

カメラ及びタッチパネルを有する スマートフォンを用いた 遠隔地間作業指示支援

阪本 敦哉[†] 鈴木 直義[†] 湯瀬 裕昭[†] 渡邊 貴之[†]

コンピュータや産業機械などの操作、物品の並べ替えなど、特定の作業を行う際、遠隔地間作業指示支援システムを用いることによって、熟練した指導者が作業の現場まで出向く必要がないため、業務の効率化やコストの削減が可能である。一方、従来の遠隔地間作業指示支援システムでは、視覚情報を提示する方式に合わせて専用の装置を開発したり、複数のセンサやデバイスを組み合わせることで装置を実現していた。そのため、誰しもが手軽に入手したり、利用できるシステムとはなっていない。本報告では、カメラ及びタッチパネルを有するスマートフォンを用いて遠隔地間作業指示支援システムを構築し、システムを用いた場合の作業効率の変化について評価を行った。

Supporting the Remote Instruction by Smart Phones Having the Camera and the Touch Panel

Atsuya Sakamoto[†], Naoyoshi Suzuki[†], Hiroaki Yuze[†] and Takayuki Watanabe[†]

The supporting system for remote instructions enables to reduce costs and raise business efficiency, because skilled technicians do not have to visit to remote workplaces. Because previous systems were composed of several separated devices, it is not easy to use for the general user outside the research community. In this report, we present an useful supporting system for remote instructions on consumer smart-phones.

1. はじめに

コンピュータや産業機械などの操作、物品の並べ替えなど、特定の作業を行う作業員に対して、指導者が遠隔地から作業の指示を行う遠隔地間作業指示では、熟練した指導者が作業の現場まで出向く必要がないため、業務の効率化やコストの削減が可能となる。

その一例としては、臨床看護師に対して、初級レベルで数十時間にも及ぶ実技を伴うフィジカル・アセスメント教育を行う場合、国内で数少ない技術指導者のもとに多忙な看護師が通学するのは困難であり、またその逆も難しい[1]。筆者らは、文献[2]において、高精細なテレビ会議システムを用いて、遠隔地間でフィジカル・アセスメント教育指導を行う実験を行い、カメラで撮影した映像及び音声をそのまま伝送した場合、言葉と身振りのみでは詳細な表現・伝達が難しい指導内容が多く存在することが明らかとなった。そのため、学習者や指導者の手技によって対象に働く圧力分布を、圧力センサを用いて測定し、AR (Augmented Reality) の手法を用いて分布を映像に合成して提示するシステムを試作している[3]。しかし、この提示システムは、圧力センサを接続したノート PC や DV カメラ、ディスプレイなど複数デバイスから構成された据置型のインタフェースを採用しており、システムの可搬性は考慮されていなかった。

一般に、作業場所や撮影する対象物が固定されない作業に対して、遠隔地から作業の指示を行う場合には、作業員側は据置型でなく、ウェアラブル型やハンドヘルド型のインタフェースを採用したシステムが適している。また、本研究で扱う作業指示支援システムは、分散同期型の CSCW (Computer Supported Cooperative Work) の範疇に含まれるが、ClearBoard2[4]や Agora[5]とは異なり、作業の指示は指導者から作業員への片方向に限定されるため非対称なシステムとなる[6]。

既存のウェアラブル型やハンドヘルド型インタフェースに対応した遠隔地間作業指示支援システムに関する研究としては、WACL[7][8]や WSLP[9]、NaviCam[10][11]などがある。これらのシステムは、主に作業員に視覚情報を提示する装置、作業対象を撮影するカメラ、ネットワークに接続されたコンピュータから構成されている。また、作業員が目視している作業対象の映像や、対象物そのものに対して、指導者が入力した指示を重ね合わせて作業員に提示している。提示方法としては、透過型 HMD (Head Mount Display) やハンドヘルドディスプレイ、実空間へのレーザポインタやプロジェクタでの投影などが用いられている[12]。これらのシステムでは、視覚情報を提示する方式に合わせて専用の装置を開発したり、既存のセンサやデバイスを組み合わせることで装置を実現している。そのため、誰しもが手軽に入手したり、利用で

[†] 静岡県立大学 経営情報学部 経営情報学科
School of Administration and Informatics, University of Shizuoka.

