

三次元物体学習過程に依存した眼球運動の変化

上田 祥行

京都大学人間・環境学研究所
日本学術振興会

ueda@cv.jinkan.kyoto-u.ac.jp

齋木 潤

京都大学人間・環境学研究所

In the 3-D object recognition task, the patterns of eye movements during the learning phase were different depending on the encoding strategy. In order to reveal the temporal process of 3-D object encoding, we measured eye movements during 3-D object learning repeatedly. In experiment, participants were assigned to either the rotation or non-rotation condition. An unfamiliar 3-D object was presented for 10 seconds at the study phase, during which participants' eye movements were recorded. In the beginning of experiment, participants fixated the center of components of objects more frequently in the rotation condition, suggesting that objects were encoded more categorically. The proportions of global saccade were the same at first, but, after a few trials, they changed depending on the test condition. In the beginning of a trial, the arising proportion of global saccade was significantly higher in the rotation condition than in the non-rotation condition. These results suggest that participants may encode 3-D objects more categorically in the rotation condition, and after a few trials they made detailed representations based on stored information.

Keywords: 3-D object recognition, eye movement, volume-based encoding, image-based encoding.

問題・目的

我々は新奇な三次元物体であっても、比較的短時間の学習によって、視点に関係なく同一物体であるかどうかを判断することが可能になる。新奇物体を初めて見たときからこのような判断ができるようになるまでに、観察者が持つ内的な物体表象がどのように変化しているのかは明らかになっていない。我々は、どのような学習過程を経ることで物体の視点独立的な認識を可能にしているのだろうか。

三次元物体の認識は知覚した物体と記憶に蓄えられた物体の表象を統合することで行われると考えられている(Bülthoff & Edelman, 1992)。観察者は、内的な物体表象に不完全な部分があればその部分を重点的に学習することで、その物体に関してより精緻な内的表象を得ていることが推測される。

本研究の目的は、三次元物体学習中の眼球運動の測定を通じて観察者が得た物体の情報を推定し、三次元物体の学習がどのように進んでいくのかを検討することであった。これを達成するために、実験では三次元物体が様々な視点から呈示され、物体の三次元的な表象が必要となる回転条件と、常に同じ視点から呈示され、物体の二次元イメージの学習のみが必要となる無回転条件の二つの状況における物体の学習再認課題を行い、学習中の協力者の眼球運動の変化を測定した。

三次元物体学習中の眼球運動のパターンは観察者の学習ストラテジーを反映していることが示唆されている(Ueda & Saiki, under review)。特に物体を構成するコンポーネントに関して、それらのカテゴリカルな情報を取得する場合にはコンポーネントの中心を見る傾向がある。また、跳躍距離の長いサッカーボール運動は物体全体を概観している運動であると考えられている(e.g. Zangemeister, Sherman, & Stark, 1995)。本研究では、このような眼球運動のパターンに焦点をあて、三次元物

体の学習が時系列的にどのように変化していくのかを検討した。

方法

実験協力者 京都大学の大学生および大学院生8名が実験に参加した。協力者はランダムに回転もしくは無回転条件に割り当てられた。

刺激 4種類の新奇な三次元物体をBlender(<http://www.blender3d.org/cms/Home.2.0.html>)を用いて作成した。それぞれの物体は4種類のGeon (Biederman, 1987)で構成されており、各物体におけるコンポーネントの空間関係はすべて同一であった。それぞれの物体に対して妨害刺激として、コンポーネントのカテゴリーや接合の空間関係を保ったまま量的な変化が施された物体が3種類と、これらが変わるように量的な変化が施された物体が3種類作成された。刺激の大きさは視角にして縦11°-15°、横18°-22°であった。

手続き 各試行の最初にモニタの左右端のどちらかに注視点が2秒間呈示された。注視点が消失した後、モニタの中央に学習物体が10秒間呈示され、協力者は自由に目を動かして物体を学習した。学習後には、すぐに再認課題が行われた。再認時には学習と同様に、モニタの左右端のどちらかに注視点が2秒間呈示され、注視点の消失後にテスト物体が1秒間呈示された。回転条件では、テスト物体が学習と同じ視点あるいは様々な異なった視点からランダムに呈示され、無回転条件では、テスト物体は常に学習と同じ視点から呈示された。協力者は視点に関わらずテスト物体が学習した物体と同じかどうかを、できるだけ早く正確に判断してキー押しを行った。学習時には物体は常に同じ視点から呈示され、協力者が途中で自由に学習を打ち切ることも可能であった。半数の試行では学習物体と同じ物体がテスト物体として呈示され、残りの半数では

学習物体と質的もしくは量的に異なった物体がテスト物体として呈示された。

これらの手続きを1試行とし、それぞれの学習物体において直近6試行の正答率が100%になった時点でその物体の学習が完了したものとした。実験は全ての物体の学習が完了するか、各物体で75試行が行われた時点で打ち切りとした。このため、理論上は最大で計300試行が行われる可能性があった。学習中の眼球運動はEyelink I(SR Research社製)によって記録された。

