

人間の内部状態を顕在化する視覚的インタラクション

平山 高嗣^{1,a)}

概要: 視線分析に関する近年の技術進展と長年の心理物理学知見の蓄積により、視線計測を応用したインタフェース設計への関心が高まっている。心の窓と言われるように、目から人間の内部状態にアクセスすることができるが、多次元変数によって記述されるであろう内部状態とそれに比べれば低次元情報を持つ視線との関係は一意に定まらない。その関係を解きほぐす鍵はコンテキストであり、外環境と視線それぞれのダイナミクスの関係が注目されている。著者は、外環境からの視覚的な働きかけに対する人間の反応に基づいて内部状態を推定する Mind Probing を提案しており、本稿ではそのコンセプトに沿って分析、設計された人間の内部状態を顕在化する視覚的インタラクションの研究動向を紹介する。

1. 視線の魅力

人間と共生するシステムを構築することはコンピュータサイエンスにおける主要な課題の一つである。システムが人間と協調的に活動するためには、人間の振る舞いを外部から観測し、興味や意図などの内部状態を推定することが重要となる。その手がかりとなり、内部状態を反映しうる振る舞いとしては、発話、表情、姿勢、動作などが挙げられ、視線もその一つである。

古くから、「目は心の窓、口ほどにものを言う」と言われ、19世紀後半から心理物理学、工学分野を中心に研究されている。目は多くの情報を入力し、かつ内部状態を出力するインタフェースであるため、視覚情報をメディアとするヒューマンコンピュータインタラクションを設計する上で、視線は非常に重要な要素とされている。システムは時間的に連続するデータとして視線を計測でき、人間はその中に無意識的に内部状態の情報を漏らしてしまうため、コマンドやポインティングだけではない人間同士のやりとりを模した自然なインタラクションを実現できる可能性がある。しかしながら、人間を外から受動的に観察するだけで内部状態を推定することは困難である。我々人間でさえも観察だけに頼って他者の内部状態を推定することは困難なことが多い。

では、人間はどのようにして他者の内部状態を推定しているのだろうか？人間同士の対話に目を向けると、相手の内部状態を正確に理解したい場合に、積極的に働きかけを行って、その反応を観察していることが分かる。そこで著

者らはこれに類推して、システムが主導権を持って積極的に対話を行うプロアクティブな対話ストラテジーを適用し、システムがユーザに視覚的な働きかけを行い、それに対するユーザの反応から内部状態を推定するという視覚的インタラクション「Mind Probing」を提案してきた [1]。本稿では、一世紀を超えて行われてきた視線研究の進展を概観し、著者らが取り組んできた Mind Probing に関わる研究動向を、インタラクションの参与対象が人と人、人とシステム、人と環境である場合に分類して紹介する。

2. 視線研究の歴史

2.1 工学、心理物理学分野の動向

視線は19世紀後半から工学、心理物理学分野を中心に、主に計測と分析の観点で研究されている。視線を正確に計測するためには、個人差があり外部から計測することが困難な眼球形状を推定する必要がある。また、注視行動は不確定性を多く含む人間の内部状態の影響を受けやすい。そのため、その進展はゆっくりしたものとなっている (図1)。

現在の視線計測技術の主流である赤外光照射法の原型といえる光源角膜反射法が1901年に Dodge & Cline によって提案された [2]。また、対話分析で現在良く用いられている頭部装着型のものについては、その半世紀後に Hartridge & Thomson によって開発された [3]。視線分析への視線計測装置の導入は1950年代に入ってから一般的になり、その多くは眼球運動の分類やコックピットなどのユーザビリティの評価を目的とした。1970年以前は、学習、記憶、注意、認知的負荷といった認知要因に入り込まない研究が主であった。

1960年代後半から工学分野での高まりが一旦落ち着く。

¹ 名古屋大学大学院情報科学研究科
愛知県名古屋市千種区不老町, 464-8603
^{a)} hirayama@is.nagoya-u.ac.jp

その一方で、心理物理学分野において認知要因を考慮した研究が盛んに行われ、神経科学分野ではサルを用いた視覚分析の研究が活発化した。Yarbus は、注視行動が観察者に課された心的課題に敏感に反応することと、顔などの視覚対象においては、自由視で注視箇所がほぼ共通する場合があることを明らかにした [4]。また、Just & Carpenter は、物体や文字の比較という認知要因に注目し、注視行動の理論的モデルを提案した [5]。共同注意という対話現象が研究対象として注目され始めたのも 1970 年代のことである。

視線計測のリアルタイム処理が可能になった 1970 年代後半から、視線を入力モダリティとするインタフェースが開発され、ヒューマンコンピュータインタラクションへの視線の介入が注目され始める。心理物理学分野においても、視覚的注意をめぐって 2 つの大きな進展があった。まず、Posner によって提案された「先行手がかり (pre-cueing)」を用いた実験パラダイム [6] が視覚的注意に関わる研究を活発化した。そして、工学分野にも影響を及ぼす 2 つのモデルが提案された。Treisman & Gelade による特徴統合理論 (Feature integration theory) [7] と Koch & Ullman による Saliency map [8] である。

1990 年代に入ると、ユーザビリティ評価への取り組みが再燃する。特徴統合理論に対抗するモデルや Saliency map に基づく視覚的注意の計算モデルも多く提案される。対話研究においては、心の理論獲得へと導かれるモジュールとして、視線によるインタラクションを内包するモデルが Baron-Cohen によって提案された [9]。この期間は、各分野における盛り上がりのピークと言えるかもしれない。

近年は、視線計測装置の高精度化、小規模化、低廉化と各分野での知見の蓄積が相まって、学際分野での研究が活発化している。実験室から外に出た環境での視線計測や多人数対話における視覚的インタラクション分析、視覚的注意の一人称視点モデル設計、霊長類の視線分析、機械学習を導入した視覚的注意や注視行動の計算モデル構築といった研究も盛んに行われている。

