

スマートウォッチを用いたモノづくりマニュアル 作成支援システムの提案

会津 慎弥^{1,a)} 塚田 浩二^{1,b)}

概要：FabLab 等のファブコミュニティにおいては、ユーザは 3D プリンタ等を用いた作品だけでなく、それを他者にシェアするための手順書の製作も期待されている。しかし、こうした手順書製作の負担は大きいため、実現されないことも多い。また、画像やテキスト中心の従来の手順書では細かい手技等は表現できないため、ユーザに一定の知識や技術を暗に要求してしまう問題もある。そこで本研究では、スマートウォッチとスマートフォンを組み合わせることで、机上での組み立て操作を手軽に記録／編集して、動画マニュアルの製作を支援するシステムを提案する。まず、モノづくり中の組立動作を両腕に付けたスマートウォッチと机上に固定したスマートフォンで記録し、モーションデータ／動画を Web 上に保存する。次に、モーションデータを元に特徴的な動作に対して、動画に手動／自動でタグ付けを行うことで、自動的にチャプターや字幕が作成され、動画マニュアルのドラフトとして活用できる。本稿では、提案システムのコンセプト／実装、及びコンテンツ事例を紹介し、課題や今後の展開について議論する。

Support System for Creating Manufacturing Manual using Smartwatches

SHINYA AIZU^{1,a)} KOJI TSUKADA^{1,b)}

1. はじめに

近年、fabble^{*1}や instructables^{*2}などのモノづくりの作業工程を記述する、ユーザ参加型のマニュアル作成の Web サービスが普及している。こうした Web サービスを用いれば、非専門家であっても、3D プリンタ等のデジタル工作機械を用いたモノづくりに比較的気軽に取り組むことが出来る。FabLab^{*3}等のファブコミュニティにおいては、上記のようなサービスで、ユーザはデジタル工作機械を用いた作品だけでなく、作り方を記した手順書を製作し、他者にシェアすることも期待されている。しかし、これらの Web サービスのマニュアルは写真やテキスト中心に構成されて

おり、写真の選択や文章の執筆など、作成に大きな手間がかかるため、実現されないことも多い。また、細かい手技等は表現できないため、詳細な組み立て工程が伝わりにくいという問題もある。一方、モノづくりの過程を動画で記録した動画マニュアルも存在し、前述した静的なマニュアルよりも作業手順や詳細を伝えやすい利点がある。ただし、わかりやすいマニュアルにするためには、撮影後の動画編集が必要であり、知識や経験のない初心者には多大な労力と時間がかかる。そこで本研究では、マニュアルの作成者（以下、作成者）が動画マニュアルを撮影する際に、スマートウォッチを装着し腕や手の動作を記録／活用することで、上記問題の解決を目指す。例えば、スマートウォッチに搭載された加速度センサ等で工具の使い方等の「手技」を定量的に記録したり、作業工程を記録する際の動作を分類して「タグ」等の付加情報として活用する。これにより、特徴的な動作を目印にすることで作成者の動画編集作業を簡易化させたり、タグ等を用いて特定の作業を検索／閲覧するといった使い方ができる可能性がある。本稿では、机上

