

# 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

青木美奈 勝本道哲

独立行政法人情報通信研究機構

視線を用いたユーザーインターフェースの実用化に向けて、様々な研究開発が行われている。その多くは、マウスやキーボード操作の代わりに、視線を用いるというものであるが、最近では、人の興味を視線を用いて判断してユーザを密かに支援するユーザーインターフェースへも開発されてきている。筆者らは視線によって観察者の興味を判断し、それに応じてストーリーが変化していくインタラクティブデジタルシネマを開発した。そこで本報告では、その有効性を確認するために、視線の動きと観察者の興味と、どのような関係にあるのかを実験的に調べ、視線を用いた興味の判定がどのようにあるべきかを検討する。

How do eyes reflect interest?

- for interest detection during watching an interactive digital cinema -

Mina AOKI and Michiaki Katsumoto

National Institute of Information and Communications Technology

For practical use of user interface with eye-gaze, various research and developments had been done. Most of them are applications which use eye-gaze instead of a mouse and keyboard operation. Recently, some user interfaces that detect user's interest by using eye-gaze to implicitly support the user. Authors had developed an interactive digital cinema in which the story can be changed according to his interest detected by his eye-gaze. In this report, we will investigate the relationship between eye-gaze and interest, and will discuss about how the detection of the interest using eye-gaze should be.

### 1. はじめに

筆者らが開発してきたインタラクティブデジタルシネマは、単にデジタルな手法によって制作した映画であるだけではない。従来の映画館で一方的に流される映画とは異なり、鑑賞者自らの映画への興味のあり方がストーリー進行の鍵を握る、個別対応型の新しい種類の映画である。鑑賞者の興味は、鑑賞者の視線の動き方によって判断され、それに応じて切れ目なく異なったストーリーに誘導される。

インタラクティブデジタルシネマのコンテンツは、2つのタイプに分かれる。1つは観察者が自分の意図を明確にストーリー展開に反映させるものであり、もう1つは観察者が見た画面中のオブジェクトによって潜在的な興味を判断して

いくものである。特に1つ目のタイプには、観察者がストーリーの主人公になりきって、作品の中の道を進んでいくものが考えられる。すなわち、主人公が行きたいと思った方向を観察者が見ることで、その方向に主人公が進んでいくのである。

このような第1のタイプのコンテンツの場合、観察者が興味を持った場所が視線と一致しているということを前提としている。しかしながら、興味をもって見ていたはずのところから発展するべきストーリーにうまく進まないという事例が少なからず見られている。視線追跡がうまくできていないという可能性以外に、観察者は必ずしも興味の対象を見ていないという可能性がある。そこで、本報告においては、視

線と興味に関する研究の歴史的な背景を述べたあと、「視線と興味」の関係を詳細に検討するための実験の手法を提案し、その実験結果について検討する。

## 2. 視線と興味

人は外界を認識するために視覚から多くの情報を取り入れている。視覚によって得られる情報量が多いことは、脳内の視覚野の大きさからも推測される。網膜上では、視力が良いのは視角にして $2^{\circ}$ 程度以内の部分であり、中心窩と呼ばれる。周辺視の部分は視力が悪いため、外界を詳細に把握するためには、視線をあちこちに動かして中心窩でとらえる必要がある。このように外界を認識するために生じる随意的な眼球運動は、高速で視線を移動するとき生じるサッカード（衝動性眼球運動）、低速で動くものを追跡するとき生じるスームスパシユート（滑動性眼球運動）、遠近方向に焦点を移動させるとき生じる輻輳性眼球運動に分類される。この中で、サッカードとスームスパシユートは両眼が同時に同方向に運動する（共同性眼球運動）。また、頭部の動きにより注視点が視対象からずれないように頭部の動きと反対方向に眼球を反射的に動かす眼球運動に前庭動眼反射がある。このような眼球運動は日常的に行っているものであるが、そのことを意識することはほとんどない。

複雑な物体を観察しているときの眼球運動についての研究は古くから行われている。Yarbus[1]が計測した人の顔写真を観察しているときの眼球運動の様子を図1に示す。図1の眼球運動の様子から、次々と視線移動しながら写真を観察していることがわかる。また、特に目や口といった特徴点に長く停留しており、スキャナーのように順序よく全体を細かくとらえているわけではない。被験者によって、この視線移動の様子は異なり、また、同じ被験者でも異なっ

