

視覚認知における知覚と記憶

齋木 潤

京都大学大学院情報学研究科

Visual cognition is a product of dynamic interaction between focused and distributed attentional mechanisms. Our group has been investigating these two mechanisms by focusing on the following issues; visual working memory for object representations, and underlying mechanisms of search asymmetry. We are trying to take “multidisciplinary approach”, where computational, psychophysical, and neuroimaging approaches are integrated. I report our current progress and future directions.

Keywords: visual cognition, visual working memory, visual search, focused and distributed attention.

はじめに

認知科学に最も大きなインパクトを与えた仕事のひとつであるMarr (1982)を引くまでもなく、視覚認知という脳の高次機能を理解するためには、異なるレベルのアプローチを有機的に組み合わせた研究手法が不可欠である。このことはMarrが強調した説明のレベルということだけでなく、既存の方法論がそれぞれ持つ限界にも大いに関係する。我々のグループでは、このような問題意識から、心理物理実験、機能的脳イメージング、計算論的モデリングを組み合わせたアプローチによって、視覚認知の過程を解明することを目指している。現時点では必ずしも目的が達成されたわけではないが、本講演では、2つのトピックを取り上げ、progress reportとして、我々がどこまで到達したのかを紹介したい。

フレームワーク

我々が現在研究している2つの主要なトピックは、「物体の一時的記憶における特徴統合」と、「分散的注意と視覚探索」である。

我々は、視覚認知が2つの過程の相互作用によって成立すると考えている。1つは、分散的注意、Vision at a glanceなどと呼ばれる過程で、視野全体にわたる視覚情報からすばやく注視すべきものを検出し、定位したり、シーン全体の統計構造を用いてシーンのカテゴリ化などを行なう。もうひとつは、第1の過程で選択された位置に選択的注意を向けることによってそこにある物体の詳細構造を認知するメカニズムで、選択的注意、vision with scrutinyなどと呼ばれている。古典的な見方では、この2つの過程が協働して、外界の詳細な内部モデル（デカルト的な意識の劇場モデルのようなもの）を生成すると考えられてきたが、我々はそのような統合は存在しないと考える（Figure 1）。選択的注意メカニズムは計算した物体の構造情報を蓄積していくことはなく、分散的注意メカニズムが効率的に処理すべき情報に誘導することで、あたかも見たい時に見たい情報がそこにあるかのような「幻想」を持つことが出来る。

このような見方の妥当性を検討するために、我々は上述した2つのトピックを取り上げている。「物体の一時的記憶における特徴統合」に関しては、我々が特徴を統合して一時的に保持できる情報は極めて少ない

ことを明らかにすることを目指している。「分散的注意と視覚探索」では、限定された物体の一時記憶だけで外界と相互作用するために不可欠な効率的な注意の移動と、シーンの全体情報の把握のメカニズムの解明を目指している。

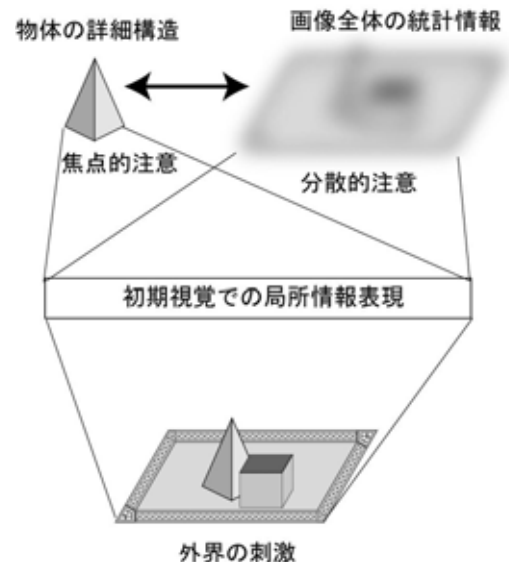


Figure 1. General framework of our approach to visual cognition. Modified from Saiki (2004).