¹ 公立はこだて未来大学
Future University Hakodate

a) g2117001@fun.ac.jp

b) tsuka@acm.org

*1 <https://fabble.cc/>

*2 <https://www.instructables.com/>

*3 <http://fablabjapan.org/>

での組み立て過程を手軽に記録／編集して、動画マニュアルの製作を支援するシステムを提案する。

2. 関連研究

本章では、本研究に関連する研究事例について、「モノづくりの支援」「動画編集／視聴の支援」と「ウェアラブルデバイスをを用いた動作認識」の3つの領域から紹介する。

2.1 モノづくりの支援

ファブ時代の到来に伴い、モノづくりの支援を目的とした研究が盛んに行われている。

FabNavi[10]は、モノづくりの組み立ての工程を記録/共有/再生することでモノづくりを支援するシステムである。机上部にはカメラとプロジェクタが設置されており、カメラでモノづくりの組み立ての工程を画像や動画ベースで記録し、プロジェクタで記録したモノづくりの組み立ての工程を投影する。ユーザはプロジェクタで投影された映像を真似しながらモノづくりをすることで手軽にモノづくりに取り組むことができる。WeBuild[2]は、複数人における組み立て作業のタスクをを分配するシステムである。メンバーのスキル/サブタスクの依存関係を考慮し、動的にタスクの割り振りを行う。これにより、タスク調整の時間を短縮することができる。SketchChair[8]は、専門的な知識を持たない人でも、オリジナルの椅子を設計できるシステムである。ユーザが設計したい椅子の輪郭をスケッチすると、システムは自動的に3Dの椅子を生成する。また、自動的に、レーザーカッターで出力できる部品データを生成でき、切断して組み立てることで椅子を製作できる。

2.2 動画編集／視聴の支援

動画の編集は、知識や経験のない初心者には多大な時間と労力がかかってしまう。この問題を解決するために、動画編集の負担を軽減することを目的とした研究が盛んにおこなわれている。DemoCut[1]は、ユーザが指定したフレームを解析し、高速再生や、不必要な部分のカットなどを半自動的に行うことで有益なチュートリアルビデオを作成するシステムである。このシステムを利用することで、初心者でも容易の高い動画が作成できる。QuickCut[9]はユーザが映像シーン内に関連する説明を音声でアノテーションすることで、重要箇所を判別し不要箇所をカットするシステムである。中江ら[6]は、モノづくりワークショップの振り返り動画の作成を支援するシステムである。動画編集機能には、シーン探索機能を備えており、画像認識を用いて、笑顔や現在使用している工具などを推測し、必要なシーンを効率的に探索できる。Egoscanning[3]は、一人称視点動画の高速視聴を支援するシステムである。手の動きや、人との会話などの特徴的なシーンをタイムライン上に強調表示し、ユーザが興味のあるシーンを指定すると、

対応するシーンが再生速度を低速で再生され、長時間の一人称視点の映像における、興味のあるシーンを素早く見つけることができる。CinemaGazer[4]は、字幕付きの動画に対し、字幕のない箇所は高速再生し、字幕のある箇所は、字幕を読むことができるに再生することで、動画の内容を把握しつつ、短時間で視聴できるシステムである。

2.3 ウェアラブルデバイスをを用いた動作認識

ウェアラブルデバイスをを用いた動作認識に関する研究を3つ紹介する。Magnifisense[11]は、手首に装着した3つの磁気センサとA/Dコンバーターを用いて、電子機器から発生する電磁波を分析することで使用している電子機器を判別するシステムである。Object Hallmarks[7]は、スマートウォッチに搭載されている9軸IMUセンサを利用して、使用した物と人物を特定することができるシステムである。前川ら[5]は、工場のライン生産方式の、繰り返し作業のサイクルタイムを推定するために、作業者の手首に装着した加速度計から得られるデータに対し、事前学習を必要としないサイクルタイムを推定する手法を開発した。

2.4 本研究の特徴

このように、多くのモノづくり支援をする研究や動画編集／視聴する研究、ウェアラブルデバイスをを用いた動作認識に関する研究が存在するが、モノづくり中の「動作」を認識してマニュアル作成等に活用する試みはあまり行われてこなかった。本研究では、動画に加え、スマートウォッチ搭載のモーションセンサを用いてモノづくり中の動作を記録/検出することで、モノづくりの動画マニュアル製作を支援する。

3. 提案

本章では、提案するシステムの概要、構成を述べる。システムの構成を図1に示す。本システムは、主に、モノづくりの過程を記録するスマートフォンアプリ、データを統合して格納し、機械学習を用いてタグ付けするWebサーバー、手動のタグ付けや、マニュアルを提示するブラウザベースのアプリケーションで構成される。記録システムは、作業台上部に取り付けたスマートフォンカメラと作成者が両手に装着するスマートウォッチで構成され(図2)、俯瞰視点の動画とユーザの動作データを記録する。具体的には、スマートウォッチに搭載した3軸加速度/角速度センサからのモーションデータを用いて、組み立て作業中の作成者の大まかな手の動作を記録する。次に、これらのデータを機械学習を用いて分類し、作業内容を自動的に判断する「タグ付け」に利用する。また、付けたタグをもとに、チャプターや字幕を搭載したマニュアルのアウトラインを自動生成する。ユーザは生成されたアウトラインを基に、編集することで、マニュアルを制作する。本システムでは、

