

## 人工網膜術後のリハビリテーション法の開発 視線移動の研究

研究分担者 北澤 茂

大阪大学大学院生命機能研究科・脳神経工学講座 /医学系研究科・脳生理学 教授

研究要旨：人工視覚を使った対象認知のリハビリテーション法開発のために、健常被験者の視線移動の時系列の規則性を解析し、定量的なモデルとして表現できるかどうかを検討した。その結果、我々の脳は視界の中の顔に強く注目し、顔の動き、大きさ、新規性、口の動き、の4項目の線形和で定義される「顔の顕著性」が最も大きい顔を選んで視線を移動させることが明らかになった。この成果は「顔を認知して見る」という行動を回復するためのリハビリテーションに応用可能である。

### A．研究目的

視覚を失うことによって生じる困難は多岐にわたる。相手の顔の位置を知り、適切なタイミングで見ることができないこともその一つである。人工視覚は、顔を適切なタイミングで見るという行動を回復するためにも応用することができるはずである。

健常被験者が映画やテレビを見るとき視線の移動のタイミングと対象は、登場人物がいる場面では驚くほど類似している(Nakano et al., 2010a; Shepherd et al., 2010)。しかし、自閉症者はこのような定型的な視線の移動を示さない(Nakano et al., 2010b)。また、Kochらはコントラストや明度や色の变化などの物理的な特徴量で定義される目立つ場所（顕著性の高い場所）を見するというモデル(Itti and Koch, 2001)を提唱してきたが、このモデルでは全く説明できないことも報告されている(Shepherd et al., 2010)。一体、どのような規則に従って、われわれは複数の顔から一つを選んで眼を向けているのだろうか。本研究では、映画やテレビを見る際の、定型的な視線の移動パターンを、定量的なモデルとして表現することを目的とした。

### B．研究方法

対象：24名の健常自発参加者を対象として研究を行った。参加者には規定の謝金を支払った。

刺激と課題：被験者は顔を顎台に載せた状態で、

前方の画面に提示されるビデオ画像を視聴した。およそ6秒の短いビデオクリップ12本を0.5秒のブランクをはさんで編集した約77秒（2237フレーム）のビデオ画像を刺激に用いた。被験者は異なる4条件（順方向再生・音あり、順方向再生・音なし、逆方向再生・音あり、逆方向再生・音なし）の4条件でビデオを視聴した。

計測：刺激を視聴する際の視線を50Hzで計測した（Tobii, X50, Tobii Technology AB）。

解析：各フレームにおいて、すべての顔の眼、口、鼻の位置を登録し、被験者ごとに視線位置と顔の距離に基づいてどの顔を見ているかを判別した。その上で、それぞれの顔を見ている被験者の割合（顔の視聴率）を計算した。

多次元尺度法による視線パターンの比較：24被験者の4条件の視線パターン（96個）Kochの顕著性モデルが予測する視線パターン（7個）視聴率が最も高い顔を選んで作った視線パターン（ピーク顔パターン）を定量比較するために、お互いの距離を成分とする距離行列を作り、多次元尺度を適用して2次元平面状にプロットした。

顔顕著性モデル：それぞれの顔が持つ顕著性を1) 大きさ、2) 顔の動き、3) 口の動き、4) 新規性の4項の和として表現し、さらにすべての顔の顕著性の和で規格化した「顔顕著性」を定義した。どの顔を見るかの確率は顔顕著性を入力とするシグモイド関数で表現した（図1）。実際の顔の視

聴率の時間変化を最もよく近似するように線形和のパラメータを調整した。

(倫理面への配慮)

大阪大学大学院生命機能研究科に設置された外部委員を含む生命機能研究科倫理委員会において、研究内容の安全性・必要性に関して、審査・承認を受けて行った。各参加者には事前に研究の内容を説明し、書面で同意(インフォームドコンセント)を得てから実験を行った。