きるシステムとはなっていない。

一方、近年急速に普及してきたスマートフォンと呼ばれる高機能携帯電話では、カメラ、タッチパネル、GPS、電子コンパス、加速度センサ、通話機能などがコンパクトに集積されている。そのため、新たに装置を開発したり、別々のデバイスを組み合わせずとも、スマートフォン単体のみで遠隔地間作業支援指示システムのプラットフォームとしての活用が期待できる。実際に、文献[13]では携帯電話を用いてビデオ透過型 AR システムの試作が検討され、携帯電話の処理性能の低さやカメラの解像度の低さが、AR におけるマーカ読み取り精度や秒間表示コマ数に与える影響について初期検討が行われている。さらに、携帯電話やそのタッチパネルをプラットフォームとした AR システムの検討が近年活発に行われている[14][15]。

本報告では、カメラ及びタッチパネルを有するスマートフォンを用いて遠隔地間作業指示支援システムを構築し、システムを用いた場合の作業効率の変化について評価を行う。

2. 関連研究

指導者が遠隔地にいながら視覚情報や音声情報を用いて作業の指示を行う遠隔地間作業指示支援に関する研究は、これまでに様々に報告されている。

WACL (Wearable Active Camera/Laser) [7][8]では、作業者が肩上にウェアラブルなカメラとレーザポインタユニットを装着する。遠隔地の指導者は、レーザポインタを使って作業の対象物体を強調することができる。また、音声や映像、カメラやレーザポインタのパン・チルト制御データなどは無線 LAN で送信されるため、作業範囲を机上に限定せず、より広い空間での自由度の高い作業を可能としている。また、作業者が送信したカメラ映像に、指導者側で線画を書き込み、その映像が作業者の HMD もしくは、CWD (Chest Worn Display) に表示される機能を持つ。文献[7]では、HMD よりも CWD の方が作業者にとって身につけた際の違和感が少なく、作業効率にも差がないことが報告されている。また、WSLP (Wearable Scanning Laser Projector) [9]では、作業対象となる物体に指導者からの指示情報をレーザースキャン方式で直接投影するシステムを提案している。

一方、NaviCam[10]は、遠隔地間作業指示支援を主眼に開発されたシステムではなく、初期の AR 用端末として試作されたものであり、ハンドヘルドディスプレイを有する PDA (Personal Digital Assistant) の背面にカメラを接続した形態を取る。また、映像処理は据置型のワークステーションで行い、PDA とは有線で接続されている。文献[11]では、NaviCam の幾つかの応用例が示されており、その一つとして指導者が遠隔地から映像に書き込んだスケッチが作業者側のハンドヘルドディスプレイに表示される遠隔地間作業支援指示について紹介されている。

3. 提案システム

既存の遠隔地間作業指示支援システムでは、視覚情報を提示する方式に合わせて専用の装置を開発したり、既存のセンサやデバイスを組み合わせることで装置を実現する例が多く、誰しもが手軽に利用できるシステムではない。本報告では、カメラ及びタッチパネルを有するスマートフォンを用いて遠隔地間作業指示支援システムを構築する。

本研究において開発するシステムは、先に示した NaviCam に類似しているが、近年高性能化が著しいスマートフォンをプラットフォームとして採用することにより、音声通話機能との統合や、指導者側におけるタッチパネルを用いた直観的な指示インタフェースを実現する。本システムでは、作業者と指導者が音声通話中に、煩雑な設定なしに数秒以内に遠隔作業指示支援アプリケーションの起動と利用開始が可能である。

