

複数の属性を持つ視覚作業記憶の情報表現に関する研究

古徳 雅史
諸上 茂光
森田 昌彦

筑波大学大学院システム情報工学研究科
筑波大学大学院システム情報工学研究科
筑波大学機能工学系

人間の視覚系では、視覚オブジェクトを色、形、運動、テクスチャなどの属性ごとに処理した後、属性の異なる特徴を統合している。視覚作業記憶では、こうした特徴統合が行われた状態で情報が保持されていることが示されているが、すべての特徴が統合された表現が保持されるかどうかは疑問である。これに関して、3つの属性をもつオブジェクトの各特徴が変化しただけかを答える実験を行ったところ、2つの特徴が同時に変化したときの正答率が最も高かった。また、実際には1つの特徴しか変化していないときに2つの特徴が変化したと誤答する率が高いのに対し、その逆はほとんどないこともわかった。この結果は、2つの特徴を統合した3種類の表現が作業記憶として保持されることを示唆する。

キーワード：特徴統合、ワーキングメモリ、2属性モデル

問題と目的

Luck & Vogel(1997)の実験により、人間は少なくとも4つ程度の視覚オブジェクトを作業記憶に保持することができること、そしてそのオブジェクトが色と形状のような異なる2種類の属性をもつ場合も、それらを統合した形で保持していると考えられることが明らかになっている。では、オブジェクトが3種類以上の属性を持つときも同じであろうか。つまり、すべての属性を統合した表現となって記憶されるのであろうか。もしそうだとすれば、属性の数がいくつでも記憶容量は変わらないはずであるが、実際にLuckとVogelの実験を3および4属性で行ったところ、記憶容量以外にも、このようなモデルに合わない結果が得られた。一方、統合されない形、すなわち属性ごとに独立に記憶されるというモデルは否定されている。この疑問を解くために、以下の実験を行った。

方法

被験者：視力0.7以上で正常な色覚を有する18～24歳の大学生男女10名。

刺激：被験者前方114cmの位置に置かれたディスプレイ画面上に提示される、視角1.6度、等面積、等速で移動する図形。

手続き：実験は暗室内でおこなわれる。図1のように、3通りの色(C)、形(S)、運動方向(M)を組み合わせた3アイテムのサンプル刺激を提示し、1000ms後、1アイテムのテスト刺激を提示する。被験者は、テスト刺激のアイテムとその位置にあったサンプル刺激のアイテムとを比較し、応答画面において各属性が変化しただけかをチェックする。テスト刺激には、600試行中、変化する属性が、なし(150)、C(50)、M(50)、S(50)、CM(50)、MS(50)、SC(50)、CSM(150)の8条件を用意しランダムにおこなった(括

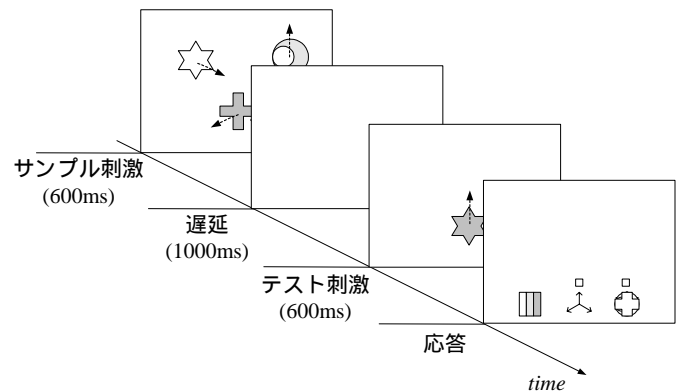


図1 実験の流れ：動きと形が変化した場合

弧内は条件ごとの試行数)。また、変化した属性を報告してもらうセッションと、変化していない属性を報告してもらうセッションとを前半と後半で入れ替えた。

結果

被験者が変化すると報告した属性を、条件ごとにまとめた結果を表1に示す(10名の被験者の平均)。左上から右下へ網掛けした部分が正答した割合を表している。

実際には変化する属性の数は等確率であるにもかかわらず、被験者は3つの属性のうち2つが変化したと感じる場合が最も多かった。そして、ほとんどの被験者について、2つの属性が変化したときの正答率が最も高く、全ての被験者で、1つの属性だけが変化した場合よりも、2つの属性が変化したときの方が正答率が高かった(図2)。

興味深いことに、1つの属性が変化したときの誤りは、ほとんどが2つの属性が変化したというものであったのに対し、2つの属性が変化したとき、1つだけ変化したと誤って答えることはほとんどなかった。

