

低価格な VR ステレオカメラによる低遅延配信システムの一検討 Study of Low-latency Streaming System using Low-price VR Stereo Camera

浦野 健太¹⁾ 米澤 拓郎¹⁾ 河口 信夫¹⁾²⁾
Kenta Urano Takuro Yonezawa Nobuo Kawaguchi

1 はじめに

アーム型・ヒト型など様々なロボットの普及に伴い、遠隔操作システムを介した作業が注目されている。既存の重機や医療分野における遠隔操作システムでは、専用の装置や回線を使った高価なものが多い。

我々は、作業場所での労働力不足と地方での過剰な労働力をバランスするために、遠隔操作ロボットとメタバース環境を活用した、リモートでの物理タスクを行うメタワーク [1] を構想している。この構想の実現には、作業用のロボットや現地映像を送るカメラ、遠隔操作に使うデバイスが低価格で容易に入手できる必要がある。また、遠隔地のロボットを操作してタスクを実行するにあたっては、ものの配置を理解しづらい単眼のカメラではなく、立体的な映像を実現できるステレオカメラと VR180 形式 [2] の映像の利用が望ましい。さらに、ステレオ映像を視聴しながら作業するにあたって VR ヘッドセットを使うことを考えると、前方視界 180 度を可能な限り高精細で低遅延に伝送できなければならない。

我々は現在、VR180 映像を取得するカメラにキヤノン社のミラーレス一眼と専用レンズを使用して出力を得ているが、一式が高額である。そこで、安価なカメラとコンピュータを用い、低遅延 VR180 映像配信システムの構築を目指す。

2 関連機材・関連研究

近年は、360 度カメラや VR 機器の普及に伴って全天球映像や VR180 形式が利用できるようになり、映像配信の多様性が増している。

2.1 市販 VR180 カメラ

周囲 360 度を一括で記録する 360 度カメラがスポーツや不動産内見で利用される一方で、VR180 で撮影できるカメラは VR コンテンツ制作向けが主用途となっている。キヤノン社は、ミラーレス一眼カメラと専用レンズを利用した VR 撮影システムを提案 [3] している。カメラとレンズの組み合わせにもよるが、一式 40 万円程度からとなる。また、リアルタイム配信ではさらに映像キャプチャおよび PC が必要となる。VR180 に対応するカメラは他にも、Kandao 社 VR Cam や、CALF 社 CALF x VISINSE 3D VR180 CAMERA が販売されている。一方で、ライブ配信向けの低価格な VR180 対応ウェブカメラとして、VR.cam 02 が VR.cam 社より販売 [4] されており、片目 1920×1920 の解像度・160 度の FoV を持つカメラを 200 ユーロで購入できる。

2.2 VR180 映像記録・配信に関する研究

全方位カメラを用いた 360 度の VR 映像を用いた研究は多数行われている。Konrad ら [5] は、ステレオカメラを回転させて 360 度全方位のステレオ映像を撮影する SpinVR を開発している。Matzen ら [6] も 360 度ステレオを撮影する低価格な手法として、リコー社の 360 度カ

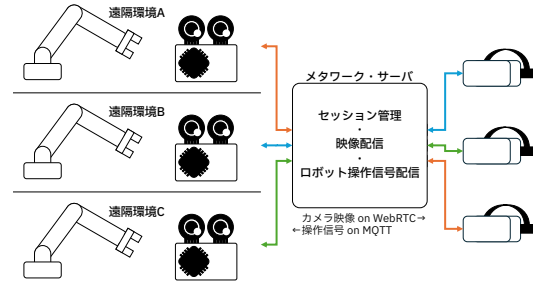


図 1 配信システムの全体図

メラである THETA S を 2 つ並べたカメラシステムを提案している。また、360 度映像はその処理や配信についても研究が行われており、Fan ら [7] は取得・伝送・表示について、Yaqoob ら [8] は低遅延で高品質な伝送を行うための様々な手法についてまとめている。

一方で、VR180 を利用した研究はまだ少ない。Scheuerpflug ら [9] は、2 台の Android スマートフォンを並べて写真を撮って VR180 形式にする方法を提案している。Smolyanskiy と Gonzalez-Franco [10] は、ドローンに 2 つのカメラモジュールを搭載し、nVIDIA Tegra TK1 ボードで映像処理をして地上の VR ヘッドセットへ伝送するシステムを提案している。Urueta と Ogi [11] は、市販の VR180 カメラ (Lenovo 社 Mirage) と Google Daydream ヘッドセットを用いた教育向けのシステムを提案している。このシステムでは 30 秒を超える遅延があったが、重要な問題ではないと結論づけられている。