2.2 視線と内部状態の関係

人間の内部状態は多様な認知要因を基底とする一方で、視線は低次元情報であるため、それらの関係を一意に定めることは困難であるが、制限された状況や特定のタスクに対しては、関係が明らかになりつつある。

その中で、知見が蓄積されているものとして、まず、インタフェースのユーザビリティ評価を目的とした注視行動の分析が挙げられる。インタフェースから情報を抽出する難しさが固視の持続時間で説明できるという結果が得られている。また、インタフェース全領域での固視回数がモジュールの検索効率と相関し、各モジュール領域に対する固視回数が各モジュールの重要性を表すとされている [10]。これらの関係に関わる重要な認知要因はインタフェースへ

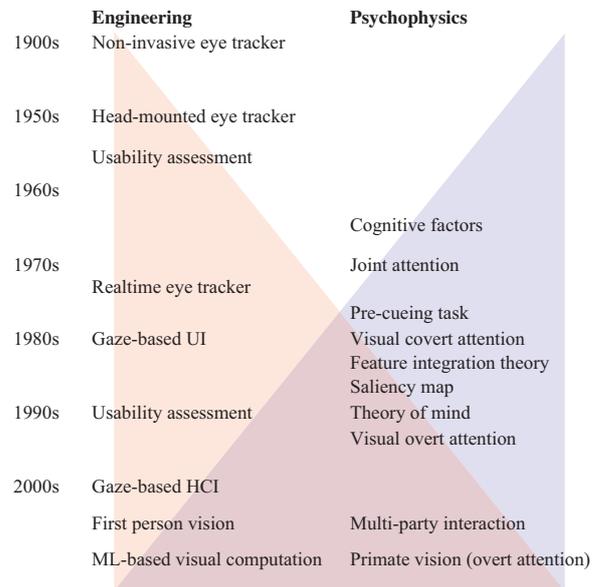


図 1 工学、心理物理学分野における視線研究の動向

の慣れである。

次に、自動車などのシステム操作時の安全性評価を目的として、認知的負荷 (cognitive distraction) と注視行動との関係が盛んに分析されている。例えば、自動車運転などの主タスクと並行して暗算やヘッドセットを用いた音声遠隔対話といった副タスクを被験者に課し、主タスクのみの場合との注視行動の差異が分析される。主タスクより副タスクに多くの注意資源が割り当てられるほど注視行動が単調になり、自動車運転状況であれば注視方向が視野前方に集中するという関係が一般的知見とされている [11]。

視覚探索課題においては、注視点を固定して被験者に視覚刺激を探索させ、探索達成時間を計測することで潜在的注意 (covert attention) の分析を行うことが主流であった。視線計測とパターン認識の技術進展とともに、近年は注視点を固定せずに、視線滞留頻度から生成される注視点マップや注視点の時系列パターン (scan path) に基づいた顕在的注意 (overt attention) の分析が多くなっている。いずれの場合にも、探索の先行手がかり (事前知識) の提示の有無と注視行動との関係が比較される [12]。また、注視点マップやスキャンパスをシミュレートする計算モデルも多々提案されている [13]。

上記ほどの知見の蓄積がないものの、視覚対象への興味とその対象への累積注視時間 [14] や注視頻度 [15] が関連するという分析結果がいくつか得られている。

3. Mind Probing

外環境からの視覚的な働きかけに対する人間の反応に基づいて内部状態を推定する Mind Probing では、外環境と視線それぞれのダイナミクスの関係、特に各々のイベント間の時間関係 (タイミング) に注目する。先行手がかりを用いた視覚探索実験パラダイム [6] と自動車運転時の認知

的負荷を評価するための Peripheral Detection Task[16] において、外部イベントに対する反応時間が計測されることや、対話における発話交替潜時から相手の気持ちを読み取れることから、内部状態を反映するタイミングという人間の時間感覚が重要であることが分かる。内部状態とタイミングの関係をモデル化できれば、あうんの呼吸で人間とやりとりするシステムを設計できる可能性がある [17]。

3.1 人間同士の視覚的インタラクション

3.1.1 視線投げかけが同意・不同意応答に及ぼす影響 [18]

人間は対話相手の意図や興味といった内部状態が表面化していない場合に、それを理解するために相手に対して働きかけを行って、反応を観察している。働きかけは基本的には言語を発話することによって行われるが、対面対話においては、視覚に作用する身体動作にも重要な効果があると考えられる。そこで、働きかけに頻繁に付随する相手への顔向け（視線の投げかけ）に注目した。人間は相手に視線を向けることで自分の行動に対するフィードバックを求める傾向がある。従って、顔向けは、働きかけによる内部状態の確認を明示化する信号であると言える。従来研究において、対話の発話交替潜時が、発話の韻律などの非言語情報と同様に話者の内部状態と関係が強いことが示されている [19], [20]。この関係は、先行話者による働きかけに顔向けが付随するか、しないかによって変化する可能性がある。そこで、対話相手の内部状態を理解する必要がある対話として、働きかけと応答が繰り返される合意形成対話を観測し、音声と映像を解析した (図2)。

その結果、顔向けを伴う働きかけに対する不同意の応答タイミングが同意よりも有意に 500msec ほど遅くなり、それらのタイミングの差が顔向けを伴わない場合より大きくなることが明らかとなった。顔向けは、対話者に対話の時間的構造を強く意識させるトリガになっていると考えられる。さらに、顔向けを伴う場合は伴わない場合より、同意応答を表す頷きなどの身体動作が促進され、不同意応答を表す首振りなどの身体動作が抑制されるという結果が得られた。顔向けによって、身体動作という対話チャネルが開かれたことで、このような対話特徴が現れたと考えられる。これらの結果より、顔向け、つまり相手に視線を投げかける行為が相手の心を反映した振る舞いを誘発することが確かめられた。この知見を応用し、システムが情報提示や質問などの働きかけをユーザに対して行う際に、顔向けのようなユーザを観る行為を共に用いることで、ユーザの反応をより強められることが期待できる。

