

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

和田 理¹ 片桐 滋¹ 大崎 美穂¹

概要: 多くのテレビ会議システムに見られるように部屋を仮想的につなぐのではなく、部屋を仮想的に重ね合わせることで同室感（メディアの対称性）の向上を目指す *t-Room* の考え方を引き継ぎつつ、そこでは実現が困難であった部屋空間全体の利用を可能にするための新しい遠隔コラボレーション支援システムを提案する。本稿では、特に、この開発の基盤となる実験システムの設計と実装例、そして人物等の映像オブジェクトを適切に抽出、伝送、再生するための画像処理技術を紹介する。多数のカメラに囲まれる本システムの部屋空間においては、部屋の中心部に位置するオブジェクトは複数のカメラに同時に撮影され、その単純な再生は同室感を著しく損なってしまう。本稿では、こうしたオブジェクト映像にかかわる問題の解決法にも焦点を合わせる。

キーワード: マルチメディアシステム, 遠隔コラボレーション支援

Usable Area Widening for Room Overlay Type Remote Collaboration Support System

OSAMU WADA¹ SHIGERU KATAGIRI¹ MIHO OHSAKI¹

Abstract: In this paper, we propose a novel remote collaboration support system that basically has the same design concept as *t-Room*, i.e., virtually overlaying one room onto the other distant rooms instead of connecting them. Inheriting the advances of *t-Room*, our new system additionally aims to solve its problem that a usable space is limited to the narrow area just in front of wall-type displays. In the paper, we introduce the design policy and implementation example of our proposed system, and also detail the system configuration and procedures for processing such visual image objects as persons. Then, we specially discuss a way for transforming an object that moves around in the system's room to an image object that naturally moves in the wall displays.

Keywords: Multi-media system, Remote collaboration support

1. はじめに

大型ディスプレイの普及や通信回線の品質向上により、等身大映像を用いた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]。実際、遠隔地の人物などの等身大表示は、遠隔地点間のコラボレーション参加者どうしの心理的な壁（「こちら側」と「向こう側」）

を低減する [4]。しかし、そうした遠隔コラボレーション支援システムの多くはディスプレイを境界として遠隔地の部屋等をつなぐ空間構成になっており、ディスプレイ面は地点間を隔てる物理的（ディスプレイを越えた人の動きを制約する）かつ心理的な壁として残り続けている。

こうした問題の解決を目指し、空間をつなぐことに代え、空間を仮想的に重ね合わせることで視聴覚メディアの対称性を確保し、「同室間」（共存在感）の向上を目指す遠隔コラボレーション支援システム「*t-Room*」 [5], [6] が提案された。*t-Room* では、部屋空間は大型のディスプレイで囲ま

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

れ、基本的に、そのディスプレイで囲まれた部屋空間全体が共有される。部屋を隔てるディスプレイを取り去り、「こちら側」と「あちら側」の利用者が部屋の壁面内において行き交うことを可能とした t-Room は、部屋間の物理的あるいは心理的境界の一定程度の低減に貢献した [7]。しかしながら、カメラとディスプレイ間における比較的単純な映像伝送の仕組みを採用したため、t-Room における実際の利用空間は壁面（ディスプレイ面）付近に制約されている。

部屋をつなぐのではなく部屋を重ねるという t-Room の特徴を考えると、その特徴を活かし、それが持つ利用空間の狭さを改善する意義は大きい。こうした観点に立ち、本稿では、共有空間が壁付近に制約される t-Room の欠点を解消する新たな遠隔コラボレーション支援システムを提案する。t-Room 部屋空間の中心部を共有空間として使用するためには、その中心部にある被写体（利用者等の映像オブジェクト）を遠隔地点のディスプレイ上に自然な形で投影する必要がある。そこでは、2次元の広がりがある部屋内のオブジェクトを1次元の壁面ディスプレイ上に、コラボレーションに必要な情報等を失うことなく表示する、困難な処理が求められる。本稿は特に、そうした新システムの構想と、設計、基盤部の実装、そして部屋中心部のオブジェクトを壁面に投影する画像処理の考え方について詳述する。

2. 関連研究

本稿で扱うシステムは、その基本を t-Room に置いているため、元々の t-Room [5], [6] に関する説明は後節に譲ることとする。本節では、それ以外の関連技術について振り返っておく。

等身大映像を用いる最も初期のシステムの例として CAVE [8] を挙げることができる。しかし CAVE は没入型の仮想現実環境を実現するためのものであり、遠隔コラボレーションを対象としていない。また、実時間での映像処理に対応していない。CAVE と類似の環境で実時間での映像の取得と表示に対応したものとして blue-c [9] を挙げることができる。しかし、blue-c で遠隔コラボレーションを行うことを考えると、空間を接続する構成となり、ディスプレイ面が地点間の境界になる問題が残る。

