

複数カメラのスイッチング半自動化機能をもつ 遠隔ピアノレッスン支援システム

松井遼太 竹川佳也 平田圭二*

概要. 近年、通信技術・AR/VR技術などICT技術の発展に伴い、予備校におけるサテライト授業や会議システム、語学学習などの分野において遠隔・通信教育が一般化してきている。一方で、楽器演奏やスポーツに関しては依然として、対面環境における指導が一般的である。その理由として、音楽ではその楽器特有の演奏法や指導が求められたり、スポーツにおいてはスポーツの種類ごとに身体動作が異なることが挙げられる。視界や、身体への接触などの動作が制限される遠隔環境においては、これらの指導を円滑に行うことは難しい。そこで、本研究では、物理的に離れた2地点間においても円滑なピアノレッスンを実現するための、遠隔ピアノレッスン支援システムの構築を目的とする。提案システムは、演奏の詳細を相手に伝達するために複数台のカメラから構成される。また、複数カメラ映像の中から最適な映像を提供するカメラスイッチング自動化機能や、ユーザ自身が手動で選択できる機能をもつ。さらに、アテンションを共有するために、打鍵個所や楽譜の書き込みを共有できる機能をもつ。

1 はじめに

ピアノ演奏においては、譜読み、指示されている鍵への正確な打鍵、適切な運指（指使い）、リズム感覚、打鍵の強弱、テンポなど、さまざまな技術が求められ、それらの修得には長期間の基礎的な訓練を必要とする。ピアノ演奏には多大な時間と労力を必要とするため、敷居の高さに利用を断念したり、習熟効率の低さから挫折してしまう演奏者が後を絶たない。特に独学の場合、初級者や上級者を問わず、読譜した音と実際に打鍵した音が誤っている場合や、運指やリズムの間違いなどは気づきにくく、悪癖が身についた際の修正が難しい。そのため、うまく弾けない箇所の攻略方法を発見できずに練習のモチベーションが低下してしまい、挫折する可能性が高い[11]。したがって、効率的なピアノ演奏技術の向上のためには教師からレッスンを受講することが不可欠である。しかし、自宅から定期的に通える範囲にはレッスン教室の数が限られており、通学時間や教室の待ち時間なども発生するため、時間に余裕がある人でなければレッスンに通うことは難しい。大人のピアノ学習者の多くは、仕事や家事、育児などに追われてしまいレッスンに通う時間を確保できない。ピアノレッスンは基本的にマンツーマンで実施されるため、教え方や性格の合う教師から指

導されることが上達やモチベーションの維持には重要である。本研究で提案している遠隔レッスンを可能になれば、最適な教師を見つけやすくなる。

一方、近年、ICT技術の発展に伴い、様々な場面において遠隔コミュニケーション技術が用いられるようになった。身近な例ではスカイプなどを用いたビデオ会議や、遠隔地の作業者にロボットで直接指示を出す遠隔作業システムなど、その例は多岐にわたる。しかし、単にビデオ通話を利用した遠隔地同士におけるピアノレッスンでは効果的な指導が難しいといえる。

そこで本研究では、物理的に離れた2地点間で対面環境のように円滑なピアノレッスンを行うことを支援するためのシステムの構築を目的とする。

なお、本研究は[15]にて発表済みである。

2 関連研究

2.1 ピアノ用 e-ラーニング教材

これまでもピアノ学習支援に関する試みはいくつか行われている。自宅に居ながら教師のアドバイス付きで独学で学習できるピアノ用のe-ラーニング教材がある[1][2]。これらは動画内で教師が楽曲全体の解説や演奏で注意すべき点などの説明をするにとどまっている。動画による解説は、学習者の身近にピアノがない環境でも、練習中の楽曲の学習を補助できる。しかし、これらは教師が生徒の楽曲の理解度や演奏の進捗を確認できず、生徒も教師への質問が出来ないために、一方通行的なシステムであり、生

徒個人の實力や弾き癖にあった指導が出来ず、効率的な演奏技術の上達は難しい。

2.2 従来のピアノ独習支援システム

独習用に特化したシステムでは、蓄積された演奏データから演奏者の苦手な奏法を割り出し集中的にトレーニングするシステム[3][4][5][6][7]や演奏を自動的に評価しアドバイスや誤りを譜面上に提示[8]するシステムがある。これらは、取得した打鍵情報から打鍵ミスや打鍵の強さなどを評価している。Piano Tutor[9]は演奏追従認識による自動譜めくり機能や、ビデオや音声による模範演奏の提示や、演奏者の演奏データを解析し改善点をテキストなどで指示する機能などをもつ。これらの研究で使用しているリアルタイムな打鍵情報からの打鍵の正誤判定や演奏追従認識などの技術は本研究においても要素技術として活用できる。

