

視線追跡による好みキャラクターの判定

加藤木 健太[†] 藤堂 英樹[†] 柿本 正憲[‡][†]東京工科大学 メディア学部 [‡]UEI リサーチ

1. はじめに

近年、視線を利用した分析は盛んに行われており、ヒートマップやスキャンパスが分析に利用されている。しかし、これらの手法は、テレビゲーム動画のように動き、かつ興味が更新されるものでは分析が難しい。我々は文献[1]において、テレビゲームにおけるプレイ中の興味対象に応じた映像分岐方法を研究した。視線追跡を利用することでプレイ中の邪魔にならない分岐が可能になり、興味のあるキャラクターへショットが移動するために有効な視線情報を考察した。しかし、実験はキャラクターやカメラが静止した条件で行っており、各視線情報も総注視時間などの合計データしか扱っていなかった。視線の即時性や直観性はさらに精密であり、リアルタイム性が高いことが確認されている。本研究は、さらに細かく時系列で視線情報を計測し、動きに対する視線情報の特性を分析した。その特性から、実際のテレビゲーム動画に近い形で利用可能な興味対象の判定方法を検討する。

2. 関連研究

視線情報を使った興味対象の分析として、青木らのインタラクティブシネマの研究[3]があるが、対象は静止画のみである。本研究では、動画を対象として分析を行うため、視線から得られる情報の特性を考慮する必要がある。Carpenterらは、注視点が注意している場所とは限らないという心眼仮説を提唱している[2]。本研究では、注視情報から指標として有効な情報を得るため、提示動画中のモーションや視線計測時間の設定で工夫を行っている。視線入力による定量的な分析は大きく2つの方法が一般的である。1つ目は、ヒートマップ[4]である。視線の滞留時間をサーモグラフィのように視覚化した分析方法であり、全体の傾向や特徴を把握することを得意としている。2つ目はスキャンパス[5]である。視線の順序と停留時間を番号付きマーカーと軌跡で概略化した分析方法であり、時系列で変化する視線情報を考慮できる点が特徴である。本研究でも動画を対象にしているため、視線情報の時系列変化に着目して分析を行う。

3. 視線情報に基づく好みキャラクターの判定

3.1 興味対象分析の前提条件

本研究で提案する好みキャラクターの判定では、図1のような映像分岐のシーンで判定を行う。分岐元画面はキャラクターが全員収まっているロングショットである。分岐先の画面では、興味対象と判定されたキャラクターに次のショットでカメラを分岐させる。図中の矢印は該当キ

ャラクタに映像が分岐したことを表している。

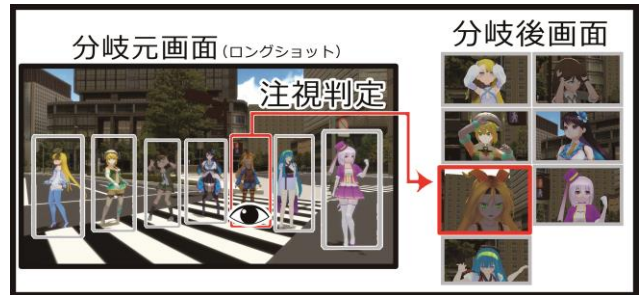


図1: 分岐の流れ。 © UTJ/UCL

3.2 注視判定を利用した興味対象分析の指標

注視判定では、視線入力情報をもとに視線と3Dシーン上の注視対象との交差判定を行う。レイトレーシング法と同様に、スクリーン上の注視点から3Dシーン中の各キャラクターとの視線交差状況を検出する。交差判定は視線計測時間中に随時行い、視線情報を更新する。興味対象の分析の指標として、文献[1]と同様に総注視時間、注視回数、注視持続時間に着目する。既報告では視線計測時間全体での総計データのみを計測していたが本提案では時系列で各指標の状態を記録する。ここで、総注視時間はキャラクターを注視した総合計時間、注視回数はキャラクターを注視した回数、注視持続時間は注視された状態が継続される時間である。注視回数を安定的に計測するため、1秒以上注視した場合のみ、カウントしている。

4. 実験

4.1 実験環境

視線追跡はTobii Technology社の“Tobii EyeX”を使用する。3Dエンジンはシーンの構築や映像分岐に必要な基本機能が揃っている点や、アイトラッカー用ソフトウェア開発キットであるTobii SDK (Tobii Software Development Kit) がサポートされている点でUnityを採用した。実装では、視線と注視対象との交差判定で、SDKに含まれている機能(Gaze Aware Component)を使用した。また、交差判定を簡略化するため、各キャラクターを囲むバウンディングボックスと交差判定を行っている。

4.2 実験設定

被験者は、20代の大学生男女26名(男性22名、女性4名)を対象として実験を行った。被験者は正常な視

“Analysis of Favorite Character by Eye Tracking”
Kenta KATO[†], Hideki TODO[†], Masanori KAKIMOTO[‡]
[†]Tokyo University of Technology [‡]UEI Research

力（矯正含む）及び色覚を有している。精密な視線データを計測するために一人一人に対し、キャリブレーションを行った。明らかに視線情報の取得に失敗している3名のデータは除外し、23名分の解析データで分析を行った。全体で1分半の動画を視聴させ、実験後に興味キャラクターを申告してもらった。どこで視線情報を取得しているのか分からないようにするため、動画後半の30秒間を視線計測区間に設定した。また、動画冒頭部分でキャラクターの顔を覚えてもらい、初見での視線誘導を防いだ。キャラクター数は7体を使用し、毎回ランダム配置を行った。

