

人物の自由視点画像作成時におけるカメラアレイ最適化設計の検討

三石 悠叶 小池 崇文
法政大学情報科学部

1. はじめに

今後、VR/AR の技術発展に伴い、自由視点画像 (Free viewpoint image 以下、FVI) は、会議や遠隔コミュニケーションに用いられると考える。FVI の作成には、観察者が見る範囲での撮影が必要となる。滑らかに変化する FVI を作成するには、カメラ間隔を狭くした多視点画像が必要である。よって、撮影に用いるカメラを取り付けた装置 (以下、カメラアレイ) に多くのカメラを取り付ける必要があるため、カメラアレイにかかる費用が高価になる。加えて、多くの画像を処理する必要があり、データ量が多いという問題がある。そこで本研究ではこの問題を解決するため、カメラアレイのカメラ配置の最適化手法を検討する。本研究では、観察者が着席している場合のみについて考え、人物を被写体とした際の手法の検討を行う。最適化は、カメラ範囲とカメラ間隔の2つの観点から行った。

2. 関連研究

Taguchi らは 64 台のカメラを用いたカメラアレイ [1] を、Wilburn らは 100 台のカメラを用いたカメラアレイを提案している [2]。一般に、FVI の移動視点範囲を広げるには、カメラアレイのサイズを大きくする必要がある。Chai らは Plenoptic Sampling を提案しており、これによりエイリアシングが起きないカメラの間隔を理論的に求めることができる [3]。Plenoptic Sampling によると、カメラアレイサイズを大きくした場合にも、カメラの間隔は狭いまま、つまりカメラの配置密度を下げることはできない制限がある。

本研究は、実際に人を被写体とした時に、観察者の頭部の移動や FVI の認知特性を元にカメラ間隔や配置の最適化が可能かを試みるものである。

3. 提案手法

本来、複数のカメラ画像とその画像を利用して作成した補間画像を用いて FVI を生成するが、今回はすべて CG を用いて FVI を生成するため、補間画像を作成せず、FVI を作成する。本研究では Taguchi らや Wilburn らの研究と同様に、カメラを被写体の正面に格子状に配置したカメラアレイのみを対象とする。今回は、被写体によって、カメラアレイの最適配置がどうか確かめるため、2つの 3D モデルを仮想空間内に用意し、それぞれの FVI を作成し、最適化を行う。用意した1つ目のモデルは、実在する人物をキャプチャし生成した 3D モデルである。2つ目のモデルは、バーチャル YouTuber である、キズナアイの 3D モデルである。

3.1. FVI の作成方法

FVI の作成方法の流れを以下に示す。

- 1) 仮想空間内に 3D モデルを配置
- 2) 仮想空間内に仮想カメラアレイ (以下、カメラアレイ) を配置
- 3) カメラアレイを用いて 3D モデルを撮影、FYI 群を取得
- 4) 観察者がヘッドマウントディスプレイ (以下、HMD) を装着

- 5) HMD の位置と向きを取得し、対応する FYI を HMD 上に表示

3.2. 画像取得

図 1 に、今回検討するカメラアレイにおけるカメラ配置とパラメータを示す。カメラの間隔を δx , δy とし、カメラアレイの大きさを X , Y とする。図 1 のように、被写体の前面にカメラアレイを設置する。

この配置で FVI に用いる画像を取得する。カメラ配置範囲の最適化では、 X , Y を十分大きな値にして、広範囲から画像を取得する。カメラ間隔の最適化では、 δx または δy を 1mm にし、合わせて、 X または Y を 999mm にして、狭間隔で画像を取得する。

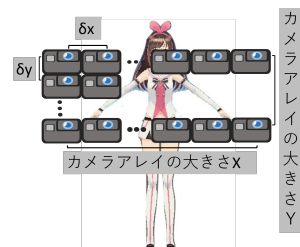


図 1. カメラアレイのカメラ配置とパラメータ

3.3. カメラ配置範囲の最適化手法

撮影する範囲の最適化を行うため、FVI の被写体となる 3D モデル観察時に、どの視点から観察されているか明らかにする。人の着席時の可動域は、座面真ん中を中心に、座高を半径とした円弧の一部であると考えられる。学校保健統計調査報告書によれば、日本人の 17 歳の座高の 95 % 以上は 80cm~98cm である [4]。また、全頭高の平均は 22.5cm である [5]。目は、顔の上下方向中心に位置するため、頭頂から 11.25cm 下に位置する。よって、着席時に椅子の座面中心から 68.75cm~86.75cm の範囲内にほとんどの人の目があるはずである。また、着席時の体の傾きは左右へ 45 度以内に収まると想定する。よって、日本人の目の位置 (以下、視点とする) は図 2(a) に示される範囲内にあると考えられる。例えば、被写体として人体をスキャンした 3D モデルや CG モデルを閲覧する場合には、図 2(b) の閉曲線の範囲内に、観察者の目があるはずである。

