

定点カメラ映像を用いた市民による動物行動観察における 着眼点の提示効果の検証

吉田信明¹ 塩瀬隆之² 田中正之^{3,4}

概要: 動物園では飼育動物の行動を撮影記録するため、動物舎内に多くの監視カメラを設置している。これらのカメラで自動的・継続的に記録された飼育動物の映像は、動物を観察するための素材として、動物園で提供される様々な教育プログラムでの活用が期待できる。また、インターネット経由で映像配信することで、より多くの人に教育プログラムを提供可能ともなる。しかし、このような監視カメラの定点記録映像はメリハリに欠ける傾向がある。特に、ズームや視点移動などといった、観察の着眼点を強調する撮影・編集は行われていない。一般市民向けの教育プログラムにおいて、このような映像で効果を得るには、映像のどの箇所を見るべきかを「着眼点」として提示することが必要と考えられる。また、記録映像は膨大な量になるため、着眼点の抽出は自動化されることが期待される。そこで、著者らは、このような着眼点の提示手法を検討するため、クラウドソーシングサービスを通じて依頼した実験協力者らに、動物園で記録した監視カメラ映像を配信する実験を行った。この実験では、飼育動物（アジアゾウとツキノワグマ）の5分の映像を続けて2本、インターネットを介して実験協力者らに見てもらった。その上で、それぞれの映像に写っている動物の、気になった行動・面白いと思った行動を映像ごとに最大5件まで、自由記述してもらった。着眼点は、着目すべき時刻を配信画面のシークバー上のマーカーとして提示することとし、その提示の有無による回答内容の違いを、言及された行動の具体性により比較した。提示する着眼点として、画面の動きが変化する時刻を機械的に算出して使用した。時刻は、密なオプティカルフローから画面の動き量を算出し、その時系列データの変化点とした。その結果、1本目の映像において、着眼点を提示したインターフェースを用いた実験協力者の過半数が、最大5件の回答項目のすべてで具体的な行動に言及した。機械的に抽出した着眼点によって、一般市民が具体的な行動に着目する手がかりを提供できる可能性が示唆される。一方、2本目の映像では、1本目での提示の有無にかかわらず、感想などの具体的な行動以外のコメントが多くなる傾向が見られた。このことは、今回提示した情報のみでは、繰り返し提示される定点記録映像において、行動への着目を持続させるには不十分であることを示唆している。

キーワード: 動物園, 教育普及活動, インターネット映像配信, 行動観察, 着眼点

Evaluation of the effect of presenting points of interest in animal behavior observations with fixed-point camera videos

NOBUAKI YOSHIDA^{†1} TAKAYUKI SHIOSE^{†2} MASAYUKI TANAKA^{†3,4}

Abstract: Many surveillance cameras are installed in animal houses to observe and record animal states in zoos. Videos recorded with such cameras will be available as materials for zoo educational programs. However, videos recorded with fixed-point cameras usually lack zoom-in or camera movement, and it is challenging for citizens to observe animal behaviors with such videos. For this problem, we propose presenting videos with computer-generated points of interest. In December 2021, we conducted a video distribution experiment to evaluate the effect of presenting such points and their effective presentation interface. We distributed two videos of elephants and bears recorded in Kyoto City Zoo with points of interest generated based on dense optical flow. Participants collected through crowdsourcing service watched them and reported a maximum of five descriptions about observed behaviors of the animals. We evaluated differences with or without points of interest on the descriptions by the specificity of mentioned behaviors. The result suggests that videos with points of interest presentation will lead citizens to focus on more specific behaviors. On the other hand, it also suggests that such attention may decrease according to repeated presentations.

Keywords: Zoo, Educational and outreach activities, Internet video distribution, Behavior observation, Points of interest

1. はじめに

動物園では、飼育動物の行動を撮影記録するため、動物舎内などに多くの監視カメラを設置している。このようなカメラで撮影された映像には、開園時間帯以外も含めた飼育動物の様子が継続的に記録されている。飼育員や獣医師といった動物園職員は、このような映像を飼育個体の状況

の確認・記録や、繁殖や出産などといった動物に関わる重要なイベントのモニタリング・記録に用いている[1]。特に、飼育動物の福祉の観点から長期的な行動観察を行うために、監視カメラは有力なツールとなっている[2]。

また、動物園の持つ機能として、教育普及活動があり、これらの記録映像は、動物の観察を通じた教育プログラムへの活用なども可能と考えられる。特に、一人一台端末や

1 京都高度技術研究所
ASTEM RI / Kyoto

2 京都大学総合博物館
The Kyoto University Museum

3 京都市動物園生き物・学び・研究センター

Center for Research and Education of Wildlife, Kyoto City Zoo

4 京都大学野生動物研究センター
Wildlife Research Center, Kyoto University

オンライン授業の急速な普及により、教育分野での映像活用が加速度的に進んでいる現実も後押ししている。さらに、本研究で素材としている監視カメラ映像は、元々自動的に撮影される映像素材であるため、教育プログラム向けに別途映像コンテンツを制作しなおすなど担当者に新たな負担を課す必要がない。

しかし記録映像の再活用は負担軽減につながる一方で、その監視カメラ映像は児童生徒の注意を惹くにはメリハリに欠ける。このようなカメラは、通常、飼育員が動物の状態をモニタリングするために、記録したいエリア全体を網羅できる動物舎の天井やグラウンドを一望できる壁の端などに広角レンズを用いて固定設置される。このような位置のカメラで定点撮影された映像は目的をもって見なければ単調な映像が続き、テレビのドキュメンタリー映像を見慣れた視聴者にとってみれば、ダイナミックな動きが少なく、ズームや視点移動などの演出効果もない退屈な映像と判断される危惧がある。このようなメリハリに欠ける映像をガイドランスなしに視聴してしまうと、その視聴操作の記録から、日中の明るく見やすい時間帯に集中し、さらに動作が大きい場面にばかり再生操作が集中してしまう傾向がみられた[3]。一般に映像の見やすさ・動きの多さが必ずしも教育プログラムで観察してもらいたい場面とは一致しないため、このギャップをどのように埋めるかが課題となる。

