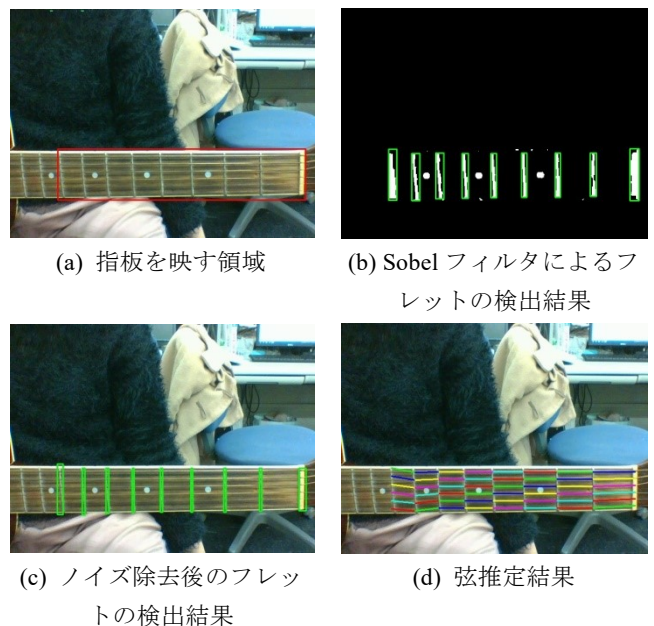


図 4 フレット検出と弦推定の結果例

の動きを運指という。

2.2 フレット検出&弦推定

フレットの初期位置を設定するため、まず画像上に赤い四角を描画する。利用者には、この四角を自由に動かしたり大きさを変えたりしてもらい、この中に指板が映るようにしてもらう。現在は、指板の 1 フレット側の端から 10 フレット目あたりまでが映ることを想定している。領域の位置が決定したらこの領域に Sobel フィルタをかけ、縦方向のエッジを検出する。エッジを検出した結果から輪郭を

抽出することで大まかにフレットを検出する．さらにノイズを除去することで，正確なフレットの位置を検出することができる．

このフレットの検出結果から弦の位置を推定することができる．各フレットの領域を縦に6分割し，それぞれの中心を弦の位置とする．フレットの検出と弦の推定の例を図4に示す．

2.3 指板トラッキング

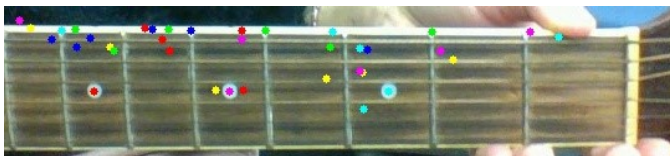
フレットの初期位置の設定時と演奏中の違いは，利用者が指板を含む領域を指定できないこと，ギターを水平に保つことができないこと，押弦中は指板に指が重なり，フレット検出にオクルージョンが発生してしまうことである．演奏中の弦の位置を推定するため，次の2つの手法を利用する．

(1) SURF 特徴量を用いた特徴点マッチング

まず，SURF 特徴量[6]を用いた手法について説明する．SURF 特徴量は回転やスケール変化に対してロバストな特徴量として知られている SIFT 特徴量を高速化した特徴量である．指板から SURF 特徴量を抽出し，次のフレームの特徴量とマッチングした特徴点を取得する．得られた特徴点のうち，ユークリッド距離の小さい3点を用いてアフィン変換行列を作成する．この行列を用いて弦位置の座標変換を行う．アフィン変換行列に利用する3点の距離が近すぎることや，マッチングにわずかな誤差があることで，座標変換の結果が大きすぎてしまう場合があるため，行列は複数の点から作成し，大きなずれを生じるとされる行列は座標変換に使用しないようにする．最終的には，正解の結果が得られると思われる行列の中央値を採用して座標変換を行う．特徴点マッチングの結果例を図5に示す．



(a) フレーム N の特徴点



(b) フレーム N+k の特徴点

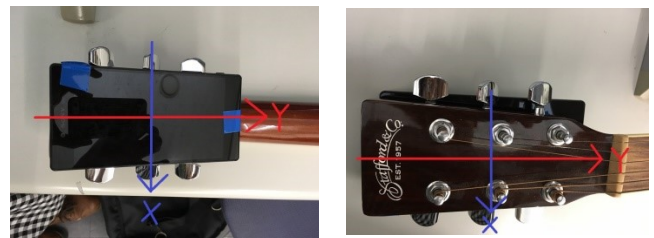


(c) フレーム N の弦検出結果 (d) フレーム N+k の弦推定結果

図5 特徴点マッチングの例

(2) センサ情報を用いたギターの傾き検出

トラッキングの精度を向上するため，センサ情報を用いてギターの動きを追跡する．スマートフォンには加速度センサが内蔵されており，図6のような軸に対する加速度を取得できる．それぞれの軸の加速度の変化を取得することで，スマートフォンを取り付けたギターの傾きを知ることができる．また，地磁気センサを組み合わせるとギターの傾きの角度が得られる．ギターの傾きを知ることによって指板の領域の傾きと Sobel フィルタをかける方向を知ることができ，フレット検出が可能となる．センサ情報と SURF 特徴量を組み合わせることは今後の課題である．