他に、遠隔コラボレーション支援を目的としたシステムとして、超鏡 (Hypermirror) [10] や WSCS (Waseda Shadow Communication System) [11] を挙げることができる。しかし、これらのシステムは遠隔地点で出力される内容に制約が存在する。超鏡はディスプレイが一面のみであり、カメラが固定されているために向きを変更することが困難である。また WSCS は、遠隔地点の情報に影で表現されるために、コラボレーションに重要な情報となる表情をうかがうことが困難である。

t-Room の研究に関連して、元々 Microsoft 社の Win-

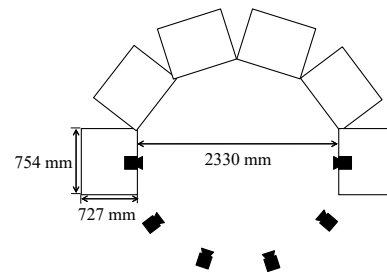


図 1 提案するシステムの空間（部屋）構成及びモノリスの配置例（平面図）。白抜き四角がモノリスであり、部屋側の面にディスプレイが設置されている。黒色の装置はカメラを表している。

Fig. 1 Implementation example of proposed system (top view). A white box indicates a monolith, and a black box indicates a camera.

dows 用に作られていた t-Room 制御用ソフトを、新たにオープンソースの Linux 用に変更する開発が行われている [12], [13], [14], [15]。それは、Windows がプロプライエタリ・ソフトウェアであるために、OS のカーネルに係る変更を行うことができず、Windows のカーネルに起因するメディア処理の遅延を軽減できないためである。音声伝送処理における遅延を従来の数分の 1 程度に抑えることができるなど [16], [17], Linux 上で音声・映像メディア処理用のソフトウェアを開発することの基本的な効果が示されている。

3. システムのハードウェア構成

基本的には従来と同様に [5], [6]、本システムの t-Room は、モノリス (Monolith) と呼ばれる液晶ディスプレイなどのマルチメディア機器から成る壁によって部屋を構成する。図 1 は、その実装例を図解している。部屋の直径は 2.3 メートルであり、モノリスは正十角形の各辺に対応する。なお、この実装例は 6 面のモノリスを持ち、正十角形の残りの 4 辺は部屋の入り口として開放されている。ここでも従来と同様に、カメラはディスプレイに対峙するように配置される。なおカメラは、撮影映像が横長であるのに対し、ディスプレイが縦長であることを考慮し、時計回りに 90° 回転するように設置する。また図 2 には、モノリスの構成を図解している。モノリスの壁面は 60 インチの縦長液晶ディスプレイで構成され、その上部にカメラを設定している。

カメラは、それに対峙するモノリスの周辺を撮影し、カメラで撮影された映像は、遠隔地の対応する（部屋を重ねた際と同じ位置にくる）モノリスのディスプレイに表示される。このカメラとディスプレイの位置関係が、空間を重畳するという t-Room の特徴を実現する基本的な要素となっている。しかしこの構成により、カメラは、人物や物体などのオブジェクトと共にその背後にあるディスプレイも撮影することになる。その結果、人物が何重にも表示される映像エコーが発生する。これは明らかにコラボレーション

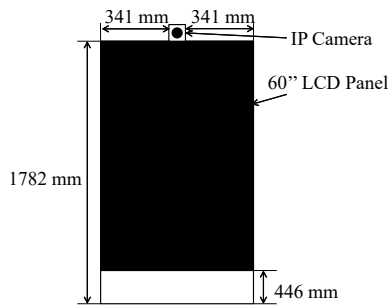


図 2 モノリスの構成 (正面図).

Fig. 2 Configuration of Monolith (front view).

を阻害するものであり、映像エコーの除去が求められる。

本システムの現在の実装例では、従来の t-Room と同様に、偏光フィルターを用いて液晶ディスプレイの光を遮ることにより映像エコーを除去する。これは、液晶ディスプレイの内部に偏光板が存在することを利用するものであり、極めて簡単なエコー除去法である。しかしながら、その光の通過を制約する性質によって、カメラに入射する光の量が減少し、撮影映像が暗くなってしまふ。この映像が暗くなる問題を防ぐためには、カメラの感度を上げるか、もしくは露光時間を長くすることが求められる。しかしながら、カメラの感度を上げると映像の S/N 比が小さくなり、その結果映像の品質が低下し、また露光時間を長くすると被写体ぶれが発生し、ここでもまた映像の品質が低下する。このため、現在、背景差分法や深度マップを用いた前景部の抽出法などの、画像処理技術を活用した映像エコー除去法 [18], [19] も並行して開発を進めている。

本来、t-Room では、映像と音声の両メディア信号の取得・伝送・表示が行われなければならない。しかし図 1 に示すように、現在の実装例では、データ量が大きく、その処理や伝送用ネットワークの構成等に工夫が求められる映像メディアの処理に焦点を合わせている。音声の取得と再生に関しては、現在は t-Room 毎に単一マイクと単一スピーカを配置する単純な方法をとっているが、将来的には並行して進めている音響反射板方式による高臨場感音場再生法 [20], [21] などを導入する予定である。