そこで、このような監視カメラ映像を教育プログラムで利用するにあたって、参加者をプログラムのねらいに則した観察に誘導する映像の提示手法やインターフェースが必要となる。特に、先に述べたように、映像を教育プログラム向けに編集する作業は担当者らにとって負担が大きいことから、自動化可能な手法が期待される。

これに対し、著者らは、映像の解析結果から「着眼点」（見るべき箇所）を機械的に生成し、映像とともに提示することで、適切な視聴に誘導する手法を提案する。現在、映像を解析し、物体や動きを抽出するための多くの手法が開発されている。また、近年のハードウェア性能の向上に伴い、ビデオ処理にかかる時間も短縮されている。適切な解析手法を用いてプログラムに合った着眼点を自動的に抽出することで、日々動物園で蓄積される膨大な映像を、教育に有効活用できるようになると考えられる。

このような手法の検討の一環として、著者らは、行動観察を行う教育プログラムを想定して、適切な監視カメラ映像の提示手法やインターフェースを検討するため、以下の2点の検証を目的とした実験を行った。

- ・ 着眼点の提示の観察結果への影響
- ・ 映像の反復的な提示の観察結果への影響

実験では、クラウドソーシングサービスを通じて依頼した一般市民の実験協力者に、京都市動物園で記録されたアジアゾウとツキノワグマの監視カメラ映像各5分を提示し、映っている「気になった行動・面白いと思った行動」を最

大5件、回答してもらった。協力者の半数には、機械的に算出した着眼点を画面上に提示した。その上で、回答から得られる観察傾向を調べることで、上記検証を行った。

本稿では、この結果について述べる。2節で関連研究について述べる。3節で本検証の方法を説明し、その結果得られた結果を述べる。5節でその結果について、監視カメラ映像の提示手法の観点から考察する。

2. 関連研究

動物園からの動画配信は、広報を目的として行われることが多い。特に、新型コロナウイルス感染症の流行に伴い閉園を余儀なくされる中、多くの動物園がSNSなどを通じた動画配信に取り組んだ[4]。こういった取り組みでは、編集を施した映像が配信されることが多いが、これに対し、本研究では、動物園職員が利用しているような無編集の映像を一般市民を対象とした教育普及活動に利用することを想定している。

このような一般市民を想定した検証を行うに当たり、本研究では、クラウドソーシングサービスを通じて実験協力者に依頼した。現在、このようなサービスを利用した調査が広くなされる一方、収集するデータの質をどのように担保するかが課題となっている。特に、*satisfice* と呼ばれる、協力者が調査に十分な注意を払わない回答行動がみられ、三浦ら[5]は、この問題について、オンライン調査におけるスクリーニング調査の重要性を指摘している。これに対し、注意チェックテストを課題に入れる手法[6]や、操作ログから *satisfice* を検出する手法[7]などが提案されている。また、*satisfice* を防止するために、不真面目な回答を減らす設問順序[8]なども提案されている。一方、本研究では、一般市民による多様な回答行動の存在を前提とし、そのような幅広い行動を適切な方向に誘導する手法を検討することを目的としている。そのため、調査への参加意志の確認以外のスクリーニングは行っていない。

中でも、映像を提示する場合、飛ばしながら視聴する「見飛ばし」がある。著者らの先行調査でも、無編集の監視カメラ映像に対してこういった見飛ばしが観察された[3]。ニュース映像のような構成を持った編集済み映像を提示する調査では全体の視聴が望まれることから、見飛ばしは *satisfice* と見なされる。その場合、*satisfice* とする判断基準として、映像長の半分未満のみの視聴にとどまった場合とする基準も提案されている[9,10]。一方、ここで対象とする監視カメラ映像には明確な構成・構造はない。観察者自らが、映像の中から必要な情報の抽出・構成を行う必要がある。この場合、かいつまんでの視聴は、むしろ、全体像を理解しやすくなる手段とも考えられる。そこで、本研究では、観察の手がかりとして着眼点を提示することで、「よりよい見飛ばし」を誘発することを目的としている。

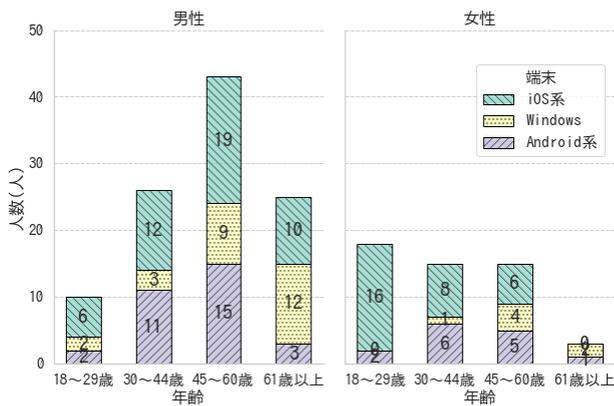


図 1 実験協力者らの属性(性別・年齢)
Figure 1 Attributes of participants (sex & age).

3. 方法

3.1 実施概要

検証は、2021年12月28日11:36~13:11に行った。回答は、webによるオンラインアンケートサービスA社を通じて収集した。提示する映像は、映像配信サービスB社を使用し、配信画面をアンケート画面に埋め込んで提示した。協力者も、A社のクラウドソーシングサービスを通じて依頼した。検証を行った時間帯に同サービスから参加リクエストが298名に送信された。後述する意思確認に対して参加意志を示したのは192名だった。このうち、途中で離脱せずに最後まで回答したのは158名だった。これらの協力者の所要時間は、平均8分9秒、中央値は4分35秒だった。回答は空欄でも可としたが、提示したすべての映像に対して1件以上回答したのは155名だった。

1件以上回答した協力者らの属性を図1に示す。男性は104名、女性51名であった。このうち、男性は45~60歳が最も多く、43名であった。女性は18~29歳が最も多く、18名であった。利用端末は、iPhone/iPadの利用者がもっとも多く、77名でほぼ半数を占めた。Android系端末は45名であり、男性の30~60歳の年齢層に多く見られた。Windows PCの利用者は33名であり、男性の45歳以上の年齢層に多く見られた。

3.2 映像

3.2.1 監視カメラ映像

京都市動物園で撮影した、アジアゾウとツキノワグマの映像を使用した(図3)。映像は、いずれも一般の観覧エリアとは異なる場所に固定的に設置された監視カメラで撮影したものである。アジアゾウのカメラ(Axis M1065-L)は、放飼場の端にある壁の上部に設置されている。ツキノワグマのカメラ(Axis M2025-LE)は、ケージ天井から下方向に向けて設置されている。いずれの映像でも、飼育個体の様子を継続的に観察できるよう、放飼場のほぼ全域が写っている。



a) アジアゾウ



b) ツキノワグマ

図 3 映像の例

Figure 3 Example images of videos.