物体の一時的記憶における特徴統合

定説

視覚情報の一時的記憶に関しては古くから多くの研究があるが、最近物体情報の一時的記憶に関する多くの研究が行なわれている。概念的にはKahneman, Treisman, & Gibbs (1992)のオブジェクトファイル、実験研究としてはLuck & Vogel (1997)の変化検出課題を用いたものを初めこれらの研究に共通する点は、視覚的ワーキングメモリが特徴の統合された物体表象を機能単位としそれを3 - 5個保持できると考えていることである。

計算論的考察

特徴が統合された物体を複数同時並列的に記憶内に一時的に保持することは計算論的にはtrivialな問題ではない。古くからバインディング問題として知られる問題を解決する必要があるからである。例えば、Luck & Vogel (1997)では神経発火の同期による属性のバインディングを用いて複数物体表象の保持が実現されている可能性に言及しているが、計算論的に見てこれが脳で実現されている可能性は極めて低いと考えられる。実際、最近の研究では、発火相関の機能は、属性のバインディングよりもむしろ注意による情報のゲーティングである可能性が示唆されている。このような考察から、脳が特徴を統合した表象を3 - 5個並列に保持するメカニズムを考えるのではなく、そもそもその仮説の妥当性を厳密に検証する可能性があると考えた。

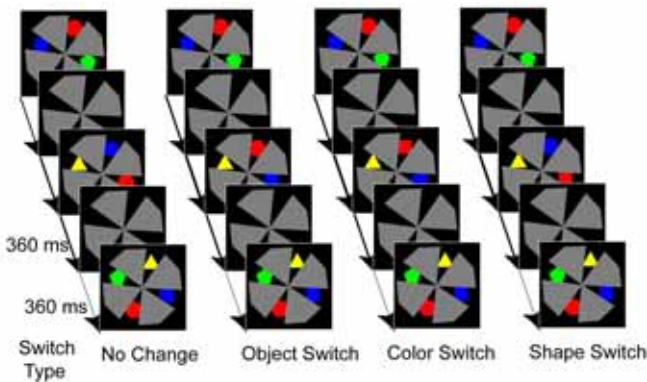


Figure 2. Schematic illustration of multiple object permanence tracking (MOPT) task.

心理物理実験 (MOPT)

物体の視覚的ワーキングメモリに関する計算論的考察から3 - 5個という容量の妥当性に疑問があること、さらに従来の研究が必ずしも特徴が統合された表象の記憶を調べるデザインになっていないことから、われわれは多物体恒常性追跡法(MOPT)と呼ばれる実験パラダイムを開発し、物体表象の視覚ワーキングメモリをより厳密に検討した。MOPTは遮蔽版の裏で回転により見え隠れする4 - 6個の物体の遮蔽時の属性の組合せの入替えを検出、或いは入れ替えタイプの同定をする課題である。属性の入れ替えの判断は属性ごとの記憶ではなく属性の組合せが不可欠なこと、また入れ替えタイプの同定では単なる顕著性の変化では課題が遂行できないことから、従来の変化検出デザインよりも厳密に特徴統合の記憶を評価できる。MOPTによる一連の実験から以下のことが明らかになった(Saiki, 2003a, 2003b)。(1)物体が静止している場合でも保持容量は1 - 2個程度に過ぎない、(2)物体が運動していると課題はさらに困難になる、(3)変化後手がかりによる成績向上が観察されないことから課題の

困難さは単なる記憶検索のボトルネックではなく、物体表象の保持そのものの問題と考えられる。

機能的脳イメージング

心理物理実験から、物体表象の保持容量は定説よりもずっと少ない可能性が示唆された。物体表象の保持の脳内メカニズムを明らかにし、計算論的モデルを構築する上での制約を明確にするために、MOPT課題を用いた機能的脳イメージング研究を並行して行っている。まず、一番単純な位置と色の結合が必要な課題を用いブロックデザインで課題時の脳活動のグローバルな状況を解析した。その結果、多物体追跡法や視覚的注意や眼球運動の制御で共通して観察される前頭頂ネットワークの活動に加えて、これらの課題では通常活動の見られない前部前頭前野(BA10)の有意な活動が観察された(Imaruoka, Saiki, & Miyauchi, in press)。また、この2つの下位領域の活動のパターンに質的な差異が見られることがわかった。現在、事象関連デザインを用いて、物体情報の保持と、変化の検出にかかる活動の分離を進めている。