竹川らのシステム[10][11]では、ピアノの上部にプロジェクタを設置し演奏を支援する情報を鍵盤やその周囲に提示することで、演奏初期段階における打鍵位置や運指の習熟を高める学習支援が可能である。これらの研究では演奏初期段階の学習者にとって、鍵盤への補助情報の提示などの直観的な方法は有用だと判明している。本研究においては打鍵情報などのレッスン補助情報を同様の方法で提示することで、遠隔環境におけるピアノレッスンにおいても効果的な支援を提供できる可能性がある。

2.3 遠隔演奏・操作

遠隔演奏に関するものでは、NETDUETTO[12]という楽器間を光回線で接続することができる技術(株)ヤマハが開発し、教育現場や遠隔セッションの場で導入されてきた。NETDUETTOは、遠隔地における音源の送受信を、人間が遅延を感じることはない速度で送受信することができる。

NETDUETTOを利用することで、遅延を解消できる見込みがあるため、本研究では、遅延対策はスコープ外とする。一方、人間が演奏に支障をきたさない音源の遅延は約20ms以内といわれており、これを超える遅延が発生すると、打鍵から発音までがずれ、正常な演奏が難しくなる。

また、遠隔地から指示を出し、作業を行う遠隔作業において、葛岡ら[13]は、実際に作業を行うユーザを「作業者」、遠隔地から指示を出すユーザを指示者と呼び、指示者によるハンドジェスチャを作業者のヘッドマウントディスプレイに提示することで両者の協調的作業を促すシステムを構築した。しかしピアノレッスンの場合、演奏者にとってヘッドマウントディスプレイやウェアラブルカメラなどの機器の装着は、演奏に支障をきたす可能性が考えられるため、本研究では用いていない。

3 設計と実装

3.1 想定環境

本システムは、生徒および教師の自宅やレッスン室に設置、両環境をインターネット接続して使用することを想定している。レッスン自体は通常の対面環境でのピアノレッスンと同様、マンツーマンを想定している。使用する生徒のスキルは、初級者や中級者を想定している。上級者になればなるほど、音楽表現の指導が中心になり、音質が重要になってくる。楽器の演奏音のクリアな伝達には限界があるため、上級者の支援は、本研究ではスコープ外とする。

3.2 設計要件

3.2.1 手指の視覚的把握の必要性

対面環境のピアノレッスンにおいては図1のように教師が打鍵方法を教示して、それを生徒が目で見模倣するといった直観的な指導が行われる。ピアノ指導においては、図2のように手指や身体動作を視覚的に把握しなければ指導できない要素が多く存在する。そのため、遠隔環境でピアノレッスンを行う際には配置するカメラを増やして、多数の視点から生徒および教師の手指を確認できることが必要となる。

3.2.2 ジョイントアテンションの必要性

遠隔環境では相手の楽譜が直接見えないため、書き込みをしたり、楽譜に指をさして「ここから弾いて」などと指示を出すことが困難である。楽譜には各小節に小節番号が振り分けられているが、単に教師が小節番号で指示を出しても、生徒は一見するとこの小節を指示しているのかわからない。指示語や指さし動作によって「この小節」という方法が最も一般的な指導方法であるといえる。遠隔環境において、このような指示語を用いた簡潔なコミュニケーションを実現するにはジョイントアテンションが必要である[14]。ジョイントアテンションとは他者と同じものに注意を向けることである。遠隔環境では、ジョイントアテンションの要素が多くなるほど、対面環境に近くなるといわれている。そのため、遠隔ピアノレッスンにおいても対面環境と同様に教師と生徒のアテンションの共有が必要である。

3.3 システム構成

図3に示したように、提案システムは、電子鍵盤楽器、投影用プロジェクタ、複数カメラ、タッチパネルディスプレイなどから構成される。これらの機材は、教師側および生徒側環境の両方に設置される。生徒、教師側の双方に配置されたPC間では鍵盤の打鍵位置情報、打鍵速度情報が送受信され両者