3.2 人間とシステムの視覚的インタラクション

3.2.1 Gaze Probing: イベント提示に基づく注視オブジェクト推定 [21]

情報提示システムのディスプレイ上に提示されたオブ

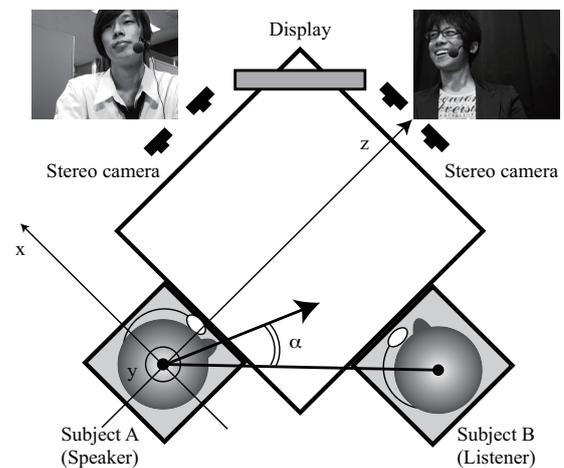


図2 対話相手に視線投げかけを行う状況設定

ジェクトのうち、ユーザが実際にどのオブジェクトを注視しているかという情報は、ユーザの興味や関心を知る上で重要な手掛かりである。しかしながら、その計測精度はユーザの立ち位置や顔向きなどの自由度とのトレードオフであるため、計測された注視座標をそのまま用いるだけでは、ディスプレイ上に提示できるオブジェクト数を少なくしなければ、ユーザの注視オブジェクトを推定することが困難となる。そこで、提示するオブジェクトに移動、停止といった特徴的な動き（イベント）を持たせた動的コンテンツを用いて、注視対象と同調した眼球運動を引き出し、そのタイミングを分析することによって注視対象を推定する新たな手法 Gaze Probing を提案した。

Gaze Probing では、眼球運動が「いつ」発生したかという時間的な情報を抽出し、各オブジェクトのイベント発生時刻との同期の程度を計測する。そのために、計測誤差が大きな視線計測機器においても眼球運動の反応を観測できる図3のような特徴的な短時間のイベントをデザインし、ユーザの注視オブジェクトを探るためのプローブとして、各オブジェクトの動きの一部に組み込む。このような動的コンテンツを提示されたユーザは、無意識に注視オブジェクトのイベントに反応し、システムはその眼球運動の発生時刻を検出できる。計測された注視座標を直接用いる従来手法や眼球運動の変位パターンを用いる従来手法と比較して、視線計測の精度が十分に望めない状況でも注視オブジェクトを高精度に推定できることを確かめた。

3.2.2 能動的な情報提示に対する眼球運動の反応遅延に基づいた興味推定 [22]

人間は空間に並べられた様々な視覚情報から好みのものを選択する状況において、まず、(1) 全体をざっと俯瞰し、その後に、(2) いくつかの情報に吟味する傾向がある。吟味状態では、興味がある情報に能動的に注意を向ける。情報提供システムが提示コンテンツに対するユーザの興味を推定する場合、まずユーザが吟味状態であることを認識する必要がある。静的なコンテンツに対しては、それを構

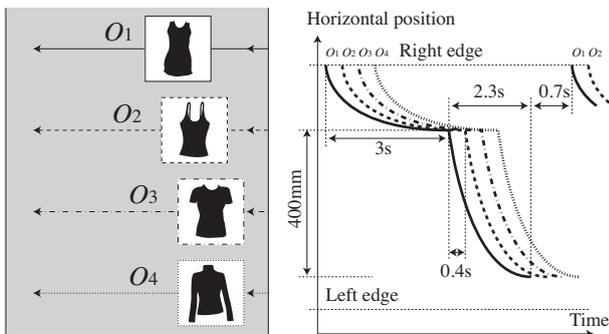


図3 Gaze Probingを行うオブジェクトに導入したイベントの一例

成する要素オブジェクト間の意味的關係に基づいてコンテンツレイアウトを階層構造で表現し、視線運動をこの構造上での注視パターンとして記述することで、吟味状態を70%程度の精度で識別することができた [23]。一方で、動的なコンテンツ設計によって、ユーザを吟味状態に強制的に誘導し、興味がある対象に能動的に注意を向ける内因性サッケード (endogenous saccade) を誘発することができると考えた、下記がそのカラクリである (図4)。

複数のページで構成されるコンテンツを複数提示する状況を想定する。システムはまず、(1) コンテンツを1つ表示して短い時間間隔でページを切り替え、続いて、ディスプレイの別の位置に他のコンテンツを同様に表示する。これを全てのコンテンツについて繰り返すことで、ユーザは一通りの情報を俯瞰することができ、脳内の認知地図にはコンテンツの空間情報が記録され、続くフェーズで吟味状態が生じやすくなることが期待できる。次に、(2) これまでに提示したコンテンツを同じ場所に再び表示する。ただし、ユーザが興味のあるコンテンツを十分に読み取ることができるよう、ページ切り替えの時間間隔を長く設定する。コンテンツ間のページ切り替えの順番は、ユーザに予想させないためにランダムとする。ユーザは、認知地図を参照することで、興味があるコンテンツのページ切り替えに持続的に注意を向けることができる。これによって、サッケードは主に内因的に表出されることが期待できる。