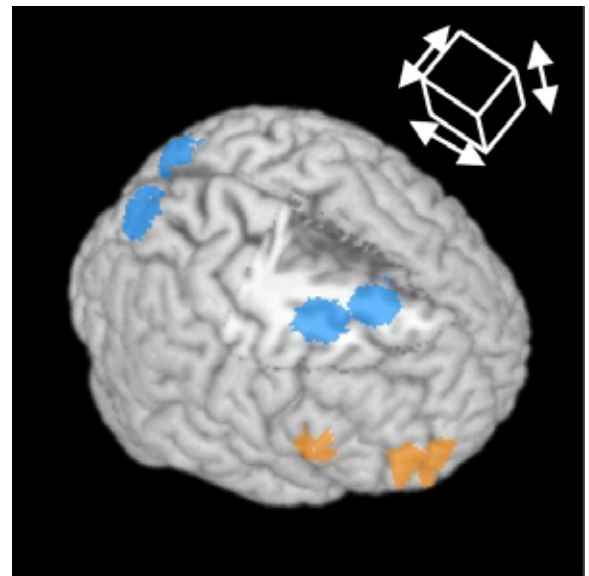


Figure 3. The regions of interest (ROIs) revealed by an fMRI experiment with MOPT task. Blue and orange regions are likely to reflect tracking and maintenance of color-location binding during MOPT task, respectively.

まとめ

物体表象の一時的保持におけるバインディング問題という計算論的な考察から発想を得て、心理物理実験の開発とデータ収集、それに対応した機能的脳イメージング研究の結果から、従来の定説とは異なり、我々の視覚認知システムが一時に保持できる物体情報は極めて限られていることが明らかになった。計算論的にはバインディング問題を回避するには一時に1個ずつの物体のみを保持することが必要であるが、現時点ではこの点についての説得的な証拠は得られておらず、今後の重要な課題である。

分散的注意と視覚探索

何が必要か？

MOPTを用いた一連の研究は、我々の視覚的作業記憶は定説よりもはるかに厳しい容量限界を持っているらしいことを示している。しかし、我々が一度に1-2個の物体情報しか保持できないとすると何故、我々は日常的にたくさんの物体を同時に認識しているように思えるのだろうか？この問いに対する有望な答えとしては、Rensink (2000)が述べているように我々は“just in time”で必要な視覚情報を抽出する極めて効率的なメカニズムを持っているというものであろう。しかし、問題はそのような効率的なメカニズムを脳が実装するためには何が必要なのかということである。

