

令和 6 年 5 月 7 日現在

機関番号：12601

研究種目：若手研究

研究期間：2021～2023

課題番号：21K17783

研究課題名（和文）遠隔映像コミュニケーション支援に向けた高速ビジョンシステムの開発

研究課題名（英文）Development of high-speed vision system for online video communication

研究代表者

金 賢悟（Kim, Hyuno）

東京大学・生産技術研究所・特任助教

研究者番号：60817328

交付決定額（研究期間全体）：（直接経費） 2,800,000円

研究成果の概要（和文）：遠隔映像コミュニケーション支援に向けた高速ビジョンシステムを構築し、複数の映像間の時間軸上の整合性を評価するための同期精度評価手法を提案した。さらに、撮像の高速性を利用して、複数のカメラ視野にまたがる物体の同定手法を開発した。また、顔姿勢の迅速な検出を目指し、比較的時間がかかるAIによる特徴マッチングを省略し、各カメラからの画像入力から顔に対するカメラの姿勢を直接求める手法を評価した。さらに、ビジョンシステムの使用毎にキャリブレーションが不要となるよう、カメラシステムの動的構造化を行なった。これらの要素技術を基盤に、仮想空間内の複数の会議参加者を想定した最適な映像提示手法を提案した。

研究成果の学術的意義や社会的意義

本研究では、従来の遠隔映像コミュニケーションで利用されていたwebカメラに置き換え、臨場感向上が期待できる高速ビジョンシステムを提案し、選択的映像配信手法を確立した。提案システムでは、仮想空間内における会議参加者の位置関係や視線などを考慮し、ある参加者周りのマルチモニター環境において最適な映像を選択的に生成して配信することができる。画一的な正面方向からの映像配信から脱皮して、よりリアルな仮想会議を再現できるため、会議参加者間の感情交流も強化されると期待される。さらに、多方向からの任意姿勢の映像が生成および記録できるため、デジタルアーカイブ用の入力装置としても応用展開が期待できる。

研究成果の概要（英文）：I have developed a high-speed vision system for remote video communication and proposed a method to evaluate the accuracy of video frame synchronization on the timeline. Additionally, I introduced a method to match and identify a target object across multiple camera views.

Furthermore, I assessed an AI-based method for classifying the pose between the face and camera directly from camera images, bypassing conventional AI feature matching. I also redesigned the camera system by implementing a movable structure that eliminates the need for pre-calibration of the vision system before usage.

Finally, based on these findings, I proposed a method to display appropriate video scenes for multiple participants in a virtual conference.

研究分野：情報理工学

キーワード：遠隔映像コミュニケーション 高速ビジョンシステム 高速画像処理

1. 研究開始当初の背景

新型コロナウイルス感染症の世界的流行により、遠隔コミュニケーションを支援するオンライン会議・会話システムの重要性が高まり、音声と映像によるコミュニケーション支援技術が注目されていた。さらに、複数の会議参加者が仮想空間内で形成する位置関係や視線・ジェスチャーなどの情報交換を実時間反映できるリアルな画像提示技術の必要性が増していた。本研究では、1,000fps の高速分散型スマートカメラネットワークと AI 基盤の顔認識技術を発展させ、オンライン会議向けの最適映像取得と低遅延提示が可能な高速ビジョンシステムの実現を目指した。

2. 研究の目的

現行のオンライン会議において、顔映像生成に用いられる Web カメラを置き換え、参加者同士が交わす視線やジェスチャーなどの非音声的情報も立体的に伝達できるよう、多視点映像生成と低遅延映像提示が可能な高速ビジョンシステムを構築する。その要素技術として、高速動作環境下で顔姿勢を検出する AI 技術の最適デザインを行い、さらに高速分散型スマートカメラシステムの動的構造化を行う。加えて、複数の会議参加者が仮想空間内で形成する位置関係や視線情報などを考慮した映像提示手法を確立し、遠隔映像コミュニケーションを支援する高速ビジョンシステムを実現する。

3. 研究の方法

高速な顔映像生成に必要なビジョンシステムを構築し、複数の映像間の時間軸における整合性を評価するための同期精度評価手法を提案した。さらに、撮像の高速性を利用して、複数のカメラ視野に跨がる物体の同定手法を開発した。顔姿勢の迅速な検出を目指し、比較的時間がかかる AI による特徴点抽出やマッチングを省略し、各カメラからの画像入力から顔に対するカメラの姿勢を直接求める手法を検討した。さらに、ビジョンシステムの使用毎にキャリブレーションが不要となり、安定な撮像と姿勢推定が可能となるよう、カメラシステムの動的構造化を行なった。これらの要素技術を基盤に、仮想空間内の複数の会議参加者を想定した最適な映像提示手法を提案した。