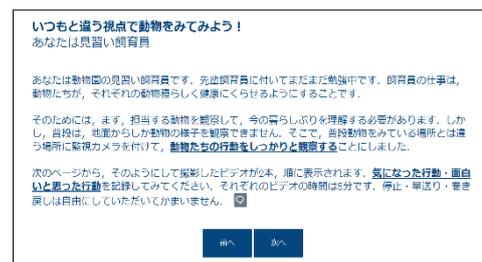


図 2 イントロダクション画面

Figure 2 Introduction screen.

これらのカメラで撮影した映像から、1本につき連続した時間帯の5分の映像を抽出した。行動観察のため、以下のような箇所を選んで抽出した。

- ・ 動きを伴う行動の時間帯がある
- ・ 空間を全体的に使用している
- ・ 複数の行動をしている

3.2.2 着眼点

映像のフレームごとに算出した画面の動き量の時系列データから、その変化点を10点抽出し、映像の冒頭を加えた11点を着眼点として提示した。

動き量は、各フレーム間の密なオプティカルフローから算出した。算出にはOpenCV[11]によるFarnebackのアルゴリズムの実装を使用した。算出結果は2次元ベクトルの配列になっているため、各ベクトル長の総和を当該フレーム間の動き量とした。全フレームについて算出した動き量を、映像のフレームレートに基づいて時間と対応づけて動き量の時系列データとした。

この時系列データに対して、動的プログラミングに基づく変化点検出アルゴリズムを適用し、1本あたり10カ所の変化点を抽出した。算出にはオフライン変化点検出のPythonライブラリであるruptures[12]を用いた。

3.3 検証の流れ

実験協力者ごとに、以下の流れで検証を行った。

(1) 参加意志の確認

実験への参加意志を確認した。選択肢として「参加する」「参加しないがビデオを見てみる」「参加せずこのまま退出する」の3つを提示した。

(2) イントロダクション

観察の背景として、協力者らに「見習い飼育員」という設定を示し、仕事に習熟するため「しっかりと観察」する必要性を示した上で、「気になった行動・面白いと思った行動」を記録する旨を依頼した(図 2)。

(3) 映像の提示・回答の入力

協力者らに提示する2本の映像は、先に述べた動物種(アジアゾウ・ツキノワグマ)・着眼点提示の有無の映像のタイプから、以下の組み合わせで提示した。

- ・ 映像①：
 - a) アジアゾウ・着眼点提示なし
 - b) アジアゾウ・着眼点提示あり
- ・ 映像②：
 - a) アジアゾウ(①と別映像)・着眼点提示あり
 - b) ツキノワグマ・着眼点提示あり

映像①・②いずれも、協力者ごとに a, b それぞれ 50% の確率でどちらかを提示した。この結果、協力者らは 4 パターンのいずれかの組み合わせで、映像①・②の順に視聴した(表 1)。

表 1 映像の提示パターンごとの協力者数

Table 1 Participants per combination of videos.

		映像①		計
		a)アジアゾウ	b)ツキノワグマ	
映像②	a)着眼点なし	45	35	80
	b)着眼点あり	36	42	78
計		81	77	158

これらの映像の提示画面を図 4 に示す。画面上から、説明文、提示映像、回答欄である。協力者は、映像視聴時に自由に再生・停止・早送り・巻き戻し可能とし、説明文には、そのための画面操作の方法を示した。

映像に着眼点を表示する場合は、映像の下部に表示されているシークバー上の着眼点の時刻の箇所に、青いマーカーを表示することとし、そのマーカーの説明を映像の上に示した。ただし、一部の端末・webブラウザではマーカーではない異なる方式で着眼点の表示がなされた可能性があり、この点については 5.3 で述べる。



図 4 実験画面

Figure 4 Experiment screen.

この画面で映像を視聴した協力者は、映像中で「気になった行動・面白いと思った行動」を 0~5 件、回答欄に入力した。以降、ある協力者が一つの映像に行った回答をまとめて「観察結果」とする。

3.4 分析

3.4.1 回答の分類

観察結果には、実験協力者が「気になった行動」や「面白いと思った行動」について、最大 5 件、回答として含まれている。行動観察による教育プログラムでは、より具体的な行動(食べている、水浴びをしているなど)に参加者が注目することを期待すると考えられる。そこで、より具体的な行動に言及した回答かどうかを基準に、個々の回答の分類を行った。なお、回答に記された行動の正しさは考慮しなかった。観察の正しさの評価が目的ではなく、具体的な行動に注目してもらう映像の提示手法の検討を目的としているためである。

分類は、回答で使用されている語彙に基づき以下のように行った。分類には、KH Coder[13](バージョン 3.Beta.04a)を使用した。

① 具体的行動

具体的な行動を示す語・示唆する語を含む回答
語彙例：水浴び、座る、拗ねる、うろうろする

② 抽象的行動

行動に関連しているが、具体的な行動示す語・示唆する語を含んでいない回答
語彙例：動き、行動する、活動する、個別行動

③ 行動以外

上記のいずれにも含まれない回答
回答例：仲がいい、鼻が長くてももしろい、特にない

3.4.2 観察傾向の判定

観察結果ごとに、含まれる回答の分類結果に基づき、以下の基準でその観察傾向を判定した。

- 具体的行動のみ
すべての回答が具体的行動に分類される場合
 - 非具体的行動あり
具体的行動に分類されない回答を1つ以上含む場合
- 以降の分析は、主にこの観察傾向を対象として、着視点の有無や協力者の属性との関連を調べるにより行った。