3 配信システムの検討

上記より、市販機材は高品質でも高価格・低自由度なため、研究では市販のカメラやモジュールを利用しつつ独自のシステムを作る例が多い。本研究が狙うメタワークでは、安価で入手性のよいデバイスをカメラとして利用したい。まず、配信システムの全体像を図 1 に示す。配信システムでは、カメラ側から時雨堂 momo というソフトウェアを使って WebRTC で映像を配信し、WebXR を使った Web ページ上で表示する。WebXR を利用することで、専用アプリ無しで、VR ヘッドセットのブラウザを使い映像視聴・ロボットハンド操作を行える。

ステレオカメラに求められる要件として、左右映像の同期と広い画角がある。VR ヘッドセットでの視聴時に違和感を覚えたり 3D 酔いが起きたりしないように、VR180 映像を構成する左右の映像は同期している必要がある。また、左右それぞれの映像は縦横 180 度に満たない場合に黒枠を付けることになるため、市販のウェブカメラの画角 (参考: ロジクール社 C920 は 78 度) では不十分である。そこで、以下の構成のカメラを試作した。

1. Jetson Orin Nano Super + Raspberry Pi HQ カメラ (一式約 7.5 万円)
2. Intel N100 搭載ミニ PC + VR.cam 02 (一式約 6 万円)

1) 名古屋大学大学院 工学研究科

2) 名古屋大学 未来社会創造機構

表1 計測結果

構成	解像度	fps	平均遅延 (ms)
1 (Jetson)	3840×1080	30	194.7
2 (N100)	3840×1920	30	180.4
参考 (Canon)	1600×1200	30	287.5

Jetson Orin Nano Super は、2 台の CSI カメラモジュールをサポートしており、nVIDIA 社の gstreamer プラグインにより左右の映像を容易に 1 フレームに合成できる。Intel N100 搭載ミニ PC は 2 万 5 千円程度から販売されており、非常に安価に入手できる。またどちらも我々が使用する配信ソフトウェアの時雨堂 momo を実行できる。

構成 1 では、Raspberry Pi HQ カメラから 4K (3840×2160) の映像を取得できるため、両目でハーフ 8K (7680×2160; side-by-side 時) の高精細な映像を取得できる。しかし、ハードウェアエンコーダを備えず、処理能力も限られるため、安定した処理・配信が難しい。構成 2 では、映像処理にハードウェア支援を利用できるため、安定して映像を送出できる。一方で、利用できるカメラが USB 接続のものとなるため、選択肢や解像度が限られる。

4 各構成での配信実験

構成 1・2 ならびに、参考としてキャノン社ミラーレス一眼でそれぞれ配信を実施し、解像度やフレームレート、遅延を確認した。Jetson からの配信では、V4L2Loopback によって生成した仮想デバイスに、gstreamer で取得・合成した映像を流し込み、仮想デバイスを momo の簡易版である sumomo で読み取り配信を行った。N100 搭載ミニ PC からの配信では、USB 接続の VR.cam 02 を momo で読み取り配信を行った。参考の構成は Canon EOS R5 + Canon RF5.2mm DUAL FISHEYE + IO データ GV-HUVC-4KV とし、N100 搭載ミニ PC に接続して momo で読み取り配信を行った。配信先は時雨堂の WebRTC SFU である Sora のプライベートのインスタンスとし、Sora のデバッグページ上のメディア情報で解像度とフレームレートを読み取り、遅延についてはカメラにストップウォッチを映して確認した。

実際に記録を取る前にパフォーマンス調査を行った。Jetson からの配信では、片目 4K で映像を取得すると両目でハーフ 8K の映像となる。片目映像は 30fps で取得されるが、gstreamer で合成を行った際に約 20fps となった。この設定で配信すると、Jetson が十分な速度でエンコードできず、解像度が落ちたため、調整の結果、安定して 30fps で動作する片目 FHD・両目ハーフ 4K で取得・配信を行うこととした。一方、Intel N100 の場合は、設定の調整なく両目 4K で安定して配信を実施できた。