興味の推定には、吟味フェーズにおけるコンテンツのページ切り替えイベントとサッケードの発生イベント間のタイミング構造に注目する。コンテンツ選択を行う被験者実験を行ったところ、興味があるコンテンツを注視している場合に、他のコンテンツのページ切り替えへの反応が遅れる傾向があることを確かめた。この反応遅延に基づいて、最も興味のあるコンテンツを推定したところ、従来研究で用いられている注視時間と注視頻度より高い精度が得られた。注視時間や注視頻度は、静的な情報提示に対して見比べの注視行動が頻繁に起こった際に良い精度が得られる。提案した動的な情報提示法は、ユーザに外因的な刺激を無視させ、興味のある情報を能動的に獲得させることを促すと考えられる。

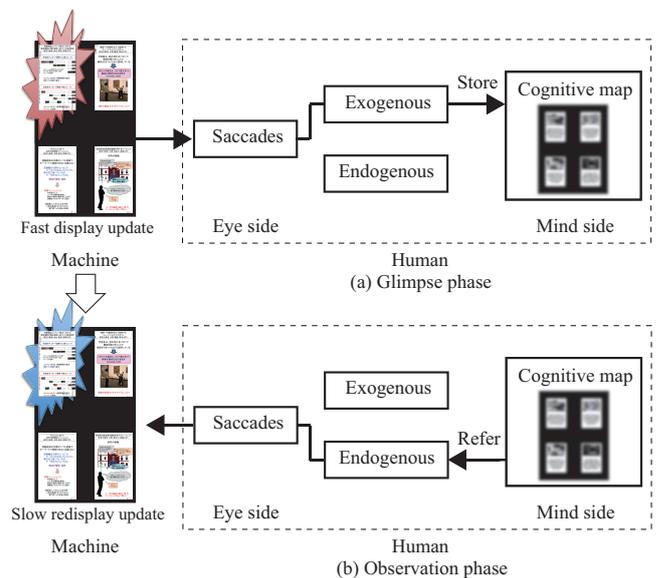


図4 興味を反映した眼球運動の反応遅延を引き起こす (a) 俯瞰フェーズと (b) 吟味フェーズにおけるインタラクション

3.2.3 Gaze Mirroring: ユーザの興味を顕在化させる注視模倣 [24]

プロアクティブな対話戦略は、お節介さや落ち着きの無さを受け手に感じさせてしまう面がある。リアクティブな対話戦略でユーザの内部状態を反応に引き出すことはできないだろうか? 言語を用いたリアクティブインタラクションは、ユーザが発話しない限り対話が進まないが、視線はユーザが目を閉じ続けられない限り常にどこかに向けられているため、それに対してシステムは何らかのリアクションを取ることができる。そこで、ユーザの反応を引き起こすリアクションとして、擬人化エージェントがユーザと共同注意を実現する注視模倣 Gaze Mirroring を設計した。

共同注意は、他者の心を理解するプロセスの一つであり、関心を共有する事物へ他者と共に注意を向けるように行動を調整することである。これに基づけば、提示情報へのユーザの注視行動を、システムが擬人化エージェントを用いて模倣することで、ユーザの注視行動に影響を与え、興味を顕在化させられる可能性がある。ユーザが注視している対象をシステムが同調かつ同期して注視すると、ユーザがその対象に興味を持っている場合は共同注意を行うことに抵抗がなくそのまま注視を続け、興味を持っていない場合には注視を外す回避行動を取ることが期待できる。つまり、注視対象のコンテンツ領域への注視滞留時間や頻度にユーザの内部状態が顕われることが予想できる。

情報提示システムが複数のコンテンツをディスプレイに提示し、ユーザの注視行動に応じてディスプレイ上の目玉エージェントの視線を変化させることで Gaze Mirroring を行ったところ (図5)、興味があるコンテンツ領域への注視持続時間を Gaze Mirroring を行わない場合に比べて

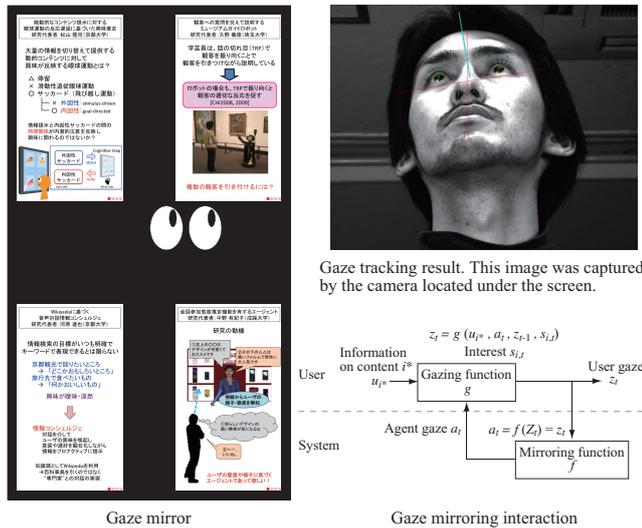


図 5 Gaze Mirroring とそのインタラクションモデル

600msec 増加させた。また、各コンテンツ領域への累積注視時間に注目し、最大の累積注視時間が割り当てられたコンテンツと興味があるコンテンツとの一致率を求めたところ、Gaze Mirroring を行った場合で 80%程度、Gaze Mirroring を行わない場合で 40%程度となった [25]。

ユーザの視線の可視化や動作の模倣によって、行動や興味への気づきを促進している研究事例がある [26]。Gaze Mirroring においても、セルフフィードバックによる自己の興味への気づきをユーザに与えている可能性がある。

3.2.4 情報コンシェルジェ：Mind Probing に基づくマルチモーダルインタラクションシステム [27]

量的に爆発し、質的にも複雑さを増大する情報空間において、人は望みの情報に辿り着くことに苦難している。明確な目的を自覚し、キーワードやコマンドベースのインタフェースを使いこなさなければ、心理的負担が増大する。このような場合、高級ホテルのコンシェルジェのように、情報に精通しているエージェントによるインタラクティブなサポートが有効に働くと考えられる。そこで著者らは、ユーザが持つ潜在的な興味を探り、そしてユーザにそれを気づかせ、さらにはそれに応じてさりげなく、気の利いた情報提供を行うコンシェルジェシステムの実現を目指し、上述の視覚的インタラクションに加え、音声発話ベースのプロアクティブインタラクションを導入した情報コンシェルジェを構築した [28] (図 6)。