3.4.3 視聴人数の推移

映像配信サービス B 社から得られた視聴人数の時間推移データ(映像単位)について、着視点との関連等を検討した。

4. 結果

3.4 で述べた分析の結果を示す。なお、検定を行う際の有意水準は、いずれも $\alpha = 0.05$ とした。

4.1 回答数

提示した映像ごとの回答数の分布を示す。すべての映像で、おおよそ半数程度の協力者が、2 件以上回答した。ただし、映像②では、3 件以上回答した協力者はいなかった。

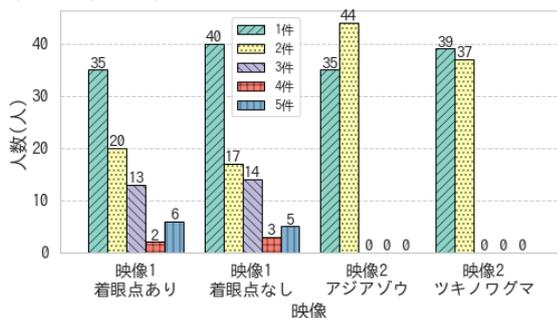


図 5 映像ごとの回答数

Figure 5 Number of answers per video type

4.2 着視点の提示

映像①での着視点提示の有無による観察傾向の違いを図 6 に示す。着視点を提示した場合、具体的行動のみの観察傾向が多くなった。着視点を提示しなかった場合、非具体的行動も含む観察傾向が多くなり、有意な差が認められた ($\chi^2(1) = 4.676, p = 0.031$)。

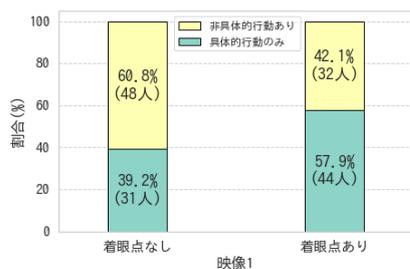


図 6 映像①の観察傾向

Figure 6 Answer tendency on video 1.

4.3 着視点提示効果の持続性

映像①での着視点の有無による、映像②の観察傾向を図 7 に示す。クマ・ゾウいずれの場合でも、映像①で着視点ありの場合(グラフ 2,4)のほうが、具体的行動のみの観察傾向がやや多く、差はクマの方が大きかった。ただし、いずれも有意差はなかった(ツキノワグマ: $\chi^2(1) = 0.322, p = 0.564$), アジアゾウ: $\chi^2(1) = 0.063, p = 0.802$)

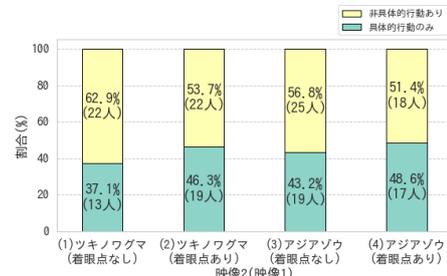


図 7 映像②の観察傾向(映像①のタイプ別)

Figure 7 Answer tendency on video 2 by video 1.

4.4 観察傾向の変化

映像①と②で提示された映像の組み合わせごとに、観察傾向を変えた人の比率を図 8 に示す。映像①着視点なしと映像②アジアゾウ、映像①着視点ありと映像②ツキノワグマの組み合わせでは、多くの協力者が傾向を変えなかった(グラフ 2,3)。その他の組み合わせでは、半数近い人が傾向を変えた($\chi^2(3) = 8.05, p = 0.045$)。

より詳細に、②でアジアゾウを提示した場合について、①での提示映像と観察傾向の組み合わせ別での結果を図 9 に示す。有意差はなかったが(Fisher の正確検定, $p = 0.289$)、①で着視点の提示がなかった場合(グラフ 1,2)、7 割程度が①と②で同じ観察傾向を示した。一方、提示ありの場合、いずれの場合も半数程度の人の傾向が変わった。

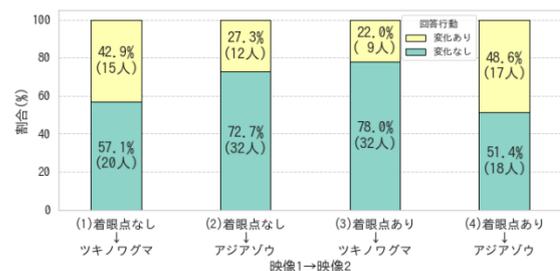


図 8 映像①・②の間での観察傾向の変化

Figure 8 Change in answer tendency between videos 1 and 2.

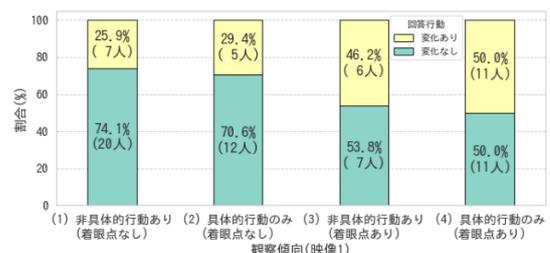


図 9 観察傾向の変化(映像②アジアゾウの場合)

Figure 9 Change in answer tendency in video 2 (elephant case).

4.5 協力者の属性

協力者らの属性(4.1)ごとの観察傾向を比較した。

4.5.1 性別

有意差は見られなかった($\chi^2(1) = 0.162, p = 0.687$)。

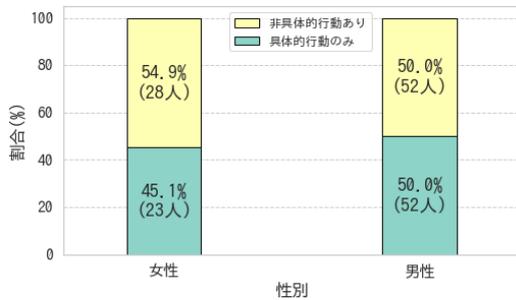


図 10 映像①の観察傾向(性別)

Figure 10 Response tendency on video 1 by sex.

