Vol.2022-GN-116 No.11
2022/3/14

遠隔コミュニケーションにおける視線一致を実現するためのディスプレイシステムの開発

和田 理[1]　岩元 和大[2]　片桐 滋[1]　大崎 美穂[1]

概要：カメラとディスプレイを用いる遠隔コミュニケーションにおいて，カメラをディスプレイと同一の光線上に設置することは困難である．このため，このような状況において視線一致を実現することは困難である．しかしながら，人と人とのコミュニケーションにおいて視線は重要な要素であり，遠隔コミュニケーションにおいて視線を一致させるということの意義は大きい．本研究では，ディスプレイデバイスの前面に格子状に複数台のカメラを設置して，その中から最も人物に近いカメラを一つあるいは複数選択し，取得した画像データより視線を一致させる映像を生成し，遠隔地点のディスプレイに出力するディスプレイシステムを提案し，その実装を行う．

Development of Display System to Realize Eye-Contact in Remote Communication

1. はじめに

遠隔コミュニケーションにおける映像伝送に関して，その映像を撮影するカメラの設置場所は映像を出力するディスプレイの周辺部，多くはディスプレイの上部もしくは下部に設置される．このような条件下では，人物とカメラ，及び人物とディスプレイの方向は一致しない．このため，遠隔地点に出力される映像において，人物とカメラ，及び人物とディスプレイの間に生じる角度に伴う映像の方向のずれが生じる．このため，遠隔コミュニケーションにおいて視線情報を利用することは困難である．またこのことは，対面でのコミュニケーションに対する遠隔コミュニケーションの制約となる．

近年では，大型のディスプレイやスクリーンに対して等身大サイズの映像を出力する遠隔コミュニケーション支援システムが多数提案されている．このようなシステムにおいては，出力される映像のサイズも大きくなる．それに伴い，出力映像における視線の方向のずれの影響も大きくなる．

本稿では，大型のディスプレイの内部に小型のカメラデバイスを格子状に複数台埋め込むディスプレイを持つシステムを提案し，その開発を行う．また，ディスプレイに出力する映像は等身大のサイズとする．カメラをディスプレイ面上に設置することにより，人物とカメラ，及び人物とディスプレイの間に生じる角度に伴うずれの縮小された映像をカメラで撮影できるようにする．また，それらの画像からよりずれの小さいカメラの映像より，それらの画像の中間の視点の映像の生成をすることによって，視線一致性を高める技術の開発についても検討する．

2. 関連研究

等身大映像を用いる遠隔コミュニケーションにおいて視線一致が試みられた例として，半透過型のスクリーンとプロジェクターを用いる [1]，スクリーンの一部を切り欠いてそこにカメラを設置する [2]，ハーフミラーを利用してディスプレイの方向を向いた人物の像をカメラの設置されている方向へ反射する [3] ことにより，人物とカメラ，及び人物とディスプレイの方向を一致させた例が存在する．その他にも複数のカメラを用いて人物の像を三次元的に再構築して，仮想空間に任意視点の映像を投影することによって方向の不一致を解決したもの [4] が存在する．

また，ディスプレイの前面に格子状にカメラが設置されたシステムにおける遠隔コミュニケーションの支援がな

1 同志社大学
Doshisha University
2 同志社大学大学院
Doshisha Uniersity

© 2022 Information Processing Society of Japan

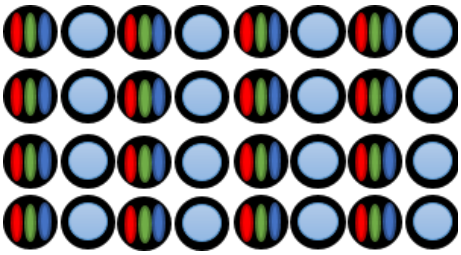

図 1 本稿のシステムにおけるカメラ配置の理想状態の前面図．1 つのピクセルごとに 1 台のカメラが存在する．

Fig. 1 Frontal view of the ideal state of camera placement in our system. There is one camera per pixel.

