

# 画像処理と MediaPipe によるギター演奏支援

押川 優<sup>a)</sup>・坂本 真人<sup>b)</sup>

## Guitar Performance Support by Image Processing and MediaPipe

Yu OSHIKAWA · Makoto SAKAMOTO

### Abstract

The guitar is one of the most popular musical instruments. In a questionnaire survey conducted by Yamano Gakki Co., Ltd.<sup>1)</sup>, 8.7% of respondents answered acoustic guitar or classical guitar when asked what kind of instrument they would like to try if they were to start playing an instrument or singing, which was the second highest overall ranking. From this, it can be said that the guitar is sometimes one of the most popular instruments among many instruments.

However, it is extremely difficult for a person with no experience in musical instruments to learn the guitar on his own. A study by Fender, a well-established guitar brand<sup>2)</sup>, found that 90 percent of guitar beginners quit guitar within a year or 90 days. From this, it can be seen that it is very difficult for beginners to learn guitar playing techniques.

I think one of the reasons why so many people quit playing guitar is that they don't have anyone to play the guitar with. In a questionnaire survey of 25 men and women registered with CrowdWorks who have been able to continue playing the guitar for more than 3 years<sup>3)</sup>, when asked why they were able to continue playing the guitar, the most common answer was that they had relatives and friends who were close to them who were playing music, accounting for 32% of the total. From this, it can be said that one of the reasons why guitar beginners stop playing the guitar is that there is no one to play the guitar with. In this study, in order to reduce the number of people who stop playing the guitar because there is no one to play the guitar with, we created a video comparing the performance of guitar learners and advanced guitar players, and asked them to use it for guitar training. This paper describes the creation of a guitar performance support video and its method to effectively approach guitar beginners.

**Keywords:** Guitier, Image processing, MediaPipe, PyCharm, Python,

## 1. はじめに

楽器初心者にとって、ギター演奏技術の習得は非常に難しいものである。筆者自身、大学2年時から軽音部に所属し、それから3年弱ギター演奏を行ってきたが、ギター初心者にとって、ギター演奏は非常に難しいものであり、軽音部という周りに楽器に詳しい方々が多くいる環境にしなければ今までギターを続けることができたかどうか定かではない。しかし、新しくギターを始めるすべての人の周りに、筆者がギターを始めた時のように楽器に詳しい方やアドバイスを頂ける方が存在しているわけではない。新しくギターを始める人のうち、多くの人が人からギターを教えて頂くのではなく、教則本や動画をもとにギターを練習し

ているのである。

近年では、楽器初心者にとってのギター演奏技術の習得のハードルを下げるために、情報処理技術を応用したギタートレーニングの研究がいくつか行われている。

例えば、元川らによる研究<sup>4)</sup>では、ギターの構造特徴追跡を利用したマーカレス AR の技術を用いて、ギタートレーニングシステムの提案をしている。このシステムは、あらかじめギターのコードを押さえている手の形の 3D モデルを作成しておき、それをマーカレス AR 表示で自分のギターの上に表示し、モニターを通して自分のギターの上にお手本の手の 3D モデルが表示されていることを視認できるというものになっている。

また、その他のギタートレーニングシステムとして、ゲーム形式でギターのトレーニングを行うことが可能な Rocksmith<sup>5)</sup>がある。

このように近年では、様々な視点から情報処理

a) 工学専攻先端情報コース大学院生

b) 工学基礎教育センター教授

技術を応用したギタートレーニングの研究が行われているが、初心者が自分の演奏が他の上級者と比べてどう違うのかを視覚的に判断することのできるシステムはまだ見受けられていない。

そこで、本論文では映像を通じて自分の演奏とギター上級者の演奏の違いを確認することができるように、画像処理技術と MediaPipe を用いたギター演奏支援映像の試作と操作について述べる。

## 2. 画像処理と MediaPipe を用いたギター演奏支援映像の作成

### 2.1 MediaPipe

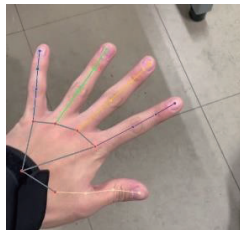


図 1. MediaPipe での手情報取得の例.

まずは MediaPipe について述べる。今回の研究では、手情報の取得のために MediaPipe<sup>6)</sup>を使用した。MediaPipe とは、Google 社が提供するフレームワークで、機械学習による画像認識によって顔認識、姿勢推定、手の検出などを行うことができるフレームワークである。今回の研究では、数ある機能の中でも、手の検出を行うために MediaPipe を Python 上で使用した。

図 1 は、MediaPipe での手情報取得の例である。今回の研究では、MediaPipe を用いて手や指の関節の 20 のポイントの位置座標を画像から取得し、それぞれのポイントの画像中での x 座標、y 座標、z 座標を csv ファイルに保存した。

### 2.2 ギター演奏支援映像の概要



図 2. 作成した映像の一部.