なお、この t-Room どうしの接続は、それぞれの接続地点に同じ形状の t-Room が設置されていることを前提として行われる。

4. システムのネットワーク構成

図 3 に、映像信号の処理を担うサーバコンピュータを含む、本システムの 1 地点分のネットワーク構成を示す。なお、前節と同様の理由で、ここでも映像信号の処理・伝送に焦点を合わせている。また、通信相手となる遠隔地点の t-Room も同様のネットワーク構成を持つものとする。

図中のネットワークは、6 台の IP カメラと 6 台のディス

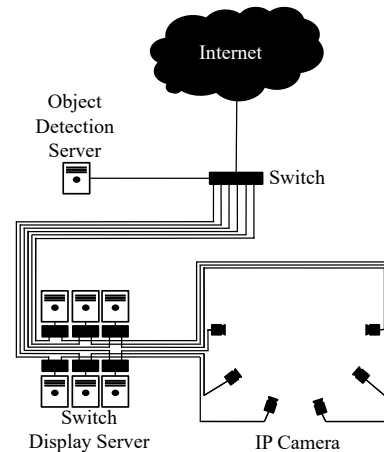


図 3 本システムの映像伝送部のネットワーク構成。

Fig. 3 Network configuration of visual data transmission in proposed system.

プレイサーバ*1 (Display Server)、1 台のオブジェクト抽出サーバ (Object-detection Server) から構成される*2。なお、ディスプレイサーバは、画像処理サーバ (Image-processing Server) も兼ねている。ディスプレイサーバとオブジェクト抽出サーバ、画像処理サーバのそれぞれの詳細な役割については後述する。

本システムでは、高性能カメラがしばしば持つ OS あるいはデバイスドライバの制約を避けるために、カメラデバイスとして IP カメラを用いている。IP カメラは、直接、対応する t-Room の LAN に接続され、IP カメラから出力された映像は、遠隔地に伝送される前の画像処理等に供される。なお後述する画像処理を伴うものの、基本的には 1 台の IP カメラで撮影された映像は対応する 1 台のディスプレイに表示される、1 対 1 の関係にある。

従来の t-Room のディスプレイサーバ及びカメラサーバのソフトウェアは、Microsoft 社の OS である Windows に対応していた。それに対し本システムのディスプレイサーバ等のソフトウェアは、Linux 用に新たに開発した t-Room 用制御ソフトウェアの例に倣い [14]、Linux 用に開発した。開発したディスプレイサーバとオブジェクト抽出サーバに用いているコンピュータの仕様を表 1 に示す。なお、ディスプレイサーバとオブジェクト抽出サーバは同一の仕様のコンピュータを使用している。

5. カメラデバイス: IP カメラ

本システムではカメラデバイスとして IP カメラを用いている。これはカメラデバイスのドライバに起因する OS の制約を回避するためである。また、撮影時間や色温度といった撮影条件の変更が可能な機種を採用することにより、映像品質の向上を図った。IP カメラの仕様及び設

*1 本実装におけるサーバは、汎用パーソナルコンピュータである。
*2 IP カメラとディスプレイサーバの台数はモノリスの面数と等しく設定される。

表 1 ディスプレイサーバおよびオブジェクト抽出サーバ用コンピュータの仕様.

Table 1 Computer specifications of display and object-detection servers.

Machine	Epson Endeavor MR7300
OS	CentOS 6.6
Linux Kernel	2.6.32.504.16.2.el6.x86_64
CPU	Intel® Core® i7-4790 3.6GHz
RAM	8.0GB
GPU	NVIDIA® GeForce® GTX 750

表 2 IP Camera の仕様.

Table 2 Specifications of IP cameras.

IP Camera	SONY IPELA SNC-EB630
Lens	Spacecom S4.8 1.8
Coding format	Motion JPEG
Image resolution (pixels)	1280 × 720
Frames per second	30
Exposure time (seconds)	1/4
Color balance	Manual adjustment

定した撮影条件を表 2 に示す。撮影条件のうち、画像解像度 (Image resolution) とフレームレート (Frames per second) については伝送遅延を回避できるように、明るさ (Exposure time) については白とびの発生を防止するように、ホワイトバランスについては使用する室内環境に対応するように予備実験を通して設定した。なお、画像解像度について表 2 には 1280 × 720 と記載しているが、本システム上では 720 × 1280 の解像度の画像として扱われる。これは、ハードウェア構成の節で述べた通り、カメラ映像とディスプレイの縦横比を合わせるためにカメラ映像を 90° 回転するようにカメラを設置しているためである。

6. 映像伝送サーバ

本システムで映像処理に関わるサーバは、ディスプレイサーバとオブジェクト抽出サーバ、画像処理サーバの 3 種類である。なお、前述の通り、ディスプレイサーバと画像処理サーバは同一のコンピュータ上で稼働するようにしている。図 4 に、これらの 3 種類のサーバの配置と各サーバ間で伝送される情報を示す。