表1 被験者の回答分布(括弧内はモデルの出力分布)

| | | 変化したと回答した属性 (%) | | | | | | | |
|--------------|-----|-----------------|--------|--------|--------|--------|--------|--------|--------|
| | | なし | C | M | S | CM | MS | SC | CSM |
| テスト刺激で変化した属性 | なし | 49(53) | 13(10) | 5(10) | 13(10) | 4(4) | 6(4) | 9(4) | 1(5) |
| | C | 1(7) | 59(53) | 1(2) | 3(2) | 11(15) | 2(0) | 20(15) | 2(6) |
| | M | 3(7) | 1(2) | 52(53) | 2(2) | 16(15) | 19(15) | 3(0) | 4(6) |
| | S | 6(7) | 3(2) | 3(2) | 49(53) | 1(0) | 14(15) | 22(15) | 2(6) |
| | CM | 1(1) | 5(8) | 2(8) | 1(2) | 71(63) | 5(3) | 5(4) | 11(12) |
| | MS | 0(1) | 1(2) | 8(8) | 6(8) | 10(4) | 61(63) | 4(3) | 10(12) |
| | SC | 0(1) | 9(8) | 0(2) | 3(8) | 4(4) | 2(3) | 70(63) | 11(12) |
| | CMS | 1(0) | 2(3) | 2(3) | 1(3) | 19(12) | 7(12) | 9(12) | 59(56) |
| | 計 | 13(15) | 26(29) | | | 42(37) | | | 18(20) |

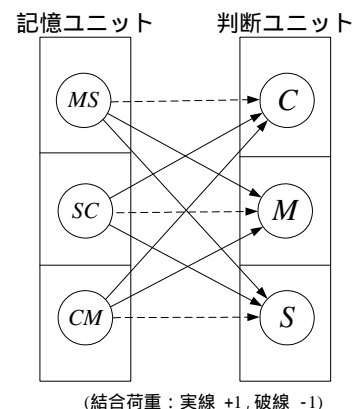


図3 2属性仮説に基づくモデル

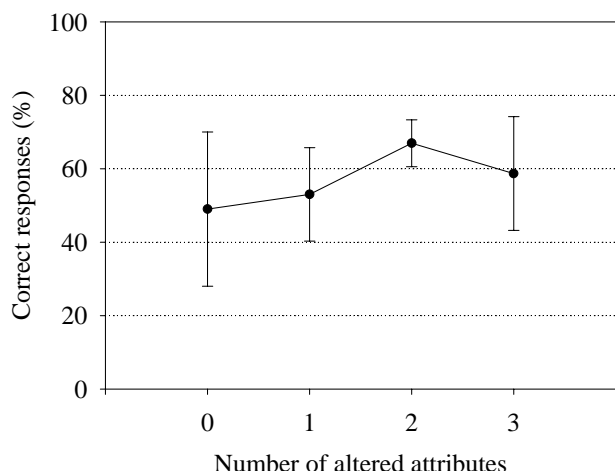


図2 変化した属性の数ごとの正答率(エラーバーは標準偏差)

考察と結論

この結果は、3つの特徴を属性ごとに独立して記憶している(1属性仮説)というモデルでも、3種類の特徴を統合した形で記憶している(3属性仮説)というモデルでも、両者の組み合わせでも、説明するのは非常に難しい。これに対して、2属性を統合した3種類の表現が保持されている(2属性仮説)とするならば、比較的容易に説明できる。例えば、2属性が変化した時の正答率が1属性のみ変化したときより高いのは、1属性のみ(例えばS)の変化を正しく検出するには2つの表現(SCとMS)の両方を正しく保持していなければならないのに対し、2属性が変化した場合には、その2属性の組だけ保持していれば正答可能だから、という説明ができる。

より具体的に、図3のようなニューラルネットモデルを考えよう。入力層の3つのユニットは、それぞれMS, SC, CMの統合した表現に対応しており、

保持している情報とテスト刺激の提示によって入力される情報とを比較してミスマッチ信号を出す。但し、2特徴の両方が異なるとき2が、一方のみ異なるとき1が出力されるとする(両方一致しているときは0)。3つの出力ユニットは、それぞれの属性に対応し、入力ユニットの荷重和 u に基づき、確率 $f(u)=1/(1+\exp(-cu))$ で1を出力、すなわちその属性が変化したという反応を示す。記憶が完全であり、かつ出力ユニットのしきい値 $h=1$ のとき、このモデルはどの場合も同じ正答率(c が大きいほど高い)を示す。

さて、このモデルにおいて、各ユニットの記憶はある一定の確率 p で失われるとする。このとき、そのユニットの出力は無条件で0となると同時に、 h が1だけ(記憶が2つ失われると2)下がると仮定する。このとき、例えば $c=2, p=0.1$ においてモデルの反応分布を計算すると、表1のようになる。誤答の生じ方を含めて、ほぼ実験データと一致することがわかる。

このモデルが脳における特徴変化検出のモデルとしてどの程度妥当であるかはわからないが、少なくとも2属性仮説に基づくならばごく単純なモデルでデータを再現でき、それ以外の仮説ではできないという事実は、2属性仮説を強く支持するものと言えよう。

また、2属性モデルによれば、1属性および2属性のオブジェクトの容量は等しいが、3属性オブジェクトはやや低下し、4属性になるとかなり低下することになる。この点についても2属性モデルの妥当性を示唆する結果が得られている。

引用文献

Steven J. Luck, Edward K. Vogel. 1997: The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279-281.