複数カメラのスイッチング半自動化機能をもつ遠隔ピアノレッスン支援システム



図1 対面環境におけるピアノレッスン

境のディスプレイに提示され、実際のレッスンと同様に教師が生徒の指使いを確認したり、生徒が教師の演奏方法を模倣できる。ディスプレイ上の楽譜に音声やタッチペンでメモや気づきを記入する機能、それらを保存し相手と共有する機能をもつ。

3.4 提示コンテンツ

図4を用いて提案システムの提示コンテンツについて説明する。図中の番号は、以下の箇条書き番号に対応している。

i) 書き込み共有楽譜

レッスン中の気づきや指摘を直観的に記録する楽譜共有機能をもつ。ディスプレイにはタッチパネルを採用しているため、タッチペンを用いて直接的な書き込みや、複雑な音楽記号などもすぐに書き込んで共有することができる。

ii) 仮想鍵盤

相手の打鍵位置をディスプレイ上に表示された仮想鍵盤で把握することができる。仮想鍵盤の打鍵情報は、実際の両者の鍵盤上にも投影される。投影機能については3.6.1にて後述する。

iii) 相手側カメラ視点

相手側のカメラ映像を表示する。演奏に合わせた最適な視点が常に1つ表示される。詳しくは3.5で後述する。

iv) カメラ手動切り替え候補

iiiで表示した視点以外のカメラ映像を、手動切り替え候補として表示している。詳しくは3.5.2で後述する。

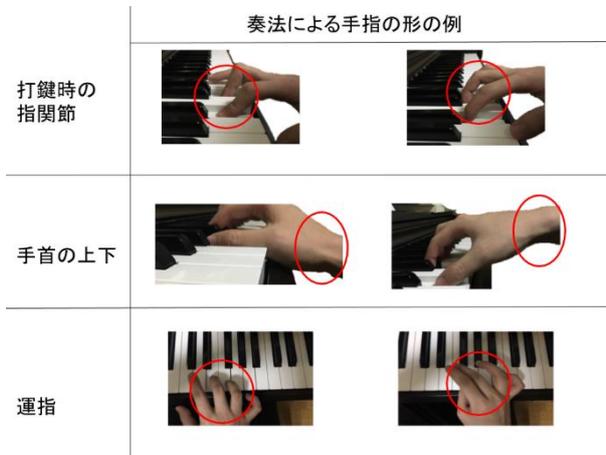


図2 手指の視覚的な把握



図3 システム構成

のディスプレイ上に可視化して表示される。

また、これらの打鍵情報は鍵盤の上方に設置されたプロジェクタによって実際の鍵盤上にも表示される。プロジェクタを鍵盤の真上に設置することで、鍵盤上やその周囲にさまざまなレッスン補助情報を提示できる。教師や生徒の表情・指示・指や手の形などを視覚的に確認するためにアングルの異なる複数台のカメラを設置する。カメラの映像情報は相手側環

3.5 手指の視覚的把握

ピアノを指導するうえで、手指や身体動作の視覚的把握は非常に重要となる。しかし、生徒の奏法や教師の指導指標によって、手指の形は異なる。手首を例にとると、打鍵主体の奏法であれば手首を高く保つ形、腕の重みで弾く圧鍵が主体の奏法では手首が上下する、などの奏法がある。これらは図2のように、実際に相手の手指を目で見なければ指摘は難しい。そのため、遠隔環境においては、様々なアングルから手指を見ることができるよう、複数のカメラを配置する必要がある。しかし、図5のように同時に複数の視点を表示すると、視点が多分見逃す可能性は低くなるが、教師または生徒はどこを見るべきかわからずに混乱してしまう。さらに、視点数が多ければ多いほど画面を圧迫し、個々の視点が小さくなってしまいうため、視認性は悪くなる。一方で、図6のように視点が一箇所であれば、広い表示



図4 スクリーンナップショット



図5 表示する視点数が多い場合

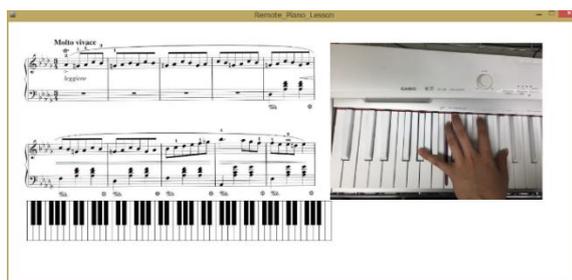


図6 表示する視点数が少ない場合

領域を確保でき、当該視点を詳細に閲覧できるが、表示していない視点における重要箇所を見逃してしまう。したがって、視点数と画面の視認性にはトレードオフが存在する。

そこで、本研究ではカメラスイッチングを自動化することで、この問題の解決を目指す。ニューラルネットワークを用いて小節ごとの最適な視点を予測し、鍵盤周辺に設置した複数カメラの中から、常に演奏に合わせた最適な視点を提示する。