従来の t-Room において、映像伝送を担っていたサーバはカメラサーバとディスプレイサーバの 2 種類であった。そのカメラサーバの役割は本システムでは IP カメラに置き換えられており、従来の t-Room におけるディスプレイサーバの役割は、本システムでは前述の 3 種類のサーバが担っていることになる。従来の t-Room では、カメラで撮影された映像 (カメラ映像) は対応するディスプレイサーバに送信され、そのディスプレイサーバによってディスプレイ表示が行われていた。ここでは、人物や物体などのオブジェクトの認識や抽出などの画像処理は必要とされず、

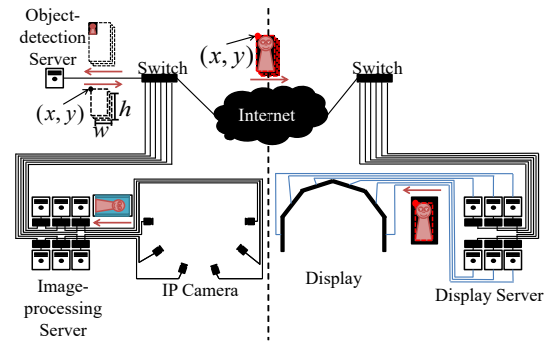


図 4 映像処理に関わるサーバの配置と各サーバ間の伝送データ.
Fig. 4 Image data-related server settings and data transmitted between these servers.

ディスプレイサーバ以外の映像処理用のサーバも必要とされなかった。

これに対し本システムでは、地点全体 (本システムにおいてはディスプレイ 6 面分) のカメラで撮影された画像全体を対象とした画像処理が要求される。それは、本システムでは、壁面付近だけでなく、t-Room の部屋空間の中心部も共有空間として利用することを目指すためである。空間の中心部にオブジェクトが存在する場合、そのオブジェクトは複数のカメラで撮影される。その結果、なんらかの画像処理を施さない限り、撮影されたオブジェクトの映像は、複数のディスプレイに、かつ (一般に) 不適切に大きな映像ととして表示されてしまう。部屋中心付近にあるオブジェクトを壁面の適切なディスプレイ上へ写像する、即ち 2 次元の部屋の広がりをもつ 1 次元のディスプレイ空間に適切に写像してオブジェクト映像を再生するため、本システムでは、全てのカメラ映像を対象としてオブジェクト映像の検出とその加工を行う。

上記の画像加工や写像の必要性がある一方で、大容量の複数画像データの伝送などの処理が限られた台数のサーバ (例えば 1 台のオブジェクト抽出サーバ) に集中することは避けたい。過度の集中は大きな伝送遅延などの困難を引き起こす。従って、画像データの伝送は、基本的には 1 台の IP カメラと 1 台のディスプレイサーバとの間で行うことが望ましい。こうした要件を同時に満たすことをめざし、システムでは図 4 に示すようなディスプレイサーバとオブジェクト抽出サーバ、画像処理サーバの配置を考案した。各サーバの役割については後述する。

6.1 仮想共有面

オブジェクトの動きは複数のカメラでとらえられ、その映像再生は隣接ディスプレイ間の境界を越えて複数のディスプレイ上で行われる。従って、そうしたオブジェクト映像の抽出も、全カメラで撮影された大きな映像空間で行う必要がある。本システムでは、この要請に応えるため、全ディスプレイによって構成される映像空間全体に対応する

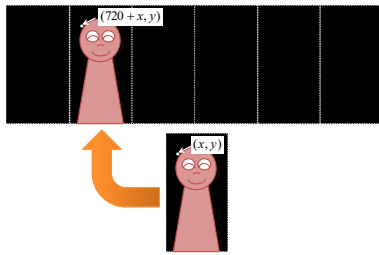


図 5 仮想共有面の考え方 (6 枚のモニリスの場合).

Fig. 5 Concept of virtual plane (for 6-Monolith case).

座標系, 仮想共有面を導入する.

仮想共有面は元々, 従来の (Windows 用) t-Room[22] においても採用されていた. しかしながら, その最新版の (Windows 用) t-Room においては事実上使われていなかった. これは, 前述の通り, カメラ映像が直接遠隔地点のディスプレイに出力され, 特に, 複数のカメラ映像あるいはディスプレイ映像を跨ぐ処理が必要とされなかったためと考えられる.

図 5 に, 6 枚のモニリスから成る本システムにおける仮想共有面の考え方を図解する. 6 枚のカメラ映像を実際のモニリスの配置に対応するように水平方向に繋げ, 単一の座標系を設定する. 水平方向 (x 軸) の座標は, 1 台ずつのカメラ映像のそれとは異なり 6 台分全体の画像空間に設定する. 一方, 垂直方向 (y 軸) の座標は元のカメラ映像のそれに等しい. 図 5 において, 人物が 2 枚目のディスプレイの前に立っている. ここで, 人物が立つ映像におけるある点 (x, y) を考える. このとき, 720 ピクセルのカメラ映像幅に基づき, この (x, y) の仮想共有面における座標は $(720 + x, y)$ となる.

