

t-Roomにおける俯瞰カメラを用いた映像出力壁面選択機構の開発

和田 理¹ 片桐 滋¹ 大崎 美穂¹

概要：部屋を重ね合わせるようにして遠隔コラボレーションの支援を目指す t-Room では、壁面ディスプレイ上における不適切な映像再生を防ぐため、人物などの撮影対象の利用可能域が壁面付近に制約されてきた。この問題を解決するため、部屋天井部に設置した俯瞰カメラを用いて部屋床面における撮影対象の位置情報を抽出し、通信先の部屋において、その位置情報から対象を映像として出力すべき壁面ディスプレイを適切に選択する機構を開発する。本稿では、その機構の概要と開発結果を報告する。

キーワード：マルチメディアシステム, 映像出力壁面選択

Development of an Overhead-Camera-Based Mechanism for Selecting Wall-Type Displays in t-Room

OSAMU WADA¹ SHIGERU KATAGIRI¹ MIHO OHSAKI¹

Abstract: In *t-Room*, which aims at supporting remote collaboration by providing a work environment constructed by virtually overlaying remote rooms, such objects as users are forced to stand near its wall, i.e., monolith, to avoid an inappropriate reproduction of their images in monolith displays. To solve this problem, we develop a new mechanism for automatically selecting a target monolith, to which an object image should be produced. The mechanism analyzes the positional relation between an object and its nearest monolith, using the positional information measured with an overview image. Based on experimental evaluation, we demonstrate that our proposed mechanism correctly selects a target monolith and appropriately outputs an image of the object standing at the central area of t-Room.

Keywords: Multi-media system, Wall-type displays selecting

1. はじめに

通信回線の品質の向上や大型ディスプレイの普及に伴い、高い解像度で等身大サイズの映像を出力する遠隔コラボレーション支援システムの研究が精力的に行われている [1], [2], [3], [4], [5]. 遠隔コラボレーション支援システム「t-Room」 [6], [7] は、従来の多くの遠隔コラボレーション支援システムに見られる、ディスプレイを介して空間を接続する構造と異なり、空間を仮想的に重ね合わせる構造を持つ。この特徴的な構造により、視聴覚メディアの対称性

を確保し、利用者に同室感 (共存在感) を提供することによる共同作業の支援が目指され、また、その特徴が地点間の心理的な壁を緩和する効果を持つことなどが検証されてきた [8].

しかしながら、t-Room にはカメラとディスプレイの位置関係に起因する、部屋空間の中心部に存在する映像オブジェクト (例えば人物) が遠隔地点のディスプレイにおいて実際より大きなサイズに拡大され、かつ複数のディスプレイに重複して出力されるという問題があった。そのため、t-Room における実質的な利用領域は、壁面ディスプレイ付近に制限されていた。コラボレーションを効果的に支援するため、この制約が解決されるべきことは明らかである。

¹ 同志社大学大学院 理工学研究科
Graduate School of Science and Engineering, Doshisha University



図 1 部屋空間中心部に人物がいるときの t-Room の映像出力例。一人の人物が、複数のディスプレイに多重にかつ拡大されて映し出される。

Fig. 1 An example of projected (human) object images when it is located at the center area of t-Room. The object is projected on multiple displays with inappropriate (enlarged) sizes.

本研究は、上記の問題意識に基づいて、t-Room 内の(人物などの)映像オブジェクトを抽出し、その部屋内の位置情報に基づいてオブジェクトを映像再生すべきディスプレイを適切に選択する機構を開発するものである。映像オブジェクトの位置情報の取得のため、t-Room の天井面に俯瞰カメラを追加する。俯瞰カメラの出力画像における壁面部の座標と映像オブジェクトの座標を比較することにより、映像オブジェクトから最も近いモノリスを選択する。本稿では映像出力選択機構の機能について述べ、また動作確認の結果を報告する。