情報コンシェルジェは大画面にいくつかの視覚情報を表示し、スピーカーを介して音声情報を提示し、ユーザの視線と発話をカメラとマイクを用いて観測する [29], [30]。一方でユーザは知識を欠如している状態でそれらと比較し、好みのものを選ぶという状況を想定する。ユーザとのインタラクションは、図 6 に示すように、初期提示 (ザッピング支援, 注視位置キャリブレーション), 気づき支援, 説明 (興味推定, 意図確認), 推薦のフェーズで構成される。注

視位置キャリブレーションには Gaze Probing (3.2.1 節), 気づき支援には Gaze Mirroring (3.2.3 節) を応用する。

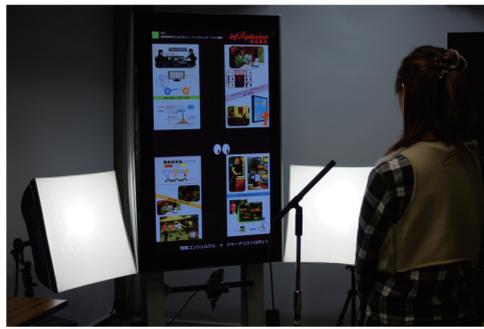
ザッピング支援フェーズは、3.2.2 節で述べた俯瞰フェーズに対応する。ここで情報コンシェルジェは、いくつかのコンテンツカテゴリを主導的に 1 つずつ列挙する形で提示し、ユーザがその中から 1 つを選んで、それを言語的に伝えることを促す。この列挙型インタラクションは、ユーザが「それ」といった参照表現での発話タイミングによってその選択をシステムに伝えることができるため、キーワード発話に対する音声認識のみの場合に比べて信頼性が高いインタラクションが可能である [31]。

興味推定フェーズでは、システムが提示コンテンツについて音声説明を行いながら、ユーザの興味を推定する。基本的には、3.2.2 節で述べた Mind Probing によって興味を推定するが、視覚情報の切り替えはシステムによる説明に対するユーザの参加度が低下した際に行う。Nakano らは、ユーザがエージェントと共同注意を行わない場合に参加度が低く、相互注視とよそ見を行わない場合に高いという分析結果を示した [32]。そこで、参加度が大幅に低下した際に、システムは説明を止め、説明を行っていたコンテンツの視覚情報を切り替える。ユーザの注視がすぐにそのコンテンツに復帰した場合には説明を再開し、そうではない場合には注視しているコンテンツの説明を始める。

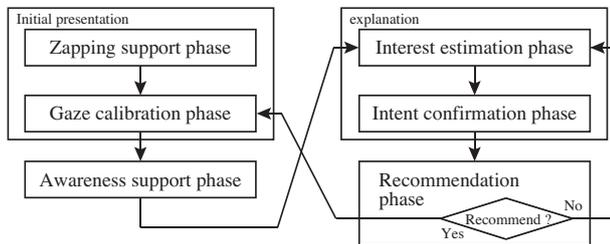
システムの働きかけに対するユーザの注視反応だけに基づいて確実に興味を推定することができるわけではない。前のフェーズまでで推定した興味に基づいてユーザの選択意図を確認するために、より明確な反応を引き出す必要がある。Kuno らは、ロボットが音声説明の意味的な区切れなどの Transition Relevance Places (TRPs) でユーザに視線を投げかけることで、ユーザのうなずきなどの反応を引き出しやすくなることを示した [33]。その知見を応用し、意図確認フェーズでは、視覚情報と音声説明を全て提示し終えたコンテンツについて、システムは選択意図の確認のためにユーザに言語的に問いかけ、その TRPs でユーザに視線を向けて反応を促す。反応が得られた場合には、それに応じて関連する情報を提供し、反応が得られなかった場合には、別のコンテンツを推薦するなどして問いかけを続ける [34]。

3.3 人間と環境の視覚的インタラクション

対話状況以外の日常活動シーンにおいても動的な視覚刺激が溢れているため、人間は外環境から Mind Probing を仕掛けられているようなものである。人間の内部状態を振る舞いに引き出すことができている視覚刺激がどういったものであるか特定できれば、環境設計などに应用することができる。そこで、種々に人間の状態を統制した状況下で外的環境変化と注視行動との関係の分析を行っている。



(a) Info-concierge



(b) Interaction flow

図 6 情報コンシェルジェとそのインタラクションフロー

3.3.1 映像の顕著性変動と視線運動の時空間相関分析に基づいた集中状態推定 [35]

あるタスクを遂行している人間の集中状態を推定することは、システムがサポートするタイミングを計る上で重要である。著者らは、人間が映像を視聴している状況を取り上げ、視聴映像から視線運動に関連するプローブとなる映像ダイナミクスを抽出し、それと視線運動との間の時空間構造から視聴者の集中度を推定する手法を提案した。

映像を視聴する際に表出される視線運動は、映像中に存在する顕著な領域（生理的に視線が向きやすい領域）の振る舞いに影響される。そこで、(1)Saliency map を算出して映像中から顕著領域を抽出、(2)画面中の顕著領域数およびそれらの動きに基づいて、顕著領域の時空間的変動のパターンを分類、(3)被験者に映像視聴のみをタスクとして課した場合（集中状態）と副タスクとして暗算を課した場合（注意散漫状態）の視線データを蓄積、(4)各パターンに対して表出された視線運動から特徴量を抽出し、ナイーブベイズ識別器を学習し、この識別モデルに基づいて集中度の推定を行った（図7）。