6.2 画像処理サーバ

画像処理サーバは, 1) IP カメラからの画像の受信と, 2) 受信した画像に対する画像処理, 3) オブジェクト領域画像の遠隔地点のディスプレイサーバへの伝送とを担当する. なお, オブジェクト領域とは, 人物や物体などのオブジェクトが含まれる矩形の領域であり, 後述するオブジェクト抽出サーバで検出されるものである. 1 台の画像処理サーバが 1 台の IP カメラの映像の処理を担うため, モニリスと同じ台数の画像処理サーバを準備する.

初めに, 対応する IP カメラから撮影画像を受信する. 次に, 受信した画像を反時計回りに 90° 回転させる. 回転した画像を, 射影変換を用いてディスプレイ領域に対応させる. カメラレンズの焦点距離の関係上, IP カメラはディスプレイ領域よりも広い範囲を撮影する. これにより, 部屋の中心部に存在するオブジェクトについても全体を映すことが可能になるため, 部屋の中心部のオブジェクトをディスプレイ上に表示するには有用であるが, その一方でディスプレイ面付近に存在する映像オブジェクトが遠隔地

点のディスプレイにおいて実際の大きさより小さく表示される. またカメラはディスプレイ上端付近に設置されているため, オブジェクトは斜め上部から撮影されることになる. 射影変換は, こうした画角などの補正を行い, 実際の映像オブジェクトの位置や大きさに対応するディスプレイ表示を実現するよう適用される. 射影変換に用いる位置変換テーブルは, 実装したカメラとディスプレイの環境において予め作成しておく. 位置変換テーブルには射影変換前 (カメラ画像) の座標と射影変換後 (ディスプレイ面) の座標の対応関係が保存されている.

次に, 射影変換を適用した画像を $1/4$ の大きさに縮小し, オブジェクト抽出サーバに送信する. この画像圧縮を行う理由は, オブジェクト抽出サーバの節で説明する.

圧縮画像をオブジェクト抽出サーバに送ったのち, 画像サーバはさらに, オブジェクト抽出サーバから戻されるオブジェクト領域の情報を用いて, 圧縮前の画像からオブジェクト領域画像を抽出する.

そして, 抽出したオブジェクト領域の画像及びオブジェクト領域の始点座標の値を, 遠隔地点の対応するディスプレイサーバに向けて WAN あるいは LAN を介して伝送する. オブジェクト抽出サーバが画像処理サーバにオブジェクト領域の情報を送らない場合 (その画像処理サーバに該当するカメラ画像に抽出すべきオブジェクトが含まれていない場合) は, 画像処理サーバはオブジェクト領域内の画像の抽出及びその領域画像の伝送を行わない. この仕組みにより, 伝送する画像データを削減することができる. また本システムにおいては, 部屋空間の中心部付近に存在する映像オブジェクトを表示する際にカメラ画像そのものではなくカメラ画像に対し画像処理を施した画像を伝送することが考えられる. この点からも, 従来の t-Room のようにカメラ映像全体を遠隔地点に送信するのではなく, オブジェクト領域の映像及びオブジェクト領域の始点座標を伝送することにより, 画像処理を適用した後の映像オブジェクトの遠隔地点への伝送と表示を行う意義は大きい.

6.3 オブジェクト抽出サーバ

オブジェクト抽出サーバは, 自地点の全ての IP カメラで撮影された映像 (画像処理サーバで射影変換を行ったもの) を対象とする映像オブジェクトの抽出, および映像オブジェクトが抽出された際の対応するモニリスに付属する画像処理サーバへのオブジェクト領域の通知を担当する. このため, 1 地点あたり 1 台のオブジェクト抽出サーバを設置する.

複数のモニリスに跨って存在する映像オブジェクトを抽出するため, カメラ単位ではなく, t-Room 内の全撮影画面を対象としてオブジェクト抽出を実行する必要がある. このためにオブジェクト抽出サーバは, 自地点の全ての画像処理サーバ (本実装例の場合 6 台) より画像を受信する

必要がある。しかしながら、それぞれ 720×1280 ピクセルからなる 6 枚の画像を 1 台のサーバコンピュータが毎秒 30 枚ずつ受信することは、明らかに通信遅延の問題を引き起こす。この問題を回避するため、オブジェクト抽出サーバで受信する画像は、予め画像処理サーバにおいて $1/4$ の大きさに縮小されてから送信される。