次に、この研究で作成した映像について述べる。まず、演奏の様子をカメラで撮影し、撮影した動画のギター上に、同じ曲を演奏しているモデルの手を表示する。モデルの手としてギター上級者の演奏の手が望ましいが、本研究ではギター歴 3 年弱の筆者自身の手を表示させている。映像上では、モデルの手と、モデルの手と自分の手がどのように違うかという情報が表示される。図 2 は、作成した映像の一部である。ギター上に表示される点が、MediaPipe で前もって取得しておいたモデルの手の情報である。指の関節など 20 のポイントの位置情報を取得し、それを点で表示し、点と点を結ぶことで手の形を模している。このモデルの手を、位置合わせ処理を行うことで、演奏者のギターの上に表示させている。また、黄色の文字で、演奏者の手がモデルの手とどう違うのかという情報を表示させている。現段階ではアドバイス文として、自分の手がモデルの手と比べてずれている方向を表示させているだけだが、最終的には学習者とモデルの弦を押していない指の違い、鳴った音の評価、演奏の評価を踏まえての練習方法などをアドバイス文として表示する予定である。

## 3. 映像作成手法

### 3.1 映像作成手法の概要

開発環境としては、JetBrains 社が提供している Python の総合開発環境である PyCharm<sup>7)</sup>を使用した。

以下に、本手法の各処理の詳細について記述する。

### 3.2 手法の各処理

#### 3.2.1 モデルの手情報取得の処理

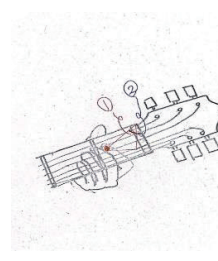


図 3. 学習者のギター上でのモデルの人差し指

先端の座標を求めている例.

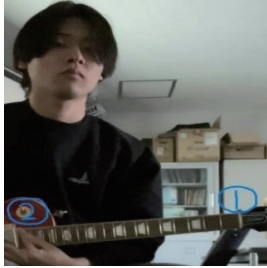


図 4. 位置合わせ処理のための工程を行っている様子.

本手法の各処理のうち、モデルの手情報取得について述べる。図 3 は学習者のギター上でのモデルの人差し指先端の座標を求めている例で、図 4 は位置合わせ処理のための工程を行っている様子である。

まず、モデルとなるギター上級者の演奏を撮影する。撮影をするにあたって 3.2.3 で後述する、モデルと学習者のギターの位置合わせの処理を行うための工程が、演奏前に存在する。撮影の流れとしては、初めに 6 弦とナットの重なる位置に人差し指の先端を置き、次に、6 弦と 22 フレットの重なる位置に人差し指の先端を置く。図 3 上の①が 6 弦とナットの重なる位置、②が 6 弦と 22 フレットの重なる位置を示している。

撮影と同時に流している演奏曲の音源の冒頭に、この工程を行うように促すアナウンスを入れており、モデルとなる演奏者はこのアナウンスに従ってこの工程を行う。撮影開始と同時に、6 弦とナットの重なる位置に人差し指の先端を重ねるように促すアナウンスが流れ、撮影開始から 10 秒後に、6 弦と 22 フレットの重なる位置に人差し指の先端を重ねるように促すアナウンスが流れる。その後、撮影開始 20 秒後に音源が始まることを伝えるアナウンスが流れ、撮影開始 24 秒後にカウントのセクションを含めた音源が流れる。映像は 1 秒ごとに 60 枚の画像ファイルに分割されるため、それぞれのポイントに人差し指の先端が重なっている時の手情報を取得するために、撮影開始から 600 枚目、1200 枚目の画像ファイルから MediaPipe を用いて手情報を取得し、それを CSV ファイルに保存する。この CSV ファイルを、以降からモデル情報 CSV ファイルと呼称する。その後、演奏の様子部分だけ手情報の取得を行うために、音源が流れ始める撮影

開始 24 秒後からの画像ファイルだけ、つまり撮影開始から 1440 枚目以降の画像ファイルだけ MediaPipe を用いて手情報を取得する。取得した手情報は、モデル情報 CSV ファイルに保存する。