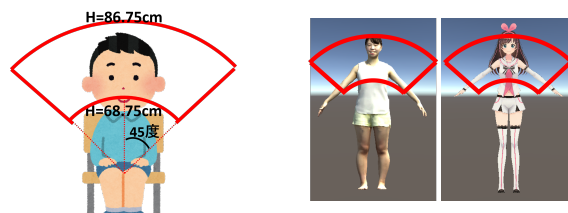


図 2. (a) 着席時視点 (目の位置) の想定移動範囲 (b) 実験に用いる被写体と、観察者の視点の想定移動範囲

3.4. 間隔の最適化手法

Chai らの研究により、理想的な最適カメラ間隔は求められるが、本研究では、実観察者による実験に基づいたカメラ間隔を求める。カメラを 1mm 間隔に一行で配置し、1000 枚

の画像を用いて FVI を作成した。観察者に、FVI を観察してもらう。次に観察者には伝えず、カメラ間隔が 2mm の FVI の観察を行ってもらった。このようにして、提示する FVI のカメラ間隔を x mm づつ広げながら、観察を続けてもらう。FVI のカメラ間隔を広げていくと、観察者の視点移動による画像の変化(運動視差)が滑らかでなくなる。この運動視差が滑らかでないと、観察者が感じたカメラ間隔を実験により求め、この間隔を、FVI 生成時のカメラ間隔の下限の閾値とする。

4. 実験

提案手法の有用性を確認するために、3つの実験を行った。1つ目はカメラ範囲の最適化、2つ目は、1つ目の実験で求めたカメラ範囲が適切であるかの評価、3つ目はカメラ間隔の最適化の実験である。本研究で用いた HMD は Oculus Rift S で、6 自由度方式である。

4.1. 範囲の最適化

カメラの間隔 $\delta x=6$ cm, $\delta y=4$ cm とした。このカメラ間隔は、市販の小型アクションカメラのサイズを想定したものである。最適化していないカメラアレイの画像取得範囲はモデルにより異なる。人物の 3D モデルでは、横-90cm $\sim$ 90cm、高さ 50cm $\sim$ 140cm の範囲の画像を取得する。また、キズナアイの 3D モデルでは、横-60cm $\sim$ 60cm、高さ 80cm $\sim$ 140cm の範囲の画像を取得する。

被写体観察時に、図 2 に示した視点範囲内に、被験者の視点があるのか計測した。被験者には HMD を装着してもらい、自由に被写体を観察してもらった。被験者の視点位置は HMD の位置座標から求めた。被験者は、身長 152cm $\sim$ 185cm、座高 81.5cm $\sim$ 94.5cm の男女 10 人である。被験者の座高別に平均をとった、視点の軌跡を図 3 に示す。水平方向を横軸(x 軸)に、鉛直方向を縦軸(y 軸)に示した。そして、図 2 で示した範囲もあわせて示す。

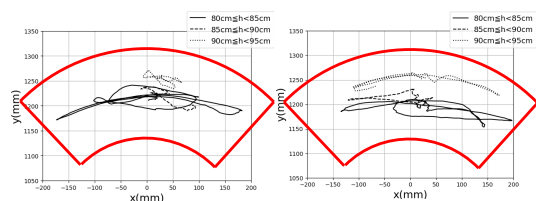


図 3. 観察時の視点の軌跡とカメラアレイ最適化後の視点範囲 (a) 人物の 3D モデル (b) キズナアイの 3D モデル

図 3 より、観察時に左右に揺れる被験者が多く、観察時の軌跡は図 2 に示した閉領域内に収まっている。よって、この領域がカメラアレイの最適配置であると考え、その範囲内にカメラアレイを配置すれば良い。結果として、カメラの台数を最適化前から 57.3 %、キズナアイの 3D モデルを被写体とした場合は、39.8 %削減することができた。