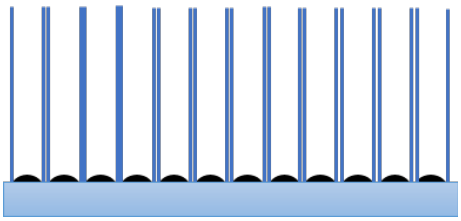

図 2 本稿のシステムにおけるカメラ配置の理想状態の上面図．カメラの画角はカメラから直線の方向に限定される．

Fig. 2 Top view of the ideal camera layout for our system. The angle of view of the camera is limited to the direction of a straight line from the camera.

された例として CamRay[5] が挙げられる．しかしながら，CamRay では映像の出力地点は数か所に限定されており，また出力される映像は等身大のサイズではない．

3. ディスプレイシステムにおける視線一致映像を取得するためのカメラの配置について

遠隔コミュニケーションにおいて視線を一致させるには，一方の地点で撮影されるカメラ映像ともう一方の地点に出力されるディスプレイ映像の方向が一致していることが必要となる．

これを実現するための手法として，ディスプレイデバイスにカメラデバイスを埋め込むことが考えられる．近年販売されている多くのスマートフォンにおいて，映像を表示する領域の拡大を目的として，ディスプレイの一部を切り欠いて人物撮影用のカメラが配置されている．

人物とディスプレイ，及び人物とカメラの方向を一致させるにあたり，**図 1**，**図 2** の状態が理想の状態となる．ディスプレイの 1 つのピクセルあたり 1 台のカメラを設置して，カメラの撮影範囲をカメラから直線の方向へと限定する．そして，これらのカメラで撮影されたものを遠隔地点のディスプレイの対応のピクセルへ出力する．これにより，人物がディスプレイ上のどの地点の方向を向いた場合でもカメラとディスプレイの方向が一致して，遠隔地点のディスプレイに出力される映像における視線情報の対称性が確保される．

しかし，当然ながらカメラの大きさはディスプレイのピクセルと比較すると大きく，このようなことを実現することは困難である．

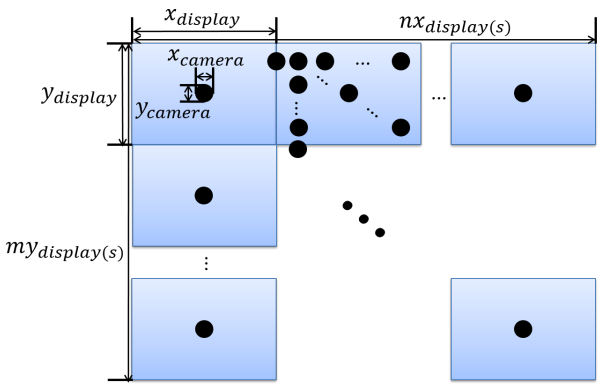

図 3 カメラをディスプレイデバイスに埋め込むユニットの構成図．カメラ及びディスプレイがそれぞれグリッド状に配置される．

Fig. 3 Configuration diagram of a unit that embeds a camera into a display device. The camera and display are arranged in a grid.

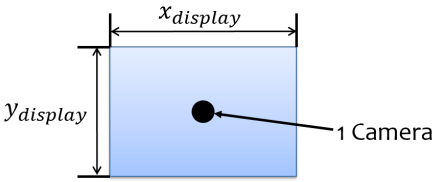

図 4 カメラをディスプレイデバイスに埋め込むユニットの最小の構成図．1 枚のディスプレイの中心部にカメラが埋め込まれている．

Fig. 4 A minimal diagram of a unit that embeds a camera into a display device, with the camera embedded in the center of a single display.