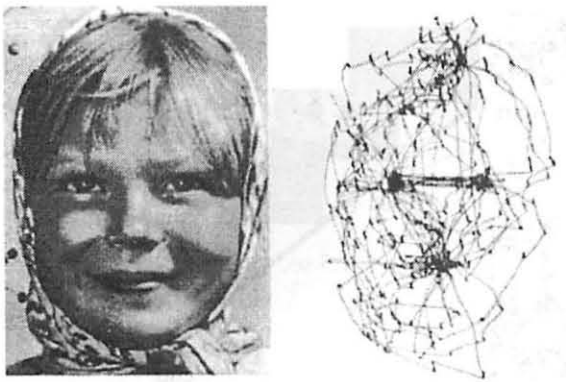


図1 顔写真を観察しているときの眼球運動

3分間、自由に左の絵を観察しているときの、眼球運動を右に示す。(文献[1]より引用)

日に行った実験では似た結果を示すことをYarbusは示している。また、被験者への観察の仕方の指示によっても、視線移動が大きく異なることも示されている。このようなことから、思考が物の見方に影響を与えることが想像される。

また、Inhibition of Return（復帰抑制）という現象も知られており、直前に着目した部分への視線の移動が抑制される。これは、新しい場所に注意を向けやすくするためのメカニズムであると考えられる[2]。

視線の移動は主にサッカードによって行われるが、非常に高速な眼球運動であり、最高速度は毎秒 $500^{\circ}$ 程度に達することもある。しかしながら、サッカードを行っている間に高速に移動する網膜上の画像を通常は知覚することはない（サッカード抑制）[3]。サッカードは、振幅は通常は $20^{\circ}$ 以下で、1秒間に数回の頻度で行われる。視線が停留している注視時間は各 $150 \sim 400\text{ms}$ 程度であり、我々は多数の次々と注視した画像を脳内で統合して全体を把握していると考えられる（サッカード統合）[3]。

眼球運動の近代科学的な観察・測定は1826年にMullerによって初めてなされたといわれている[4]。当初は外部から観察したり、残像の動きを被験者に報告させる方法であったが、1897年にHueyにより、コンタクトレンズ状のものにレバーのついたものを装着して、装置による計測が始まった。その後、20世紀に入り、角膜反射光法と呼ばれる、角膜に光源を照射してその反射光から光学的に眼球位置を算出する方法が開発された。また強膜反射法という黒目と白目の反射光量の違いを利用した方法、サーチコイル法といわれる細いコイルを埋め込んだコンタクトレンズが磁界内で動いたときに生じる誘導電流により計測する方法、EOG法と呼ばれる眼球の角膜と網膜側の帯電の違いを利用した方法、ビデオ撮影する方式などが開発されてきた。当初は眼球運動そのものを計測するのが主目的であったため、アーチファクトとなる頭部運動をさせないために頭部の固定を必要とするタイプが多かったが、近年は頭部の動きによる眼球運動成分を計算によって除外することにより、一定程度までの頭部運動を許容するタイプが主流となってきている。また、被験者の負担を軽くする目的で、非接触で計測できる装置も開発されてきて現在に至っている。

眼球運動計測装置の開発の進展に伴い、視線をユーザーインターフェースとして使用する研究は1980年代から行われてきた[5]–[8]。近年では、手が使えない場面ではマウスよりも速く有効な手段として使えるようなレベルになってきている。しかし、視線を一般の入力用のユーザーインターフェースとして使用するには、困難な側面が存在する。例えば眼をマウスやキーボードを操作する手の代わりに使用する場合には眼は運動器官であるが、一方で眼は前述したように重要な感覚器官でもあり、無意識な眼球運動が

多数含まれる。見たものすべてが入力されてしまっはユーザーインターフェースとしては役に立たないものになってしまう [9]。入力に特化させると外界を把握するための自由な眼球運動を行わない努力が必要となってくる。そのため、この点を克服した装置の開発が必要である。

一方、視線を直接の入力インターフェースとして使用するのではなく、視線を利用した意図の推定に利用する研究も行われている [10]。人間の視線を向けるパターンの知識を用いて、ユーザの興味を知るというものである。特定の用途に限られるが、実用化に向けて進んできている。筆者らが開発してきたインタラクティブ・デジタルシネマもその1つである。

視線を利用して意図を推定するためには、視線は興味と相関がある、という仮定が正しくなければならない。視線はどこまで興味を反映しているのだろうか。視線を向けたところを興味を持ったところであると判断するシステムを有効に利用するためには、実験的に視線と意図がどの程度の相関を持つか、どのようなタイミングで判断するべきかを検討する必要がある。時節では、この目的で行った実験について述べ、さらに、その結果について考察する。