以上が、モデルの手情報取得の処理手順である。

### 3.2.2 学習者の手情報取得の処理

本手法の各処理のうち、学習者の手情報取得について述べる。学習者の手情報取得の処理の流れとしては、モデルの手情報取得の処理の流れと同じである。モデルの時と同じように、撮影時に位置合わせの処理を行うための工程を行い、映像を画像ファイルに分割し、撮影開始から 600 枚目、1200 枚目、また 1440 枚目以降の画像ファイルから MediaPipe を用いて手情報を取得し、その情報を CSV ファイルに保存する。この CSV ファイルを、以降から学習者情報 CSV ファイルと呼称する。

以上が、学習者の手情報取得の処理手順である。

### 3.2.3 位置合わせの処理

本手法の各処理のうち、学習者とモデルのギターの位置合わせの処理について述べる。初めに、3.2.1 と 3.2.2 で作成したモデル情報 CSV ファイルと学習者情報 CSV ファイルを読み込み、モデルの手情報、学習者の手情報を取得する。次に、画像中の学習者とモデルのギターのナットから 22 フレット間、すなわち指板の長さを求める。以下の式(1)が双方の指板の長さを求める式である。以下の式(1)を、学習者の画像ファイルとモデルの画像ファイル双方に対して行う。

$$\text{ギターの指板の長さ} = \sqrt{(\text{図3①の}x\text{座標} - \text{図3②の}x\text{座標})^2 + (\text{図3①の}y\text{座標} - \text{図3②の}y\text{座標})^2} \quad (1)$$

MediaPipe は、z 軸取得の精度があまり高くないため、x 座標、y 座標の値のみを使って指板の長さを求めた。それに従い、撮影時はギターの向きが、カメラに向かって正面を向くようにして撮影を行った。次に、ギターの傾きの角度を求める。以下の式(2)が双方のギターの傾きの角度を求める式である。w は画像の横の長さ、h は画像の縦の長さである。以下の式(2)を、学習者の画像ファイルとモデルの画像ファイル双方に対して行う。

$$\text{ギターの傾きの角度} = \tan^{-1} \frac{\text{図3①のy座標} - \text{図3②のy座標}}{(\text{図3①のx座標} - \text{図3②のx座標}) * \frac{w}{h}} \quad (2)$$

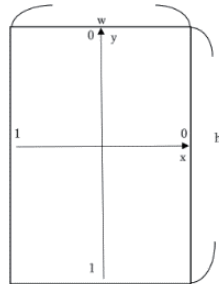


図5. MediaPipe の座標.

なお、式(2)において、分母、分子の数値、もしくはギターの傾きの角度がそれぞれ負の値になった場合、プログラム内で正の値に変換する処理を行っている。分母に  $w/h$  の式を組み込む理由は、MediaPipe の座標取得の仕様にある。MediaPipe の座標を図5に示す。MediaPipe では、画像の大きさに関係なく、画像の右端の  $x$  座標の値が  $0$ 、画像の左端の  $x$  座標の値が  $1$ 、画像の上端の  $y$  座標の値が  $0$ 、画像の下端の  $y$  座標の値が  $1$  となっている。このことから、角度を求める際に、それぞれの座標の差だけで計算してしまうと、画像の大きさによる  $x$  座標、 $y$  座標の長さの違いが加味されなくなってしまうため、式(2)に  $w/h$  の式を組み込んでいる。

次に、学習者とモデル間でのギターの距離、角度の違いを、位置合わせの処理の際に修正するために、双方のギターの角度の差、また、距離の違いの倍率を求めた。以下の式(3)、式(4)が、双方のギターの角度の差  $D\_DEG$ 、距離の違いの倍率  $M$  を求める式である。

$$D\_DEG = \text{学習者のギターの角度} - \text{モデルのギターの角度} \quad (3)$$

$$M = \frac{\text{学習者のギターの指板の長さ}}{\text{モデルのギターの指板の長さ}} \quad (4)$$

次に、学習者のギターの上にモデルの手を投影するために、学習者のギター上でのモデルの手の各ポイントの座標を求める。座標を求める処理の流れとしては、初めに、図3中の①、②の角度を求める。次に、図3で、中央に赤色で表している直角三角形の斜辺の長さ、つまり①のポイントから求めている手のポイント間の距離を求める。次に、求

めた2つの角度を足し合わせた値、斜辺の長さを使って、ギターを地面と水平にした際の各ポイントとナット間の距離、各ポイントと6弦間の距離を求める。また別に、ギターを地面と水平にした際の各ポイントと指板表面間の距離も求める。次に、求めた距離を使って、学習者の画像上におけるモデルの手の各ポイントの座標を求める。