したがって，本稿で開発するシステムは**図 3** のように 1 枚，あるいは複数枚のディスプレイの内部に複数台のカメラを埋め込むものとなる．最小の構成として**図 4** のように 1 枚のディスプレイの中心部に 1 台のカメラを埋め込む例があげられる．この場合，人物の移動可能な範囲はカメラで撮影可能な範囲に制限される．

そこで，本稿のシステムでは人物の移動可能な領域を拡大するべく，カメラを複数台，格子状に配置することとする．カメラを複数台配置する場合，人物の存在する位置に従って出力するカメラを適宜選択していく必要がある．このとき，設置するカメラ間の距離，及びカメラの画角によって人物の像の欠損が生じる領域が存在する．カメラの設置条件と映像に欠損の発生する領域の関係を**図 5** に示す．図 5 における黄色の三角形の領域に人物が存在するとき，人物像の一部あるいは全体がいずれのカメラの撮影領域にも含まれない．したがってこの範囲に存在する人物の像に関して，遠隔地点に出力することはできない．この黄色い領域を縮小することによって，移動範囲の制約は少なくなる．三角形の領域を縮小する方法としてカメラ間の距

© 2022 Information Processing Society of Japan

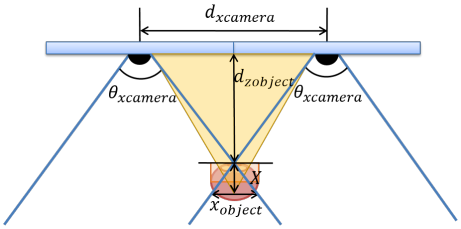

図 5 複数のカメラの映像を同時出力する際に映像の欠損が発生する位置．カメラの焦点距離とカメラ間の距離によって決まる．

Fig. 5 The position where video loss occurs when images from multiple cameras are output simultaneously. It is determined by the focal length of the camera and the distance between the cameras.

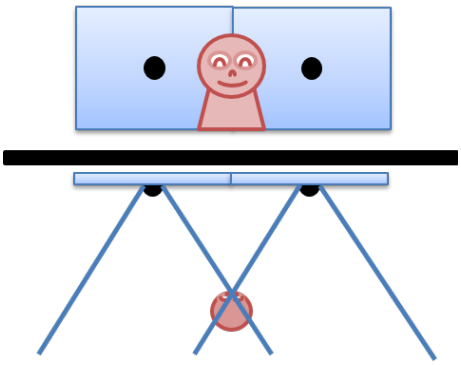

図 6 ディスプレイと人物との距離が特定の距離になったとき，人物の映像は等身大で，また映像の境界面において像の重複も欠けも発生しない．

Fig. 6 When the distance between the display and the person is a specific distance, the image of the person is life-size, and there is no overlapping or missing of images at the boundary of the image.

離 ($d_{xcamera}$) を縮小することや，カメラの画角 ($\theta_{xcamera}$) を拡大することが挙げられる．

その一方で映像を等身大のサイズで出力することにおいても，カメラの画角は重要である．カメラで撮影された画像において人物が等身大サイズとなる場所はカメラと人物との距離，及びカメラの画角によって決定される（**図 6**）．カメラと人物との距離が小さいとき（**図 7**），人物の像の大きさは等身大よりも大きくなる．また，人物の像の一部がカメラの撮影範囲から外れるために出力映像において人物の像の欠けが発生する．一方でカメラと人物との距離が大きいとき（**図 8**），人物の像の大きさは等身大より小さくなる．また，人物の像が複数のカメラに重複して撮影されるため，出力映像において人物の像の重複が発生する．カメラの配置，撮影，及び出力の条件に関して，これらのことを考慮して決定する必要がある．また，必要に応じて撮影された映像の全体を出力するのではなく，映像の一部の領域を抽出して出力するなどの処理を行うことが求められる．

また，遠隔コミュニケーションにおける視線情報の利用

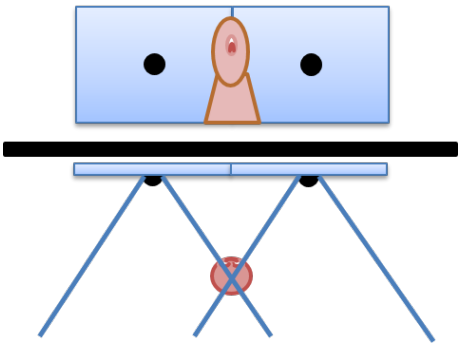

図 7 ディスプレイと人物との距離が小さいとき，人物の映像のサイズが等身大より大きくなり，また映像の境界面における像の欠けが生じる．

Fig. 7 When the distance between the display and the person is small, the size of the person's image becomes larger than life-size, and there is a lack of image at the boundary of the image.