解像度・フレームレート・遅延を 1 分ごと計 10 回記録を取った結果を表 1 に示す。解像度とフレームレートは安定して同じ値であった。遅延は 10 回の平均を示す。参考構成では解像度が低かったが、キャプチャが 4K を出力できても momo の仕様で低解像度になる場合が生じるためである。構成 1 と 2 はほぼ同じ遅延 (30fps ではフレーム間が 33.3ms のため、それ以下なら誤差となる) であり、参考構成より明確に少ない遅延で配信を行えた。以上より、本検討の範囲では、Intel N100 搭載ミニ PC + VR.cam 02 の構成がよいという結果になった。

遅延についてさらに調べると、Sora のデバッグページ

で読み取れる情報ではおよそ 70ms 程度となっていた。したがって、参考構成、構成 1・2 のいずれも、カメラ内の処理やコンピュータでの読み込みなどで 120 から 200ms 程度の遅延が生じていることになる。これは今後解決すべき課題である。

5 おわりに

本項では、VR 機器で遠隔地に接続し、ロボットアームを用いて物理タスクを行うメタワークを実現するための、高画質・低遅延な配信を行える低価格なステレオカメラの構築について検討した。市販の VR180 カメラは高画質でも高価格で自由度が低いため、入手性に優れた安価なコンピュータで VR180 映像の配信を実現したい。

検討では、Jetson Orin Nano Super ならびに Intel N100 搭載ミニ PC を利用するパターンを試作した。結果として、Jetson では、自由度は高いものの高解像度・高フレームレートの両立が難しく、Intel N100 搭載ミニ PC に VR.cam 02 を接続する方が安価で安定的に配信を実施できる結果になった。今後は、適切なカメラの継続的な検討ならびに、数値的な品質の評価を行い、実際に作業環境へカメラを設置しての遠隔作業実験を行う。

謝辞

本研究の一部は、内閣府 SIP3 (JPJ012495) および JST CREST (JPMJCR22M4) に支援いただきました。

参考文献

- [1] バーチャルエコノミー拡大に向けた基盤技術・ルールの整備. METAWORK, Access 2025.05.28. <https://sip3-metawork.com/>.
- [2] Google Spatial-Media. VR180 Video Format, Access 2025.05.28. <https://github.com/google/spatial-media/blob/master/docs/vr180.md>.
- [3] Canon. VR レンズの世界, Access 2025.05.28. <https://personal.canon.jp/product/camera/rf/rf-lensworld/vr-lens>.
- [4] VR.cam. VR.cam 02, Access 2025.05.28. <https://vr.cam/product/vr-cam-02/>.
- [5] Robert Konrad, Donald G. Dansereau, Aniq Masood, and Gordon Wetzstein. SpinVR: towards live-streaming 3D virtual reality video. *ACM Trans. Graph.*, Vol. 36, No. 6, pp. 1–12, 2017.
- [6] Kevin Matzen, Michael F. Cohen, Bryce Evans, Johannes Kopf, and Richard Szeliski. Low-cost 360 stereo photography and video capture. *ACM Trans. Graph.*, Vol. 36, No. 4, pp. 1–12, 2017.
- [7] Ching-Ling Fan, Wen-Chih Lo, Yu-Tung Pai, and Cheng-Hsin Hsu. A Survey on 360° Video Streaming: Acquisition, Transmission, and Display. *ACM Comput. Surv.*, Vol. 52, No. 4, pp. 1–36, 2019.
- [8] Abid Yaqoob, Ting Bi, and Gabriel-Miro Muntean. A Survey on Adaptive 360° Video Streaming: Solutions, Challenges and Opportunities. *IEEE Communications Surveys & Tutorials*, Vol. 22, No. 4, pp. 2801–2838, 2020.
- [9] Felix Scheuerpflug, Christian Aulbach, Daniel Pohl, and Sebastian von Mammen. Experimental Immersive 3D Camera Setup for Mobile Phones. In *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, pp. 881–882, 2024.
- [10] Nikolai Smolyanskiy and Mar Gonzalez-Franco. Stereoscopic First Person View System for Drone Navigation. *Frontiers in Robotics and AI*, Vol. 4, pp. 1–10, 2017.
- [11] Steven H Urueta and Tetsuro Ogi. 3D VR180 Livestreaming System for Semi-Synchronous Distance Learning in TEFL Applications. *Computer Assisted Language Learning Electronic Journal*, Vol. 22, No. 1, pp. 186–206, 2021.