3.5.1 カメラスイッチング自動化機能

FFNN (Feedforward neural network) を用いて、小節ごとにそのとき見るべき最適な視点を予測する。その後、予測したカメラ位置をもとに、ルールベースに従ってスイッチングを自動で行う。

現在の演奏、過去の演奏、楽曲構造、カメラ位置を、それぞれ入力パラメータとし、合計13個のパラメータが存在する。入力ノード数は対象楽曲の小

節数 × 入力パラメータ数で、出力ノード数は使用するカメラの個数とする。出力については出力値が最大のもを1とし、それが当該小節で使用するカメラとなり、それ以外のものを0として扱うことで出力値と使用カメラを制御する。あらかじめ使用する楽曲に合わせた各カメラの優先度を定めておくことで、ほぼ同値の出力が得られた場合でも、カメラを予測することが出来る。また、各カメラの使用率が、大きく差が出ない場合に判別できるよう、中間層のノード数は6とした。

入力パラメータ

現在の演奏、過去の演奏、楽曲構造、カメラ位置の4種類の入力パラメータがカメラを予測する際の入力となる。

現在の演奏に関してはDPマッチング (Dynamic Programming) を通して現在の打鍵位置の同定を行い、予測する小節と合致しているかを判定する。

過去の演奏に関しては、レッスンごとの演奏データを記録し、それを入力パラメータとして用いる。具体的には、当該小節での前回演奏時のミスの有無と、前後の小節でのミス頻発小節の有無がある。ミス頻発小節とは、過去の演奏データのうち、小節ごとのミスなく弾けた回数を、今まで課題曲を弾いた回数の合計で割り、閾値0.7を下回ったものをミス頻発小節と設定する。

楽曲構造に関しては、当該小節の最高音、最低音、小節頭の音の傾き、休符長、音数、最短音長、最長音長を入力とする。

カメラ位置に関しては、予測する小節の前後の、前回予測時のカメラ位置を入力する。これは、小節前後のカメラ位置が、その後の小節のカメラ位置に影響を与えるパラメータとなりえるため、設定している。

カメラスイッチングのルールベース

カメラをスイッチングする際に、以下のルールベースを適用する。

- ① 繰り返し部分は全く同じスイッチングを使用
 - ② カメラ切り替えタイミングの候補は毎小節の頭
 - ③ スwitching後、最低2小節は切り替えない
- ①に関して、曲中の類似部分ではなく、リピートなどの全く同様の繰り返し部分に関して適応される。
- ②に関して、小節の途中で切り替えるのではなく、毎小節の頭にカメラを切り替えるタイミングが用意される。
- ③に関して、毎小節ごとにカメラが切り替わると、目が追い付かなくなる可能性があるため、一度切り替えたならその小節を含めて2小節間は、カメラの切り替えを行わない。なお、カメラ切り替え後の小節が前小節と同じカメラに切り替わった場合は、前小節の同カメラから小節数をカウントする。

性能評価

F. F. Chopin の子犬のワルツの中間部を除く全 113 小節でカメラの切り替え予測を行った。カメラは合計 7 個使用し、配置場所は鍵盤と演奏者を取り囲むように設置した。過去の演奏データには 8 回分の課題曲演奏データを用意した。各小節ごとに 7 台のカメラから 1 台を割り当て、図 7 に示す主観的な基準で課題曲におけるカメラ選択の正解データを作成した。なお、これは基準の概要で、実際には言語化できない基準は存在する。

8-fold cross validation でカメラ予測の正解率を算出したところ、68.1%の正解率が得られた。

課題曲内には繰り返しでなくとも和声的に類似した構造が多く、現在入力としているパラメータのみでは分類が難しいと思われる箇所が存在するため、正解率は低くなったと考えられる。