以下、図3中の①、②それぞれの角度の求め方について述べる。①の角度は、①のポイントの座標と求めている手のポイントの座標から導き出すことが可能である。以下の式(5)が、①の角度を求める式である。 $p$  は、手の各ポイントの座標である。また、式(2)の時と同様に、分母、分子がそれぞれ負の値になった場合、正の値に変換する処理をプログラム内で行っている。式(5)を学習者の画像ファイルとモデルの画像ファイル双方で行う。

$$\text{①の角度}_i = \sum_{i=1}^{20} \tan^{-1} \frac{\text{図3①のy座標} - p_i \text{のy座標}}{(\text{図3①のx座標} - p_i \text{のx座標}) * \frac{w}{h}} \quad (5)$$

一方、図3中の②の角度は、ギターの傾きの角度と等しいため、式(2)で求めた値をそのまま使うことができる。

また、図3中の①のポイントから手の各ポイントの間の距離を求める。以下の式(6)が、図3中の①のポイントから手の各ポイントの間の距離を求める式である。 $DIS$  は、図3の①のポイントから手の各ポイントの間の距離である。式(6)を学習者の画像ファイル、モデルの画像ファイル双方で行う。

$$DIS_i = \sum_{i=1}^{20} \sqrt{(\text{図3①のx座標} - p_i \text{のx座標})^2 + (\text{図3①のy座標} - p_i \text{のy座標})^2} \quad (6)$$

次に、ギターを地面と水平にした際の各ポイントとナット間の距離、各ポイントと6弦間の距離を求める。まずは、ギターを地面と水平にした際の、各ポイントとナット間の距離の求め方について述べる。以下の式(7)が、ギターを地面と水平にした際の、各ポイントとナット間の距離を求める式である。 $DIS\_X$  は、ギターを地面と水平にした際の各ポイントとナット間の距離、 $DIS$  は、モデルの画像ファイルから求めた値である。

$$DIS\_X_i = \sum_{i=1}^{20} DIS_i * M * \sin(\text{図5①の角度}_i + \text{図5②の角度}_i - D\_DEG) \quad (7)$$

一方、以下の式(8)がギターを地面と水平にした際の各点と6弦間の距離を求める式である。 $DIS\_Y$  は、ギターを地面と水平にした際の各点と6弦間

の距離、DIS は、モデルの画像ファイルから求めた値である。

$$DIS_{Y_i} = \sum_{i=1}^{20} DIS_i * M * \cos(\text{図5①の角度}_i + \text{図5②の角度}_i - D\_DEG) \quad (8)$$

次に、学習者の画像上でのモデルの手の各ポイントの座標を求める。以下の式(9)、式(10)が、学習者の画像上でのモデルの手の各点の座標を求める式である。図3中の①の座標は、学習者の画像ファイルの座標を使用している。POS\_X は、学習者の画像上でのモデルの各点の x 座標、POS\_Y は、学習者の画像上でのモデルの各点の y 座標である。

$$POS_{X_i} = \sum_{i=1}^{20} DIS_{X_i} + \text{図3①のx座標} \quad (9)$$

$$POS_{Y_i} = \sum_{i=1}^{20} DIS_{Y_i} + \text{図3①のy座標} \quad (10)$$

次に、学習者の画像上でのモデルの手の各ポイントの z 座標を求める。以下の式(11)、式(12)が、学習者の画像上でのモデルの手の各ポイントの z 座標を求める式である。DIS\_Z は、ギターを地面と水平にした際の各ポイントと指板の表面間の距離、POS\_Z は、学習者の画像上でのモデルの各ポイントの z 座標である。また、式(11)での図3中の①の座標は、モデルの画像ファイルの座標を使用しており、式(12)での図3中の①の座標は、学習者の画像ファイルの座標を使用している。

$$DIS_{Z_i} = \sum_{i=1}^{20} (p_i \text{のz座標} - \text{図3①のz座標}) * M \quad (11)$$

$$POS_{Z_i} = \sum_{i=1}^{20} DIS_{Z_i} + \text{図3①のz座標} \quad (12)$$

この処理で求めた学習者の画像上におけるモデルの手の各ポイントの座標のデータは、CSV ファイルに保存している。以降、この CSV ファイルのことを、位置合わせ後の座標 CSV ファイルと呼称する。

以上が、位置合わせの処理の手順である。

### 3.2.4 学習者の画像ファイルにモデルの手を描画する処理

本手法の各処理のうち、学習者の画像ファイルにモデルの手を描画する処理について述べる。本処理は、3.2.3 の処理で作成した位置合わせ後の座

標 CSV ファイルを用いて行われる。

初めに、音源が流れ始める撮影開始 24 秒後からの画像ファイル、つまり撮影開始から 1440 枚目以降の画像ファイルの情報を取得する。次に、位置合わせ後の座標 CSV ファイルを読み込み、学習者の画像上でのモデルの手の各点の座標を取得する。取得した座標の情報を使って、学習者の画像ファイルにモデルの手を描画していく。