3.1 システムの構成

提案システムを実現する通信方式としては、C/S 型と P2P 型が考えられる。Apple 社の iPhone[16]や Google 社の Android OS[17]を採用したスマートフォンでは、携帯電話の通信回線だけでなく無線 LAN 経由の通信も可能である。そのため、単純な P2P 型では、プライベート IP アドレス空間に設置された無線 LAN 及び NAT を経由してインターネットに接続する際に、インターネット側から端末の IP アドレスを特定することが困難となるという問題がある。一方、IP ベースの通話アプリケーションの一つである Skype[18]では、プライベート IP アドレス空間内の端末同士で、NAT を経由した通話が可能である。これは、グローバル IP アドレスを持つ他のユーザの端末が、スーパーノードと呼ばれる中継器の役割を担うことにより問題を解決している。従って、グローバル IP アドレス空間内に多数のアクティブユーザが存在することを前提としたシステムであり、本システムには合致しない。

以上から、本システムでは、グローバル IP アドレスを持った Web サーバを経由して作業側側の端末と指導者側の端末が映像や指示情報をやり取りする C/S 型を基本構成とした。本研究では、作業側側と指導者側の端末で動作する専用のアプリケーションを開発した。また、本システムは Web サーバを経由してやり取りが行われるため、指示のみを行う指導者側は専用のアプリケーションでの対応だけでなく、Web ブラウザを用いた対応も容易である。従って、指導者側端末で動作する Web アプリケーションについても構築した。

提案システムの構成を図 1 に、システムの流れを図 2 に示す。本システムでは、まず、作業側側のカメラの撮影画面を一定時間 (秒間 2 コマ) ごとに自動でキャプチャし、その画像を jpeg に変換して Web サーバに送信する。スマートフォンのハンドヘルドディスプレイは自在に回転できるため、撮影時の画面の回転情報についても Web

サーバに合わせて送信する。

次に、指示者側は、作業側がアップロードした Web サーバ上の画像を自動でダウンロードし、ディスプレイに表示する。指導側が専用アプリケーションを用いている場合には、作業側が撮影した画像の回転位置が自身のディスプレイの回転位置と異なっていた場合でも、ディスプレイを簡単に回転させることで対応できる。一方、指導側が Web アプリケーションを用いている場合には、Web アプリケーションはスマートフォンのみならず通常の PC でも動作するため、画面の回転に容易に対応できない可能性も考えられる。そのため、Web アプリケーションでは、作業側がアップロードした撮影時の回転情報をもとに、撮影画像を適切に回転して表示する。

次に、指示者は画像上の任意の座標をタッチすると、その座標にマークがオーバーレイ表示される。また、座標情報は Web サーバに送信される。次に、作業側の端末は、指導側がアップロードした座標情報を Web サーバからダウンロードし、自身のカメラ画像上の当該座標にマークをオーバーレイ表示する。

