

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 3 日現在

機関番号：34406

研究種目：基盤研究(C)

研究期間：2010～2012

課題番号：22500101

研究課題名（和文）視覚的共同注意が可能な遠隔コミュニケーション環境の実現

研究課題名（英文）A remote communication environment in which visual joint attention is possible

研究代表者

西口 敏司 (NISHIGUCHI SATOSHI)

大阪工業大学・情報科学部・准教授

研究者番号：80362565

研究成果の概要（和文）：本研究課題では、遠隔地間コミュニケーションシステムにおいて、相手の映像を複数の深度センサを用いて三次元データとして獲得する技術や、ネットワーク上で互いにリアルタイムに送受信し、視点位置の変化に応じて提示する技術、及び、相手が指示している物体上の位置を遠隔地間で共有する技術を開発した。これらの技術によって、遠隔地間の対話者が、相手が見ているものや位置を把握しやすい遠隔対話インタフェースを構築した。

研究成果の概要（英文）：In this research, we developed three methods on a remote communication system. First, a method for acquiring remote dialogist's figure using multiple depth sensors as 3D data is proposed. Second, a method for real-time transmission and rendering of remote dialogist's figure transmitting on a network according to change of an observing point is proposed. Finally, we developed a method for sharing the location which a remote dialogist is pointing. Using these methods, we constructed a remote dialog interface that enables the users to grasp what the remote dialogue is seeing.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	1,700,000	510,000	2,210,000
2011年度	1,000,000	300,000	1,300,000
2012年度	600,000	180,000	780,000
年度			
年度			
総計	3,300,000	990,000	4,290,000

研究分野：センサ情報処理

科研費の分科・細目：情報学・メディア情報学・データベース

キーワード：情報システム，遠隔コミュニケーション，視覚的共同注意

1. 研究開始当初の背景

(1) 遠隔コミュニケーションを実現しようとする製品開発や研究は国内外で多数行われており、高精細な映像を伝送する機能を持つ製品や、電子的なホワイトボード画面などの資料を対話者間で共有することが可能な製品が市場に登場している。また、対話者のアイコンタクト映像を合成する手法についても活発に議論されている。これらの研究・開

発の特徴は、相手の姿や資料自体を対話者間同士で共有することにより臨場感を向上させようとする点である。しかしながら、実際の対話型コミュニケーションでは、相手の姿や資料だけではなくそれらに対する指示やインタラクションの様子を把握できることが対話中の意図を伝えるという観点から重要である。例えば、テーブルを挟んだ対話型コミュニケーションでは、相手がテーブル上

の何を見て話しているのか、あるいは、どの資料を指しているかが分かれば、何について議論しているかが明らかとなり相手の意図が把握しやすい。従来の遠隔コミュニケーションシステムでは、このような指示やインタラクションの様子に関する情報を遠隔地間で共有することについて考慮していなかったため、相手の意図が把握しにくいという問題があった。

(2) 一方、複数の人間が同じオブジェクトに対して視覚的な注意を向けることを、視覚的共同注意(visual joint attention)と呼ぶ。視覚的共同注意は、心理学的には情報共有のための手段として人間に備わった能力であると考えられており、工学的にはロボット工学の分野で、人間とロボットとの視線共有の実現や、ロボット同士が共同注意する手法に関して様々な研究が行われている。しかしながら、人と人の遠隔コミュニケーションを支援するために、遠隔地間での視覚的共同注意を可能にしようとする研究は国内外を含めてほとんどなされていない。

2. 研究の目的

本研究では、意図を伝えやすい遠隔コミュニケーション環境の構築を目的とした遠隔地間での視覚的共同注意を可能とするために、対話者間の「視線追従」と「指さし理解」を遠隔地間で実現する技術を提案する。具体的には、まず、自分の実空間のオブジェクト(ローカルオブジェクト)と相手の実空間のオブジェクト(リモートオブジェクト)を重畳提示するオブジェクト共有テーブルを、対話者それぞれの環境に構築する。そして、この共有テーブル上のオブジェクトが自分のテーブル上のオブジェクトであり、かつ相手のテーブル上のオブジェクトでもあること、つまり鏡像関係にあることと、対話者のテーブルに対する姿勢推定に基づいて、遠隔地間の視線追従、及び、遠隔地間の指さし理解を実現する。遠隔地間の視線追従とは、遠隔地の相手が見ているオブジェクトが何か、把握できるということであり、遠隔地間の指さし理解とは、遠隔地の相手が指さしているオブジェクトが何か、把握できるということである。さらに、これらの要素技術を適用し、双方向に視覚的共同注意が可能な遠隔コミュニケーション環境を試験的に構築し、実験によって提案手法の有効性を検証する。