以上が、学習者の画像ファイルにモデルの手を描画する処理の手順である。

### 3.2.5 フレット情報取得の処理

表 1. 本研究で使用したギターのフレットの長さの一覧。

スケール	628.7 (mm)	【24.75インチ】	
	フレット間の距離(mm)	ナット(0F)からの距離(mm)	
0F~1F	35.3	0F~1F	35.3
1F~2F	33.3	0F~2F	68.6
2F~3F	31.4	0F~3F	100.0
3F~4F	29.7	0F~4F	129.7
4F~5F	28.0	0F~5F	157.7
5F~6F	26.4	0F~6F	184.1
6F~7F	24.9	0F~7F	209.1
7F~8F	23.5	0F~8F	232.6
8F~9F	22.2	0F~9F	254.9
9F~10F	21.0	0F~10F	275.8
10F~11F	19.8	0F~11F	295.6
11F~12F	18.7	0F~12F	314.3
12F~13F	17.6	0F~13F	332.0
13F~14F	16.7	0F~14F	348.6
14F~15F	15.7	0F~15F	364.3
15F~16F	14.8	0F~16F	379.2
16F~17F	14.0	0F~17F	393.2
17F~18F	13.2	0F~18F	406.4
18F~19F	12.5	0F~19F	418.9
19F~20F	11.8	0F~20F	430.6
20F~21F	11.1	0F~21F	441.8
21F~22F	10.5	0F~22F	452.2
22F~23F	9.9	0F~23F	462.1
23F~24F	9.3	0F~24F	471.5

本手法の各処理のうち、フレット情報取得の処理について述べる。この処理は、学習者の画像上のギターの各フレットの長さを求める処理で、学習者の指の位置がモデルの指とずれているかどうか調べるために必要な処理である。

初めに、学習者情報 CSV ファイルを読み込み、学習者の手情報を取得する。次に、学習者のギターの指板の距離を求める。その際、式(1)を用いて学習者のギターの指板の距離を求めた。次に、以上の処理で求めたギターの指板の距離を用いて、各フレットの長さを求める。以下の式(13)が、各フレットの長さを求める式である。FRET\_DIS が各フレットの長さである。表 1<sup>8)</sup>は、本研究で使用したギターのフレットの長さの一覧である。使用ギターの各フレットの長さは表 1 を参考にしている。また、452.2 という数値は、使用ギターの各フレットの長さをすべて足し合わせた数値である。

$$FRET_{DIS_i} = \sum_{i=1}^{22} \frac{\text{学習者のギター}の\text{指板}の\text{長さ} \times \text{使用ギター}のi\text{フレットの長さ}}{452.2} \quad (13)$$

式(13)は、使用ギターの各フレットの長さを全体の長さで割った数値に、学習者の画像ファイルから求めた学習者のギターの指板の長さをかけることによって、画像上での各フレットの距離を求めることができる。求めた各フレットの長さは、CSVファイルに保存している。以降から、このCSVファイルをフレット情報 CSV ファイルと呼称する。以上が、フレット情報取得処理の手順である。

### 3.2.6 学習者とモデルの手のずれを計測する処理

本手法の各処理のうち、学習者とモデルの手のずれを計測する処理について述べる。初めに、各画像ファイルでの、位置合わせ後のモデルの座標、学習者の手の座標の差を求める。次に、前もって作成しておいた、音源の小節数、弾く弦、フレット、音価の情報、その弦、フレットを弾く際に使用する指の情報を格納する CSV ファイルを読み込む。以降、この CSV ファイルのことを演奏情報 CSV ファイルと呼称する。演奏情報 CSV ファイルから読み込んだ小節数の情報を用いて、1小節に使われる画像ファイルの枚数を求める。以下、式(14)が1小節に使われる画像ファイルの枚数を求める式であり、MEASURE\_IMAGE\_N は1小節に使われる画像ファイルの枚数で、ALL\_MEASURE\_IMAGE\_N は演奏全体で使われる画像ファイルの枚数で、MEASURE\_N は小節数である。

$$MEASURE\_IMAGE\_N = \frac{ALL\_MEASURE\_IMAGE\_N}{MEASURE\_N} \quad (14)$$

次に、式(14)で求めた1小節に使われる画像ファイルの枚数、演奏情報 CSV ファイルから読み込んだ音価の情報から、演奏中に出てくる特定の音を奏でる間に使われる画像ファイルの枚数を求める。以下、式(15)が演奏中に出てくる特定の音を奏でる間に使われる画像ファイルの枚数を求める式である。SOUND\_IMAGE\_N は演奏中に出てくる特定の音を奏でる間に使われる画像ファイルの枚数で、SOUND\_VALUE はその音の音価である。音価は、