2. t-Room の課題：部屋空間中心部の映像オブジェクトの取り扱い

t-Room は遠隔地点に存在する同一の構成をもつ部屋空間どうしで通信するシステムである。t-Room の部屋空間は、液晶ディスプレイやスピーカなどのマルチメディアデバイスの付いた壁面、すなわちモノリスと、モノリスと対峙する位置に設置されるカメラによって構成されている。モノリス(ディスプレイ)とカメラが対峙する位置関係は、空間の仮想的な重畳を実現する上で重要な要素である。しかし一方で、この対峙関係は、部屋空間中心部に映像オブジェクトが存在する際に、図 1 のような不自然な映像出力を遠隔地のディスプレイに生成する原因となる。

図 1 のディスプレイ部に注目すると、人物が等身大よりも拡大され、かつ複数のディスプレイ上に多重に出力されていることがわかる。従来の t-Room の部屋空間中心部では視覚メディアの対称性が確保されておらず、t-Room の利用可能領域はディスプレイ付近に制限されていたと言わざるを得ない。

t-Room の部屋空間中心部に存在する映像オブジェクトが拡大されて生成される原因を図 2 に、また複数のディスプレイに多重生成される原因を図 3 に図解する。

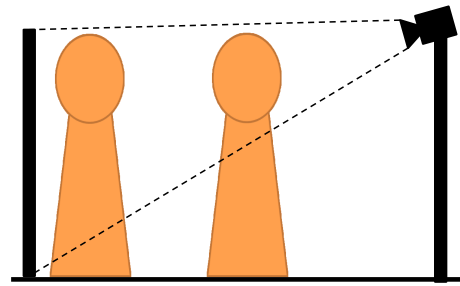


図 2 部屋空間中心部の人物が拡大して生成される原因。

Fig. 2 A reason why a human object is projected with enlarged sizes when it is located at the center area of t-Room.

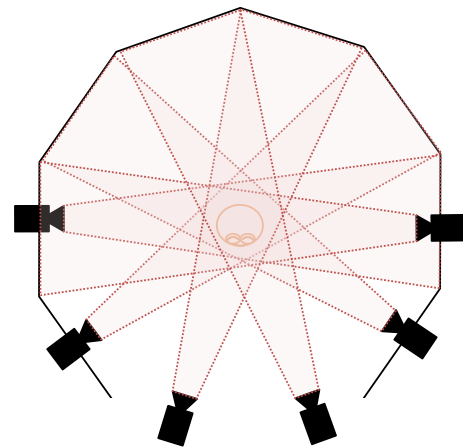


図 3 部屋空間中心部の人物が複数のディスプレイに多重に生成される原因。

Fig. 3 A reason why a human object is projected on multiple displays when it is located at the center area of t-Room.

図 2 において、カメラから、それに対峙するモノリス(ディスプレイ)に向けて、カメラの画角を表す 2 本の点線を記す。即ち、これらの点線に挟まれる領域がカメラの撮影範囲である。図から、ディスプレイ面付近では人物の全身がカメラの撮影範囲に入るものの、部屋空間中心部では人物の上半身のみがカメラの撮影範囲に入ることがわかる。また明らかに、ディスプレイに映像として出力する際には、点線に挟まれた領域が出力映像の上下端となる。このため、部屋空間中心部に人物が存在する場合、ディスプレイにはその上半身のみがディスプレイ映像全体に拡大され、出力される。

図 3 に、部屋を囲むように配置されている各カメラの撮影範囲を図解する。図中、各撮影範囲は、カメラのレンズの先端面を上底とし、カメラと対峙するモノリス(ディスプレイ)面を下底とする台形である。図から、部屋空間中心部は全ての台形の範囲に含まれることがわかる。即ち、部屋空間中心部は t-Room 内の全てのカメラの撮影領域に含まれる。一方、これまで t-Room では、カメラが撮影した映像を、遠隔地 t-Room のディスプレイにそのまま出力していた。このため、部屋空間中心部に存在する映像オブジェクトは遠隔地 t-Room の全ディスプレイに多重に出力