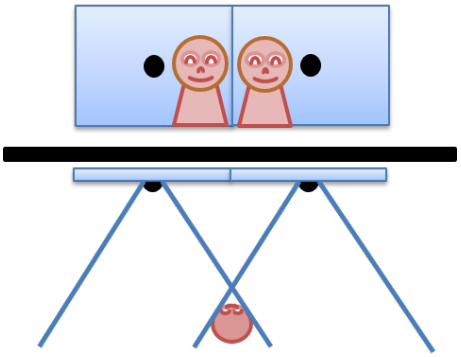

図 8 ディスプレイと人物との距離が大きいとき，人物の映像のサイズが等身大より小さくなり，また像の重複が発生する．

Fig. 8 When the distance between the display and the person is large, the size of the person's image becomes smaller than life-size, and overlapping of images occurs.

を実現するには，カメラで撮影される映像とディスプレイに出力される映像において，人物の位置，および視線の方向が対応の関係であることが求められる．例えば，両方の地点において人物がカメラから直線の方向に存在し，またカメラを直視している場合は，出力される映像においても，人物は相手の方向を直視する（**図 9**）．

一方の地点の人物が隣接するカメラから直線の位置へと移動し，そのカメラを直視するとき，対応するカメラの直前に人物の映像を出力するのであれば，遠隔地点においても出力場所の移動が行われ，出力する方向も一致する（**図 10**）．

この条件下でお互いの地点の人物がディスプレイに出力されている遠隔地点の人物を直視したとき，ディスプレイ出力においても人物は相手地点の人物を直視することとなる（**図 11**）．このように，人物の位置と視線方向の対称性が確保されるようにディスプレイに出力される映像の位置と方向について考慮することが求められる．

© 2022 Information Processing Society of Japan

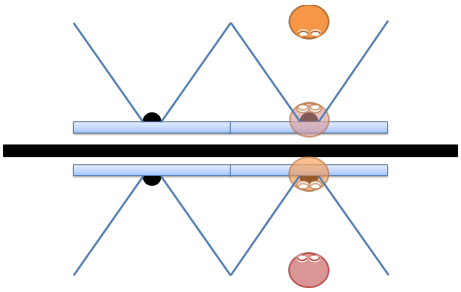

図 9 人物が正対しているときのディスプレイの出力映像．ディスプレイに出力される映像においてお互いの人物を直視していることがわかる．

Fig. 9 Output image on the display when the person is facing each other. It can be seen that they are looking directly at each other in the image output on the display.

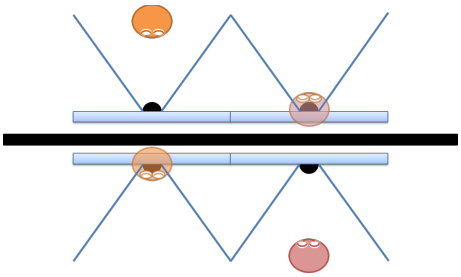

図 10 一方の人物が隣接するカメラの前に移動してカメラを正視したときのディスプレイの出力映像．ディスプレイに出力される映像において人物が隣接するカメラの正面の方向を向いていることがわかる．

Fig. 10 Output image on the display when one person moves in front of the adjacent camera and looks squarely at the camera. It can be seen that the person is facing the front of the adjacent camera in the image output on the display.