4分音符であれば4、8分音符であれば8として格納されている。

$$SOUND\_IMAGE\_N = \sum_{i=1}^{\text{演奏で使われる音の数}} \frac{1}{SOUND\_VALUE_i} \times MEASURE\_IMAGE\_N_i \quad (15)$$

式(15)で求めた演奏中に出てくる特定の音を奏でる間に使われる画像ファイルの枚数分、CSVファイルに画像ファイルの番号、指の情報、フレットの情報、弦の情報、位置合わせ後のモデルと学習者の指の座標の差の情報を保存する処理を行う。以降、このCSVファイルのことを、座標差 CSV ファイルと呼称する。

以上が、学習者とモデルの手のずれを計測する処理の手順である。

### 3.2.7 動画中に文を表示する処理

本手法の各処理のうち、動画中に文を表示する処理について述べる。初めに、座標差 CSV ファイル、フレット情報 CSV ファイルを読み込む。次に、画像ファイルごとに、座標差 CSV ファイルから取得したモデルと学習者の手の座標の差、演奏フレットの情報、フレット情報 CSV ファイルから取得した各フレットの長さの情報から、学習者の指の位置が、モデルの指の位置と比べて、ずれている表示を行うか決める処理を行う。学習者の指の位置が、モデルの指の位置と比べてずれているかどうか判定する基準として、ギターの向きを水平にした際に、x座標のずれが当該フレットの長さの半分以上ずれているか、y座標のずれが弦間の長さ以上ずれているか、という基準を設けた。基準に用いている弦間の長さを求めるために式(16)を用いる。

$$\text{弦間の長さ} = 1\text{フレットの長さ} \times 0.2439 \quad (16)$$

この0.2439という値は、実験で使用したギターのナットの長さ43.05mmを実験で使用したギターの1フレットの長さ35.3mmで割り、その値を5で割った値である。この値に画像中での1フレットの長さをかけることで、画像中での弦間の長さを求めることができる。x座標のずれが当該フレットの長さの半分以上ずれている場合、手がブリッジ側、もしくはヘッド側にずれているという旨を文で表示し、y座標のずれが弦間の長さ以上ずれている場合、手が6弦側、もしくは1弦側にずれているという旨を文で表示した。読み込んだ画像ファイルが、基準以上ずれている場合以上の文を画像ファイルに書き込み、基準以上ずれていなかった

場合画像に対して何も処理を行わないようにした。  
以上が、動画中に文を表示する処理手順である。

### 3.2.8 画像ファイルを動画ファイルに変換する処理

本手法のうち、ここまで処理を行ってきた画像ファイルを動画ファイルに変換する処理について述べる。画像ファイルを動画ファイルに変換するにあたって、Python の OpenCV ライブラリを使用した。また、画像ファイルを取得して、その情報から動画ファイルを作成するだけでは、音のない動画ファイルが出来上がってしまうため、前もって、学習者の演奏動画から、演奏の部分だけの音声ファイルを抽出しておいた。作成した動画ファイルに、Python の moviepy ライブラリを使用して、学習者の映像から抽出した音声ファイルを付与し、動画ファイルを作成した。音声ファイルの抽出には、Wondershare Technology 社が提供している動画編集ソフト Wondershare Filmora<sup>9)</sup>を使用した。

以上が、画像ファイルを動画ファイルに変換する処理の手順である。以上 8 つの処理工程を踏むことで、ギター演奏支援映像を作成することができる。

## 4. 評価実験

### 4.1 実験内容

本研究で作成した映像の有効性、手法の精度を確かめるために評価実験を行った。評価実験は、20 歳から 24 歳の男女 20 人に協力して頂いた。実験参加者 20 名のうち、11 名がギターを習得した経験がなかったが、ギター習得経験のない 11 名のうち 4 名は、ピアノや琴などの楽器の習得経験があった。

実験では、初心者でも習得可能な簡単で短いフレーズをカメラの前で弾いてもらい、その様子を撮影した。撮影した映像を用いてギター演奏支援映像を作成し、実際に映像を見てもらい、アンケートに答えてもらった。初心者でも習得可能な簡単で短いフレーズとして、“かえるのがっしょう”を課題曲に設定した。

### 4.2 アンケート結果

実験で得られたアンケート結果について述べる。アンケートでは、実験参加者の年齢や性別、ギター習得経験の有無のほかに、評価実験に関する質問として、以下の 10 の質問を行った。