複数のモノリスに跨ったオブジェクト抽出を実行するため、オブジェクト抽出サーバは、受信したすべての画像処理サーバの画像をモノリスの位置関係に従って 1 枚の画像に統合する。このことから、統合画像は圧縮された小さな仮想共有面に対応するといえる。受信した映像は送信元の画像処理サーバ毎に独立して格納されているため、画像処理サーバの属するモノリスの位置関係に従って 1 枚の画像に統合することは容易に実現できる。そして、統合した画像に対して映像オブジェクト抽出を実行する^{*3}。

サーバは続いて、抽出された映像オブジェクトに基づき、映像オブジェクトを含む矩形の領域を求め、その矩形領域の始点の座標と幅、高さの数値データを画像処理サーバに送り戻す。なお、圧縮画像が統合された仮想共有面から求められるオブジェクト領域の座標と幅、高さの値は、画像処理サーバのオブジェクト領域抽出の対象となる画像におけるオブジェクト領域の値とは異なる。そこで、本オブジェクト抽出サーバは、画像処理サーバがそのまま処理を進められるよう、以下のように変換した値を送信する。なお、括弧内の 4 つの数値は左から順に始点の x 座標値と y 座標値、矩形領域の幅、高さを指す。

$$(4x \bmod W, 4y, 4w, 4h), \quad (1)$$

式 1 において、 x は検出されたオブジェクト領域の始点の x 座標値、 y は始点の y 座標値、 w は矩形の幅、 h は矩形の高さ、 W はオブジェクト領域抽出の対象となる画像の幅（本実装の場合 720）を指す。オブジェクト抽出サーバはオブジェクト領域抽出の対象となる画像を $1/4$ の大きさに縮小されたものを画像処理サーバより受信しているため、矩形領域の始点の座標値と幅、高さの値についてそれぞれ 4 倍することによってオブジェクト領域抽出の対象となる画像に対応させている。また、抽出の対象となった画像はモノリスの配置に従って画像処理サーバから受信した画像を並べたものであるため、オブジェクト領域抽出の対象となる画像における x 座標値は検出結果の x 座標値を 4 倍したものをオブジェクト領域抽出の対象となる画像の幅で除したときの剰余で求めることができる。

なお、映像オブジェクトの抽出結果がモノリスの境界を越える場合がある。このときは抽出した矩形領域をモノリスの境界を境として複数の矩形領域に分割し、それぞれ対応する画像処理サーバに通知する必要がある。このときに

^{*3} 但し、オブジェクト抽出の機能については現在未実装であり、オブジェクト抽出の実装は今後の課題となる。

それぞれの画像処理サーバに通知される値は以下のようになる。なお、 i は抽出されたオブジェクトが属する（複数の）モノリスを指し、1 は最初（仮想平面において左端）のモノリス、 n は最後（仮想平面において右端）のモノリスを指す。その他の記号及び括弧内の数値の意味は式 (1) に準じる。

$$\begin{cases} (4x \bmod W, 4y, W - 4x, 4h) & (i = 1), \\ (0, 4y, W, 4h) & (2 \leq i \leq n - 1), \\ (0, 4y, 4(w + x) - (n - 1)W, h) & (i = n). \end{cases} \quad (2)$$

6.4 ディスプレイサーバ

ディスプレイサーバは、遠隔地点から受信したオブジェクト領域の画像と始点座標値を基に、ディスプレイに表示する画像を生成し、生成した画像に対応するディスプレイに出力する。1 台のディスプレイサーバが 1 台のディスプレイを制御するため、モノリスと同じ台数のディスプレイサーバが必要となる。

従来の t-Room ではディスプレイサーバはディスプレイに出力する画像の全体を受信していた。これに対し本システムでは、映像オブジェクトが存在する矩形領域の画像のみを受信する。即ち、ディスプレイに出力する画像の一部のみを受信する。結果として、受信した画像以外の部分をディスプレイサーバが用意する必要がある。ここで、未受信部分の画像とは映像オブジェクトの存在しない部分である。t-Room においては映像オブジェクトの存在しない部分の画素の色は黒（映像出力なし）としている。このためディスプレイサーバは、初めに、ディスプレイに出力する画像と同じ解像度を持つ黒一色の画像を生成しておく。そして、遠隔地点の画像処理サーバから映像オブジェクトの存在する矩形領域の画像及び矩形領域の始点座標を受信し、受信した画像及び始点座標の情報を基に、黒一色の画像上の対応する座標に受信画像を上書きする。こうしてサーバは、生成した画像をディスプレイに出力する。

7. 共有空間の拡張：投影規則の構築

従来の t-Room に残されていた問題である共有空間の制約を解決するにあたり、中心となる課題は、空間の中心部に存在する映像オブジェクト（人物など）を自然な形で壁面部のディスプレイに投影する規則の構築である。特に、ここで自然な投影とは、利用者に違和感を生じさせないことに加え、コラボレーション作業を支援し得るものでなければならない。

（従来の）t-Room 空間において、カメラは、対面位置のディスプレイをやや上部からではあるもののほぼ正面から、しかもそれが撮影画像のほぼ全てを占める程度に撮影する。このため、t-Room 空間の中心部は、そこに設置さ