人間の目を引く領域が適度にデザインされていると考えられるTV コマーシャルを用いた評価実験の結果、提案手法が比較手法より20%ほど高精度の集中度推定を実現した。比較手法には、顕著領域への累積注視時間を用いたものと、臉が閉じられている累積時間（PERCLOS）を用いたものを採用した。この提案手法はその後、データ駆動的なものに改良され、顕著領域の時空間変動を線形システムで表現し、同定したシステムパラメータに基づいたパターン分類を行うものへと発展している [36]。

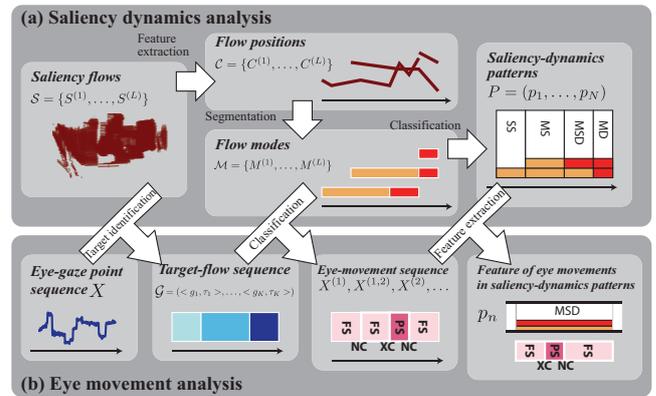


図 7 映像の顕著性変動と視線運動の時空間相関分析に基づいた集中状態推定

3.3.2 周辺車状況変化に対する注視反応タイミングに基づいた注意散漫状態検出 [37]

集中状態を推定することが有用である状況としては、自動車を運転する状況も挙げられる。そこで、周辺車状況の変化に対する運転者の注視行動の時間的関係について、運転集中状態と注意散漫状態における差異を大規模運転行動信号コーパス [38] を用いて分析した。追い越し車両を外部刺激として注目したところ、通常走行時（運転集中状態）は音声対話による楽曲検索走行時（注意散漫状態）に比べ500msec程度、有意に早く注視反応が起こることを確認した。また、追い越し車両への注視の時間分布を尤度分布とし、ナイーブベイズに基づいて集中状態と注意散漫状態の識別を行ったところ、従来研究で提案されている前方注視割合（Percentage Road Center, PRC）を応用した識別に比べ20%程度の精度向上を達成した。この研究では、誘目性が高い追い越し車両に注目して分析を進めたが、周辺車状況と注視行動と運転者の状態との間の関係を機械的に抽出するアプローチを現在検討している。

また、運転スキルは集中状態に比べてかなり恒常的ではあるが運転者の内部状態であると考えられる。運転者のスキル差と、運転車両と周辺車両間における危険度に応じて、運転者の視認行動に違いが表れる可能性がある。そこで、大規模運転行動信号コーパスを用いて、視線方向情報と周辺車両情報を基に模範運転者（教習所教官）と一般運転者による周辺車両の視認行動に差異が表れるか分析した。分析手法としては、視線方向情報および周辺車両情報から求められる注視度（周辺視野を考慮した注視方向の分布を算出）および周辺車両危険度（車間距離と相対速度に基づいて算出）を定義し、これらの一致係数を求める手法を提案した。その分析結果として、右車線変更を行っている時区間における一致係数の平均値に差が現れ、一般運転者に比べ模範運転者が大きな平均値を示す傾向にあることが確認された。これは、提案した一致係数が周辺車両への注意度合いを反映している可能性を示唆している [39]。

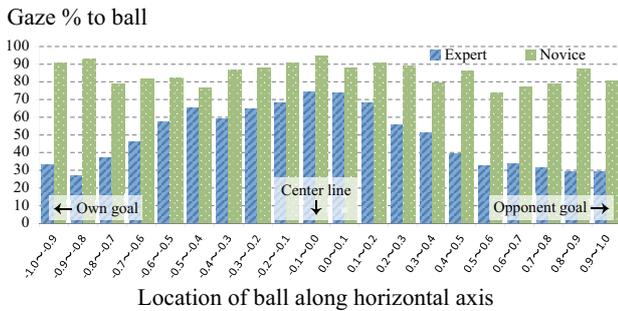


図 8 ボールの存在領域ごとの視線配布率

3.3.3 サッカー指導の熟練者と未経験者の 注視行動比較 [40]

スキル差に関わる注視行動の違いは、スポーツを対象に盛んに分析が行われている。著者らは、サッカーの試合映像に対する熟練指導者（S級ライセンス保有者）と指導未経験者の注視行動の差異を、映像内で重要度が高い動的対象であるボールと視線運動の関連性に基づいて分析した。まず、視線配布を中心窩視野（中心窩から2度の範囲）を対象を取めた状態として定義し、ボールへの視線配布率を算出したところ、熟練指導者が約5%、指導未経験者が約50%となった。また、傍中心窩視野（中心窩から5度の範囲）を含めた領域を視線配布の範囲としたところ、熟練指導者が約50%、指導未経験者が約80%となった。熟練指導者は事後アンケートでボールへの理想的な視線配布率は4割から5割と答えていたため、理想に近い視線配布を行っていた可能性がある。これらの結果から、指導未経験者はボールを視野の中心で捉えて追跡している時区間が多く、一方で熟練指導者は周辺視で捉える時区間が多いことが分かる。また、フィールド上でのボールの存在領域と視線配布率の関係を分析したところ、指導未経験者がボールの存在領域によって視線配布率をほとんど変化させないことに対して、熟練指導者は、センターライン付近では視線配布率が高く、ゴールラインに近くなるほど低くなる傾向があった（図8）。つまり、熟練指導者は、時々刻々と変化する状況に応じて、注視対象を効率良く変化させていると考えられる。今後は、個性や戦略、試合展開を考慮し、選手への注視を含めた分析を行うことで熟練指導者の視技能のコツを抽出したい。