### 3. 実験

#### 3.1 実験方法

##### 3.1.1 被験者

被験者は2名の成人女性である。被験者の視力は眼鏡等による矯正も含め、正常である。実験開始前に実験内容を十分に説明した上で、被験者同意書を提出してもらった。

##### 3.1.2 実験装置

被験者は椅子に座り、眼球運動計測装置を装着し、モニターに呈示される画面を指示に従って観察した。装置の



図2 実験装置

配置を図2に示す。

視線の計測には、(株) ナックイメージテクノロジー社製の EMR-8HM を利用した。これは、頭部にかぶった帽子に装着されたカメラにより撮影された眼球画像から求めた頭部と瞳孔の位置により視線位置を計測をする方式で、サンプリング周波数は 60Hz である。

被験者の眼前 70cm の位置には、23 インチ液晶モニター (SONY 製 SDM-P234B, 1920 × 1200 pixels) を設置してある。被験者が興味を持ったと感じた範囲を回答する手段としてジョイスティックを用いた。被験者の目の位置から見たモニターの左右の広がり視角で 38.9° となる。EMR-8HM の計測可能領域は水平・垂直それぞれ 40° であり、モニターはその領域内にある。

##### 3.1.3 実験課題

被験者が行った実験の流れを図3に示す。被験者がジョイスティックのボタンを押してから1秒後に5秒間の画像が呈示される。被験者は「興味のあるところを探してじっと見る」という指示に従って画面を観察する。1.5秒の画面消去のち、さきほどと同じ画面に赤い丸印が表示されたものが呈示される。被験者は5秒間の観察時間中に最も興味を持って見ていた場所に赤丸をジョイスティックによって移動させ、決定ボタンを押すことで1回の試行が終了する。これを動物や景色の写真画像 20 画面、アニメーションから抜粋した静止画像 44 画面について別の実験セッションとして行った。

ジョイスティックとともに移動する赤円のサイズは固定で、視角にして 1.2° であった。

##### 3.1.4 データの取得と解析

実験終了後に各試行ごとに5秒間の画面呈示時間中の視線データを解析した。被験者がジョイスティックにより回

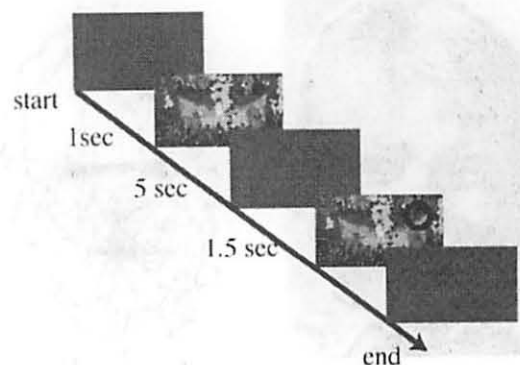


図3 実験課題の流れ

1回の試行の開始と終了は、被験者がジョイスティックのボタンを押すことで決定される。

表1 実験結果の分類

1. 最初から興味のある場所付近のみを見ている。
2. 興味のある場所とその他の場所を何度か行き来して、最後は興味のある場所を見ている。
3. 他の場所を見たあとに、興味のある場所に到達し最後まで見続けている。
4. 興味のある場所とその他の場所を何度か行き来して、最後は他を見ている。
5. 最初は興味のある場所を見ていたが、他に移って最後は他を見ている。
6. 興味があると回答した場所に全く視線を向けていない。
7. その他

答した赤丸の位置に、視線をどのようなタイミングで向けたか、ということに特に着目した。

ところで、視線は実際に見ている物体と必ずしも一致せず、視角で1°程度のずれがあることが知られている[11]。また、注意を惹く刺激が複数あると、その重心位置に視線が向けられるという現象[12]もある。そこで、ジョイスティックで回答した場所の周辺半径2.5°程度の範囲内に停留点が入っていた場合を視線と回答の場所の一致とした。

実験結果を1秒ごとに回答した場所に視線が向けられていたかどうかを判定した。

### 3.1.5 実験結果

5秒間の被験者の視線の移動は、表1のような7通りの傾向に分類される。

被験者への指示は「興味のあるところを探してじっと見る」であるので、1, 2, 3が、ジョイスティックによって回答した場所と視線の位置が一致していて、指示に従って視線の移動が行われたとみなせるものである。

解析したデータを表1の分類に従って整理すると、表2のようになる。ここで、眼球運動が正しく計測できていなかったものは除外した。

表2から、被験者の報告した場所を一度は注視していることがわかる。また、大半は何度かその場所を注視して

いたこともわかる。このように被験者の報告と視線の動きは、かなり一致しているものの、中には、回答した位置を全く見ていない事例もあった。また、複数の注視点の中間位置に回答した場所があったものもあった(7.その他に分類)。