汎用 3D プリンタやレーザーカッター等で出力した部品を机の上で組み立てる作業を主に想定している。この作業環境における動作(部品嵌め合わせ, 工具とネジによる固定等)をモーションデータとして記録することで, 動画だけでは判別しにくい「手技」を記録/伝達できる可能性がある。また, 動画中の作業内容をモーションデータから推定出来れば, 動画に対して自動的にタグ付けを行い, マニュアル作成の負担の軽減につながる可能性がある。

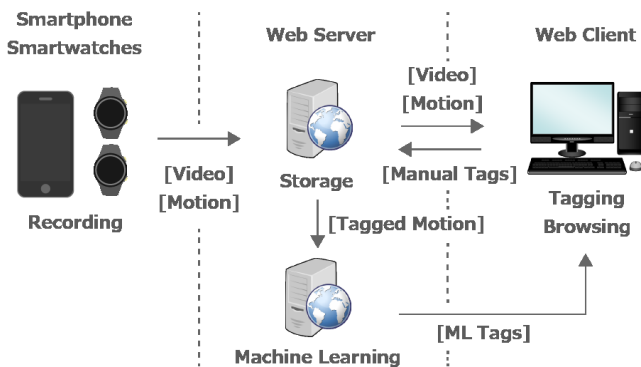


図 1 システム構成



図 2 記録システムの外観

4. 実装

本章では, 「記録スマートフォンアプリ」「タグ付け Web アプリ」「自動タグ付けシステム」「動画マニュアル編集 Web アプリ」の実装について述べる。

4.1 記録スマートフォンアプリ

前述したように, 記録スマートフォンアプリは, スマートフォンと 2 台のスマートウォッチで構成される。Android OS は 1 台のスマートフォンで複数台のスマートウォッチと接続できるため, スマートフォンとして Android 端末 (MotoG5 Plus), スマートウォッチとして Huawei Watch2 を採用し, 実装を行った。スマートウォッチアプリでは, ユーザがモノづくり中の動作の記録を開始/停止するためのボタンが用意されている。記録が開始されると, スマートウォッチはセンサデータ (3 軸加速度/3 軸角速度/心

拍数) を記録し, これらのデータをスマートフォンに送信する。なお, データ取得の周期は 50Hz とした。スマートフォンアプリは, 動画 (mp4 形式) を撮影し, スマートウォッチから受信したセンサデータ (CSV 形式) とともに, Web サーバーにアップロードする。

4.2 タグ付け Web アプリ

動画とセンサデータが Web サーバーへ格納された後, 作成者はタグ付け Web アプリを用いて, 組み立て過程の各動作に「タグ」を付加することができる。タグ付け Web アプリの外観を図 3 に示す。本アプリは Web サービスとして実装した。本アプリは, 動画再生画面とセンサデータのグラフを描画している。グラフ上では, センサデータを折れ線グラフとして可視化し, 動画の操作とタグ付けをすることができる。センサデータ上の任意の位置をクリックすると, 指定した位置の動画が再生される。また, 現在の再生位置は灰色のバーにより確認することができる。グラフ上に表示するデータ軸は, 画面中央下に表示されているチェックボックスから変更することが可能である。次に, タグ付けの流れについて, 図 4 に示す。まず, タグ付けしたいセンサデータの範囲をドラッグアンドドロップし, 右下のテキストボックスでタグを入力して, 作成ボタンをクリックすることでタグ付けを行うことができる。また, 付けたタグは, グラフの上部に表示される。また, グラフの上部に表示されるタグをクリックし, ダイアログ上で操作することで, そのタグを削除することができる。