4.5.2 年齢層

30~44歳で過半数が具体的行動のみの観察傾向を示した。ただし、全体として有意差は見られなかった($\chi^2(3) = 1.414, p = 0.702$)。

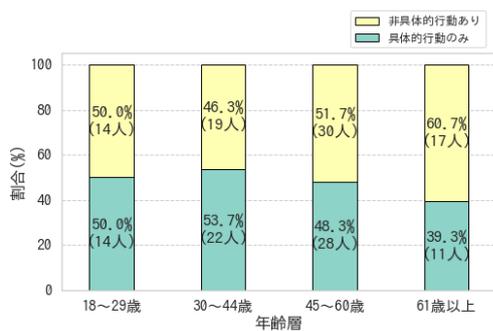


図 11 映像①の観察傾向(年齢層別)

Figure 11 Response tendency on video 1 by age group.

4.5.3 使用端末

iPhone/iPad利用者では過半数で非具体的行動への言及があった。一方、Android系端末とWindows端末では、過半数が具体的行動のみの観察傾向を示した。ただし、有意水準を超える差は認められなかった($\chi^2(2) = 4.078, p = 0.13$)。

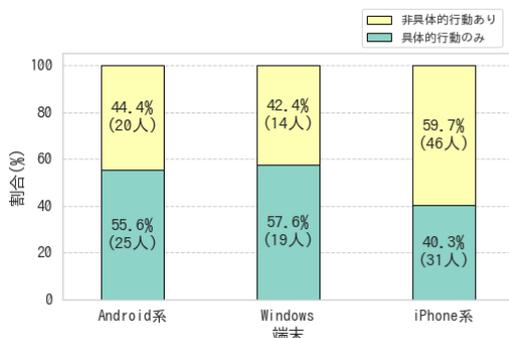


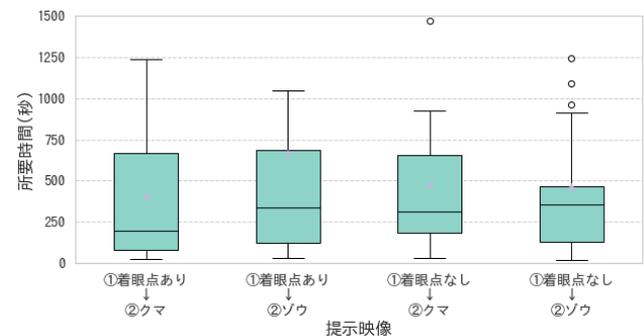
図 12 映像①の観察傾向(端末別)

Figure 12 Response tendency on video 1 by device type.

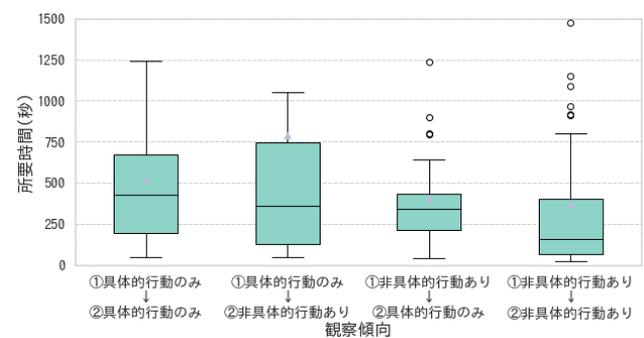
4.6 回答時間

協力者ごとの回答時間の分布について、2本の映像における観察傾向および映像の提示パターンで比較した(図13)。なお、回答時間は、3.3に示した一連の流れすべてを完了するのに要した時間であり、映像ごとの所要時間ではない。そのため、a)映像①・②で提示した映像の組み合わせ、および、b)映像①・②での観察傾向での組み合わせにより比較した。

まず、a)提示映像による比較では、組み合わせの4グループ間で有意差は認められなかった(Kruskal-Wallis検定, $p = 0.599$)。一方、b)映像①・②での観察傾向の組み合わせ別で比較すると、4グループ間での比較では、有意差が認められた(Kruskal-Wallis検定, $p = 0.003$)。そこで、多重比較を行ったところ、2つの映像を通して観察傾向が変わらなかったグループ間、すなわち、いずれの映像でも具体的行動のみの観察傾向だったグループと、非具体的行動に言及したグループ間で有意差が認められた(Wilcoxonの順位和検定, Bonferroni法で補正, $p = 0.004$)。



a) 提示映像別



b) 観察傾向別

図 13 回答時間(1500秒以上は省略)

Figure 13 Time required for answer.

4.7 視聴人数の推移

B社の映像配信サービスから得られた、視聴人数の時間推移(秒単位)を図14に示す。赤い点線は提示した着眼点の位置である。なお、映像②は、その映像を見た協力者すべてのデータであり、映像①での視聴映像にかかわらず合算されている。

映像①では、着眼点の提示の有無に関わらず、視聴人数が維持される時間帯が見られた。また、映像後半の200秒前後の着眼点が集中する時間帯では、着眼点を提示したほうが、視聴人数が維持される傾向が見られた。一方、映像②では、ツキノワグマでは映像①のように視聴人数を維持する時間帯が見られたが、①と同じアジアゾウの映像ではほとんど見られなかった。