4. ますます熱い視線

本稿では、工学と心理物理学分野における視線研究の動向を概観し、著者が過去に行った視覚的インタラクション研究を紹介した。およそ一世紀前から視線計測技術の開発が始まり、要素還元を主体としてゆっくりとしたペースで視線が分析されてきたが、近年はタブレットPCやスマートフォンに搭載可能な視線計測モジュールの開発が進展するなど、視線を計測、利用する場が実験室から日常環境へと移りつつあり、今後はデータの蓄積がこれまでとは全く異

なる上昇曲線を描き、計算機資源を活用した分析が急速に進められるものと考えられる。情報学分野における視線研究の盛り上がりは、視線を主題とした国際会議が頻りに開催されていることが示している*1。

著者は現在、動的な事象に対する熟練した見方を職業知として抽出し、技能習得支援やエンターテインメントを目的として観察者にフィードバックするインタラクション設計や、多視点映像などの人間の活動を拡張する環境内での注視行動の特性分析と注視行動の誘導方法の開発 [41], [42] を主として、これまでの研究で得た知見やノウハウを基に研究を進展させている。

謝辞 本研究の一部は、文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しいIT基盤技術の研究」における「情報爆発時代におけるヒューマンコミュニケーション基盤」(18049046)と同若手研究B「視線に基づく模倣インタラクションによる自己の興味への気づきとその評価方法の確立」(23700168)、独立行政法人科学技術振興機構CREST「共生社会に向けた人間調和型情報技術の構築」における「行動モデルに基づく過信の抑止」および総務省戦略的情報通信研究開発推進制度(SCOPE)「広視域角自由視点映像のための次世代マルチビュービデオ配信・視聴方式」の支援によるものである。

参考文献

- [1] 水口, 浅野, 佐竹, 小林, 平山, 川嶋, 小嶋, 松山: Mind Probing: システムの積極的な働きかけによる視線パターンからの興味推定, 情報処理学会研究報告 (HCI-125), Vol.2007, No.99, pp.1-8 (2007).
- [2] Dodge, Cline, *The Angle Velocity of Eye Movement*, Psychological Review, Vol.8, pp.145-157 (1901).
- [3] H. Hartridge, L. C. Thomson, *Methods of Investigating Eye Movements*, British Journal of Ophthalmology, No.32, pp.581-591 (1948).
- [4] A. L. Yarbus, *Eye Movements and Vision*, Plenum Press, New York (1967).
- [5] M. A. Just, P. A. Carpenter, *Eye Fixations and Cognitive Processes*, Cognitive Psychology, No.8, pp.441-480 (1976).
- [6] M. I. Posner, *Orienting of Attention*, Quarterly Journal of Experimental Psychology, No.32A, pp.3-25 (1980).
- [7] A. M. Treisman, G. Gelade, *A Feature Integration Theory of Attention*, Cognitive Psychology, No.12, pp.97-136 (1980).
- [8] C. Koch, S. Ullman, *Shifts in Visual Attention: Towards the Underlying Circuitry*, Human Neurobiology, No.4, pp.219-222 (1985).
- [9] S. Baron-Cohen, *Mindblindness: An Essay on Autism and Theory of Mind*, The MIT Press (1995).
- [10] R. J. K. Jacob, K. S. Karn, *Eye Tracking in Human-Computer Interaction and Usability Research: Ready*

*1 近年では、ACM Symposium on Eye Tracking Research and Applicationsを始め、IEEE Workshop on Egocentric Vision, International Workshop on Eye Gaze in Intelligent Human Machine Interaction, International Workshop on Gaze Sensing and Interactions, International Workshop on Visual Attention and Interactionなどが開催されている。