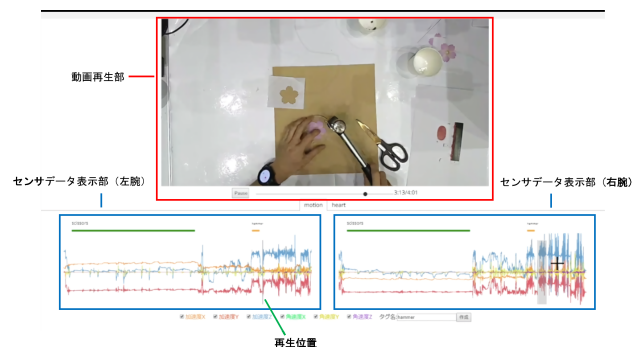


図 3 タグ付け Web アプリのインターフェース

4.3 自動タグ付けシステム

一定量のタグが格納された後, システムは機械学習を用いて自動的にモーションデータにタグを追加することができる。機械学習の方式は複数検討中だが, 現時点では, 滑り窓法と DTW (Dynamic Time Warping) 法を採用した。学習データは, 手動で付けたタグに基づいて Web サーバー上にあるモーションデータから抽出する。滑り窓を用いて, 入力データと学習データを比較し, DTW 距離が閾値を下回った場合, 学習データのタグが入力データの滑り窓

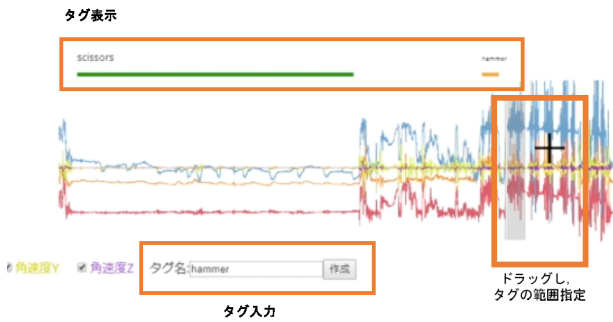


図 4 センサデータ表示部の拡大図

の範囲に付加される。なお、ウィンドウサイズは 150 サンプル（約 3 秒）とした。

4.4 動画マニュアル編集 Web アプリ

タグ付けアプリ等で付加したタグを基に、本アプリでは、チャプターや字幕を含む動画マニュアルのドラフトを生成する。マニュアル編集アプリの画面を図 5 に示す。字幕は画面中央下に表示される。チャプターは、シークバー上に赤いマーカーとして表示される。また、チャプターの一覧はリストとしても表示され、任意のチャプターをクリックすると、その再生位置まで移動することができる。これにより、どの再生位置でどのような動作をしていたかを容易に確認することができる。さらに、本アプリは字幕に基づいた要約再生機能を提供する。具体的には、字幕のある箇所は通常再生（1 倍速）し、字幕の無い箇所は高速再生（8 倍速）することができる。それぞれの再生速度は任意に設定することができ、動画を効率的に閲覧することができる。また、本アプリは字幕やチャプターの編集機能も提供しており、ユーザは、生成されたドラフトを基に編集を行い、マニュアルを製作する。

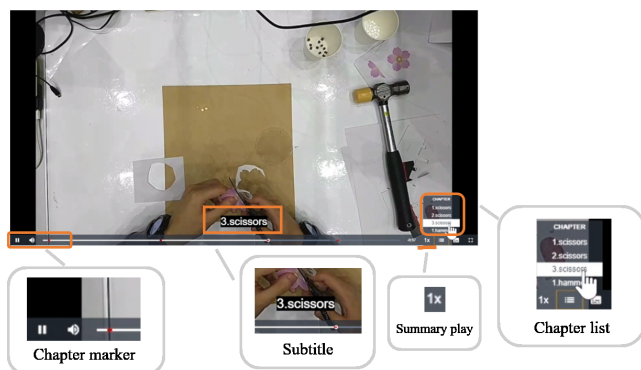


図 5 動画マニュアル編集 Web アプリのインターフェース

5. コンテンツ事例

本章では、実際に本システムで収集したコンテンツを紹介する。また、収集したデータを分析し、モノづくり中の動作の傾向を考察する。

5.1 3D プリントモンスターの組み立て

はこだて未来館*4にて開催されたワークショップにて、3D プリンターで出力したモンスターに様々なパーツをはめ、キーホルダーを作成する過程（図 6）を 7 件記録した。キーホルダー作成のデータの例を図 7 に示す。パーツをはめる穴が小さく、パーツが入らなかったため、棒やすりで穴を広げる過程があった。棒やすりをひねる動作のデータを見ると、腕をひねる動作により、利き手の角速度 X が大きく振幅していた。