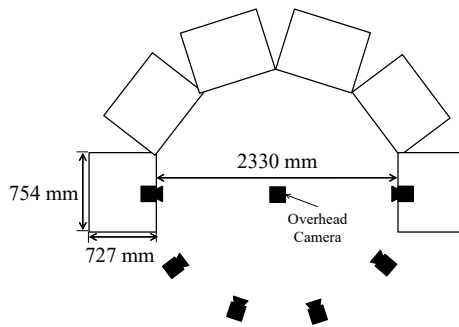


図 4 提案するシステムの平面図.
Fig. 4 Top view of proposed system.

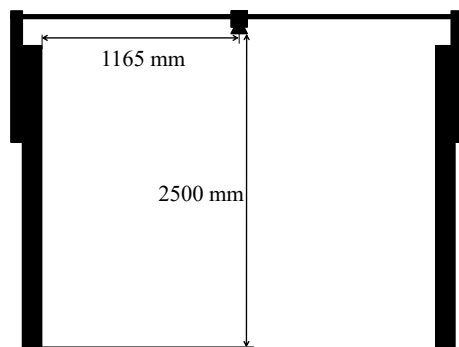


図 5 提案するシステムの側面図.
Fig. 5 Side view of proposed system.

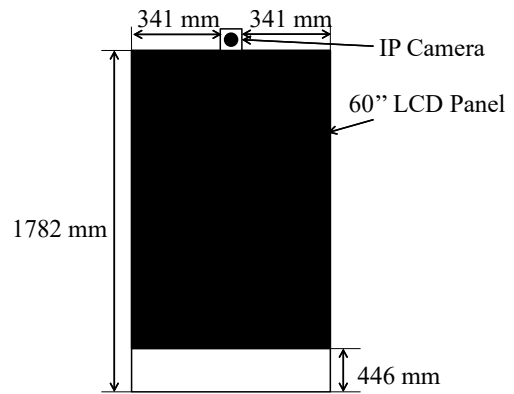


図 6 提案するシステムにおけるモノリス正面図.
Fig. 6 Frontal view of monolith of proposed system.



図 7 システムの設置例。天井部に俯瞰カメラを追加している。
Fig. 7 System implementation example. An overhead camera is added on the ceiling.

されていた。

以上を踏まえると、t-Room の利用可能領域を部屋空間の中心部まで拡大するためには、映像生成における不適切な拡大と多重化との問題を解決する必要があることがわかる。本稿では、そうした問題解決の一環として、特に、多重出力の問題の解決を目指し、映像出力先のディスプレイを適切に選択する機構を開発する。

3. 提案システムの部屋構成

本稿で提案するシステム (以下、本システムと呼ぶ) は、先行研究で開発したシステム [9] を基に部屋内における映像オブジェクトの位置情報を抽出する機能を追加するものである。本システムの平面図を図 4 に、側面図を図 5 に、正面図を図 6 に示す。なお、本稿では映像に関する処理を主に扱うため、以降の本システムの説明については映像メディアに関するものに絞って述べる。

先行研究で開発したシステムと同様に、本システムにおいても 2.3 メートルの直径をもつ正十角形の部屋を使用する。正十角形の壁面のうち、6 面にモノリスを配置し、残りの 4 辺は部屋空間の出入り口として開放する。また、モノリスと対峙する位置にカメラを設置する。また、ディスプレイ映像をカメラが何重にも撮影することを防止するため [6]、カメラのレンズに偏光フィルタを装着する。本稿で提案するシステムではそれらに加え、部屋空間の天井部に俯瞰カメラを配置する。追加した俯瞰カメラにより、部屋