1. ギター演奏支援映像の満足度
2. モデルの手の視認のしやすさ
3. モデルの手の位置合わせの精度
4. 位置合わせの手法のわかりやすさ
5. 位置合わせの手法の容易さ
6. モデルの手とのずれを文字で表示する機能の必要性
7. 今後ギターを習得する際に参考にしたいと感じたか
8. ギター演奏支援映像のよかった点
9. ギター演奏支援映像の満足いかなかった点
10. ギター演奏支援映像を見て、追加してほしい点や改善してほしい点

1 から 7 の質問については数字を選択することで回答して頂き、8 から 10 の質問は自由記述である。

図 9~15 が 1~7 の質問のアンケート結果である。評価の指標として、1 が良い、10 が悪くとした。結果を見ると、3 以外の項目で高い評価(評価の平均値が 1.3~4.1)を得たが、3 の項目の平均点は 5.2 点と高くなく、また図 11 で示すように評価にばらつきがあり、このことから映像によって精度にむらがあったことが分かる。また、項目 8 のギター演奏支援映像のよかった点に対する回答として、位置合わせの手順が分かりやすい点、ずれが可視化できる点、具体的にどこが間違っているのかを文字で表示していた点、色が見やすかった点、使いやす点などの意見があった。特にずれが可視化できる点良かったという意見が多くあり、初心者が正しい演奏をしているギター上級者とのずれを認識するうえで、この映像は良いアプローチになるのではないかと考える。

また、項目 9 のギター支援映像の満足いかなかった点に対する回答として、モデルの手がうまく反応していなかった点、表示される手の位置の精度が低い点、映像で音ずれが起きている点、動画化できなかった点などの意見があった。特に、表示される手の位置の精度が低いという意見は多く出て

おり、この手法の課題として、モデルの手の位置合わせの精度の向上が挙げられる。

また、項目 10 のギター演奏支援映像を見て、追加してほしい点や改善してほしい点に対する回答として、音程が正しければ正しいと判定するようにしてほしい、右手についての評価もしてほしい、表示される文字が見えにくかった、小節ごとに精度を表示する機能がほしい、テンポが速い曲の時に使用してもわかりやすい機能がほしい、1音、1音音があるか確認できる機能が欲しい、お手本、実際に出了音それぞれの音程の表示、映像作成処理の高速化などの意見があげられた。さらに、音程が正しければ正しいと判定するようにしてほしいとの意見があるが、筆者自身それは感じており、この映像を発展させるうえで、音声認識の技術は必要不可欠だと感じている。音声認識、モデルとの指の位置の違いを判断することで、より高い精度でギター学習者、モデルの演奏の違いを表すことができるのではないかと考えている。また、表示される文字が見えにくかったとの意見があるが、この問題は文字の周りに背景をすることで改善できるのではないかと考える。また、テンポが速い曲の時に使用してもわかりやすい機能が欲しいとの意見もあるが、これは動画変換の際に、動画の再生速度を指定して動画をつくるような処理を入れることで、対応可能ではないかと考えている。例えばスロー再生の映像を作成するなど、テンポが速い曲の時に使用してもわかりやすいようにできるのではないかと考えている。その他の意見も、この映像をより良いものにするためには必要な機能、改善点であり、今後の課題としたい。

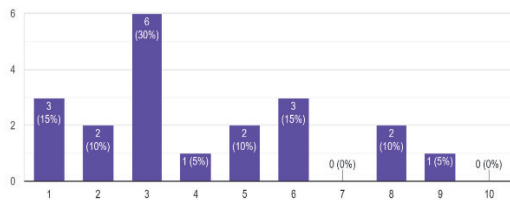


図 9. ギター演奏支援映像の満足度アンケート結果.

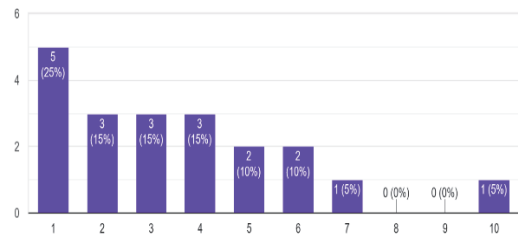


図 10. モデルの手の視認のしやすさアンケート結果.

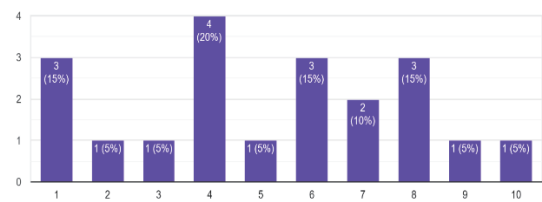


図 11. モデルの手の位置合わせの精度アンケート結果.