れる全てのカメラの撮影範囲となる。実際、カメラによる撮影映像がそのまま遠隔地点のディスプレイに出力されていた従来の t-Room では、t-Room 空間の中心部付近の映像オブジェクトは、それがただ1つであるにもかかわらず遠隔地点の複数のディスプレイ上に出力されていた。これは明らかに不適切である。また、カメラの撮影範囲に関する上述の設定は、ディスプレイ上に映像オブジェクトがほぼ等身大に映るようになるためのものでもあった。このため、t-Room 空間の中心部にある映像オブジェクトは、遠隔地点のディスプレイでは実際より大きく表示されていた。これもまた、明らか不自然かつ不適切である。従来の t-Room が利用領域をモノリス壁付近に制限していた理由は、こうした不適切な画像表示を避けるためであった。

こうした従来の t-Room が持つ制約を解消するためには、以下に示すような、選択的かつ画像加工を伴うオブジェクトの表示が必要となる。

- ディスプレイで表示される際に等身大となるよう、撮影されたオブジェクトの大きさを正規化する。
- 複数のカメラ映像のそれぞれに含まれるオブジェクト画像から1つのオブジェクトのみを適切に選択し、そのオブジェクトの映像のみを遠隔地点の t-Room に伝送する。
- 部屋空間内の2次元の動きを、ディスプレイ上の1次元の動きに自然な形で写像する。

ここで、第1項は比較的容易に実現できる。第2項の選択は、撮影されるオブジェクトを見る側の（一般に遠隔地の t-Room 内にいるコラボレーション参加者の）視点情報を基に行われる必要がある。そのためやや複雑な処理が求められるが、元々多数のカメラを持つ t-Room において参加者位置の把握は難しくない。特に天井部に鳥瞰カメラを設置すれば、その把握はさらに容易となる [23]。第3項の写像の考え方を図6に図解する。図は、t-Room 部屋内において壁付近から中心部に人が移動する様子を示している。その動きは、部屋の2次元の広がりをも前提としている。しかし、撮影されたその人映像の表示は、基本的にディスプレイ内の1次元の動きに封じられる。また、この被写体である人は、（接続された遠隔地の t-Room 内で該当する位置でもある）その正面にいる相手に向かって移動しているものとする。相手からはその移動人物の正面が見えていることが自然である。しかし、ディスプレイ上に写像された人画像は、相手にまっすぐ向かうのではなく、ディスプレイで構成される壁面を回るように移動せざるを得ない。そのとき、表示される人映像は、その移動の方向に応じて姿勢を変えなければ不自然である。図には、この姿勢の変更を図解している。最終的にはコラボレーション相手の正面に向き合うものの、回り道的に壁面を伝って移動する時にはその移動が自然に表示されるように被写体画像の姿勢を変更する。このような姿勢の変更には、t-Room の空間

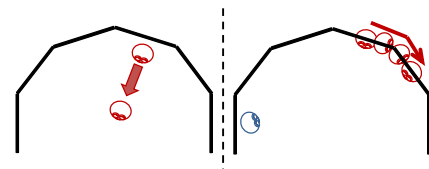


図6 壁面から中心部に2次元的に移動する人物オブジェクトのディスプレイ面内の1次元的動きへの写像の例。

Fig. 6 Example of transforming a 2-dimensionally moving human object to 1-dimensional moving in Monolith displays.

を囲むように設置している複数台のカメラの映像を利用する。なお、この写像の処理に関しては、現在プロトタイプング中である [24]。複数とはいえ、カメラの台数は限られている。求められる姿勢の映像を実現するため、Virtualized Reality 技術 [25] の導入等を進めている。

8. おわりに

従来の t-Room では困難であった部屋空間中心部の共有空間としての利用を目指して、新しい遠隔コラボレーション支援システムの開発を進め、その設計の考え方と実装例を報告した。従来は撮影したカメラ映像を全て伝送していたのに対し、新しいシステムでは映像オブジェクトの存在する領域の画像のみを抽出した後に伝送し、受信側で映像を再構成する。これにより、伝送データの量を削減すると共に、ディスプレイ上における自然な映像オブジェクトの表示を目指す画像処理の実装を容易にした。