空間内の映像オブジェクトの位置情報を抽出する。本システムでは便宜上、モノリス及びそれらと対応するカメラについて出入り口から見て一番左側に存在するモノリスより時計回りの順に 1 から 6 のモノリス番号を付与する (図 4 参照)。また、側面のカメラは、ディスプレイに対し時計回りに 90° 回転させて設置する。これは、カメラの出力映像が横長であるのに対しディスプレイを図 6 のとおり縦長に設置しているためである。なお、本システムにおいて、側面のカメラと俯瞰カメラのいずれにも IP カメラを利用する。本システムの設置例を図 7 に示す。

4. 提案システムのデータ処理に係る機器およびネットワーク構成

4.1 概要

本稿で提案するシステムにおいて映像データの処理に係る機器は、側面カメラと、俯瞰カメラ、サーバコンピュー

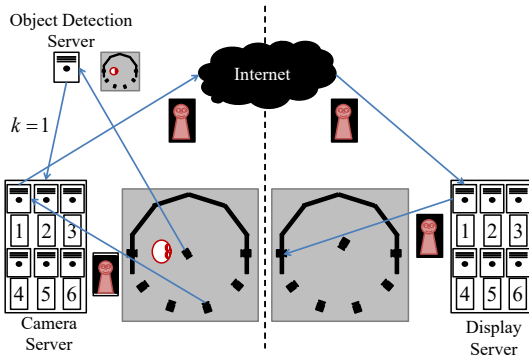


図 8 システム内の映像伝送に関する映像データ及びそれに付随するデータの入出力の流れ (1 番のモニタの近くに利用者がいる場合)。

Fig. 8 Flow diagram of image data and their related control data in proposed system (in the case where a (human) object is located near monolith No. 1).

タ^{*1}から成る。カメラとサーバコンピュータは LAN (Local Area Network) を介して接続される。本システムにおける映像データおよびそれに付随するデータ入出力の流れを図 8 に示す。なお、図中において映像データおよびそれに付随するデータの入出力に関連しないサーバ (通信制御を行うもの等) は省略している。また、図中の k は映像オブジェクトに最近傍のモニタのモニタ番号を指す。以降の説明においては、映像オブジェクトが 1 番のモニタ付近に存在するものとする。

本システムには画像データを処理するサーバとして、オブジェクト抽出サーバ (Object Detection Server)、カメラサーバ (Camera Server)、及びディスプレイサーバ (Display Server) の 3 種類のサーバが存在する。それぞれのサーバの機能について以下に述べる。

4.2 オブジェクト抽出サーバ

オブジェクト抽出サーバは、映像出力先の選択機構を実現するために必要な映像オブジェクトに最近傍のモニタの情報の抽出を実施するサーバである。各地点に 1 台ずつ存在する。オブジェクト抽出サーバの主な機能は以下の通りである。

- (1) 俯瞰カメラからの画像の受信
- (2) 受信した画像からの映像オブジェクトの俯瞰カメラ映像上における座標情報の抽出
- (3) 抽出した映像オブジェクトの座標情報と俯瞰カメラ映像上におけるモニタの座標情報との比較
- (4) 映像オブジェクトに最近傍のモニタのモニタ番号の取得
- (5) 取得したモニタ番号の自地点のカメラサーバへの伝送

*1 本稿において、サーバコンピュータとは汎用パーソナルコンピュータのことを指す。

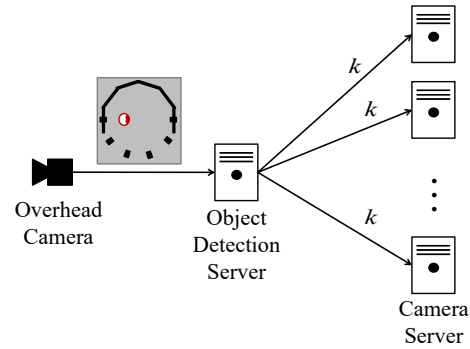


図 9 オブジェクト抽出サーバにおけるデータ入出力。
 Fig. 9 Data flow related to Object Detection Server.