一方、作業者と指導者の音声通話は、スマートフォンの通話機能を用いる。

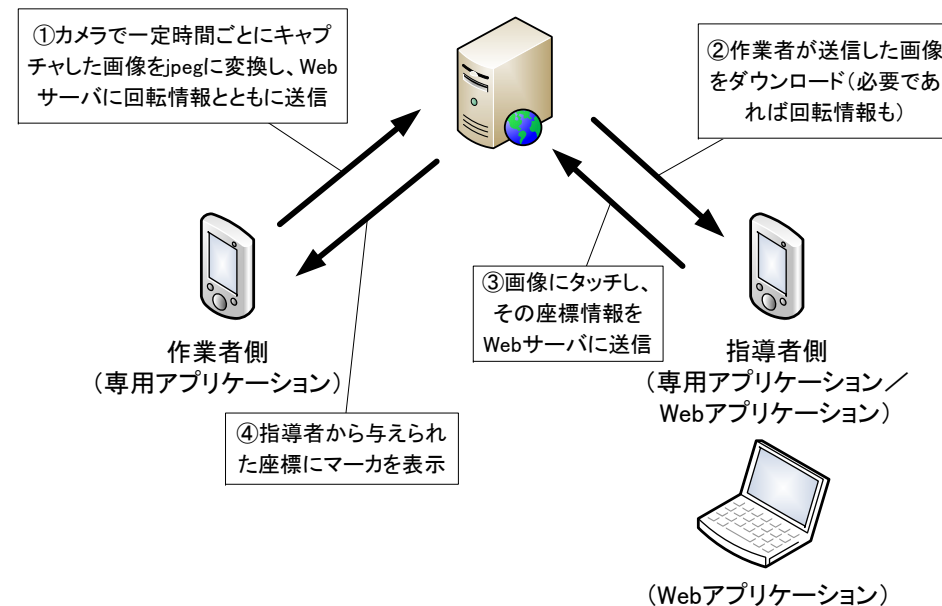


図1. 提案システムの構成

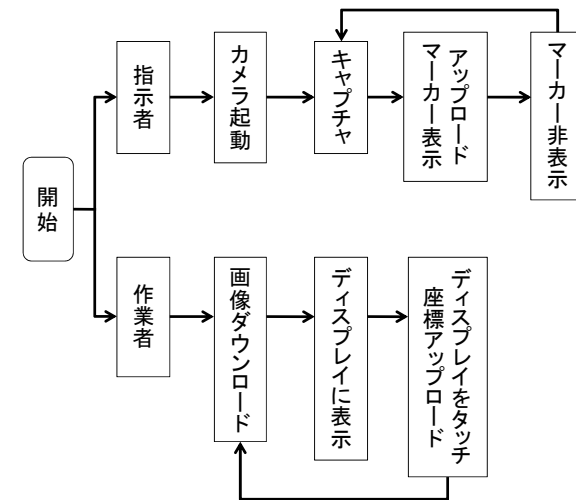


図2. 提案システムの流れ

3.2 アプリケーションの構築

本システムでは、アプリケーション開発実績の多い Apple 社の iPhone3GS を用いてシステムの開発を行った。iPhone 3GS の主な使用を表1に示す[16]。

表1 iPhone 3GS の端末性能表

画面サイズ	3.5 インチ HVGA(480x320)
CPU 動作周波数	600MHz
メモリ	256MB
Bluetooth	Bluetooth 2.1 + EDR
センサ	加速度センサ, 電子コンパス
静止画撮影	300 万画素, オートフォーカス
電池寿命	連続通話:300 分 待受時間:300 時間
重量, 大きさ	135g, 115.5×62.1×12.3mm

iPhone 上で動作する専用アプリケーションは、Objective-C により iPhone SDK 3.1.2 を用いて開発した。また、Web サーバとしては Apache を用い、Web サーバ側で動作する画像のアップロード用のスクリプトは PHP を用いて記述した。

一方、Web ブラウザ上で動作する指導者側の Web アプリケーションは、HTML5 の機能である Canvas 要素を用いて、JavaScript と JQuery[19]による Ajax (Asynchronous JavaScript + XML) 指向のアプリケーションとして構築した。

専用アプリケーションでは、起動後、図 3 に示すように、作業モードと指示者（指導者）モードの選択画面が表示される。ここで、例えば指示者を選択すると、作業者が撮影した画像が表示され、タッチした場所に赤いマークが表示される（図 4）。

一方、指導者側で利用可能な Web アプリケーションを、PC 上の Firefox で実行した例を図 5 に示す。Web アプリケーションでは、作業者の撮影画像の回転情報をもとに、画像を適切に回転し表示している。また、Web アプリケーションは iPhone 組み込みの Safari でも動作することを確認している。

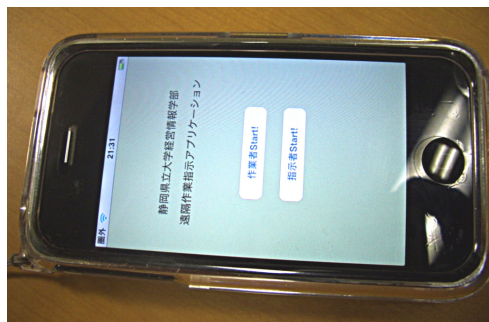


図 3. モード選択画面



図 4. 指示者側のマーカーイメージ



図 5. Web アプリケーションによる指示者側のインタフェース

4. 評価実験と考察

本章では、提案システムを使用し、遠隔地間作業指示支援に関する評価実験を行った結果を示す。

4.1 評価実験内容

本研究では、FPGA 評価ボード上を流れるクロック信号をオシロスコープを用いて測定する作業を、オシロスコープや電気回路の知識を全く持たない作業者に対して、指導者が遠隔地から作業の指示を行うシナリオについて行った。本実験では、20 代の男性を 2 つの作業者グループに分け、それぞれ下記のように作業指示を行った。