4.2. 有用性の検討

4.1 で求めたカメラアレイ範囲の最適化の有用性を検討した。カメラアレイで取得した画像の範囲内を観察している時間を計測する。最適化後のカメラアレイで作成した FVI を観察した際の範囲内を観察している時間が、最適化前のカメラアレイで作成した FVI を観察した時間と変化がない場合は、範囲を適切に最適化することができたとする。被験者は HMD を装着し、FVI の観察を 15 秒間行う。被験者には、最適化していないカメラアレイで作成した FVI か、最適化したカメラアレイで作成した FVI かを伝えずに実験を行う。今回の被験者は、身長 152cm $\sim$ 176cm の男女 7 人である。最適化したカメラアレイの画像取得範囲は、どちらの 3D モデルでも同じく、図 2 に示した円環の一部である。

被験者ごとの、取得範囲内の画像観察時間の結果を図 4 にまとめた。左が人物の 3D モデルの FVI の観察時間で、右

がキズナアイの 3D モデルの FVI の観察時間である。左の棒グラフが最適化前、右の棒グラフが最適化後である。

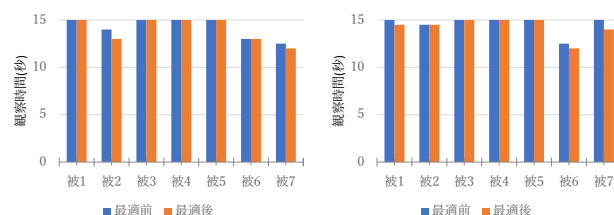


図 4. (a) 人物の 3D モデルの実験結果 (b) キズナアイの 3D モデルの実験結果

帰無仮説を、最適化前後で、取得した範囲の画像を観察している時間は変わらないとし、t 検定を行う。有意水準を 0.05 とした。人物の 3D モデルの FVI の場合は、自由度 6, $t=1.4411$, t 境界値両側=2.446 であるため、帰無仮説は棄却されず、有意差は認められない。そして、キズナアイの 3D モデルの FVI の場合は、自由度 6, $t=1.922$, t 境界値両側=2.447 であるため、帰無仮説は棄却されず、有意差は認められない。よって、提案手法によるカメラ範囲の最適化は適切である可能性が高い。

4.3. カメラ間隔の最適化

3.4 で述べた手法について実験を行った。被写体を図 2(b) に示した人物の 3D モデルとし、キズナアイの 3D モデルとし、横一列と縦一列にそれぞれ 1000 枚、合計 4000 枚の画像を取得した。被験者は HMD を装着し、FVI の観察を行う。カメラ間隔の切り替わりは被験者には伝えず、10 秒ごとに行った。被験者は 4.1. と同様である。

結果は、人物の 3D モデルを被写体とし、横一列の画像で作成した FVI では、平均 5.36mm で被験者は画像が滑らかに切り替わらず、カタついて見ると回答した。縦一列の画像のときは、平均 5.73mm だった。また、キズナアイの 3D モデルを被写体とした、横一列か縦一列の画像で作成した FVI では、どちらも平均 5.27mm だった。それより、FVI の違いがあまり見られず、約 5mm の間隔で被験者は画像が滑らかに変わらないと感じることがわかった。

5. 結論

本研究では、3D モデルを用いて、FVI 作成時に用いるカメラアレイのカメラの範囲と間隔の最適化設計を行った。範囲の最適化では、人物の 3D モデルを被写体とした場合は、カメラの台数を最適化前から 57.3 %、キズナアイの 3D モデルを被写体とした場合は、39.8 %削減することができた。範囲の最適化では、FVI の違いはあまりなく、約 5mm の間隔で観察者は画像が滑らかに変わらないと感じることがわかった。その結果、本手法は FVI を作成する際に用いるカメラアレイの最適化に有効であるといえる。

参考文献

- [1] Y. Taguchi, et al., "Real-Time All-in-Focus Video-Based Rendering Using a Network Camera Array", *IEICE Trans. Information and Systems*, Vol. E92.D, No. 7, pp. 1442-1452, July 2009.
- [2] B. Wilburn, et al., "High Performance Imaging Using Large Camera Arrays", *ACM Transactions on Graphics*, Vol. 24, No. 3, pp. 765-776, July 2005.
- [3] J. X. Chai, et al., "Plenoptic sampling", *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 307-318, 2000.
- [4] 文部科学省, 編. 平成 27 年度学校保健統計調査報告書, 統計表第 1 表、第 4-1 表、第 4-2 表, 2016.
- [5] 河内, 他. 日本人頭部寸法データベース 2001, 産業技術総合研究所 H16PRO-212, 2008.