オブジェクト抽出サーバにおける映像データ、及びそれに付随するデータの入出力の関係を図 9 に示す。

初めに、オブジェクト抽出サーバは俯瞰カメラより画像を取得する。

続いて、取得した画像を前景とし、事前に取得した部屋内に映像オブジェクトが存在しない状態の俯瞰カメラの出力画像を背景として、背景差分法を用いる撮影対象オブジェクトの抽出を行う。そして、背景差分法によって切り出された画像に対して 2 値化処理を適用する。これにより、床面 (映像オブジェクトの存在しない) の領域と撮影対象オブジェクトの領域を分離する。分離した映像オブジェクトの領域に対し、ラベリング処理を行い、映像オブジェクトを抽出する。

抽出したラベル領域の中心点を映像オブジェクトの座標位置として、映像オブジェクトに最近傍のモニタを選択する。そして、こうして選択されたモニタのモニタ番号を自地点のカメラサーバへと伝送する。

4.3 映像オブジェクトに最近傍のモニタの探索手法

映像オブジェクトに最近傍のモニタのモニタ番号を k とおくと、 k は式 (1) により求められる。

$$k = \arg \min_i d(\mathbf{m}_i, \mathbf{p}) \quad (1)$$

ここで \mathbf{p} は抽出した映像オブジェクトの中心点を、 \mathbf{m}_i は各モニタの中心点を、 $d(\mathbf{m}_i, \mathbf{p})$ は \mathbf{p} と \mathbf{m}_i とのユークリッド距離 $|\mathbf{m}_i - \mathbf{p}|$ を、 i はモニタ番号を指す。なお、各モニタの中心座標 \mathbf{m}_i について、事前にディスプレイの下部に映像を出力し、映像が出力された領域の中心の座標を \mathbf{m}_i として設定する。映像オブジェクトの中心座標 \mathbf{p} 及びモニタの中心座標 \mathbf{m}_i の位置関係を図 10 に示す。

4.4 カメラサーバ

カメラサーバは、側面の IP カメラより受信した画像の処理、及び映像出力選択を実施するサーバである。カメラサーバはモニタごとに 1 台ずつ存在する。カメラサーバは主に以下の機能をもつ。

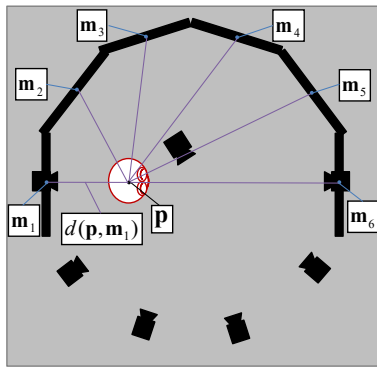


図 10 映像オブジェクトの中心点とモノリスの中心点との俯瞰カメラ映像中における位置関係。

Fig. 10 Positional relationship between the center point of an image object and the center points of monoliths in the coordinate of an overhead camera image.

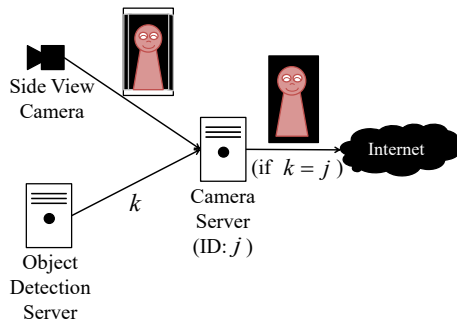


図 11 カメラサーバにおけるデータ入出力。

Fig. 11 Data flow related to Camera Server.