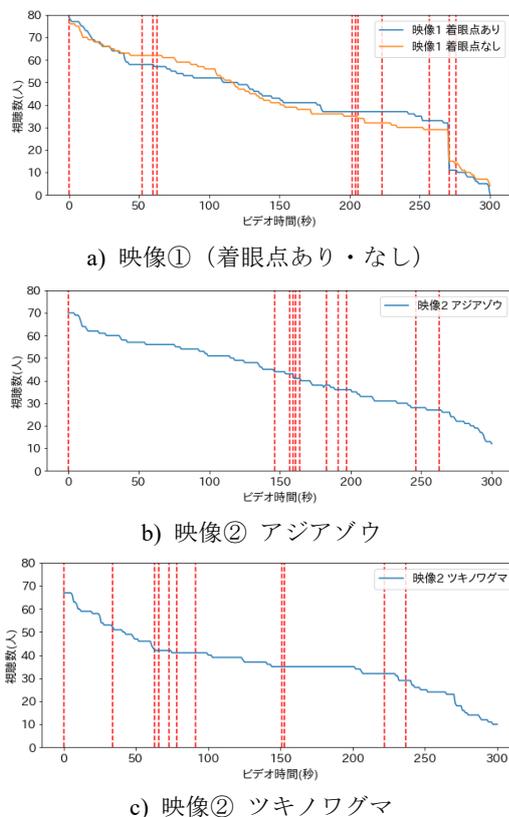


図14 視聴人数の推移

Figure 14 Transition of viewer counts

5. 考察

5.1 観察内容への影響

映像①において、具体的な行動のみに言及した協力者は、着眼点を提示した場合のほうが有意に多かった(4.2)。また、着眼点がまとまって提示された箇所でも、より視聴を維持する傾向も見られた(4.6)。これらのことから、着眼点の提示は、協力者の観察の視点をより具体的な行動に誘導し、一般市民が具体的な行動に着目する手がかりを提供できる可能性が示唆された。

本実験では、映像の画面全体から抽出した、画面の動きに関わる特徴量を着眼点として用いた。映像には、他にも、

動きの方向や速さ、場所、色など、空間的・時間的に抽出できる様々な特徴量が存在する。このような多様な特徴量の着眼点としての利用可能性に加え、それらを効果的に提示できるインターフェースの検討も必要である。

例えば、今回の実験では、一般的なビデオ配信サービスに似た画面上に、時間軸上の点として提示した。この場合、このような点は頭出しのための目印として機能すると考えられる。一方で、一定の時間継続する行動に対して着目してほしい場合には、時間区間として提示するインターフェースも考えることができる。このようなインターフェースでは、視聴行動に一定の変化が起きる可能性がある。

5.2 監視カメラ映像の反復提示の影響

冒頭で述べたように、今回使用したような監視カメラ映像は、動物の行動を網羅的に記録している一方で、メリハリに欠ける課題がある。このような映像を反復的に提示する場合、その効果として観察への習熟が期待される一方で、同じような映像を繰り返し見ることによる関心の低下や飽きなどのマイナス効果も懸念される。今回の検証では、後者がうかがわれる結果となった。

まず、2本目の映像に対する多くの参加者の反応は低下した。2本目での回答件数が減少し、3件以上の回答をした協力者はいなかった(4.1)。また、映像①での着眼点提示の、映像②における観察への影響も明らかではなかった(4.3)。所要時間についても、提示パターンによる差は明確ではなく、着眼点の提示がより慎重な観察を促したとはいえなかった(4.6)。

今回の実験では、見飛ばし操作は可能であったものの、5分の映像を2本、必ず見る必要があり、観察を続けることに対する協力者らの自由はなかった。そのため、協力者の時間的な制約や関心度合いといった要因が影響した可能性がある。回答時間は観察傾向による差のほうが映像の違いよりも明確であり、具体的な行動のみを回答した協力者の方が、それ以外の回答をした協力者より有意に長かった(4.6)ことからもうかがえる。実際の教育プログラムでは、観察への参加方法について自由度の高いシステムを用意することが望ましいと考えられる。

一方で、プログラム構成の工夫も必要と考えられる。映像①で着眼点を提示した場合、映像②で別の動物種を提示すると多くの参加者が1本目と同じ観察傾向を示した一方で、同じ動物種を提示すると半数程度が観察傾向を変えた(4.4)。また、映像②での視聴数の時間推移でも、異なる動物種を提示すると①と似た視聴傾向だったが、同じ動物種では視聴者数は下がり続けた。初めて見る映像に対しては、参加者は観察態度を維持する可能性を示唆している。複数の動物種を観察するなど、参加者が新鮮な気持ちで取り組める工夫も必要と考えられる。

5.3 端末環境の影響

iPhone/iPad 利用者の観察傾向が、そのほかの端末 (Windows や Android 系端末) 利用者とは異なる傾向を示している (4.5.3)。想定される一つの原因として、実験画面 (図 4) 上では、着眼点が青い目印として表示される旨の説明がなされていたが、端末やブラウザによっては、実際にはそのように表示されていなかった可能性がある。iPhone/iPad 系の端末では、iPad では着眼点が青い目印として表示されることを事前に確認していたが、古い OS の iPod Touch ではマーカーが表示されなかったり、iPhone ではマーカーの代わりにリストから選択する方式で表示されたりする場合があった。ただし、iPhone/iPad とそれ以外の端末それぞれの観察傾向についても、全体の傾向 (4.2) と同様、着眼点を提示した方が、具体的な行動に言及する傾向を示していた (図 15)。

このような端末ごとの差異をすべて事前に把握することは困難である。また、先に述べたような、時間区間といったより複雑な着眼点では、今回のような一般的な映像配信画面では適切な提示が難しい。そのため、教育プログラム向けの専用のアプリケーションを作成するなど、参加者にはなるべく同じ環境を提供することが期待される。

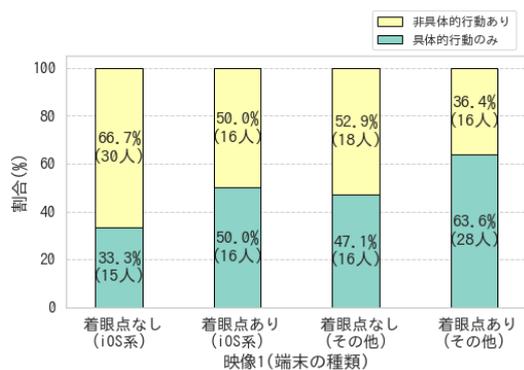


図 15 映像①の観察傾向(端末種別ごと)

Figure 15 Response tendency on video 1 by device type.