3. 研究の方法

(1)遠隔地間の視線追従のための視線方向の推定、及び空間写像モデルの構築については、遠隔地の対話者が、オブジェクト共有テーブル上に置かれたローカルオブジェクト、あるいは表示されているリモートオブジェクト

に視線を向けたときに、どのオブジェクトに視線を向けているかをローカルの対話者が把握できるようにする必要がある。

これを可能とするためには、対話者が共有テーブル上のどのオブジェクトを見ているかを推定する必要がある。対話者が共有テーブル上のオブジェクトを見るときは、共有テーブルの天板を上から覗き込む姿勢となるため、本研究では、液晶ディスプレイを天板として構築した共有テーブルの、天板付近に利用者を見上げる形で設置した2台のカメラを利用して、共有テーブルの天板に対する視線方向を推定する。推定した視線方向と、共有オブジェクトの位置から視線を向けている共有オブジェクトを推定する。

(2)遠隔地間の指さし理解の実現には、視線追従の実現の場合と同様に、遠隔地の対話者が、オブジェクト共有テーブル上の共有オブジェクトを指さしたときに、どのオブジェクトを指さしたかをローカルの対話者が把握できるようにする必要がある。これを可能とするために、共有テーブルを真上から垂直に見下ろす位置にカメラを設置し、対話者の腕の形状及び動きを観測する。このとき、観測カメラからは、対話者の手や腕だけではなく、共有テーブル上に置かれたローカルオブジェクトや、液晶ディスプレイ天板上に表示されたリモートオブジェクトも同時に観測してしまうため、対話者の手や腕の形状と動きのみを抽出して追跡する工夫が必要となる。手や腕の位置が推定できれば、その位置に存在している共有オブジェクトを指さしていると推定する。

(3)遠隔地間の対話者の存在感を高めるために、遠隔地の対話者の空間的な写像モデルを構築する必要がある。これを実現するために、本研究では、対話者の像を一般的な可視カメラで獲得するのではなく、深度センサ付き可視カメラで獲得し、ローカルで対面して対話しているのと同等の位置関係を保つような写像モデルを構築する。

(4)(1)から(3)までで構成された要素技術に基づき、どのオブジェクトを見ているかが把握しやすいような提示を行う遠隔対話システムを構築することで、視覚的共同注意を可能とする遠隔コミュニケーション環境の構築による提案手法の有効性評価と総括を行う。このために、本研究で提案する視覚的共同注意が可能な遠隔コミュニケーション環境が、従来のような視覚的共同注意ができない遠隔コミュニケーション環境に比べてどの程度意図が伝わりやすくなるかについて実験・評価を行い、その効果を実証する。

4. 研究成果

(1)遠隔地の対話者がオブジェクト共有テーブル上に置かれたローカルオブジェクトに視線を向けたときに、その状況をローカルの対話者が把握できるようにする必要があるため、液晶ディスプレイを天板として構築した共有テーブルに対する視線方向をカメラで観測した画像を用いて推定する手法を開発した。対話者の手と腕の形状及び動きを観測するために、共有テーブルを真上から垂直に見下ろす位置にカメラを設置した。このとき、対話者の手や腕だけではなく、共有テーブル上に置かれたローカルオブジェクトや、液晶ディスプレイ天板上に表示されたリモートオブジェクトも同時に観測してしまうため、手や腕の位置を推定することにより、対話者の手や腕の形状と動きのみを抽出して追跡する手法を提案した。図1に、共有テーブルを真上から見たときの様子を示す。この図は、CGによって共有ディスプレイ上に表現された15個の円状のマーカのうちの1つを、図の下から伸びた手の指先で指示している様子であり、上から伸びている手とその指先は、遠隔地の対話者が伸ばした手を真上から撮影したカメラで撮影し、ネットワーク上を伝送した上で重畳表示したものである。