4. 研究成果

(1) 高速ビジョンシステムの構築と評価手法の提案

複数スマートカメラを用いたビジョンシステムを構築し、各カメラ映像間のフレーム同期精度を評価するための手法を提案した[1]。提案手法では、直線的に振動する光点の撮像情報を用いることで、別途の計測機器を用いずに 1ms 以下の同期精度が計測可能であることを実証した。また、複数カメラの視野を跨がる観測物体の同定問題に対して、軽量かつ高速な処理アルゴリズムを提案した[2]。撮像の高速性を活かした画像処理アルゴリズムによって、複数のカメラ視野に跨がる物体の同定を物体の位置情報を基盤にして簡単に行えることを理論と実験で証明した。

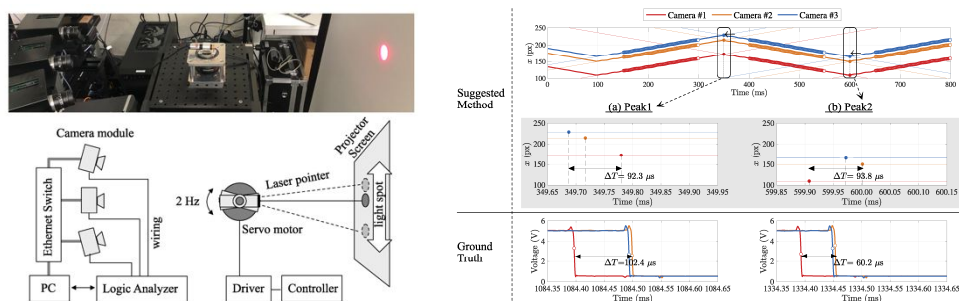


図 1 高速ビジョンシステムにおける撮像同期精度の評価 [1]

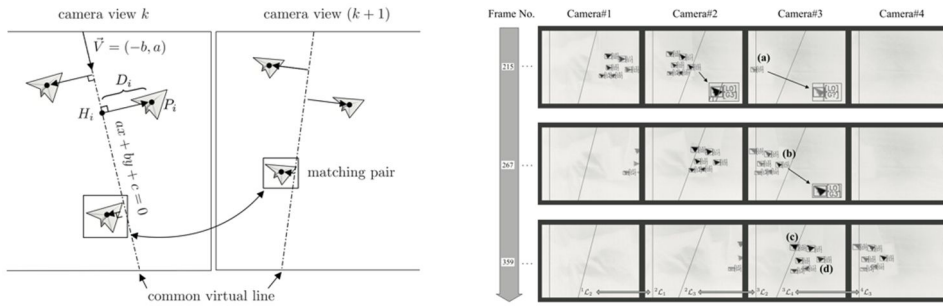


図2 高速ビジョンシステムにおける物標の同定 [2]

(2) 高速な顔姿勢推論と他視点映像生成手法の開発

顔姿勢の迅速な検出を目指し、比較的時間がかった AI による特徴点検出とマッチング過程を省略し、各カメラから得られる画像を入力して、顔に対するカメラの姿勢を直接推論できる深層学習基盤モデルを試した。その結果、区間 $[-90^\circ < \theta < +90^\circ]$ を 15° 間隔でクラス化した 13 段階の姿勢を 500fps 以上の速度で推論可能であることを確認した [3]。また、ビジョンシステムの使用毎にキャリブレーションが不要となるよう、カメラシステムの動的構造化を行なった。実装においては、回転軸を有する多リンク構造に 7 台のカメラを配置し、電動アクチュエータを用いて各カメラの姿勢が制御できるようにした。上記の推論情報を用いることで適応的に姿勢制御ができるため、顔の初期位置や姿勢にロバストに映像を生成し、提示することができるビジョンシステムを構築することができた。

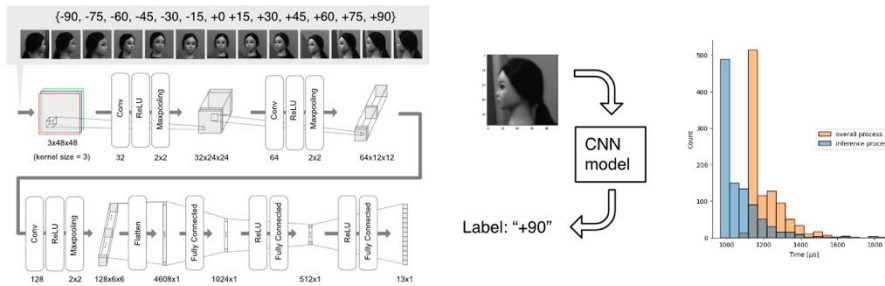


図3 CNN モデルによる顔姿勢推論の高速化と評価 [3]