## 4. 考察

実験結果から、被験者は5秒間の呈示時間のいずれかのタイミングで、かなりの確率で興味があると回答した場所を見ていることがわかった。しかしながら、「興味のある」場所だけを見ていた事例はごくわずか(0-5%)で、画面上の他の部分も含めて頻繁に視線を移動しながら観察していることがわかった。また、視線移動の順番も画面によって様々であった。さらに、被験者が事後に回答した場所を、呈示時間の間には全く見ていなかった場合もあった(5-35%)。これらの結果は、筆者らが経験した、「興味をもって見ていたはずのところから発展するべきストーリーにうまく進まない」事例の傍証となった。

これらのことから、ごく自然に向けている視線の動きだけで興味を持っている画像領域を判断するのは非常に困難であると考えられる。しかしながら、指示の与え方、画像の呈示の仕方、視線の検出方法等を工夫することによって、観察者を惹きつける領域にかなりの確率で視線が向くと考えられる。

今回行った実験では、被験者への指示が「興味のあるところを探してじっと見る」というものであったが、一般にインタラクティブデジタルシネマを鑑賞しているときには、そのようなことは考えずに、見たいところを見ている状況のはずである。その場合には、今回の実験結果よりも5秒間の最後の段階で興味をもった場所を注視している(表1の1,2,3)確率は減ると思われる。指示の与え方の実験結果への影響を調べることは今後の課題として残る。

ところで、被験者の実験後の報告の中に、計測装置の帽子的存在によって、自然に画像を観察しているのとは異なった心理状態になること、また、興味を持った領域が画像によっては広がり大きくもったものであるのに、回答は

表2 実験結果まとめ

写真課題、アニメ課題について、ジョイスティックにより回答した場所と視線の注視点の一致の様子を示す。1-7は表1の分類を示す。

	1	2	3	4	5	6	7	合計
被験者A 写真	0	5	1	6	1	7	0	20
被験者A アニメ	2	17	1	13	3	6	1	43
被験者B 写真	1	12	4	1	0	2	0	20
被験者B アニメ	0	26	2	11	0	2	0	41

小さな丸印でしか行えない，ということがあった。そこで今後，非接触型の眼球運動計測装置を用い，また，ペンタブレットにより自由に興味を持った領域を回答できる実験を行い，また被験者の数を増やし，指示の与え方も複数とりにれた実験を行う予定である。

また，その実験の結果から，インタラクティブデジタルシネマにおいて，興味を抽出するために視線の向いている場所を判断するタイミングや時間幅，視線を向けたと判断するための注視領域の大きさ，観察者への伝え方などの検討を行う予定である。

## 5. まとめ

「目は心の窓」，「目は口ほどにものを言う」というように，視線は人の感情や意志を表出するものである。その視線をインタラクティブデジタルシネマに利用するにあたって，視線が興味をどのように反映するかを本報告において調べた。その結果，興味を持った場所を見ている確率は非常に高いものの，画像によってそこを見ているタイミングは様々であった。今後の検討により，さらに観察者がストレスなくインタラクティブデジタルシネマを体験できるようになると考える。

## 参考文献

- [1] A. L. Yarbus : Eye Movements and Vision, Plenum Press, New York, pp.171-196 (1967).
- [2] 日本視覚学会編：視覚情報処理ハンドブック，朝倉書店，p. 460 (2000).
- [3] 塩入諭：サッカード抑制，視野安定およびサッカード統合，「眼球運動の実験心理学」(荻坂，中溝，古賀 編)，名古屋大学出版会，pp.101-122 (1993).
- [4] 古賀一男：眼球運動実験ミニ・ハンドブック，(財)労働科学研究所出版部 (1998).
- [5] R.A. Bolt : Eyes at the Interface, Proceedings of the 1982 conference on Human factors in computing systems, pp.360-362(1982).
- [6] C. Ware, H. H. Mikaelian : An Evaluation of an Eye Tracker as a Device for Computer Input, Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface, pp. 183-188 (1987).
- [7] R. J. K. Jacob : What you look at is what you get: eye movement-based interaction techniques, Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people, pp. 11-18 (1990).
- [8] L. E. Sibert, R.J.K. Jacob : Evaluation of Eye Gaze Interaction", Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 281-288 (2000).
- [9] 大野健彦：視線を用いたインタフェース，情報処理，Vol.44, No.7, pp.726-732 (2003).
- [10] P. Qvarfordt, S. Zhai : Conversing with the User Based on Eye-Gaze Patterns, Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 221-230 (2005).
- [11] D. D. Salvucci and J. R. Anderson : Automated eye-movement protocol analysis, Human-Computer Interaction, vol. 16, pp. 39-86 (2001).
- [12] J. M. Findlay : Global visual processing for saccadic eye movements, Vision Research, vol. 22, pp. 1033-45 (1982).