- to Deliver the Promises, In J. Hyönä, R. Radach, H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, (pp.573–605). Amsterdam: Elsevier (2003).
- [11] Y. Dong, Z. Hu, K. Uchimura, N. Murayama, *Driver Inattention Monitoring System for Intelligent Vehicles: A Review*, IEEE Trans. on Intelligent Transportation Systems, Vol.12, No.2, pp.596–614 (2011).
- [12] J. M. Findlay et al., *Active Vision: The Psychology of Looking and Seeing*, Oxford University Press (2003).
- [13] A. Kimura, R. Yonetani, T. Hirayama, *Computational Models of Human Visual Attention and their Implementations: A Survey*, IEICE Trans. on Information and Systems, Vol.E96-D, No.3, pp.562–578 (2013).
- [14] V. Tanriverdi, R. J. K. Jacob, *Interacting with Eye Movements in Virtual Environments*, Proc. of the ACM CHI Human Factors in Computing Systems Conference, pp.265–272 (2000).
- [15] P. Qvarfordt, S. Zhai, *Conversing with the User Based on Eye-Gaze Patterns*, Proc. of the ACM CHI Human Factors in Computing Systems Conf., pp.221–230 (2005).
- [16] G. L. Rupp, *Performance Metrics for Assessing Driver Distraction: The Quest for Improved Road Safety*, SAE International (2010).
- [17] 松山, 川嶋, 平山: 時間と時間感覚に対する感性の情報処理—豊かなヒューマンコミュニケーションの実現を目指して—, 信学会誌, Vol.92, No.11, pp.952–954 (2009).
- [18] 平山, 大西, 朴, 松山: 対話における顔向けを伴う働きかけが同意・不同意応答のタイミングに及ぼす影響, ヒューマンインタフェース学会論文誌, Vol.10, No.4, pp.1–10 (2008).
- [19] M. Yoshida, Y. Miyake *Relationship between Utterance Dynamics and Pragmatics in the Conversation of Consensus Building Process*, Proc. of the 15th IEEE International Symposium on Robot and Human Interactive Communication(ROMAN2006), pp.641–645 (2006).
- [20] N. Fujiwara, T. Itoh, K. Araki, *Analysis of Changes in Dialogue Rhythm due to Dialogue Acts in Task-oriented Dialogues*, Springer-Verlag Lecture Notes in Artificial Intelligence(LNAI), Vol.4629, pp.564–573 (2007).
- [21] 米谷, 川嶋, 平山, 松山: Gaze Probing: イベント提示に基づく注視オブジェクト推定, ヒューマンインタフェース学会論文誌, Vol.12, No.3, pp.125–135 (2010).
- [22] T. Hirayama, J. B. Dodane, H. Kawashima, T. Matsuyama, *Estimates of User Interest Using Timing Structures between Proactive Content-Display Updates and Eye Movements*, IEICE Trans. on Information and Systems, Vol.E-93D, No.6, pp.1470–1478 (2010).
- [23] E. Ishikawa, R. Yonetani, H. Kawashima, T. Hirayama, T. Matsuyama, *Semantic Interpretation of Eye Movements Using Designed Structures of Displayed Contents*, Proc. of the 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction (2012).
- [24] 石川, 米谷, 平山, 松山: Gaze Mirroring による注視模倣効果の分析, 情報処理学会論文誌, Vol.52, No.12, pp.3637–3646 (2011).
- [25] 平山, 朴, 松山: Gaze Mirroring: ユーザの興味を顕在化させるための注視模倣, 電子情報通信学会技術研究報告 HCS, Vol.108, No.487, pp.79–84 (2009).
- [26] Y. Sejima, T. Watanabe, M. Yamamoto, *Analysis by Synthesis of Embodied Communication via Virtual Actor with a Nodding Response Model*, Proc. of 2nd International Symposium on Universal Communication (ISUC2008), pp.225–230 (2008).
- [27] 河原, 川嶋, 平山, 松山: 対話を通じてユーザの意図・興味を探り情報検索・提示する情報コンシェルジュ, 情報処理, Vol.49, No.8, pp.912–918 (2008).
- [28] T. Hirayama, Y. Sumi, T. Kawahara, T. Matsuyama, *Info-Concierge: Proactive Multi-Modal Interaction Based on Mind Probing*, Proc. of the Asia Pacific Signal and Information Processing Association Annual Summit and Conference (2011).
- [29] 佐竹, 小林, 川嶋, 平山, 水口, 小嶋, 松山: インタラクティブな情報提示システムのための非装着・非拘束な視線推定, 情報処理学会研究報告 (HCI-125), Vol.2007, No.99, pp.9–16 (2007).
- [30] A. Lee, T. Kawahara, *Recent Development of Open-Source Speech Recognition Engine Julius*, Proc. of the Asia Pacific Signal and Information Processing Association Annual Summit and Conference, pp.131–137 (2009).
- [31] K. Matsuyama, K. Komatani, T. Ogata, H. G. Okuno, *Enabling a User to Specify an Item at Any Time during System Enumeration –Item Identification for Barge-in-Able Conversational Dialogue Systems–*, INTERSPEECH2009, pp.252–255 (2009).
- [32] Y. I. Nakano, R. Ishii, *Estimating User's Engagement from Eye-Gaze Behaviors in Human-Agent Conversations*, Proc. of International Conference on Intelligent User Interfaces, pp.139–148 (2010).
- [33] Y. Kuno, K. Sadazuka, M. Kawashima, K. Yamazaki, A. Yamazaki, H. Kuzuoka, *Museum Guide Robot Based on Sociological Interaction Analysis*, Proc. of the ACM CHI Human Factors in Computing Systems Conference, pp.1191–1194 (2007).
- [34] T. Misu, T. Kawahara, *Speech-Based Interactive Information Guidance System Using Question-Answering Technique*, Proc. of the International Conference on Acoustics, Speech, and Signal Processing, Vol.IV, pp.145–148 (2007).
- [35] R. Yonetani, H. Kawashima, T. Hirayama, T. Matsuyama, *Mental Focus Analysis Using the Spatio-Temporal Correlation between Visual Saliency and Eye Movements*, Journal of Information Processing, Vol.20, No.1, pp.267–276 (2011).
- [36] 米谷, 川嶋, 加藤, 松山: 映像の顕著性変動モデルを用いた視聴者の集中状態推定, 信学論, Vol.J96-D, No.8, pp.1675–1687 (2013).
- [37] T. Hirayama, K. Mase, K. Takeda, *Detection of Driver Distraction based on Temporal Relationship between Eye-Gaze and Peripheral Vehicle Behavior*, Proc. of the 15th IEEE Intelligent Transportation Systems Conference (2012).
- [38] K. Takeda, J. H. L. Hansen, P. Boyraz, L. Malta, C. Miyajima, H. Abut, *International Large-Scale Vehicle Corpora for Research on Driver Behavior on the Road*, IEEE Trans. on Intelligent Transportation Systems, Vol.12, No.3, pp.1–15 (2011).
- [39] M. Mori, C. Miyajima, P. Angkititrakul, T. Hirayama, Y. Li, N. Kitaoka, K. Takeda, *Measuring Driver Awareness based on Correlation between Gaze Behavior and Risks of Surrounding Vehicles*, Proc. of the 15th IEEE Intelligent Transportation Systems Conference (2012).
- [40] 岩月, 平山, 榎堀, 間瀬: サッカー指導者の注視行動の分析—動的対象との関連性—, 情報処理学会全国大会 (2013).
- [41] 間瀬, 平山, 榎堀, 丸谷: コーチング支援のための多視点カメラとマルチセンサ利用, ビジョン技術の実利用ワークショップ ViEW2012 (2012).
- [42] 朝倉, 平山, 丸谷, 加藤, 間瀬: 多視点映像の視聴履歴を用いた視点遷移予測モデルの検討, 人工知能学会全国大会 (2013).