

J-010

「網膜像刺激」の推定に基づく自然画像統計解析
The analysis of natural scene statistics based on "retinal image stimulation"

田村 隼[†], 松宮一道[‡], 栗木一郎[‡], 塩入諭[‡]
Hayato tamura, Kazumichi Matsumiya, Ichiro Kuriki, Satoshi Shioiri

1. はじめに

我々の視覚は網膜像を通して、外界から多くの情報を得ている。網膜像は物体そのものの情報全てを表現できないため、視覚処理はなんらかの推定を含む。日常生活において我々の視覚が機能している事実は、その推定がうまくいっていることを意味する。その理由は、日常的に得られる網膜像には物体や環境に依存した統計的性質があり、視覚はそれを利用して考えるからと考えられる。この統計的性質には、面の連続性など一般的な事実もあるし、場面や行動に依存した特性も考えられる。したがって視覚処理を理解するためには、日常生活で得られる視覚刺激の統計的性質を調べる必要がある。

視環境の統計的な性質を調べた研究の多くは静止画像を扱っている (Antonio and Aude[1]) が、我々の周囲は常に変化しているし、網膜像は自身の行動によっても変化するため、動画像の研究が必要である。Dong and Atick[2] は彼らが撮影した動画を画面中央の $64[\text{pixel}] \times 64[\text{pixel}]$ を各フレームで切り出し、その連続した 64 フレームを1つの segment として複数個用意し、それらの時空間周波数分布を調べた。しかし、網膜に落ちる画像情報を調べるためには、カメラで撮影した映像をただ解析するのでは不十分である。なぜなら、視線の移動により網膜像は変化するためである。例えばテニスの試合を見ている場合、固定したカメラで得られた画像と、視線を動かしている観察者の網膜像は大きく異なる。本研究では、日常生活の様々な場面における網膜情報の統計的性質を調べるために、簡易型の眼球運動測定システムを構築し、様々な環境下での被験者の網膜像を推定し、その統計的性質の評価を試みる。

2. 実験方法

2.1 眼球運動測定システム

今回、眼球運動の測定のために簡易型の眼球運動測定システムを構築した。概要としてはカメラ2台を用いて頭部運動/眼球運動に対する動画像を撮影する。その後、プログラムを用いて scene 中の視線先を決定するというシステムになる。プログラムには Starburst [3] と呼ばれるアルゴリズムを用いて眼球運動の捕捉を行っている。

図1にシステムの構成を示す。頭部カメラは観察者の頭の向いた方向の映像を撮影するもので、この映像を解析すると従来の研究と同様の統計的性質の評価ができる。眼球カメラは、観察者の眼球を写す。この映像の各フレームから、黒目部分を抽出し、それを楕円で評価することから、視線方向を測定する (詳細は文献[3]参照)。



図1 眼球運動測定システム概要

2.2 実験手続き

今回は、2つの状況での網膜像の評価について報告する。上述の眼球運動測定装置を用いて、(1)サッカーの試合をテレビで視聴している場面 (以下サッカー映像) および(2)ジャグリングを行っている場面 (以下ジャグリング映像) の映像を撮影した。眼球カメラの映像から視線方向を測定し、頭部カメラで得た映像の各フレームを視線位置が原点となるように補正した。これによって、網膜像に対応する映像 (ここでは網膜映像と呼ぶ) を得ることができる。なお実験には1人の被験者が参加し、それぞれの条件で測定を1回行った。

統計的性質として、Dong and Atick にしたがって時空間周波数特性の検討を行った。網膜映像は、視線位置によって範囲が異なるため、中心の $64[\text{pixel}] \times 64[\text{pixel}]$ の部分を切り出し、網膜中心の映像として処理した。時空間周波数の解析のために、時間方向にも 64frame を単位として、各時間間隔ごとに $64 \times 64 \times 64$ のセグメントについて、3次元フーリエ変換、 $R(\mathbf{f}, \omega)$ を求めた ((1) 式)。今回は、時間周波数と空間周波数の関連に着もするため、 x, y 方向 (f_x, f_y) については極座標変換を行い ((2) 式)、空間周波数情報 (f_r) と時間周波数について解析した (方向については考慮していない)。空間周波数 f_r 、時間周波数 ω に対する、対数パワー $P(f_r, \omega)$ の変化 ((3) 式) を、2つの状況 (サッカーの試合視聴時とジャグリング時) で比較した。 $P(f_r, \omega)$ は、64 フレーム毎に算出したが、解析には、動画(1)に対しては 946 個のセグメント (2015 秒に対応) の平均を、動画(2)に対しては 370 個のセグメント (788 秒に対応) の平均を用いた。