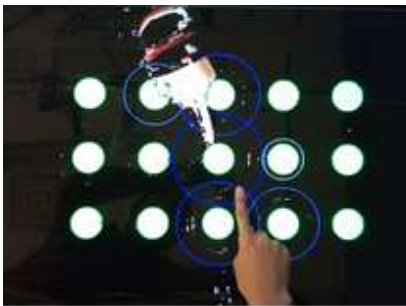


図1 指差し情報の共有機能

(2)遠隔地間の共同注意を支援するための基本システムを構築した。具体的には、相手の姿を映すディスプレイを正面に設置し、手元の情報を提示するために、ディスプレイのスクリーンを上に向けて設置してテーブル型ディスプレイとした環境を2組構築した。構築したシステムを図2に示す。相手の様子は正面のディスプレイに表示され、その様子を伺うことができる。テーブル型ディスプレイに(1)と同様に15個のマーカをCGで表示し、遠隔地の対話者の視線に関する情報をパターンマッチングにより推定し、推定結果としてのマーカの色を変えるなどして強調表示することで、どの程度正確に視線情報が伝達できるか評価した。その結果、視線推定精度は高くなかったものの、正しく視線方向が推定できていた場合は、強調提示がなかったときと比べて、よりよく視線情報が伝達できることが分かった。



図2 試作した遠隔コミュニケーションシステムの様子

(3)遠隔地の対話者の様子をより存在感を持って理解することを可能とするために、対話者の像を三次元CGとして構成し、テーブル型ディスプレイを拡張空間とローカルの実空間との間で幾何的整合性を持つような仮想現実空間を構築した。まず、対面ディスプレイの上に対話者に向けて設置した深度センサで対話者の様子を観測し、深度マップを獲得する。獲得した深度マップから三次元点群データを求める。遠隔地の対話者の環境で得られた三次元点群データは深度センサを原点とする座標系で表現されている。これを自分の実環境の延長上の仮想空間に、通信及び座標変換しながら再構成し、仮想現実空間を構築する。また、対話者に対して視覚的整合性を持った提示を実現するために、構築された仮想現実空間を対話者の運動視差に基づいて提示を行う。このような提示は、対話者の視点位置を推定し、この視点位置から観測可能なシーン映像を人物提示ディスプレイの画面上にレンダリングすることで実現した。構築したシステムについて、視覚的共同注意がどの程度支援できているかを評価するために、対話者の状況である運動視差を反映しなかった場合と反映した場合の2つ場合について実験により評価を行った。遠隔地の対話者が見ているオブジェクトが分かるかを、視線方向、指差し、の2つの指示方法で評価した。その結果、運動視差を反映しなかった提示では視線方向のみでは33%、指差しのみでは80%の正答率であったのに対して、運動視差を反映した提示では、視線方向のみ

では 47%, 指差しのみでは 100%と正答率が向上した. 深度センサで獲得した遠隔地の対話者を自分の視点位置に応じて正面のディスプレイに提示した例を図 3 に示す.



(a)視点位置を左に動かしたとき



(b)視点位置を正面に動かしたとき



(c)視点位置を右に動かしたとき

図 3 運動視差を反映した提示

(4)(3)で開発したシステムは, 深度センサによって獲得した点群データをネットワークを通して共有し, それぞれの対話者の実空間を拡張する空間に射影する機能, 及び, 仮想的に拡張された空間を自分の視点位置に応じてレンダリング表示することで運動視差に応じて提示する機能を持つが, 各対話者を観測する深度センサが 1 台である場合は, 提示時に視点位置が大きく変化したときに, 獲得側の深度センサから見えていない部分の提示が欠損し, 違和感が大きくなるという問題がある. この問題を解決するために, 各対話者を複数の方向から観測できるように深度センサを増設し, これらのセンサで得られる点群データを統合することで, 欠損の少ない映像提示ができるように改良した. また同時に, リアルタイム性の向上を行った. (3)のシステムでは, 提示する映像の更新頻度が約 1Hz であり, 特に運動視差に応じた映像を提示する場合に, 視点位置に応じた映像が提

示されるまでに大きな遅延が生じていた. これは, アンケート結果からも, 違和感が生じる要因の一つとなっている. そこで, (3)のシステムで実現している点群データの座標変換処理やレンダリング処理を, GPU (Graphics Processing Unit)を用いて超並列処理することでリアルタイム性能を向上した. 2 台の深度センサを用いて対話者の像を獲得する方式を図 4 に示す. また, 遠隔地から送信された三次元データに基づいて, ローカル環境で遠隔地の対話者の像を視点位置に基づいて再構成する方式を図 5 に示す. 実際に遠隔地の対話者の像をローカルのディスプレイに提示した映像の例を図 6 に示す.