- グループ 1：指導者は、作業者が撮影した画像を確認し、音声のみで指示を与える。
- グループ 2：指導者は、作業者が撮影した画像を確認し、作業者に指示したいディスプレイ上の座標をタッチすることで、作業者側にマーカーを表示させ、指示を行う。

本報告では、グループ 1 とグループ 2 に対し、それぞれ 3 名ずつ計 6 名の被験者に対して実験を行った。尚、指導者はすべての実験で同一とし、30 代のオシロスコープの知識を持つ男性が担当した。図 6 に作業の様子を示す。

また、より詳細な作業の内容を表 2、図 7、図 8 に示す。グループ 2 への指示では、指示位置に関しては、マーカーを活用して指示代名詞を積極的に用いた指示を行った。

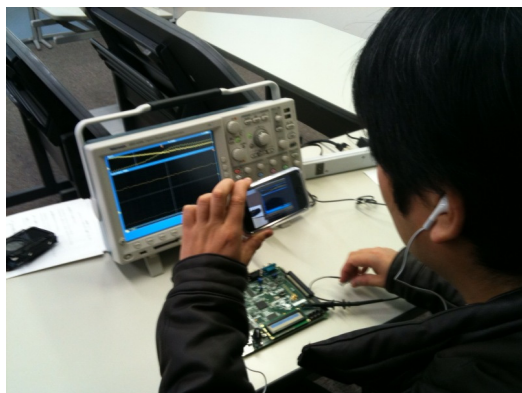


図 6．実験の様子

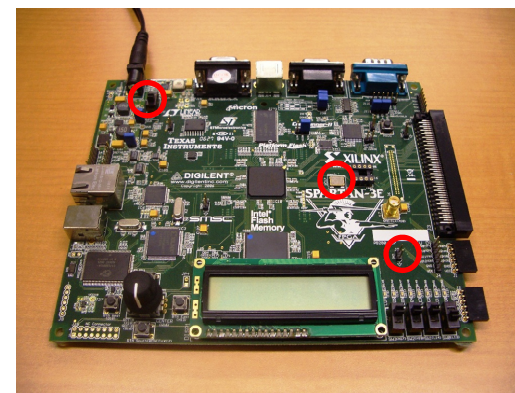


図 7．FPGA ボードと指示位置

表 2．実験における指示内容

作業内容	グループ 1 への指示内容	グループ 2 への指示内容
全体説明	作業全体の概要説明	作業全体の概要説明
電源投入	電源スイッチは、電源ケーブルのそばにある、スライド型のスイッチです。	この場所（図 7 左上の○）のスイッチを入れてください。
GND プローブの接続	プローブの端子を、ピンにひっかけます。ピンは、ボード中央から右の金色のコネクタの下、縦に 2 つ並んだピンの下側の方です。	プローブの端子を、ピンにひっかけます。ピンは、ここ（図 7 右下の○）にある 2 つ並んだピンの下側の方です。
信号プローブの接続	場所は、ボード中央すこし右の銀色の小さな長方形の部品です。その部品の左下の端子に押し当てます。	この（図 7 中央右の○）部品の左下の端子に押し当てます。
グラフの位置調整	波形の表示位置が上にずれているので、黄色いボタンの上の小さなつまみを回して、下に下げます。	この（図 8 中央右の○）つまみを回して、波形の表示位置を下げます。
グラフのスケール調整	波形を拡大してみたいので、中央の大きなつまみを右に回して調整します。	波形を拡大してみたいので、この（図 8 右上の○）大きなつまみを右に回して調整します。