[†] 東北大学大学院情報科学研究科, GSIS

[‡] 東北大学電気通信研究所, RIEC

$$R(\mathbf{f}, \omega) = \sum_0^{L-1} \sum_0^{T-1} r(\mathbf{x}, t) \exp\left(-2\pi i \left(\frac{\mathbf{f} \cdot \mathbf{x}}{L^2} + \frac{\omega t}{T}\right)\right), \quad (1)$$

$$\mathbf{x} = (x, y), \quad (2)$$

$$\mathbf{f} = (f_x, f_y), \quad (3)$$

$$f_r = \sqrt{f_x^2 + f_y^2}, \quad (4)$$

$$P(f_r, \omega) = \log|R(f_r, \omega)| \quad (5)$$

ここで $r(\mathbf{x}, t)$ は動画像を, $R(\mathbf{f}, \omega)$ はその Fourier 変換を表す. なお, $L=64, T=64$ である. \mathbf{x} は画像中の位置を表し, \mathbf{f} は Fourier 領域での位置を表す. f_r は Fourier 領域での原点からの距離 (空間周波数) を表し, $P(f_r, \omega)$ は時空間周波数領域における対数パワーを示す.

3. 実験結果と考察

サッカー映像, ジャグリング映像それぞれについて, 視線方向を考慮した時空間特性の解析結果と考慮しない頭部映像に対する解析結果を, 図2, 図3に示す. ここで縦軸は時間周波数を, 横軸は空間周波数を表す. 明暗がパワーの大きさを表す.

まず視線を考慮しない結果を見ると, どちらの動画の場合にも, 先行研究と同様に時空間的に低周波にスペクトルが集中する傾向がみられる. また, 水平, 垂直方向には比較的高い周波数までスペクトル成分が分布している様子がわかる. 映像間の比較ではジャグリング映像でより時間周波数方向で広範囲な分布が見られる. なお, 観察対象の特性の違いが反映されている.

視線を考慮した場合には, いずれの映像でも時間周波数方向に高い周波数まで分布が広がっている. この時間的に高周波の成分については, 網膜像の刺激の変化というよりも主に視線検出の誤差によるものと考えている. 短時間で変動する視線方向誤差は, すべての空間周波数成分に影響し, 図1, 2においては縦方向にぼかした効果として現れる.

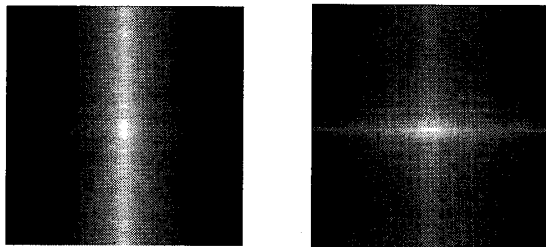


図2 サッカー映像における時空間周波数分布 (左: 視線方向考慮あり 右: 視線方向考慮なし)

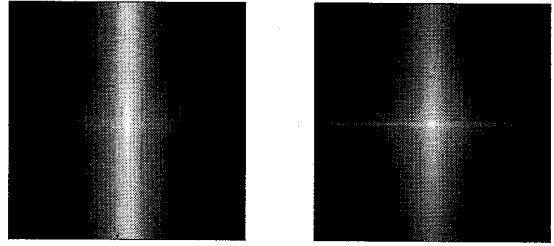


図3 ジャグリング映像における時空間周波数分布 (左: 視線方向考慮あり 右: 視線方向考慮なし)

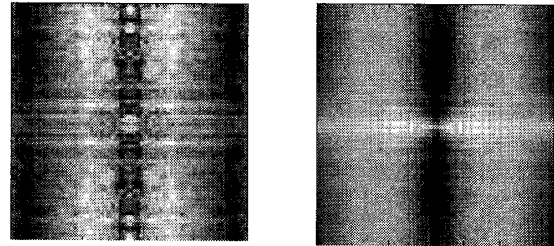


図4 時空間周波数分布の差分 (左: 視線方向考慮あり 右: 視線方向考慮なし)

4. まとめ

今回, 網膜に入力される視覚刺激の統計的性質として2つの状況における頭部カメラ映像と網膜映像の時空間周波数分布を調べた. その結果は, 先行研究で示されている. 低時間, 低空間周波数に偏った分布であることを確認した. さらに, 視線移動が網膜映像の時空間周波数特性を顕著に変化することを示す結果を得た. しかし, 今回の解析には視線検出の誤差の影響が含まれていると考えられ, 今後さらに精度の高い視線方向の検出に基づく検討が必要である. その上で, 人間の活動状況に依存した網膜映像の時空間周波数分布を調べ, 視覚系の運動検出特性などとの関連を調べていく予定である.

参考文献

- [1] Antonio Torralba, Aude Oliva, "Statistics of natural image categories", Network: Computation in Neural Systems, 14, 391-412 (2003).
- [2] Dawei W. Dong, Joseph J. Atick, "Statistics of Natural Time-Varying Images", Network: Computation in Neural Systems, Vol 6(3), 345-35 (1995).
- [3] Li, D., Parkhurst, D. J., "Open-Source Software for Real-Time Visible-Spectrum Eye Tracking", Proceedings of the COGAIN Conference, 18-20 (2006).