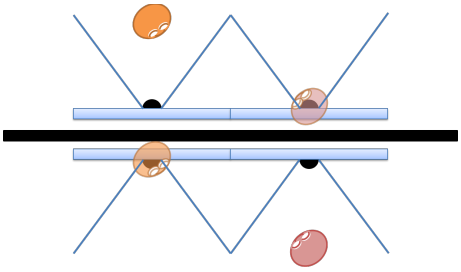

図 11 一方の人物が隣接するカメラの前に移動してディスプレイ上の人物の方向を向いたときのディスプレイの出力映像．ディスプレイに出力される映像においてお互いの人物を直視していることがわかる．

Fig. 11 Output image of the display when one person moves in front of the adjacent camera and faces the direction of the person on the display. It can be seen that both people are looking directly at each other in the image output on the display.

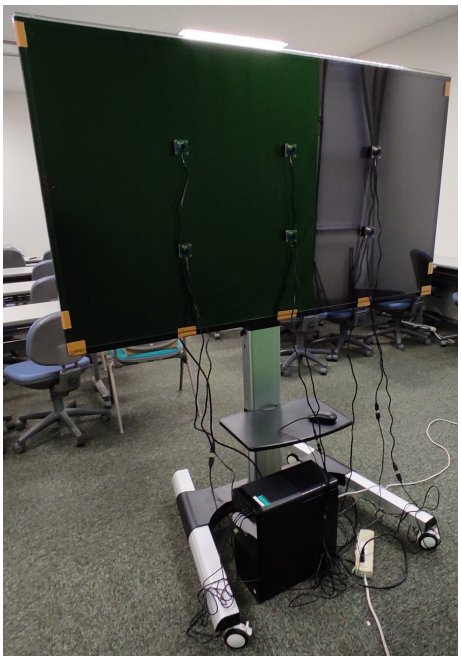

図 12 開発したディスプレイデバイスの写真．ディスプレイの前面に格子状に 6 台のカメラが設置されている．

Fig. 12 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.

表 1 ディスプレイの諸元
Table 1 Specification of display.

Screen Size	65 inches (1439 mm (H) × 809 mm (H))
Bezel Size	18 mm (each side)
Height from Ground	931 mm

表 2 カメラの諸元
Table 2 Specification of camera.

Dimensions	38 mm (W) × 38 mm (H))
Sensor Size	1/3 inches (48 mm(H) × 36 mm (H))
Resolution	2 mega pixels
Angle of View	97.6° (H), 81.2° (V), 110° (D)
Number of Cameras	6

4. 開発するディスプレイシステムについて

4.1 ハードウェア

本稿のシステムでは大型のディスプレイを 1 枚用いる．また，小型のカメラを横方向に 3 箇所，縦方向に 2 箇所，計 6 箇所格子状に配置する．作成したシステムの写真を**図 12** に示す．また，ハードウェアの諸元を**表 1**，**表 2**，及び**表 3** に示す．また，ディスプレイ面におけるカメラの設置場所を**図 13** に示す．

© 2022 Information Processing Society of Japan

表 3　サーバーマシンの諸元
Table 3　Specification of Server Machine.

CPU	Intel Core i7 8700 (3.20 GHz, 6 core, 12 threads)
RAM	16.0 GB
GPU	NVIDIA GeForce GTX 1070 (8GB RAM)
NIC	Intel Killer E2400 Gigabit Ethernet Controller

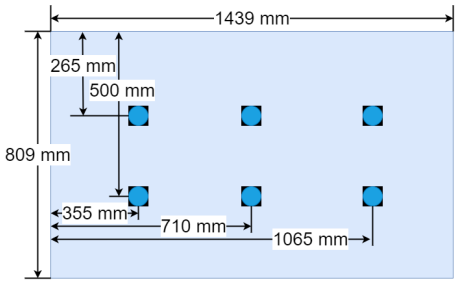

図 13　ディスプレイ面におけるカメラの設置場所．格子状に横方向に 3，縦方向に 2，合計 6 台設置されている．
Fig. 13　Location of the cameras on the display surface. A total of six cameras are installed in a grid, three in the horizontal direction and two in the vertical direction.