結果

分析では、80ms以下の短い停留点は除外された。学習は回転条件では平均36.9試行、無回転条件では18.1試行で終了した。

各物体の1-3回目の学習における、停留点から物体を構成するコンポーネントの中心までの距離の時系列的な推移をFigure 1に示す。各物体の1-3回目の学習では、いずれにおいても25回以上の停留が見られたことから、1-25回目までの停留においてANOVAを用いて検定を行った。その結果、回転条件では、無回転条件よりもより頻繁にコンポーネントの中心が見られていたことが示された($F(1,6) = 4.97, p < .10$)。これは、学習の初期の段階において、回転条件でよりカテゴリカルな物体の記録を行っていたことを示唆している。

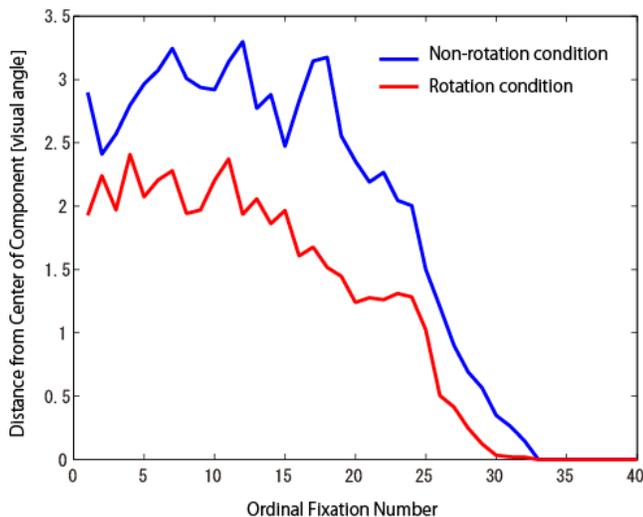


Figure 1. The distance from fixation locations to the center of components in 1st- 3rd trial. The blue plot indicates the non-rotation condition, and the red plot indicates the rotation condition.

次に、物体全体を概観するような運動がどのように推移したのかを調べた。コンポーネントの中心間距離の最小値の半分よりも大きなサッカードが行われた割合をFigure 2に示す。コンポーネントの中心間距離の半分よりも大きなサッカードはコンポーネント間を移動する動きであり、物体全体を概観していると考えられる。Figure 2Aは学習の初期段階である各物体の1-3回目の学習における長距離サッカードの割合を示して

おり、Figure 2Bでは正答率が上昇し、少し学習が進んだ状態である6-8回目の学習における長距離サッカードの割合を示している。各試行の1-20回目までのサッカードにおいてANOVAを用いた検定を行ったところ、1-3回目の試行における長距離サッカードの割合に有意な回転条件の差は見られなかったが($F(1,6) < 1, n.s.$)、6-8回目の試行では回転条件とサッカード数の間に交互作用が見られ($F(19,114) = 1.50, p < .10$)、試行開始直後のサッカードにおいて回転条件の方が無回転条件よりも長距離サッカードの割合が有意に大きかった。これは、ある程度の学習が進んだ後、回転条件では、各試行の最初に物体を概観する運動が起こっていることを示唆している。

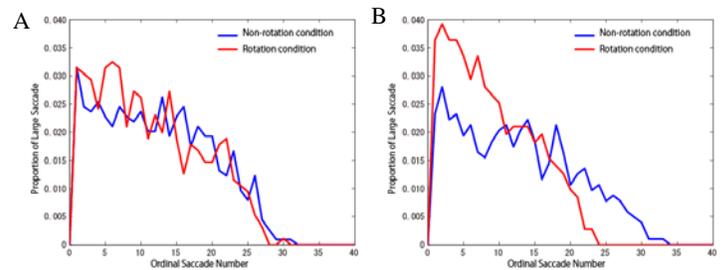


Figure 2. The arising proportion of large saccades in 1st-3rd trial (A) and 6th-8th trial (B). The blue plot indicates the non-rotation condition, and the red plot indicates the rotation condition.

考察

本研究では、三次元物体の学習過程を明らかにするために、新奇な三次元物体を繰り返し学習し、学習中の眼球運動を測定した。物体の三次元的な表象が必要になると考えられる回転条件と二次元イメージのみが必要である無回転条件での眼球運動を比較したところ、回転条件では学習過程の初期の段階でコンポーネントの中心付近により頻繁に停留しており、学習が進んだ状態では物体を概観するような眼球運動が試行の最初に見られるようになった。これに対して、無回転条件では、学習過程の初期でも学習が進んだ後も大きなサッカードが起こる割合に変化は見られなかった。これらの結果は、回転条件では無回転条件に比べて、コンポーネントを中心とした表象を作成していることを示唆している。また、このような表象が学習の早い段階で作られていることも示唆している。

引用文献

- Bülthoff, H. H., & Edelman, S. (1992). Psychological Support for a 2D view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences, USA*, **89**, 60-64.
- Ueda, Y., & Saiki, J. (under review). Distinct 3-D object encoding strategies for intra- and inter-modal recognition as revealed by the spatial distribution of fixations.