図 6 3D プリントモンスターの例と組み立ての様子



図 7 3D プリントモンスターのモーションデータ例。やすりを利用する際に、利き手の角速度 X に大きな振幅が見られる。

5.2 のこぎりで木を切る

FabLab SENDAI – FLAT に訪問し、のこぎりで木を切る過程（図 8）を 6 件記録した。のこぎりで木を切る動作のデータの例を図 9 に示す。のこぎりの動作は、押し引きするような前後方向の動きにより、利き手の加速度 X に大きな振幅があり、切断に伴う振動より、利き手/逆手の加速度 Y / Z は同程度の小さな振幅が見られた。また、図 7 と比べると角速度はあまり変動しなかった。さらに、他者のデータ同士を比較すると（図 10）、のこぎりの角度や、切り方の違いにより、利き手の加速度 X の振幅の変位量が大きく異なっていた。さらに図 11 のデータでは、のこぎり

*4 <http://hakodate-miraiproject.jp/>

の刃が木に引っかかってしまい、加速度の振幅が、途中で止まっている箇所があった。



図 8 のこぎりで木を切る様子

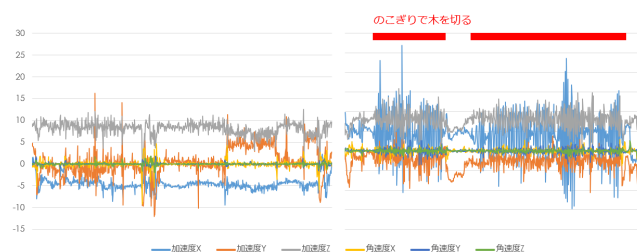


図 9 のこぎりで木を切るモーションデータ例

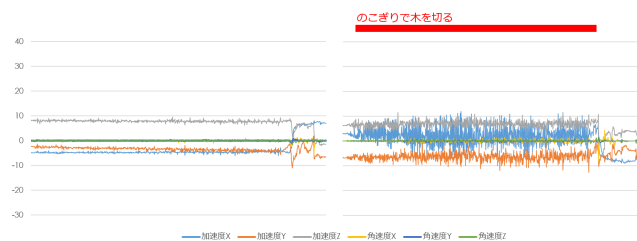


図 10 のこぎりで木を切るモーションデータの比較 (別ユーザ)。のこぎりの使い方により、利き手の加速度 X の振幅の変位量に差がでる。

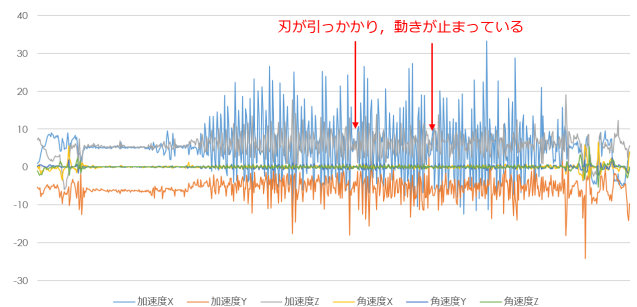


図 11 のこぎりで木を切るデータ (右手) 例。のこぎりの刃が木に引っかかってしまい、加速度 X の振幅が、途中で止まっている。

5.3 コースターの組み立て

本学研究室にて、コースターを組み立てる過程を 8 件記録した。コースターの組み立て手順は次の通りである。

- (1) はさみで花の型紙を切り取る
- (2) 2 枚の亚克力板で花をはさみ、組み立てる
- (3) 金槌でピンを打ち付ける

コースターを組み立てるモーションデータの例を図 13 に示す。このコンテンツでは、モーションデータがなだらかな箇所がはさみを使っており、変動の大きい箇所が金槌を使っている。はさみを利用する場合、加速度/角速度共に変化量が非常に小さく、その中で利き手の加速度 Z に周期的な変化が確認できた。また、金槌のような叩きつける動作は、下方向に叩く動作から、利き手の加速度 Z に大きな振幅が確認できた。利き手の加速度 X / Y には、叩く際の前後左右の腕の動きや、金槌を打ち付けた際の振動より、やや小さい振幅が見られた。また、逆手の加速度 X / Y / Z にも、金槌を打ち付けた際の振動より、小さな振幅が見られた。