6. まとめ

行動観察を行う教育プログラムを想定して、適切な映像の提示手法やインターフェースを検討するため、クラウドソーシングサービスを通じて依頼した一般市民に、動物園で記録された動物の監視カメラ映像を、着眼点を提示しつつ観察してもらう実験を行った。自由記述で得られた観察結果の傾向を分析し、このような着眼点が、具体的な行動に着目して観察を実行する手がかりとなる示唆を得た。一

方、同じような映像を繰り返し提示する場合には、参加者らの関心を持続させるための工夫が必要となることも示唆された。今後、考察で述べたような課題等を検討しつつ、より具体的なプログラムを想定した検証を行う計画である。

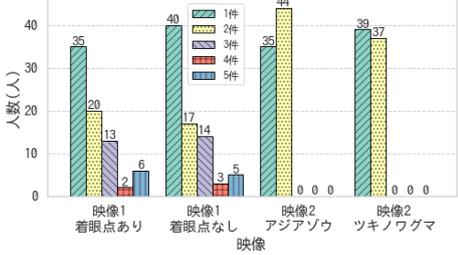
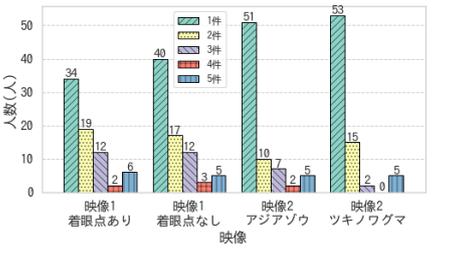
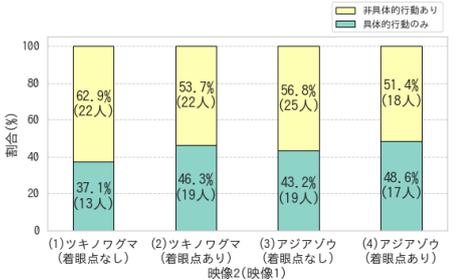
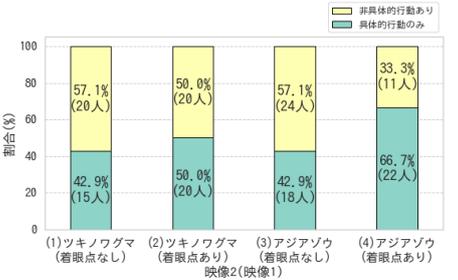
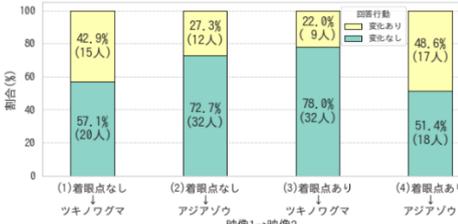
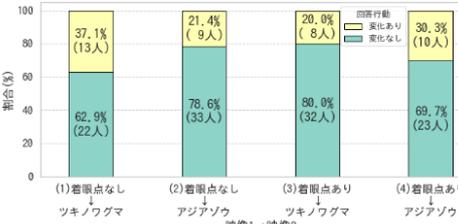
謝辞 本研究は JSPS 科研費 19K01150, 21K18387 の助成を受けたものである。本研究は、京都市動物園の研究協力の下で行われ、映像の提供を受けた(承認番号 2021KCZ-032)。

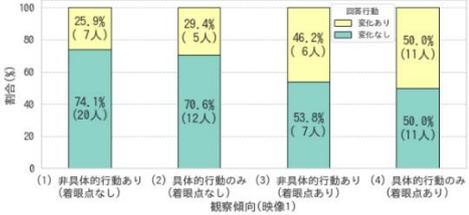
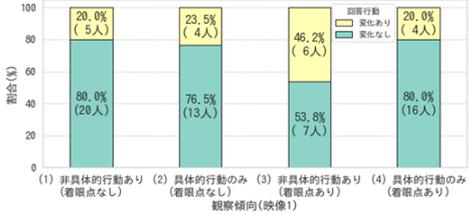
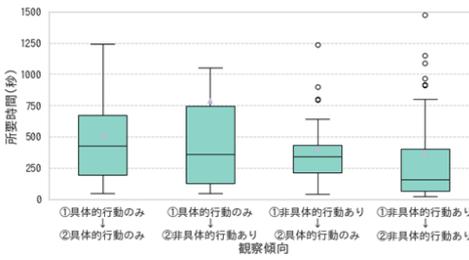
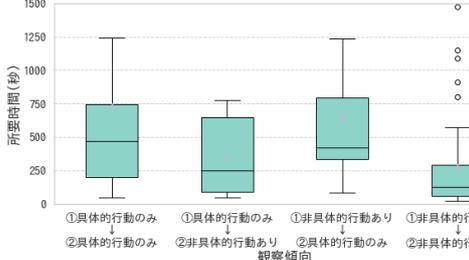
参考文献

- [1] 吉田信明, 田中正之, 塩瀬隆之. 動物園における飼育動物映像の多面的な活用に向けた監視カメラ映像システムの検討. 研究報告情報システムと社会環境 (IS), 2019-IS-150(9), 1-8 (2019-11-16).
- [2] Fazio, J. M., Barthel, T., Freeman, E. W., Garlick-Ott, K., Scholle, A., & Brown, J. L. Utilizing Camera Traps, Closed Circuit Cameras and Behavior Observation Software to Monitor Activity Budgets, Habitat Use, and Social Interactions of Zoo-Housed Asian Elephants (*Elephas maximus*). *Animals* Vol. 10(11), 2026, 2020.
- [3] 塩瀬隆之, 吉田信明, 田中正之, 荒蒔祐輔, 岩橋宣明. 動物園教育コンテンツ制作のための長期モニタリングのログ分析. ヒューマンインターフェースシンポジウム 2019, 2019.
- [4] 公益社団法人日本動物園水族館協会. “ステイホーム” <https://www.jaza.jp/about-jaza/stayhome> (参照 2022-02-06)
- [5] 三浦麻子, 小林哲郎. オンライン調査モニタの Satisfice に関する実験的研究. 社会心理学研究, 31 (1), 1-12, 2015.
- [6] 眞嶋良全, クラウドソーシングを認知科学研究に使うべきだろうか. 認知科学, 26 巻, 2 号, p. 272-281, 2019.
- [7] Gogami, M., Matsuda, Y., Arakawa, Y., Yasumoto, K. Detection of Careless Responses in Online Surveys Using Answering Behavior on Smartphone. *IEEE Access*, 9 53205–53218. 2021.
- [8] 山崎郁未, 伊藤理紗, 中村聡史, 小松孝徳. Web アンケートにおける不真面目回答予防システム実現に向けた自由記述配置の基礎検討. 研究報告ユビキタスコンピューティングシステム (UBI), 2021-UBI-72(34), 1-8, 2021.
- [9] M. R. Maniaci, R. D. Rogge. Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality*, Vol. 48, 61-83, 2014.
- [10] 三浦麻子, 小林哲郎. オンライン調査モニタの Satisfice はいかに実証的知見を毀損するか. 社会心理学研究, 31 (2), 120-127, 2015.
- [11] OpenCV. Open Source Computer Vision Library. 2015.
- [12] C. Truong, L. Oudre, N. Vayatis. Selective review of offline change point detection methods. *Signal Processing*, 167:107299, 2020.
- [13] 樋口耕一. 社会調査のための計量テキスト分析—内容分析の継承と発展を目指して— 第2版. ナカニシヤ出版, 2020.