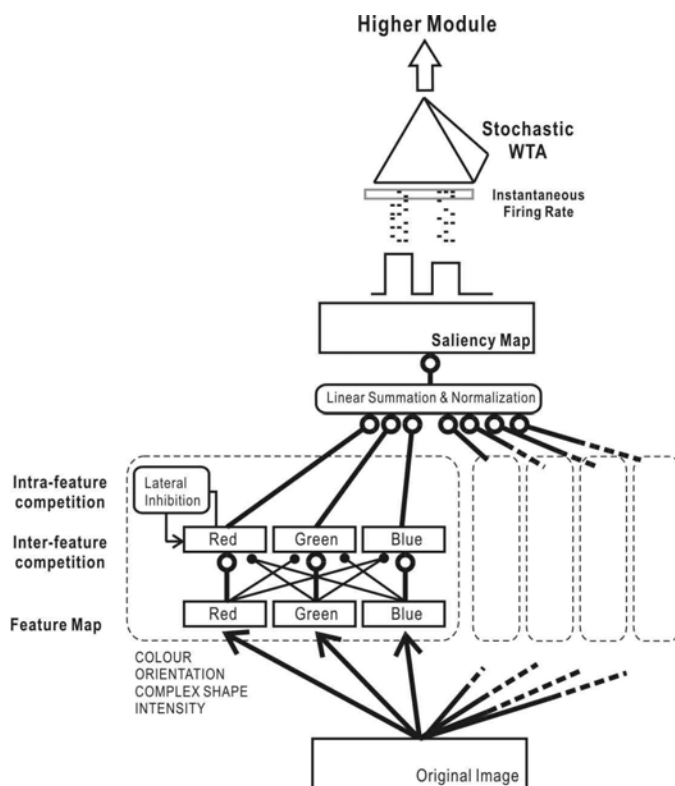


Figure 4. Schematic illustration of stochastic saliency map model by Koike & Saiki (2002).

修正顕著性マップモデルからの予測

計算論的モデル研究の立場からこの問題に迫るものとして顕著性マップモデルがある。顕著性とは、特徴情報を融合してある視野上の位置の視覚情報の顕著さの度合いを表現しているとする仮説的構成概念であり、視覚探索の計算モデルでは顕著性をマップ表現した表象を用いて注意を向ける対象を決定していると仮定している。従来のモデルは、特徴探索と結合探索の違いなど視覚探索に見られるいくつかの重要な知見をうまく説明できるが、例えば、探索非対称性はうまく説明できない。これは、顕著性マップモデルが注意を向け

る対象を決定するプロセスが顕著性の値の順序のみによって決まり、値の差分を反映しないことに起因する。このため、モデルは探索非対称性を示す刺激ペアに対し、ポップアウトと極端に非効率的な探索という二極分化したパフォーマンスを示してしまう。Koike & Saiki (2002)はこの問題を解決するため、顕著性マップモデルに2つの修正を加えた。1つは、非効率的な探索時にターゲットが常に最低の顕著性値を取らないようにするための特徴間統合のメカニズム、もうひとつは顕著性の値の分布を反映した注意の移動を実現する確率的顕著性マップの導入である。これらの修正により、探索非対称性のデータをよく再現することが可能になった。また、この修正モデルでは、探索非対称性はボトムアップ的なメカニズムのみによって生起するので、ターゲット刺激の確定性が探索非対称性の生起の必要条件ではないことを予測する。

心理実験

Wolfe (1994)の誘導探索モデルを初めとするいくつかのモデルは探索非対称性が目標刺激の変化による探索属性の変化によると説明している。これに対し、我々の修正顕著性マップモデルは目標刺激を知ることが必要条件ではないと考える。この点を検討するために、シングルTON探索課題で探索非対称性が生起するか否かを検討した。その結果、OとQ(円に垂直バーが付加されたもの)、漢字の舌とその鏡映反転、いずれの場合も、ターゲットが既知の場合、シングルTON探索の場合でほぼ同程度の強さの探索非対称性が観察された。このことは修正顕著性マップモデルの予測を支持し、目標刺激の変化による探索特徴の切り替えという説明の妥当性に疑問を呈する。

探索属性は何か？

通常の視覚探索課題を用いた一連の実験は確かに、モデルの予測どおり、標的の刺激が既知であるかに関係なく同程度の探索非対称性を示した。このことは、ある種のモデルが仮定するように探索非対称性の生起にトップダウン的な探索属性の選択は必要ないことを示唆するが、探索非対称性の生起メカニズムに迫るためには実際に我々が何を探しているのかをより直接的に知る方法が必要である。古典的な実験計画法はあらかじめ統制された刺激に対する反応傾向を見るため、網羅的、体系的な実験デザインを組まない限り刺激に含まれるどの情報が反応を導いているのか推定することが困難である。例えば、OとQの探索に用いられる可能性のある視覚属性は多数ある。これに対し、近年、逆相関法を心理物理学実験に応用したClassification Image (CI)法を用いて、反応を導いた刺激の情報をより直接的に推定する手法が脚光を浴びている。我々はCI法を用いてOとQの探索において用いられている視覚属性を推定するとともに何が探索非対称性の元になる過程かを検討した。その結果、(1)標的定義探索、シングルTON探索ともバーの有無を探索属性としていること、(2)標的がどちらであっても探索属性は同じであること、(3)エラーはほぼ全ての場合に