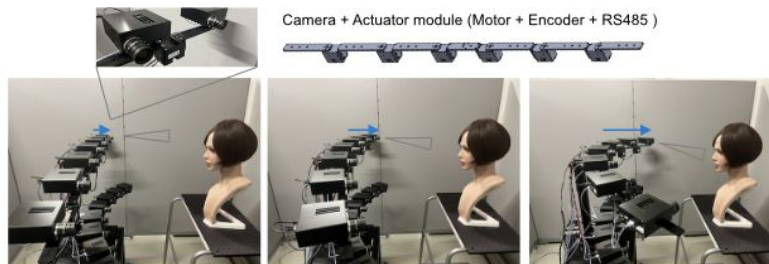


図4 カメラシステムの動的構造化

(3) 遠隔映像コミュニケーションのための映像提示手法の提案

仮想空間内の複数の会議参加者を想定して、最適な映像提示手法を提案した。提案手法では、参加者同士の仮想空間内の位置関係から計算できるある参加者に提示が必要な姿勢を求め、その姿勢のカメラ映像を選択的に転送する方法を確立した。提案手法と高速ビジョンシステムより特定姿勢の顔映像の提示が可能となり、マルチモニターを用いて仮想的な集合を再現することや、常に参加者の正面顔を見ることを選択できる。高速ビジョンシステムより複数の姿勢からの映像を高速かつ選択的に得られるため、デジタルアーカイブに活用が期待できる機能となる。

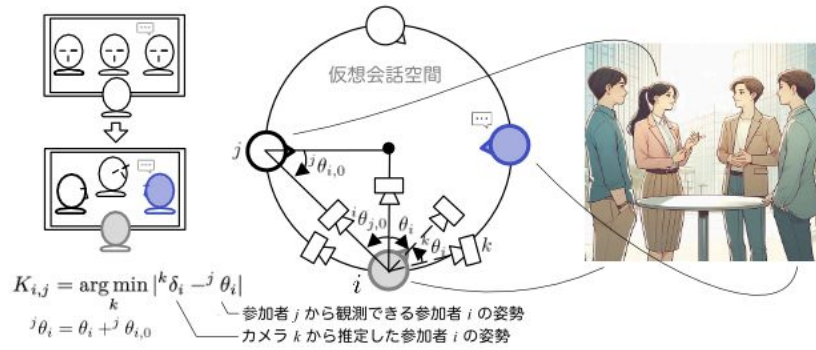


図5 遠隔映像コミュニケーションのための映像提示手法[4]

<引用文献>

- [1] Hyuno Kim, and Masatoshi Ishikawa: Sub-Frame Evaluation of Frame Synchronization for Camera Network Using Linearly Oscillating Light Spot, Sensors, Vol.21, Issue 18, Article No.6148, pp.1-14, 2021.
- [2] Hyuno Kim, Yuji Yamakawa, and Masatoshi Ishikawa: Seamless Multiple-Target Tracking Method Across Overlapped Multiple Camera Views Using High-Speed Image Capture, Journal of Robotics and Mechatronics, Vol.34, No.5, pp. 1043-1052, 2022; doi: 10.20965/jrm.2022.p1043.
- [3] Hyuno Kim, Seohyun Lee, and Yuji Yamakawa: Direct face pose estimation using multiple camera views and deep learning, SPIE PHOTONICS WEST, AI and Optical Data Sciences IV, (San Francisco, California, USA) / Proceedings, Paper No. 12449-8 (2023)
- [4] 金賢梧, イソヒョン, 石川正俊, 山川雄司: 高速マルチカメラシステムによる遠隔映像コミュニケーション支援手法, 第24回 計測自動制御学会 システムインテグレーション部門講演会 (SI2023) (新潟, 2023.12.15) / 講演論文集, pp.2058-2059 (2023)

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 Hyuno Kim, Yuji Yamakawa, Masatoshi Ishikawa	4. 巻 34
2. 論文標題 Seamless Multiple-Target Tracking Method Across Overlapped Multiple Camera Views Using High-Speed Image Capture	5. 発行年 2022年
3. 雑誌名 Journal of Robotics and Mechatronics	6. 最初と最後の頁 1043 ~ 1052
掲載論文のDOI（デジタルオブジェクト識別子） 10.20965/jrm.2022.p1043	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Kim Hyuno, Ishikawa Masatoshi	4. 巻 21
2. 論文標題 Sub-Frame Evaluation of Frame Synchronization for Camera Network Using Linearly Oscillating Light Spot	5. 発行年 2021年
3. 雑誌名 Sensors	6. 最初と最後の頁 6148 ~ 6148
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/s21186148	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計2件（うち招待講演 0件/うち国際学会 1件）

1. 発表者名 金賢梧, イソヒョン, 石川正俊, 山川雄司
2. 発表標題 高速マルチカメラシステムによる遠隔映像コミュニケーション支援手法
3. 学会等名 第24回 計測自動制御学会 システムインテグレーション部門講演会 (SI2023)
4. 発表年 2023年

1. 発表者名 Hyuno Kim, Seohyun Lee, Yuji Yamakawa
2. 発表標題 Direct face pose estimation using multiple camera views and deep learning
3. 学会等名 SPIE PHOTONICS WEST, AI and Optical Data Sciences IV (国際学会)
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------