Face 1, 2, ..., 5

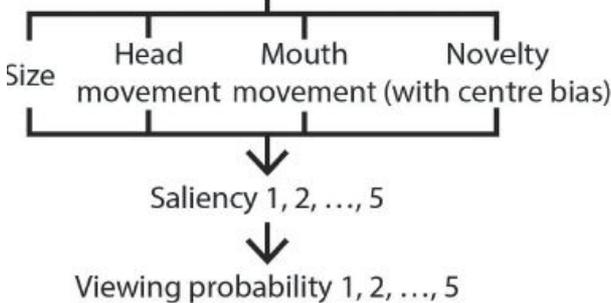


図1 顔顕著性モデル

### C. 研究結果

多次元尺度法を適用して視線の移動パターンを比較したところ、1) 音の有無はパターンを変化させない、2) 一方、再生の方向はパターンを変化させる、3) 物理的な顕著性モデルは実際の視線移動パターンとは全く異なる、4) 最高視聴率の顔をつないだ人工的な視線パターン(ピーク顔)は実際の視線パターンとよく類似している、ことが明らかとなった(図2)。

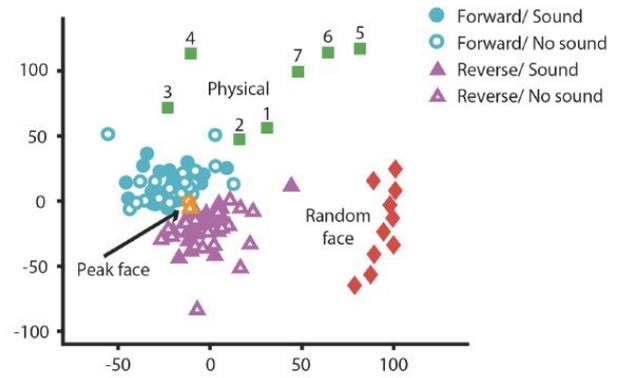


図2 多次元尺度法を用いた視線の移動パターンの比較。順方向再生( )と逆方向再生( )は全く異なる分布をする一方、音の有無は影響がない。

顔顕著性モデルをデータに適用した結果の一例を示す(図3)。このおよそ6秒のビデオクリップは、二人の少年が、交互に話す場面である。ビデオクリップが開始した直後は、正面向きで顔の面積が大きい右の少年の視聴率が高いが、左の少年が話をする時点に向って左の少年に視線が移動し、口を動かしている時(1.83秒)に80%のピーク視聴率を示した。発話終了後は右の少年に視線が移動して、右の少年が発話する4.53秒後にやはり視聴率は80%を超え、その後徐々に左の少年に視線が移動した(図3の点線)。顔顕著性モデルで予測される視聴率は図3の実線である。実際のデータとよく一致し、83%の分散を説明した( $d.c. = 0.83$ )。

決定係数の中央値は0.87と高かった(図4)。パラメータの推定を11個のビデオクリップで行い、残りの1個で検定した場合でも、中央値は0.57であった。顔顕著性モデルが視線移動の過半を説明することが示された。