(a) ギターの背面

(b) ギターの前面

図6 ギターとスマートフォンの位置関係

2.4 押弦推定&音出力

利用者の押弦を推定するため，運指中の指先の位置を検出する．爪に貼られたテープの青色を認識するため，運指中の画像を RGB 表色系から HSV 表色系に変換する．検出範囲を指板の領域に限定することで，爪のテープのみを検出することができる．検出された青色の領域と推定した弦の位置が一致している場合，その弦を押さえているものとする．

ギターのコードには，一つの指ですべての弦を押さえるバレーコードと呼ばれるコードがある．バレーコードの多くは，押さえているフレットのうち一番右側のフレットを一つの指で押さえている場合が多く，またバレーコード以外のコードでは，一番右側のフレットが第6弦を押さえている場合は少ない．そのため右端に検出された指の領域が第6弦を押さえていると認識された場合，バレーコードを弾いているとみなし，右端のフレットはすべての弦が押さえられているものとする．

推定された押弦結果から各弦の MIDI 情報を作成し，利用者が弦を弾くタイミングに合わせて音を出力する．現在は利用者がパソコンのキーボードを入力することでタイミングを制御しているが，今後は弦を弾くタイミングを自動で取得することを目標としている．

3. 評価実験

前章の提案手法について評価実験を行った．フレームレートは 30fps，空間解像度は 640x480 である．実装には OpenCV2.4 を使用した．

3.1 特徴点マッチングの精度評価

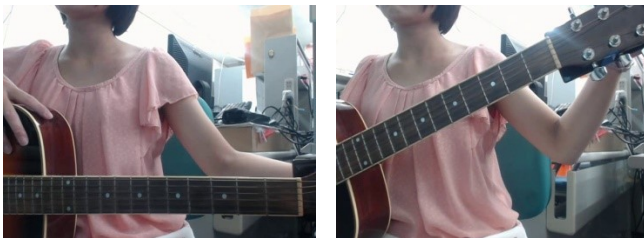
まず3つの動画像を用意し、特徴点マッチングの成功率を評価した。成功率は、各フレームにおいてマッチングにより得られた特徴点群のうち、目視での主観によりマッチングに成功していると判定した特徴点の数の割合である。それぞれの動画での特徴点マッチングの成功率を表1に示す。すべての動画像において、高い成功率が示された。

表 1 特徴点マッチングの成功率

Sequence Title	Seq. 1	Seq. 2	Seq. 3
Success Rate[%]	98.3	86.7	91.4

3.2 傾き検出の精度評価

次にセンサ情報による傾き検出の精度を評価した。図6に示すようにスマートフォンをギターに取り付け、ギターを水平の状態から徐々に上に傾けていき、センサから得られた角度と目視により計測したギターの角度を比較した。目視による計測では、1フレット側の指板の端の上部と9フレットの上部の座標を目視で取得し、2点から角度を計算した。ギターを水平にした初期位置と傾けた時の最高の位置を図7に示す。センサにより得られた傾きと目視により得られた傾きの時間変化を図8に示す。図8において初期位置を0度とし、上に傾けるほど角度が大きくなっていく。また目視により得られた傾きに対するセンサでの傾きの差の時間変化を図9に示す。図9より目視との差は10度以内であるため、ある程度問題なくセンサ情報によりギターの傾きを取得できていることがわかった。



(a) ギターの初期位置 (b) ギターの最高位置

図 7 ギターの初期位置と最高位置

3.3 押弦推定の精度評価

最後に、コード別の押弦推定の成功率の精度評価を行った。この実験では爪が正しい弦を押さえており、弦位置の推定に成功していることを前提としている。13人の被験者に協力してもらい、すべての押弦が正しく検出された場合を成功とする。コード別の押弦推定の成功率を表2に示す。すべてのコードにおいて、50%以上の成功率が示された。成功率はコードによりばらつきがある。Emコードの成功率が高い理由として、初心者にとって比較的押さえやすいコードであることが挙げられる。反対にGコードの成功率が低い理由は、押弦に使用しない指が指板上に存在し、そ

の指の位置を押弦位置として誤認識してしまうためだと考えられる。Emコード、Gコード及びFコード(バレーコード)の検出結果例を図10に示す。図10において、指板の領域内の正解の検出箇所を赤、誤検出を青の丸で囲んでいる。