動きのあるシーンでの視線情報の特性を分析するため、区間ごとに動きの特徴を変えてモーション設定を行った。図2は区間ごとのモーション設定を表し、動きの激しさにもつづいてグラフ化したものである。まず、全視線計測時間の30秒間を10秒間の微動セッションと20秒間の躍動セッションに分けた。微動セッションではほぼ静止している待機モーションの重心微動を設定し、躍動セッションでは各キャラクターがタイミングを変えてバラバラに手を振る、飛ぶなどの躍動をおこなうよう設定した。躍動モーションの前には首ふりを主とする全身微動を設定し、動きの強弱をつけた。設定した区間ごとに総注視時間、注視回数、注視持続時間の指標データを計測し、興味申告結果と合わせて分析を行った。図3に躍動セッションにおける指標データの一例を示す。この例ではキャラクター6が申告された興味キャラクターであり、総注視時間の最大値のみが興味対象と一致している。

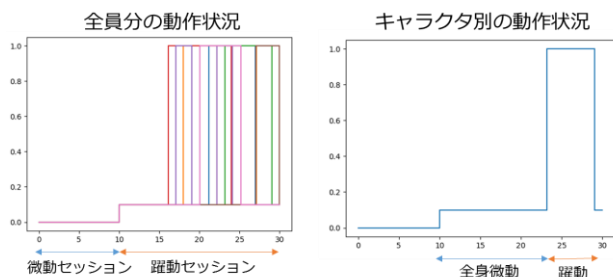


図2：各区間における動きの特徴設定。

	総注視時間	注視回数	注視持続時間	興味申告の結果
キャラクター1	0.8275218	1	0.2326125	
キャラクター2	0.528129	2	1.251609	
キャラクター3	1.170907	2	0.5282088	
キャラクター4	1.014445	1	0.1895224	
キャラクター5	1.230421	1	0.5930003	
キャラクター6	1.880012	1	0.3690501	○
キャラクター7	1.850316	2	0.5669447	

図3：躍動セッションにおける指標データ例。

4. 3 実験結果

表1に各被験者が最も興味を持ったキャラクターの視線情報について、動作設定区間ごとに興味対象判定の正解状況をまとめた。正答数は、指標による判定が興味申告と合致した数を表している。正答率は文献[1]同様に $p=T/N$ と定義して比較を行った。Tは正答数、Nは全被験者数を表す。微動セッションでは文献[1]同様に、重心微動での視線情報を取得する実験を行っている。文献[1]では3体のキャラクターを対象とした実験で総注視時間63%、注視回数37%、注視持続時間68%の正答率とな

っており、7体のキャラクターを対象とした今回の実験でも各指標について同様の傾向が見られた。したがって、人数は正答率に影響しにくいと予測される。また、動きの有無に注目すると、躍動セッションでは微動セッションと比べ注視回数の正答率が高くなり、注視持続時間の正答率が低くなる傾向が出た。各キャラクター個別の躍動では他区間との違いが顕著であり、正答率が極端に低くなっている。原因として、激しい動きをしているものに目が行き、興味が無くても指標値が高くなっているためと考えられる。一方で、全身微動では躍動セッションの中でも総注視時間と注視回数で高い正答率が得られている。これは、躍動と比べて興味の無い対象に視線が誘導されにくいと、興味のあるキャラクターを優先して見ているためだと考えられる。

表1：動作設定区間ごとの興味対象判定の正答率。

		70%以上		50%以下	
		総注視時間	注視回数	注視持続時間	注視回数
視線計測時間の全体(30秒)	正答数	15/23人	15/23人	13/23人	13/23人
	正答率	65%	65%	57%	57%
微動セッション(10秒)	正答数	16/23人	13/23人	16/23人	16/23人
	正答率	70%	57%	70%	70%
躍動セッション(20秒)	正答数	17/23人	17/23人	13/23人	13/23人
	正答率	74%	74%	57%	57%
躍動	正答数	8/23人	10/23人	11/23人	11/23人
	正答率	35%	43%	48%	48%
全身微動	正答数	18/23人	16/23人	14/23人	14/23人
	正答率	78%	70%	61%	61%

5. おわりに

本研究は、視線入力に基づく好みのキャラクターの判定を動画で分析し、効果の高い指標の計測方法を検討するための被験者実験を行った。実験では、文献[1]の3つの指標についてモーションを適応させた場合の比較をし、影響を調べた。新たに実験した躍動セッションにおいても全身微動の区間で高い正答率が得られており、実際に近いモーション有りのシーンにおいても興味対象の判定が可能である見通しが得られた。今後は、分析結果から動作設定区間、各指標に対して重みづけを行い、判定精度の向上を進める予定である。

謝辞

視線追跡システムに関して協力いただいた東京工科大学メディア学部千葉瑞希さんに感謝いたします。

参考文献

[1] 加藤木健太, 藤堂英樹, 柿本正憲, リアルタイムゲームにおける視線追跡を利用した映像分岐, NICOGRAPH 2016 ポスター発表, pp. 153-154.
 [2] Just MA, Carpenter PA, A theory of reading: from eye fixation to comprehension. Psychol. Rev 87, pp. 329-354, 1980.
 [3] 青木美奈, 勝本道哲, 視線はどこまで興味を反映するか - インタラクティブデジタルシネマにおける興味抽出のために -, 情報処理学会マルチメディア通信と分散処理, ワークショップ論文集, pp. 318-322, 2005.
 [4] 畑元, 小池秀樹, 佐藤洋一, 解像度制御を用いた視線誘導, 情報処理学会インタラクション, No8, pp. 57-64, 2014.
 [5] 杉邑洋樹, 上野秀剛, 視線を利用したユーザインタフェースに対するユーザの慣れの定量化, 情報処理学会研究報告, Vol. 150, No9, pp. 1-7, 2012.