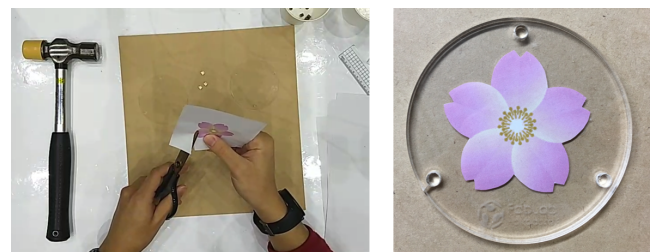


図 12 コースターを組み立てる様子と組み立てるコースター

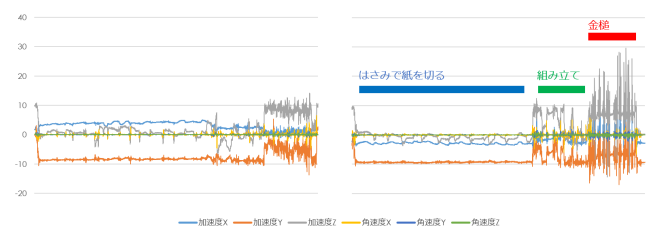


図 13 コースターを組み立てるモーションデータ例

5.4 まとめ

3D プリントモンスターやコースターの組み立て、及びのこぎりで木を切る動作を記録し、分析を行った。腕を回す、押し引きするような工具を使った動作は、工具を持っている腕の特定の加速度 (のこぎり: X 軸, ハンマー: Z 軸) に大きな変位が確認できた。また、工具の使い方でも、モーションデータに個人差が出ることや、上手くできていない作業が推定できる可能性があることがわかった。正常な動作と比較することで、上手くできていない作業を検出し、工具の使い方などの支援に活用できる可能性があると考えられる。

6. 議論

本章では、「モノづくりの動作検出手法」と「動作検出手法の比較」について議論する。

6.1 モノづくりの動作検出手法

本システムでは、DTWと滑り窓を用いた部分シーケンスマッチングにより、モノづくり中の動作検出を試みた。現在、前処理や特徴量の抽出をせずに、動作検出を行っている。定量的な評価は現在進行中だが、前章で述べたように、工具を使った動作などの特徴的な動きの検出は比較的容易にできる可能性があると考えられる。しかし、工具の使い方も個人差があるため、モーションデータに差が出てしまうため、(1) ユーザ個人毎に学習データを生成するか、(2) 他者の学習データから適切な特徴量を抽出する必要があると考えられる。また、動作検出の精度を向上させる方法として、指などに多数のセンサを取り付けたり、モーションキャプチャ等で指の動きを検出する方法が考えられる。しかし、センサの装着負荷/設置コストなどが増大し、ユーザの負担が増えるため、新たにセンサ類を取り入れることは、慎重に検討する必要がある。今後もモノづくり中の動作を分析し、適切な前処理/特徴量抽出、手法の検討を行い、精度の向上を目指していきたい。

6.2 画像/音声による動作検出手法との比較

本研究では、加速度/角速度センサを用いた動作検出を行った。他の動作検出手法として、画像認識技術を用いて、動作を検出する方法も考えられる。例えば、武田ら[12]は、動画から取得した肌色領域から手の領域/指の領域を検出し、これらの情報を用いて、手や指部分の座標等の情報からジェスチャの判定を行っている。このような技術を用いれば、手指の情報からのモノづくり中の動作認識が可能になり、細かい動作も検出できる可能性がある。また、画像認識技術を用いることで、工具や部品の検出が可能になると考える。一方で、加速度/角速度センサを用いた動作検出では、工具や部品の検出や、指の情報は取得することは難しいが、手首を回した角度や腕の動かす速さ等の動作を取得できると考える。また、音声を利用し、動画編集を支援しているシステムも数多く存在する[1][9]。音声を利用することで、重要な動作や、間違った動作などを推定できる可能性がある。今後は、この議論を基に、それぞれの手法の特徴を活かした、効果的な利用方法を検討したいと考える。