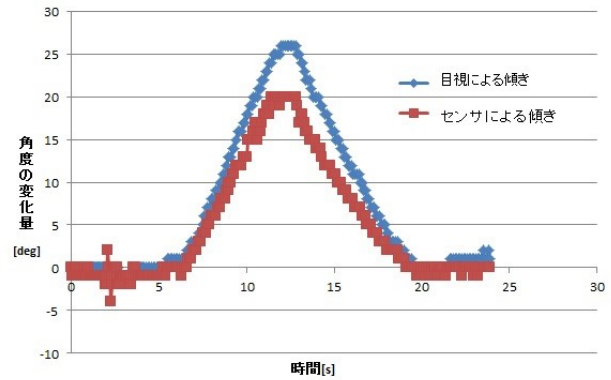


図 8 2つの手法から検出した傾きの時間変化

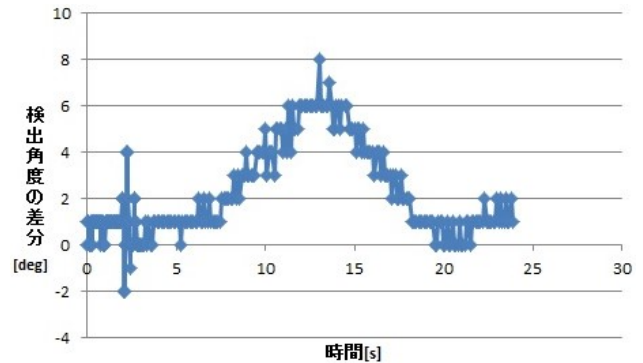
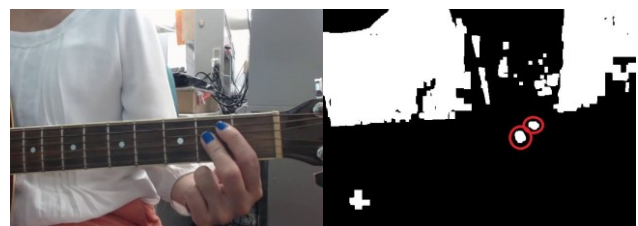


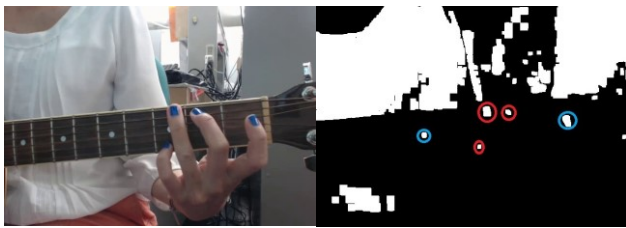
図 9 2つの手法から検出した傾きの差分結果

表 2 押弦推定の成功率

Chord	Success Rate[%]	Chord	Success Rate[%]
C	61.7	B	87.3
D	72.2	Dm	77.8
E	83.3	Em	100.0
F	57.5	Fm	83.3
F	50.0	Am	73.4
A	66.7	Bm	91.7

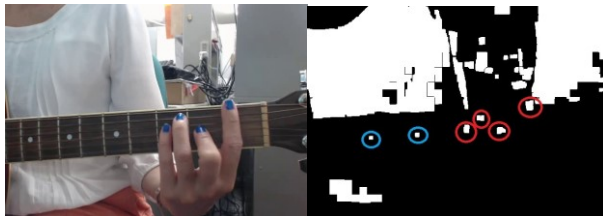


(a) Emコードの演奏例 (b) Emコードの検出例



(c) G コードの演奏例

(d) G コードの検出例



(e) F コードの演奏例

(f) F コードの検出例

図 10 押弦の検出例

4. おわりに

本稿では、初心者のための練習支援システムを提案し、フレット検出、弦推定、指板トラッキング、押弦推定の手法について説明した。評価実験では、特徴点マッチング、傾き検出、押弦推定の精度評価を行い、特徴点マッチングと傾き検出では成功率が高いことを示し、押弦推定ではコードにより成功率に違いが見られることがわかった。今後の課題は、SURF 特徴量とセンサ情報を組み合わせて指板トラッキングの精度を向上すること、弦を弾くタイミングを自動で取得すること、G コードのように押弦に使用しない指が指板に重なってしまう場合にも対応できるように、押弦推定の手法を改善することである。

参考文献

- [1] 岩見直樹, 三浦雅展: “MIDI 楽器を用いたドラム演奏練習支援システムの提案”, 情報処理学会研究報告, 2007-MUS-072, pp.85-90, October 2007
- [2] 横山裕基, 西本一志: “ピアノと押し演奏における演奏ミスによる中断を防ぐための練習支援システム”, インタラクション 2010, pp.205-208, March 2010.
- [3] 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.789-7960, July 2007
- [4] Anne-Marie Burns and Marcelo M. Wanderley: “Visual Methods for Retrieval of Guitarist Fingering”, Proceedings of the 2006 International Conference on New Interfaces for Musical Expression, NIME06, pp.196-199, June 2006.
- [5] 有賀治樹他: “デブス画像解析に基づく仮説検証型ピアノ運指認識手法”, 情処研報, 2013-MUS-100, 11, pp.1-6, August 2013.
- [6] H. Bay, T. Tuytelaars, and L.V. Gool, “SURF: Speeded-Up Robust Features,” European Conference on Computer Vision, pp. 404-417, May 2006.