バーの存在を検出しそこなうことによっており，探索非対称性はバーを持つ刺激の数で決まる，ことがわかった．

まとめ

顕著性マップモデルに関する考察から視覚探索の効率を決めるメカニズムに関していくつかの実験的な研究を展開した．特に，探索非対称性をめぐっていくつか新たな知見が得られた．

今後の展開

物体の構造の認知における視覚 - 触覚統合

我々が物体の詳細構造を認知する必要があるのは単にそれを見るためだけではなく，多くの場合その物体と相互作用するためである．従って，多くの日常場面では，物体の視覚情報と触覚情報をうまく組み合わせることによって物体との相互作用が成立している．現在，力覚呈示装置PHANToMを用い，動的な知覚場面における視覚情報と触覚情報の統合，及び複雑な物体やシーンの記憶における視覚情報と触覚情報の相互作用についての研究を始めている．

シーンのカテゴリ知覚

“Vision at a glance”のシステムの機能は，視覚探索の際に効率的に注意を誘導することだけではなく，視野全体の視覚情報の統計的構造を抽出しシーンのレイアウトやgistを知覚することがある．我々の視覚システムが実際に極めて効率的にシーンを知覚できることはBiedermanの古典的研究や最近のultrarapid categorizationの研究からも明らかである．しかし，何故それが可能なのかはほとんど解明されていない．現在，心理物理実験と，自然画像の統計構造解析の手法を組み合わせ，シーン認知で利用される画像情報を解明することを目指している．

おまけ 文化と知覚

ミシガン大学のKitayamaやNisbettとの共同研究で，従来文化心理学の文脈でよく行なわれてきた思考や推論を含むような高次の認知課題ではなく，視覚探索のようなより知覚に近い課題でも有意な文化差があることが示された．我々の視覚経験のかなりの部分を構成する人工物はまさに文化の反映であることを考えると，文化は知覚学習の結果として，我々の日常の視覚経験を通して「身体化」されていると考えることはそれほど無理な話ではないと思われる．知覚学習や視覚シス

テムの可塑性の問題と関連付けながら知覚と文化の問題にアプローチすることは興味深い課題である．

おわりに

視覚認知の理解のためには，心理実験，モデル，イメージングといったひとつの手法に閉じるのではなく，それらを有機的に組み合わせるとともに，新しい手法を積極的に取り入れることが重要である．近年，自然画像の統計解析，CIなど注目すべき手法が近接領域で開発されており，こうした手法を取り込みながら視覚認知研究の方法論を洗練しつつ，より深い理解を目指すことが肝要であると思われる．

引用文献

- Imaruoka, T., Saiki, J., & Miyauchi, S. (in press). Maintaining coherence of dynamic objects requires coordination of neural systems extended from anterior frontal to posterior parietal brain cortices. *NeuroImage*.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object specific integration of information. *Cognitive Psychology*, **24**, 175-219
- Koike, T., & Saiki, J. (2002). Stochastic guided search model for search asymmetries in visual search tasks. *Lecture Notes in Computer Science*, **2525**, 408-417.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.7).
- Marr, D. (1982). *Vision*. Freeman: San Francisco.
- Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research*, **40**, 1469-1487.
- 齋木 潤 (2004). 分析 統合から相互作用へ：自然画像を用いた視覚認知研究の展開．*学術月報*, **57**, 790-793.
- Saiki, J. (2003a). Feature binding in object-file representations of multiple moving items. *Journal of Vision*, **3**, 6-21.
- Saiki, J. (2003b). Spatiotemporal characteristics of dynamic feature binding in visual working memory. *Vision Research*, **43**, 2107-2123.
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, **1**, 202-238.