図 4 三次元データの獲得

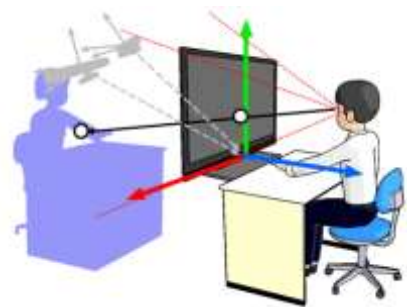


図 5 遠隔地の対話者の像の再構成



図 6 遠隔地の対話者の 3D 合成提示の例

(5)指示の共有について, 3DCG で構成した仮想物体を遠隔地間で情報共有する際に, 物体への指示に関する情報を互いに送受信して提示することが, 本研究課題の目標である遠隔コミュニケーション下での視覚的共同注

意の成立にどの程度影響を与えるかについて評価した。構築したシステムは、遠隔地の対話者の映像に、共有物体としての 3DCG を液晶ペンタブレット上に重畳表示し、物体上の注目点をペンで指示した位置を対話者それぞれのディスプレイ上にマーカー付与によって強調表示するというものである。対話者に表示される映像の例を図 7 に示す。



図 7 3DCG にマーカーを付与した例

被験者 6 名について 2 名 1 組になってもらい、システムを使用して、一方の被験者が液晶ペンタブレット上に表示した 3DCG 物体上の任意の位置をペンで指し示したときに、その点にマーカー付与しないと付与したときの、気づきの速さ、正確に位置が分かったか、の観点で、5 点法 (5 が一番評価が高い) によって検証した。この結果を表 1 に示す。

表 1 指示位置の明確化に関する評価結果

		マーカー付与	
		無し	有り
効率	気づきの速さ	3.7	4.8
質	正確さ	3.3	5.0

この結果から、3DCG で表現された物体上のペンで指示した位置に対して、マーカーを用いてその位置を相手に伝えた方が、遠隔地間コミュニケーションでは位置を把握しやすいことが確認できた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 2 件)

①森村吉貴, 角所考, 西口敏司, 八木啓介, 美濃導彦, "分節性に基づくライブ映像ストーリーミングの pre-roll 制御", システム制御情報学会論文誌, Vol.25, No. 9, pp.211 - 220, 2012-09. (査読有)

②宮川 崇, 西口 敏司, "遠隔コミュニケーションのための仮想現実空間の構築," 画像の認識・理解シンポジウム (MIRU2012) 論文集 DS-11, (2012.08). (査読無)

[学会発表] (計 3 件)

①藤川 貴史, 西口 敏司, "遠隔コミュニケーションにおける視覚的共同注意支援のための仮想三次元環境の構築," 電子情報通信学会信学技報, vol.111, no. 380, MVE2011-59, pp. 31-36, (2012.01.19), 於大阪電気通信大学 (大阪府) .

②古沢 香織, 西口 敏司, "視線推定を用いた遠隔視覚的共同注意の実現," 電子情報通信学会 信学技報, vol. 110, no. 381, PRMU2010-201, pp. 393-397, (2011.01.21), 於立命館大学 (滋賀県) .

③櫻木 秀樹・西口 敏司, "指示方向の推定による遠隔共同注意の実現," 電子情報通信学会 信学技報, vol. 110, no. 381, PRMU2010-202, pp. 399-403, (2011.01.21), 於立命館大学 (滋賀県) .

6. 研究組織

(1)研究代表者

西口 敏司 (NISHIGUCHI SATOSHI)
大阪工業大学・情報科学部・准教授
研究者番号：80362565

(2)連携研究者

豊浦 正広 (TOYOURA MASAHIRO)
山梨大学・医学工学総合研究部・助教
研究者番号：80550780