有用性

本研究では FFNN およびルールベースによるカメラスイッチングの自動化を実現している。一方、画像処理により単純に手指が大きく映っているカメラを選択するような画一的な自動化手法も考えられる。画一的な自動化手法の場合、ピアノ全体を幅広く使う本課題曲では、カメラの切り替えが頻繁に生じてしまう。また、レッスンでは生徒がうまく演奏できず打鍵ミスをする場合はよくある。このとき、教師は打鍵ミスの原因を特定するために、打鍵ミスをした手指に注目する。打鍵ミスをした直後、打鍵ミスした箇所を撮影しているカメラに突然切り替わった場合、教師は状況理解に時間を要するため分析がうまくできない。また、原因が打鍵ミスする前にあれば正しく分析できない。

本手法は楽曲構造や過去の演奏履歴を入力とすることで、楽曲中の難しい箇所やミスをしやすい箇所を推定し、最適なカメラを自動で選択できる。提案手法の性能評価実験においてこれらの問題が解決できていたことは実際に確認でき、一定の有用性はあるといえる。

3.5.2 カメラ手動切り替え機能

3.5.1 で示したようにカメラスイッチングの予測精度には向上の余地が残されているものの、カメラの切り替えには教師の指導方法が生徒の悪癖に依存するという個人差があり、予測精度には限界がある。この問題を解決するために、本システムでは自動スイッチング機能を利用していても、図 4 のカメラ手動切り替え候補の映像をタップすることで、任意にカメラ映像を手動で切り替え可能にする。手動切り替え機能によって、自動切り替えの予測精度をカバーすることができる。自動切り替えでは、予測したカメラ位置が本来見るべき最適な視点とは異なる場合がある。その

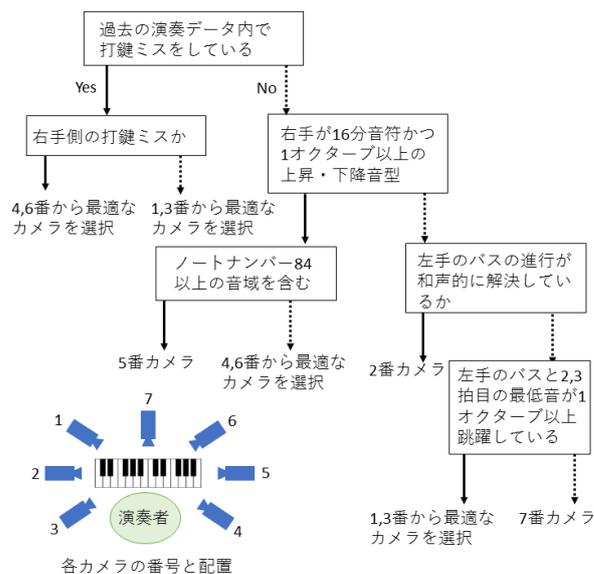


図 7 カメラの正解データ作成基準の概要

際に、手動切り替え機能を使用することで、最適な視点へとシームレスに切り替えることができる。この手動切り替え機能を例えるならば、テレビ番組でピアニストが演奏している際に、右手が映っているアングルが放映されているが、左手側を見てみたいと思った際に、即時自分で任意のアングルに変更できる機能である。さらに本機能は、手動カメラ切り替え結果をもとに、3.5.1 で説明した FFNN の入力パラメータを更新し、自己学習する。レッスンにおいて手動切り替え操作を実行する要因には、生徒がレッスン中に苦手箇所を克服したり、同じ部分で突発的なミス打鍵を繰り返すなどが挙げられる。これらはスイッチングを行う要素の一つとして考えられる。これによって、レッスンを繰り返すごとに、生徒の奏法や教師の指導法にフィッティングしたカメラの自動スイッチングが可能となる。

3.6 アテンションの共有

本システムは、3.2.2 で説明したジョイントアテンションを遠隔環境で実現するために、打鍵位置投影機能および予備動作付き楽譜ポインティング機能を提案する。

3.6.1 打鍵位置投影機能

図 8 のように生徒および教師の現在の打鍵位置をお互いの鍵盤上に投影する。これは同室レッスンにおける「教師が鍵盤に手を置き、打鍵する箇所を教える」という動作を模倣した機能である。これにより、教師または生徒は相手の打鍵位置を直観的に把握できる。