7. まとめと今後の展望

本稿では、スマートウォッチを用いたモノづくりマニュアル作成支援システムについて記述した。本システムは、スマートフォンとスマートウォッチを使用して組み立て作

業を記録する。次に、作成者/共同作業者は、動作データ内の特徴的な動作に「タグ」を追加する。一定量のタグが格納された後、システムは自動的にモーションデータにタグを追加することができる。次に、これらのタグを使用して、字幕やチャプターを含んだドラフトマニュアルが生成される。また、本システムを使用して、モノづくり過程を記録し、モノづくり中の動作の傾向の分析を行った。これにより、工具を使用する動作は、特徴的な動きになることがわかった。今後も引き続き、本システムを用いてデータ収集を行い、モノづくり中の動作の傾向を分析する。また、その分析から適切な特徴量を検討し、モノづくりの動作検出を試み、動作検出精度の向上を目指す。さらに、本システムが動画マニュアル作成支援に有効かどうかをユーザ評価実験より検証していきたい。

謝辞 本研究の一部はJST CRESTの支援を受けた。

参考文献

- [1] Chi, P.-Y., Liu, J., Linder, J., Dontcheva, M., Li, W. and Hartmann, B.: DemoCut: Generating Concise Instructional Videos for Physical Demonstrations, *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, ACM, pp. 141–150 (2013).
- [2] Fraser, C. A., Grossman, T. and Fitzmaurice, G.: WeBuild: Automatically Distributing Assembly Tasks Among Collocated Workers to Improve Coordination, *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, ACM, pp. 1817–1830 (2017).
- [3] Higuchi, K., Yonetani, R. and Sato, Y.: EgoScanning: Quickly Scanning First-Person Videos with Egocentric Elastic Timelines, *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, ACM, pp. 6536–6546 (2017).
- [4] Kurihara, K.: CinemaGazer: A System for Watching Videos at Very High Speed, *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, ACM, pp. 108–115 (2012).
- [5] Maekawa, T., Nakai, D., Ohara, K. and Namioka, Y.: Toward Practical Factory Activity Recognition: Unsupervised Understanding of Repetitive Assembly Work in a Factory, *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '16, ACM, pp. 1088–1099 (2016).
- [6] Nakae, K. and Tsukada, K.: Support System to Review Manufacturing Workshop Through Multiple Videos, *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*, IUI '18 Companion, ACM, pp. 4:1–4:2 (2018).
- [7] Ranjan, J. and Whitehouse, K.: Object Hallmarks: Identifying Object Users Using Wearable Wrist Sensors, *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, ACM, pp. 51–61 (2015).
- [8] Saul, G., Lau, M., Mitani, J. and Igarashi, T.: SketchChair: An All-in-one Chair Design System for End Users, *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '11, ACM, pp. 73–80 (2011).

- [9] Truong, A., Berthouzoz, F., Li, W. and Agrawala, M.: QuickCut: An Interactive Tool for Editing Narrated Video, *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, ACM, pp. 497–507 (2016).
- [10] Tsukada, K., Oki, M., Watanabe, K. and Akatsuka, D.: FabNavi: Support system to assemble physical objects using visual instructions, Vol. 2, Paper presented at Fab10, pp. pp.2–8 (2014).
- [11] Wang, E. J., Lee, T.-J., Mariakakis, A., Goel, M., Gupta, S. and Patel, S. N.: MagnifiSense: Inferring Device Interaction Using Wrist-worn Passive Magneto-inductive Sensors, *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, ACM, pp. 15–26 (2015).
- [12] 武田大暁, 須藤 智, 恩田憲一: 画像情報処理によるハンドジェスチャ認識アルゴリズム研究, 全国大会講演論文集, Vol. 72, pp. 699–700 (2010).