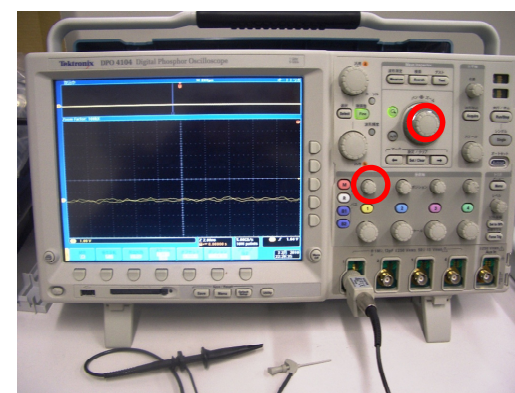


図 8．オシロスコープと指示位置

作業は、作業者及び指導者ともに、専用アプリケーションを用いて行った。また、音声通話以外の通信は無線 LAN 用いて行い、指示内容を聞き取りやすくするために、作業者及び指導者はマイク付きイヤフォンを装着している。また、画像の更新間隔は秒間 2 コマとし、jpeg 画像の品質を 20% に設定した。このとき、jpeg 画像のファイルサイズは最大で約 20kb であった。

4.2 実験結果と考察

表 2、表 3 から、グループ 1 と比較してグループ 2 の方が、作業時間が短縮されて

いることがわかる．なお、被験者 f については、途中で音声通話が途切れ、電話のかけ直しとアプリケーションの再起動を行っている．しかし、専用アプリケーションの立ち上げは数秒で完了するため、この際の時間的ロスを含めても、依然としてグループ 1 の作業時間の平均値よりも早く作業を完了していることが分かる．

グループ 1 に比較して、グループ 2 の作業時間が全体として短縮された理由は以下のように考えられる．グループ 1 では音声のみによる指示のため、指導者側は、場所を指定する際、形状、色などの情報を言葉で伝える必要があるが、グループ 2 では、マーカを表示させて指示代名詞により簡単な指示をするだけで、場所の指定が可能となる．また、作業者にとっても、マーカによって容易にアウェアネスを得ることができる．

表 2 グループ 1 の結果

	被験者 a	被験者 b	被験者 c	平均値	標準偏差
作業時間	373.44s	413.08s	407.89s	398.14s	21.54s

表 3 グループ 2 の結果

	被験者 d	被験者 e	被験者 f	平均値	標準偏差
作業時間	283.40s	325.47s	375.93s	328.27s	46.33s

次に、実験終了後に被験者に対して行ったアンケートの結果を表 4 に示す．結果より、マーカ表示の有無にかかわらず、被験者は作業者の指示が分かりやすかったと感じていることが分かる．また、スマートフォンを片手に持って作業を行うため、プローブやオシロスコープを片手で操作することが困難になる場面が度々見受けられた．指導者は、実験開始時に、必要であればスマートフォンを一旦置いて構わないと指示している．そのため、両手を必要とする作業を行う際、被験者はスマートフォンを置き、再度持つ動作が必要となる．結果として、表 4 からわかるように、被験者の多くがカメラを撮影しながらの作業に困難を感じていることがわかる．

表 4. アンケート結果（グループ 1，グループ 2）

Q2.指導者の指示はわかりやすかったか？	1. はい	6 名
	2. いいえ	0 名
	3. どちらでもない	0 名
Q3.カメラで撮影をしながら作業をするのは困難だったか？	1. はい	4 名
	2. いいえ	2 名
	3. どちらでもない	0 名

5. まとめ

本報告では、カメラ及びタッチパネルを有するスマートフォンを用いた簡易な遠隔地間作業指示支援システムを開発した．FPGA 評価ボード上を流れるクロック信号をオシロスコープを用いて測定する作業について実験を行った結果、本システムの有効性が確認された．