正誤表

下記の箇所に誤りがございました。お詫びして訂正いたします。

訂正箇所	誤	正
概要・下から3行目	1本目での提示の有無にかかわらず、	(削除)
5ページ 図5 および 説明	ただし、映像②では、3件以上回答した協力者はいなかった。 	ただし、映像②では、多くの協力者が1件のみの回答にとどまった。 
5ページ 4.2節 1行目	映像①での着眼点提示の有無による観察傾向の違いを図6に示す。	映像①での着眼点提示の有無による観察傾向の違いについて、映像①について回答した協力者を対象に集計した結果を図6に示す。
5ページ 図7 および 説明 (4.3節 4行目)	差はクマの方が大きかった。ただし、いずれも有意差はなかった(ツキノワグマ: $\chi^2(1) = 0.322, p = 0.564$, アジアゾウ: $\chi^2(1) = 0.063, p = 0.802$)。 	差はアジアゾウの方が大きかった。ただし、いずれも有意差はなかった(ツキノワグマ: $\chi^2(1) = 0.149, p = 0.699$, アジアゾウ: $\chi^2(1) = 3.307, p = 0.069$)。 
5ページ 図8 および 説明 (4.4節 1段落目 4行目)	映像①着眼点なしと映像②アジアゾウ、映像①着眼点ありと映像②ツキノワグマの組み合わせでは、多くの協力者が傾向を変えなかった(グラフ2, 3)。その他の組み合わせでは、半数近い人が傾向を変えた($\chi^2(3) = 8.05, p = 0.045$)。 	いずれの組み合わせでも、6割以上の協力者が傾向を変えなかった($\chi^2(3) = 3.686, p = 0.297$)。 

<p>5 ページ 図 9 および 説明 (4.4 節 2 段落目 3 行目)</p>	<p>有意差はなかったが(Fisher の正確検定, $p = 0.289$), ①で着眼点の提示がなかった場合(グラフ 1,2), 7 割程度が①と②で同じ観察傾向を示した. 一方, 提示ありの場合, いずれの場合も半数程度の人の傾向が変わった.</p> 	<p>有意差はなかった(Fisher の正確検定, $p = 0.3328$). ①で着眼点を提示した場合(グラフ 3,4), 1 回目で具体的行動のみの回答をした人の多くは, 2 回目でもその傾向を維持したが, 有意差はなかった(Fisher の正確検定, $p = 0.139$).</p> 
<p>6 ページ 左 2 行目</p>	<p>協力者らの属性(4.1)ごとの観察傾向を比較した.</p>	<p>映像①での観察傾向を, 協力者らの属性(4.1)ごとに比較した.</p>
<p>6 ページ 図 13b および 説明 (4.6 節 2 段落目 3 行目)</p>	<p>…一方, b)映像①・②での観察傾向の組み合わせ別で比較すると, 4 グループ間での比較では, 有意差が認められた (Kruskal-Wallis 検定, $p = 0.003$). そこで, 多重比較を行ったところ, 2 つの映像を通して観察傾向が変わらなかったグループ間, すなわち, いずれの映像でも具体的行動のみの観察傾向だったグループと, 非具体的行動に言及したグループ間で有意差が認められた (Wilcoxon の順位和検定, Bonferroni 法で補正, $p = 0.004$).</p> 	<p>…一方, b)映像①・②での観察傾向の組み合わせ別で比較すると, 4 グループ間での比較では, 有意差が認められた (Kruskal-Wallis 検定, $p = 7.596 \times 10^{-7}$). そこで, 多重比較を行ったところ, 2 本の映像いずれも具体的行動のみの観察傾向だったグループに対して, 非具体的行動に言及したグループは有意に所要時間が短かった (Wilcoxon の順位和検定, Bonferroni 法で補正, $p = 1.3 \times 10^{-5}$). また, 1 本目では非具体的行動に言及した協力者で, 2 本目で異なる観察傾向を示したグループ間でも有意差が認められた (同 $p = 7.1 \times 10^{-5}$).</p> 
<p>5.2 節 2 段落目 2 行目</p>	<p>2 本目での回答件数が減少し, 3 件以上の回答をした協力者はいなかった(4.1).</p>	<p>2 本目での回答件数が減少し, 多くの協力者は 1 件の回答にとどまった(4.1).</p>
<p>5.2 節 4 段落目</p>	<p>一方で, プログラム構成の工夫も必要と考えられる. 映像①で着眼点を提示した場合, 映像②で別の動物種を提示すると多くの参加者が 1 本目と同じ観察傾向を示した一方で, 同じ動物種を提示すると半数程度が観察傾向を変えた (4.4). また, 映像②での視聴数の時間推移でも, …</p>	<p>一方で, プログラム構成の工夫も必要と考えられる. 映像の提示の組み合わせに関わらず, 多くの協力者の観察傾向は変わらなかった (4.4). しかし, 映像②での視聴数の時間推移では, …</p>