今後は、Virtualized Reality 技術の応用などを含めて、オブジェクト抽出機能の実装、及び投影規則の構築と実装等を進めていく予定である。

参考文献

- [1] Harrison, S.: *Media Space 20+ Years of Mediated Life*, Springer Publishing Company, Inc. (2009).
- [2] Fuchs, H., State, A. and Bazin, J.-C.: Immersive 3D Telepresence, *IEEE Computer Magazine*, Vol. 47, No. 7, pp. 46-52 (2014).
- [3] Cesar, P., Kaiser, R. and Ursu, M.: Toward Connected Shared Experiences, *IEEE Computer Magazine*, Vol. 47, No. 7, pp. 86-89 (2014).
- [4] Okada, K.-I., Maeda, F., Ichikawa, Y. and Matsushita, Y.: Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design, *Proc. CSCW '94*, ACM, pp. 385-393 (1994).
- [5] 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).
- [6] 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).
- [7] Naomi, Y., Keiji, H., Toshihiro, T., Yasunori, H., Yoshinari, S. and Shigemi, A.: 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).
- [8] Cruz-Neira, C., Sandin, D. and DeFanti, T.: Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE, Proc. SIGGRAPH '93, ACM, pp. 135–142 (1993).
- [9] Gross, M., Würmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., Van Gool, L., Lang, S., Strehlke, K., Moere, A. and Staadt, O.: Blue-c: A Spatially Immersive Display and 3D Video Portal for Telepresence, ACM Trans. Graph., Vol. 22, No. 3, pp. 819–827 (2003).
- [10] Morikawa, O. and Maesako, T.: HyperMirror: Toward Pleasant-to-use Video Mediated Communication System, Proc. CSCW '98, ACM, pp. 149–158 (1998).
- [11] 三輪敬之, 石引 力, 渡辺 隆, 篠原 淳: 影を自己のエージェントに用いた共存在的コミュニケーションシステムの開発, 電子情報通信学会技術研究報告. HCS, ヒューマンコミュニケーション基礎, Vol. 103, No. 113, pp. 35–40 (2003).
- [12] 竹森幸輝, 前田佳奈, 岩原正典, 片桐 滋, 大崎美穂: ローカル・ラグ制御機能とログ同期機能を持つ音響サーバの開発, 情報処理学会研究報告, Vol. 2012-AVM-76, No. 3 (2012).
- [13] 村上 昂, 飯田卓也, 片桐 滋, 大崎美穂: 遠隔コラボレーション支援のための映像伝送システムの開発とその遅延評価, 情報処理学会研究報告, Vol. 2013-GN-86, No. 19 (2013).
- [14] 荻野裕也, 杉本直也, 片桐 滋, 大崎美穂: Linux 版 t-Room の開発: デバイス制御システムの設計と実装, 情報処理学会研究報告, Vol. 2014-GN-90, No. 13 (2014).
- [15] 李 榮宰, 片桐 滋, 大崎美穂: ローカル・ラグ制御を伴うマルチメディア伝送サーバの構築, 情報処理学会研究報告, Vol. 2015-GN-94, No. 22 (2015).
- [16] 荻野裕也: Linux 版 t-Room のためのデバイス制御システムの開発, 同志社大学大学院理工学研究科修士論文 (2014).
- [17] 大島義博, 片桐 滋, 大崎美穂: ローカル・ラグ制御機能を持った音響サーバの性能評価実験, 2014 年度情報処理学会関西支部 支部大会 講演論文集, Vol. 2014 (2014).
- [18] 小寺晋平, 片桐 滋, 原田康徳, 平田圭二, 大崎美穂: 映像フィードバックに伴うエコーのキャンセリング法に関する実験的評価, 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, Vol. 109, No. 306, pp. 291–296 (2009).
- [19] 上埜敏司, 片桐 滋, 大崎美穂: t-Room における赤外線センサを用いた動画オブジェクト抽出, 2014 年度情報処理学会関西支部 支部大会 講演論文集, Vol. 2014 (2014).
- [20] 中谷彰皓, 片桐 滋, 大崎美穂: 高臨場感を与える視聴覚ディスプレイのための音響制御システム, 情報処理学会研究報告, Vol. 2015-GN-94, No. 23 (2015).
- [21] 山下春香, 中谷彰皓, 片桐 滋, 大崎美穂: ディスプレイ上下に配置した音響反射板つきスピーカの実験的評価, 日本音響学会 2015 年春季研究発表会 (2015).
- [22] 平田圭二, 原田康徳, 高田敏弘, 青柳滋己, 白井良成, 山下直美, 大和淳司, 梶 克彦: 遠隔ビデオコミュニケーションシステムのための仮想共有面の実装方式, ワークショップ 2007 (GN Workshop 2007) 論文集, Vol. 2007, pp. 113–118 (2007).
- [23] 竹内香奈, 片桐 滋, 大崎美穂: 遠隔コラボレーション支援システム「t-Room」における複数カメラ映像を用いた利用者追跡, 電子情報通信学会技術研究報告. IE, 画像工学, Vol. 111, No. 442, pp. 319–324 (2012).
- [24] Wada, O., Ueno, S., Katagiri, S. and Ohsaki, M.: Implementation of Advanced t-Room (to appear), Proc. TENCON 2015, IEEE (2015).
- [25] Kanade, T., Rander, P. and Narayanan, P.: Virtualized reality: constructing virtual worlds from real scenes, IEEE MultiMedia Magazine, Vol. 4, No. 1, pp. 34–47 (1997).