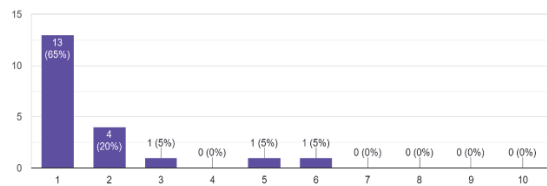


図 12. 位置合わせの手法のわかりやすさアンケート結果.

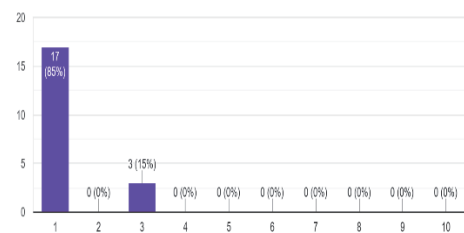


図 13. 位置合わせの手法の容易さアンケート結果.



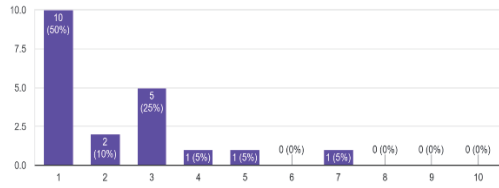


図 14. モデルの手とのずれを文字で表示する機能の必要性アンケート結果.

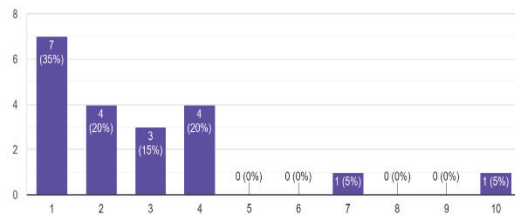


図 15. 今後ギターを習得する際に参考にしたいと感じたかアンケート結果.

## 5. 考察

本研究を通して、モデルとの手や指の位置の違いを可視化することでギター初心者にとって良いアプローチになることを確認できた。このことは、アンケートの結果からも裏付けできる。一方、位置合わせの精度の低さや様々な不具合も露呈した。これらの点を改善するためには、本研究の手法の見直しをする必要がある。例えば、ギターにマーカを取り付け、そのマーカをもとにギターの指板の長さ、角度を取得するなどである。このような手法を取り入れることで、学習者が位置合わせの処理を行う必要がなくなるため、学習者側の負担も減り、精度も現段階より向上すると考えられる。

## 6. おわりに

本論文では、画像処理技術と MediaPipe を用いたギター演奏支援映像についての概要、作成手法、評価実験、考察について述べてきた。論文の概要でも述べたが、ギターを始めたが、習得の難しさに挫折してしまい、ギター演奏をあきらめるギター初心者はたくさん存在する。このことからギター演奏のハードルが高いという認識が世間一般で広がっている。

将来的にはこの研究がさらに発展し、ギター初心者が他者からのアドバイスを受けられないこと

で、ギター演奏を途中でやめることがないように環境ができることを祈りたい。

最後にアンケートに協力して頂いた皆様に感謝の意を表したい

## 参考文献

- 1) PRWire.: “【山野楽器】「2022 年に始めたい楽器」アンケート、結果発表人気第 1 位は楽器の王様 “ピアノ”、2 位は “ギター” とおうち時間増で人気を伸ばした定番の楽器が上位 安定。さて、注目の 3 位は...” [Online].  
<https://kyodonewsprwire.jp/release/202201256509/amp> (参照 2024-02-05)
- 2) Kuroda Takanori: “老舗ギターブランド「フェンダー」が成長し続けている理由”, Rolling Stone Japan, 2019 [Online].  
<https://rollingstonejapan.com/articles/detail/32388/2/1/> (参照 2024-01-01)
- 3) 風ぺえ: “ギターを挫折せず続けるためのコツとは？アンケート結果から見える共通点”, 風ぺえギター研究所 [Online].  
<https://nagipee.com/howto/keepdoing/> (参照 2024-02-05)
- 4) 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”, TVRSJ, Vol.13, No.2, pp.267-278, 2008.
- 5) Rocksmith 公式サイト [Online].  
<https://www.ubisoft.com/ja-jp/game/rocksmith/plus> (参照 2024-02-02)
- 6) Mediapipe 公式サイト [Online].  
<https://developers.google.com/mediapipe/> (参照 2024-02-02)
- 7) Pycharm 公式サイト [Online].  
<https://www.jetbrains.com/ja-jp/pycharm/> (参照 2024-02-02)
- 8) Frontisland Custom Guitar Works: “ギターのスケール(弦長)とフレット位置” [Online].  
<https://frontisland.com/reference-book/scale/> (参照 2023-12-27)
- 9) Filmora 公式サイト [Online].  
<https://filmora.wondershare.jp/> (参照 2024-02-02)