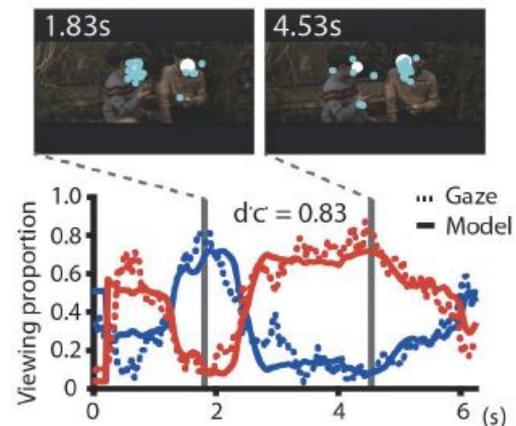


図3 視聴率変化とモデル予測の一例  
視線の視聴率変化(点線)とモデル予測(実線)はよく一致した。d.c.は決定係数を表す。

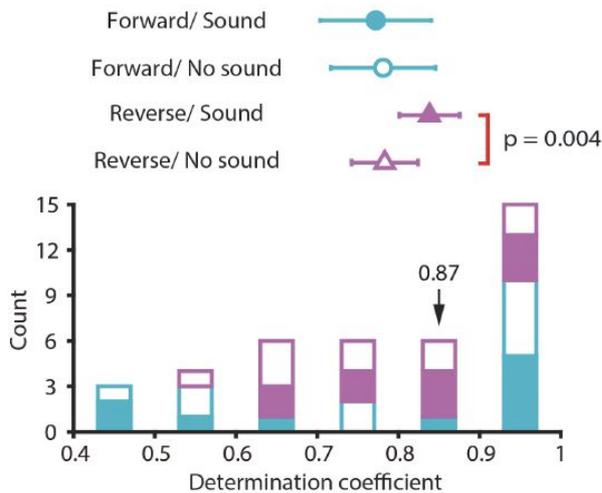


図4 顔顕著性モデルの決定係数の分布

#### D. 考察

本研究では、1) 最も注目を集める顔だけを見て、他のものは全く見ない、という人工的な視線パターンは、実際の視線パターンをよく近似すること、2) 音声の有無は視線の移動パターンに決定的な変化を与えないこと、3) 逆再生刺激の視線移動パターンは単なる順方向刺激の視線パターンの折り返しではないこと、4) 顔の大きさ、頭の動き、口の動き、新規性の4個のパラメータを使うだけで視線移動パターンの過半が説明できること、を明らかにした。

音声の有無が視線移動に影響を与えないのは驚きである。我々は、声を聞いてから眼を向けるのではなく、発声の前に生じる顔全体の上下の動きに反応して、次の話者を特定して眼を向けている可能性がある(図2)。我々とは独立したグループの研究でも、音声の有無によらず、話者に注目が集まることが最近報告されている(Coutrot & Guyader, 2014)。

本研究で提案したモデルは、1) 顔の抽出、2) 大きさの計算、3) 顔全体の動きの計算、4) 口の抽出、5) 口の動きの有無の検出、6) 新規に登場してから時間経過、の情報さえあれば、すぐに応用することが可能である。顔の抽出に関してはすでにデジタルカメラに搭載されているし、表情検出アルゴリズムを使えば口の動きの検出も容易だろう。顔の大きさや顔全体の動きの検出はさらに容易なはずである。つまり、既存の画像処理技術を応用するだけで、現在の視野のどの顔に眼を向けるのが自然か、が瞬時に計算できるはずである。

人工視覚の現在の解像度では、複数の顔の検出や、口の動きを検出することは難しい。しかし、画像処理を行った結果として得られる注目すべき顔のある場所を提示するために用いることは可能である。人の顔を見る、という社会的には大切な行動を回復するためのリハビリテーションに本研究の成果を応用することが可能だろう。

#### E. 結論

我々の脳は、顔を検出して、大きさやうごきなどの少数の情報を組み合わせて比較して、眼を向けるべき顔を選んでいることが明らかになった。この成果は画像処理技術と人工視覚を組み合わせることで、適切なタイミングで人の顔を見る、という行動を回復するためのリハビリテーションに応用できるだろう。

#### F. 健康危険情報

該当する危険なし

#### G. 研究発表

##### 1. 論文発表

Suda Y & Kitazawa S. A model of face selection in viewing video stories. Scientific Reports 5: 7666, 2015.

##### 2. 学会発表

1) Uchimura M, Nakano T, Morito Y, Ando H, & Kitazawa S. Automatic representation of a visual stimulus relative to a background in the right precuneus. Vision, Memory, Thought: how cognition emerges from neural network, International Symposium, December 6, 2014, Tokyo, Japan.

2) Uchimura M, Nakano T, Morito Y, Ando H, & Kitazawa S. Automatic representation of a visual stimulus relative to a background in the right precuneus. Neural Mechanisms of Vision and Cognition, CiNet 2<sup>nd</sup> International Symposium, March 2, 2015, Osaka, Japan.

#### H. 知的財産権の出願・登録状況

特記事項なし。