3.6.2 予備動作付き楽譜ポインティング機能

遠隔環境の場合、相手側の手元や楽譜が見えないため、指さし動作による指示出しは生徒に伝わりづらい。相手の指示した箇所がどこなのかを理解する速



図8 打鍵位置投影機能



図9 予備動作付き楽譜ポインティング機能

さが、遠隔環境における円滑なコミュニケーションには欠かせない。瞬時に指示箇所を理解するためには、指さしなどの指示を出すまでの、予備動作の観察が求められる。本機能は図9のように、指をさした箇所の周囲に円が収縮するアニメーションを付与し、そのまま指を離すとポインタのみが楽譜上に残るため、瞬時に指示箇所を共有できる。

4 まとめ

本研究では物理的に離れた2地点間における遠隔ピアノレッスンを支援するためのシステムを構築した。カメラスイッチングの自動化や、アテンションの共有機能などによって遠隔環境においても対面環境のように円滑なピアノレッスンが実施できる。

今後は、実際にレッスン形式でユーザーテストとシステムの改善を行い、よりシームレスな遠隔ピアノレッスンを可能にするためのシステムを目指す。

謝辞

本研究はJSPS 科研費 16k12560 の助成を受けたものである。

参考文献

- [1] 河合楽器製作所：簡単!ピアノマスター。
<http://www.kawai.co.jp/cmusic/products/pm/index.html>
2013-EC-27, 2013.
- [2] 一般社団法人全日本ピアノ指導者協会:ピティナ

e-ラーニング. <http://www.piano.or.jp/seminar/elearning/>.

- [3] 大島千佳, 井ノ上直己: 不得手要素を克服させるピアノ学習支援システムにむけて, 情報処理学会研究報告(音楽情報科学研究会 2007-MUS-71), Vol. 2007, No. 81, pp. 185-190.
- [4] M. Mukai, et al.: Generation of Suitable Phrases for Basic Training to Overcome Weak Points in Playing the Piano, Proceedings of International Congress on Acoustics, MUS-07-018(2007).
- [5] T. Kitamura and M. Miura: Constructing a Support System for Self-learning Playing the Piano at the Beginning Stage, Proceeding of International Conference on Music Perception and Cognition, pp. 258-262, (2006).
- [6] S. Akinaga, et al.: An Algorithm to Evaluate the Appropriateness for Playing Scales on the Piano, Proceedings of International Congress on Acoustics, MUS-07-005, (2007).
- [7] S. Akinaga, M. Miura, N. Emura, and M. Yanagida: Toward Realizing Automatic Evaluation of Playing Scales on the Piano, Proceedings of International Conference on Music Perception and Cognition, pp. 1843-1847, (2006).
- [8] 森田慎也, 江村伯夫, 三浦雅展, 秋永晴子, 田益造: 演奏特徴の強調およびアドバイス文呈示によるピアノ基礎演奏の独習支援, 日本音響学会平成 20 年度秋季研究発表会, pp. 933-934, (2008).
- [9] R. B. Dannenberg, et al.: A Computer-Based Multi-Media Tutor for Beginning Piano Students, Journal of New Music Research, 19(2-3), pp. 155-173, 1990.
- [10] 竹川佳成, 寺田努, 塚本昌彦: 運指認識技術を活用したピアノ演奏学習支援システムの構築, 情報処理学会論文誌, Vol. 52, No. 2, pp. 917-927, (2011).
- [11] 竹川佳成, 寺田努, 塚本昌彦: ピアノ演奏補助情報からの独立を促す学習支援システムの構築, 音楽情報科学研究会, 2012-MUS-96. (2012).
- [12] ヤマハ株式会社: "NETDUETTO", <http://www.y2lab.com/project/netduetto/>
- [13] 葛岡英明, 庄司裕子: 空間型共同作業の評価手法の提案とその利用, 電子情報通信学会論文誌, Vol. J 77-A No. 6 pp. 915 - 922, 1994.
- [14] B. Scassellati, Imitation and Mechanisms of Joint attention: A Developmental Structure for Building Social Skills on a Humanoid Robot, in Computation for Metaphors, Analogy and Agents, Vol. 1562 of Springer Lecture Notes in Artificial Intelligence, pp. 176 - 195, Springer Verlag, 1999.
- [15] 松井遼太, 竹川佳成, 平田圭二: 最適多視点カメラワークを自動生成する遠隔ピアノレッスン支援システムの設計と実装, 情報処理学会研究報告, Vol. 2017-MUS-114 No. 23, pp. 1 - 7, (2017年2月).