一方で、スマートフォンを片手に持つことで作業が困難になる場面が確認された．この問題の解決法としては、WACL での CWD のように、スマートフォンを胸部に固定する方法も考えられるが、カメラの撮影方向が作業者の視線と大幅にずれるため、現実的でない．また、映像の秒間コマ数が 2 コマと少ないため、H.264 などへの対応も今後の課題である．

参考文献

- [1]. 今本, 北村, "解剖生理学を基盤にしたフィジカルアセスメントの看護学教育への導入," 形態・機能, コ・メディカル形態機能学研究会 機関誌, 第 3 巻第 1 号 2004 年 9 月.
- [2]. 鈴木, 酒井, 洪沢, 森下, 岡本, 湯瀬, 芥川, 山上, 伊藤, "リモート・インストラクションによるフィジカル・アセスメントの実習訓練の試み,II," 情報処理学会研究報告, 情報システムと社会環境 2007(25), pp.91-98, 2007 年 3 月.
- [3]. 洪沢, 渡邊, 佐藤, 岡本, 米樹, 湯瀬, 松浦, 青山, 鈴木, "客観的な指標に基づく触診型手技の学習支援システムの試作," 日本 e-Learning 学会誌, Vol.9, pp.64-73, 2009 年 5 月.
- [4]. H.Ishii, M.Kobayashi, J.Grudin, "Integration of Inter-Personal Space and Shared Workspace: ClearBoard Design and Experiments," Proc. ACM CSCW '92, pp.33-42, 1992.
- [5]. 山下, 葛岡, 井上, 山崎, "コミュニケーションにおけるフィードバックを支援した実画像通信システムの開発," 情報処理学会論文誌, Vol45(1), pp.300-310, 2004 年 1 月.
- [6]. M. Billingham, S. Bee, J. Bowskill, and H. Kato, "Asymmetries in collaborative wearable interfaces", In Proc. ISWC99, pp.133-140, 1999.
- [7]. 酒井, 蔵田, 葛岡, "遠隔協調作業のためのレーザポイントと装着型ディスプレイによる視覚的アシスト," 日本バーチャルリアリティ学会第 10 回大会論文集, pp.233-236, 2005 年 9 月.

- [8]. 蔵田, 酒井, 葛岡, 興梠, 大隈, 西村, "遠隔協調作業のためのウェアラブル・タンジブルインタフェース," SICE 第 69 回パターン計測部会研究会, pp.11-18, 2006.
- [9]. 安藤, 雨宮, 前田, "AR における注釈表示のためのウェアラブル・スキャニング・レーザー・プロジェクター," 日本 VR 学会論文誌, Vol.10, No. 2, pp. 191-200, 2005.
- [10]. J.Rekimoto, "Augmented Interaction: Toward a New Human-Computer Interaction Style Based on Situation Awareness", Interactive Systems and Software II (WISS'94 Proceedings), pp.9-17, 1994.
- [11]. J.Rekimoto and K.Nagao, "The World through the Computer: Computer Augmented Interaction with Real World Environments", User Interface Software and Technology (UIST '95), 1995.
- [12]. 清川 清 "複合現実感マルチユーザ・コラボレーション ～ヒューマンコミュニケーションの視点から～", 画像の認識・理解シンポジウム (MIRU 2002), A1-1, 2002.
- [13]. M.Mohring, C.Lessig, and O.Bimber, "Video See-Through AR on Consumer Cell-Phones," Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality, pp.252-253, 2004.
- [14]. M.Hakkarainen, C.Woodward, M.Billinghurst, "Augmented assembly using a mobile phone," Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, p.167-168, Sep. 15-18, 2008.
- [15]. G.A.Lee, U.Yang, Y.Kim, D.Jo, K-H.Kim, J.H.Kim, and J.S.Choi, "Freeze-Set-Go interaction method for handheld mobile augmented reality environments," Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology, pp.143-146, Nov. 18, 2009.
- [16]. iPhone, "<http://www.apple.com/jp/iphone/>"
- [17]. Android, "<http://www.google.co.jp/mobile/android>"
- [18]. Skype, "<http://www.skype.com/intl/ja/>"
- [19]. JQuery, "<http://jquery.com/> "