4.2　ソフトウェア

本稿で開発するシステムでは，ディスプレイの前面に配置されたカメラから人物の正面に近いカメラを 1 つあるいは複数選択して，人物の正面の映像を生成して遠隔地点に設置されたシステムへ伝送する．

処理の流れは以下の通りとなる．最初にシステムに装着された全てのカメラで画像を撮影する．次に撮影された各カメラ画像より顔領域，及び顔の特徴点の抽出を行う．次に各カメラで撮影された顔の特徴点から顔の中心の座標とカメラ画像の中心の座標との距離を計測する．計測した距離の値に基づいて，顔の中心とカメラの中心の距離の小さい画像を 1 枚あるいは複数枚選択して，これらの画像より視線が一致する画像を生成する．最後に生成された視線が一致する画像を遠隔地点へ伝送し，遠隔地点のディスプレイに出力する．

顔の認識，及び顔特徴量の抽出はオープンソースの画像処理ライブラリーである OpenCV，及びオープンソースの機械学習ライブラリーである dlib を用いて実装する．最初にカメラで撮影された画像より画像の中の顔領域を取得する．次に顔領域の取得に成功した画像に対し顔特徴量の抽出を行う．抽出された顔の特徴量より抽出された顔領域の中心の座標を抽出する．そして，抽出された顔領域の中心の座標とカメラの中心の座標の距離の計測をすべてのカメラに対して行い，人物に最も近いカメラを検出する．

また，本システムではリアルタイムでの通信を行うため，伝送する映像については低遅延かつ高効率であることが求められる．これを実現するために，地点間の映像伝送において RTP (Real-time Transport Protocol) を用いる．また，伝送する映像に対して H.264 形式の圧縮を行う．これらの処理の実装にはオープンソースのマルチメディアフレームワークである GStreamer を利用する．

4.3　動作確認

開発したシステムを用いて 2 つの地点で通信したときの写真を**図 14** に示す．また，カメラの動作確認のためにカメラと同じ地点のディスプレイに出力する条件でシステムを動作させたときの写真を**図 15** に示す．図 15 の左図と右図より，人物が屈んだときに出力対象の映像を撮影するカメラが変更されていることがわかる．これは，人物が屈んだことによって人物の顔領域とカメラの中心の距離が最小であるカメラが変化したことによるものである．しかし，現在は選択されたカメラ映像を直接出力しているため，等身大よりも小さなサイズで出力される．これを等身大サイズとするためには，出力対象とする領域に制限を加えることによって大きさの正規化を行う必要がある．このとき，出力対象とする領域は人物の映る場所に基づいて決定されるため，人物の存在する位置と座標を取得することが必要である．

人物がディスプレイの中心から正面で直立し，正面の方向を直視しているときのシステムに装着された 6 台それぞれのカメラで撮影された画像を**図 16** に示す．人物がディスプレイの中心の正面で直立しているため，選択される映像はディスプレイの中央に位置する上下 2 台のカメラのいずれかで撮影されたものであることがわかる．ここでは中央上部のカメラの映像が出力対象として選択される．このときは，人物はカメラの方向を直視しているために遠隔地点間において視線情報が一致する．しかしながら，これだけでは人物の存在する位置や人物の視線によって，視線情報の一致が実現しない状況が生じる．これを解決する手法として，カメラ間の中間に仮想的な視点を生成し，その視点の映像を生成することによって視線情報の一致する位置や方向の制約を緩和するということが考えられる．

中間視点の映像の生成によって視線情報を一致させる手法として，視点とするカメラを選択するために取得した顔特徴点の値を利用して，モーフィング処理を適用し，いずれかの画像の顔領域を処理結果をカメラ画像へ置き換えるといったことが考えられる．

5.　おわりに

本稿では，遠隔コミュニケーションにおける遠隔地点間の視線情報の対称性の確保を目指し，ディスプレイの前面にカメラを格子状に複数台配置することにより，人物とカメラ，人物とディスプレイの方向のずれを縮小した．また人物に近い複数台のカメラより中間視点の映像を生成することによって人物とカメラ，人物とディスプレイの方向を一致させる方法に関する検討を行った．今後は，出力される映像の大きさの正規化のための人物位置の取得技術の開