- (1) 側面の IP カメラからの画像の受信
- (2) 受信した画像に対する画像処理
- (3) オブジェクト抽出サーバからの映像オブジェクトに最近傍のモノリスのモノリス番号の受信
- (4) 受信したモノリス番号とカメラサーバに対応するモノリス番号との比較、およびそれに伴う伝送の実施/非実施の判定
- (5) 処理した画像の遠隔地点のディスプレイサーバへの伝送

カメラサーバにおける映像データ、及びそれに付随するデータの入出力の関係を図 11 に示す。

初めにカメラサーバに対応する側面の IP カメラより画像を受信する。同時にオブジェクト抽出サーバより映像オブジェクトに最近傍のモノリスのモノリス番号 k を受信する。

次に受信した画像を反時計回りに 90° 回転させる。

続いて、回転した画像に対する射影変換を適用する。

そして、映像オブジェクトに最も近いモノリスの番号 k とカメラサーバに対応するモノリスの番号 j を比較する。 $k = j$ の場合のみ、射影変換を適用した画像を遠隔地点の対応するモノリスのディスプレイサーバへ伝送する。この伝送条件を付与することによって、遠隔地点のディスプレ

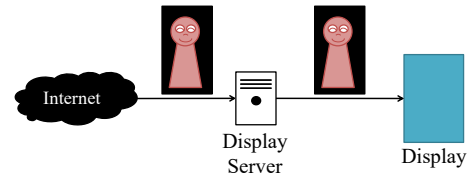


図 12 ディスプレイサーバにおけるデータ入出力。

Fig. 12 Data flow related to Display Server.

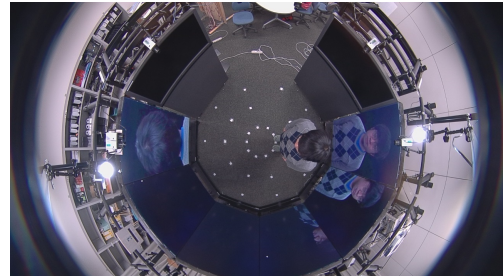


図 13 映像出力先を選択する機構を用いない場合の映像出力例。

Fig. 13 An image example produced without the proposed mechanism of selecting a target monolith (its corresponding life-size display).

イサーバに対する映像伝送が 1 面のモノリスに限定されて実施される。したがって、遠隔地点におけるディスプレイ出力が 1 面のみ限定される。これにより、部屋空間の中心部に映像オブジェクトが存在する際に単一の映像オブジェクトが複数のディスプレイに多重に出力されるという問題を解決する。

4.5 ディスプレイサーバ

ディスプレイサーバは、遠隔地点のカメラサーバから受信した画像のディスプレイへの出力を実施するサーバである。ディスプレイサーバはモノリスごとに 1 台ずつ存在する。なお、本稿の実験環境においては、カメラサーバとディスプレイサーバは同一のサーバコンピュータ上で動作させている。ディスプレイサーバにおける映像データの入出力の関係を図 12 に示す。

5. 提案システムの動作の確認

便利のため、カメラサーバの出力映像を同一地点のディスプレイサーバに伝送する設定で、本提案システムの動作の確認を行った。

図 13 に、本システムに依る出力先ディスプレイの選択を伴わない、従来型の出力の様子を俯瞰カメラで撮影した結果を示す。図では、撮影対象である人物が、部屋空間中心部からわずかに 1 番のモノリスに近い位置に立っている。そして、撮影・伝送された人物映像が、複数のディスプレイに拡大されて生成されていることを見てとることができる。

一方、図 14 に、提案した映像出力先を選択する機構を適用した場合の出力映像例を示す。人物映像が 1 番のモノ

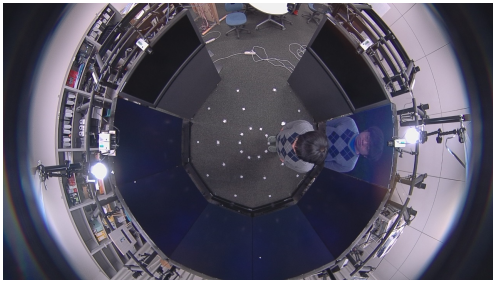


図 14 映像出力先を選択する機構を用いた場合の映像出力例。

Fig. 14 An image example produced with the mechanism of selecting a target monolith (its corresponding life-size display).