© 2022 Information Processing Society of Japan

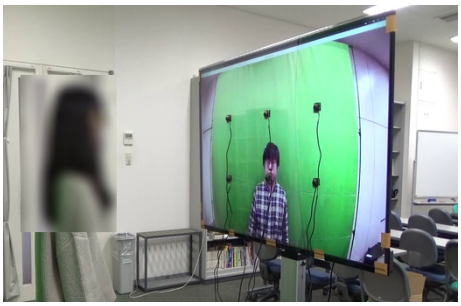

図 14 2 地点間の通信の実例.

Fig. 14 A communication example between two locations (two systems).

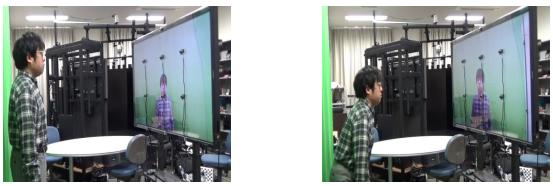

図 15 本稿のシステムを送信地点と受信地点を同一の地点として動作させたときの写真．左図では人物は直立しており，右図では人物は屈んでいる．人物の位置に応じてカメラの映像が選択され，ディスプレイに出力される．

Fig. 15 A picture of our system when the sending and receiving points are the same. In the left figure, the user is standing upright: in the right figure, the user is bending down. The image from the camera is selected according to the user's position and output to the display.

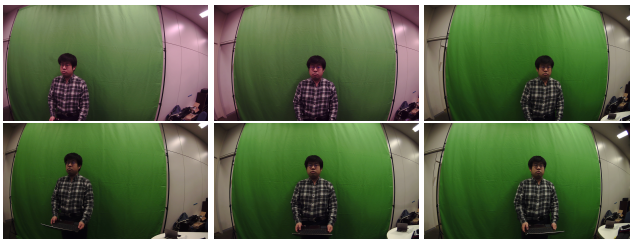

図 16 システムの中心から直線の位置に人物が立ち，ディスプレイの中心の方向を正視しているときのシステムに備え付けられている 6 台のカメラの撮影結果．このうち中央上部のカメラの映像において顔領域がカメラ画像の中心部の最も近くに存在するため，この画像が選択，伝送されて受信側のディスプレイに出力される．

Fig. 16 Images taken by the six cameras in this system when a person is standing in a straight line from the center of the system and looking straight at the direction of the center of the display. In the image from the top center camera, the face area is closest to the center of the camera image, so this image is selected, transmitted, and output to the receiver's display.

発や，映像上における視線のずれを縮小するための中間視点の映像の生成技術の開発を進めていく．これらの機能を開発することにより．等身大の映像を用い，視線が一致された遠隔コミュニケーション環境の実現を目指す．

謝辞 本研究を実施するため，本学の安田悠月氏と半下石澪氏には大変世話になった．著者一同，感謝申し上げる．

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] Kuzuoka, H., Yamashita, J., Yamazaki, K. and Yamazaki, A.: Agora: A Remote Collaboration System That Enables Mutual Monitoring, *CHI '99 Extended Abstracts on Human Factors in Computing Systems*, ACM, pp. 190–191 (1999).

[3] Bondareva, Y. and Bouwhuis, D.: Determinants of social presence in videoconferencing, *AVI2004 Workshop on Environments for Personalized Information Access*, pp. 1–9 (2004).

[4] Ogi, T., Yamada, T., Tamagawa, K., Kano, M. and Hirose, M.: Immersive telecommunication using stereo video avatar, *Proceedings IEEE Virtual Reality 2001*, pp. 45–51 (2001).

[5] Avellino, I., Fleury, C., Mackay, W. E. and Beaudouin-Lafon, M.: Camray: Camera arrays support remote collaboration on wall-sized displays, *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 6718–6729 (2017).

© 2022 Information Processing Society of Japan