リスのディスプレイ上にのみに出力されていることがわかる。即ち、t-Room の部屋空間中心部の映像オブジェクト出力の問題の一つ、単一の映像オブジェクトが複数の映像出力として生成される問題が解消されていることがわかる。

6. おわりに

遠隔コラボレーション支援システム t-Room における利用可能領域を拡張することを目指し、t-Room 部屋内の中心付近に位置する撮影対象オブジェクトが複数のディスプレイに映像再生されてしまう問題を解決するため、映像出力先ディスプレイを選択する機構を開発した。本提案法は、新たに導入した俯瞰カメラの出力映像を用いて部屋内の映像オブジェクトの位置情報を抽出し、抽出した位置情報に基づく映像オブジェクトに最近傍のモノリスを選択、その選択されたディスプレイのみに映像を生成させるものであった。評価実験の結果、提案手法によって、目標とした単一映像オブジェクトの複数ディスプレイへの多重出力の問題を解決できることを確認した。

現在のところ、出力選択機構によって指定されたディスプレイに、カメラサーバの出力画像を直接出力している。このため、部屋空間中心部の映像オブジェクトが拡大して出力される問題は解決されていない。また、現在の仕様では、モノリスの境界部付近に存在する映像オブジェクトやディスプレイを跨ぐ大きさを持つ映像オブジェクトについて、その全体を出力することができない。これらの問題について、出力サイズの正規化や隣接するモノリスのカメラ映像の利用などにより解決していくことが今後の課題となる。

参考文献

- [1] Okada, K., Maeda, F., Ichikawa, Y. and Matsushita, Y.: Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design, *Proc. CSCW '94*, ACM, pp. 385–393 (1994).
- [2] Morikawa, O. and Maesako, T.: HyperMirror: Toward Pleasant-to-use Video Mediated Communication System, *Proc. CSCW '98*, ACM, pp. 149–158 (1998).
- [3] Harrison, S.: *Media Space 20+ Years of Mediated Life*,

Springer Publishing Company, Inc. (2009).

- [4] Dou, M., Shi, Y., Frahm, J., Fuchs, H., Mauchly, B. and Marathe, M.: Room-sized informal telepresence system, *Virtual Reality Short Papers and Posters (VRW)*, IEEE, pp. 15–18 (2012).
- [5] Noda, S., Ebara, Y., Ishida, T., Hashimoto, K. and Shibata, Y.: Implementation of High Presence Video Communication System for Multiple Users Using Tiled Display Environment, *2015 IEEE 29th International Conference on Advanced Information Networking and Applications Workshops*, pp. 494–499 (2015).
- [6] Hirata, K., Harada, Y., Takada, T., Aoyagi, S., Shirai, Y., Yamashita, N. and Yamato, J.: The t-Room - Toward the Future Phone, *NTT Technical Review*, Vol. 4, No. 12, pp. 26–33 (2006).
- [7] Hirata, K., Harada, Y., Takada, T., Aoyagi, S., Shirai, Y., Yamashita, N., Kaji, K., Yamato, J. and Nakazawa, K.: t-Room: Next Generation Video Communication System, *Proc. IEEE GLOBECOM 2008*, IEEE, pp. 1–4 (2008).
- [8] Yamashita, N., Hirata, K., Takada, T., Harada, Y., Shirai, Y. and Aoyagi, S.: Effects of Room-sized Sharing on Remote Collaboration on Physical Tasks, *情報処理学会論文誌*, Vol. 48, No. 12, pp. 788–799 (2007).
- [9] 和田 理, 片桐 滋, 大崎美穂: 空間重畳型遠隔コラボレーション支援システムにおける利用可能領域の拡大, *情報処理学会研究報告*, Vol. IPSJ-GN-96, No. 22 (2015).