

視覚と触覚を用いた三次元物体認識における2つの学習ストラテジー

上田 祥行
齋木 潤

京都大学人間・環境学研究所
日本学術振興会
京都大学人間・環境学研究所

The patterns of recognition performance were different between intra- and inter-modal three-dimensional (3-D) object recognition tasks; the recognition performance was not affected by observer's viewpoint in inter-modal condition, whereas the learned viewpoint showed an advantage in intra-modal recognition (Ueda & Saiki, 2007). The first purpose of this study was to show the different encoding strategies between intra- and inter-modal recognition, which may lead to the different recognition performance. The second was to reveal what feature observer used in each recognition. In order to achieve these purposes, we measured observer's eye movement during encoding 3-D objects. The patterns of eye movement during encoding were significantly different depending on prespecified test modality. Moreover, clustering of fixation data showed different patterns, possibly reflecting different learning strategies; participants focused on segment points in intra-modal recognition, whereas they focused on the center of each component in inter-modal condition. These different patterns of eye movements may reflect different strategies in learning 3-D objects, leading to different recognition performance.

Keywords: intra-modal object recognition, inter-modal object recognition, eye movement, vision, haptics.

問題・目的

我々は眼で覚えた三次元物体の形状を比較的容易に触って再認することができる。視覚と触覚を用いた物体認識の実験では、モダリティ内で物体の学習 - 再認を行ったときには視点依存的な再認成績のパターンが見られるのに対し、モダリティ間で物体の学習 - 再認を行ったときには視点独立的な再認成績のパターンが見られることが報告されている(Lacey, Peters, Sathian, 2007; 上田・齋木, 2007)。これらの結果は、モダリティ内の物体の学習 - 再認とモダリティ間の物体の学習 - 再認は異なった情報に基づいて行われることを示唆している。

このような再認成績の違いは、学習時に用いる情報に基づくものか、再認時に用いる情報に基づくものか明らかではない。上田・齋木(2007)は、学習時における再認感覚の教示の違いによって再認成績が変化したことを報告しているが、再認成績を用いたアプローチでは、実際に学習時に用いていた物体部位の特徴の違いを明らかにすることが困難である。そこで本研究ではこれらの違いを明らかにするために、三次元物体学習中の眼球運動を測定した。本研究の第1の目的は、モダリティ内とモダリティ間の物体認識で学習時に注視される領域が異なることを示すことであった。

第2の目的は、新奇な三次元物体の学習時の眼球運動を測定することで、学習のストラテジーを推測し、間接的に学習時に生成される表象を推定することであった。先行研究で見られた再認パターンから、モダリティ内の物体認識では視点に依存的な特徴に基づいて、モダリティ間の物体認識では視点に独立的な特徴に基づいて、物体が学習されている可能性が考えられる。これらの仮説を検討するため、実験では新奇な三次元物体を視覚で学習し、その後、視覚再認(Intra-modal condition)もしくは触覚再認(Inter-modal condition)を行った。三次元物体学習中の協力者の眼球運動は記録されており、協力者には学習前に再認するモダリティが

教示された。この2つの条件下で、再認モダリティの変化が三次元物体学習時の眼球運動に与える影響を検討した。

方法

実験協力者 大学生および大学院生4名が実験に参加した。

刺激 実験には10種類の新奇な三次元物体を用いた。それぞれの物体は複数のGeon (Biederman, 1987)型のコンポーネントで構成されており、コンポーネント同士は1つの面で接していた。作成された刺激は、視覚ではBlender (<http://www.blender3d.org/cms/Home.2.0.html>)を用いた三次元グラフィックスとして呈示され、触覚では木で作られた実物が呈示された。

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、注視点が呈示された側と反対側に学習物体が2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。Intra-modal conditionでは学習と同様に、モニタの左右端のどちらかに注視点が2秒間呈示され、協力者は注視点の消失後にモニタに呈示されたテスト物体の再認を行った。Inter-modal conditionでは協力者は目を閉じ、体の前に呈示されたテスト物体を触って再認した。どちらの条件でも、テスト物体は学習と同じ視点あるいは様々な異なった視点から呈示され、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。テスト物体は協力者が反応するまで呈示された。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異なる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、36試行を1ブロックとして、各々の条件で10ブロック計720試行を行った。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

すべての協力者の再認率は90%以上であった。分析では、刺激・条件ごとにすべての協力者の停留点が集められ、80ms以下の短い停留点は分析から除外された(Intra-modal conditionの8.3%、Inter-modal conditionの7.8%)。カーネル密度推定によって推定した停留点分布の一例をFigure 1に示す。Hotellingの T^2 検定を用いて停留点の分布の違いを検討したところ、10種類の物体すべてで2つの条件間の分布が有意に異なった(e.g. object 2; $F(2, 1320) = 6.30, p < .01$)。

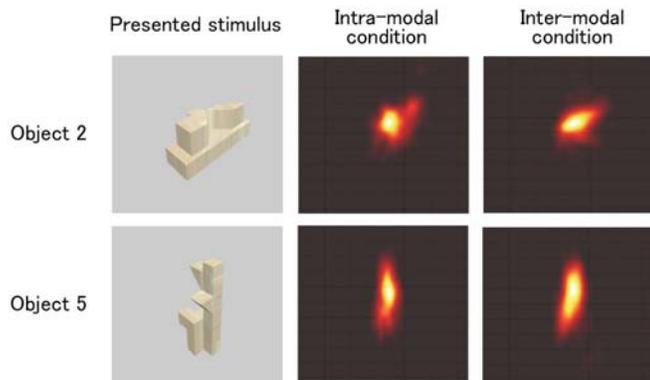


Figure 1. Kernel density estimation of fixation distributions.

次に、どのような情報を含んだ領域により頻繁に停留されていたのかを調べるために、Mean Shift (平均移動法)と呼ばれる方法を用いて空間位置の近い停留点をクラスタリングした(Santella & De Carlo, 2004)。結果をFigure 2に示す。クラスタリングの結果、Intra-modal conditionではコンポーネント同士の接続部を中心にして停留点が分布していたのに対し、Inter-modal conditionでは各コンポーネントの中心部を中心にして停留点が分布していた。コンポーネントの接続部と最も近いクラスタの中心との距離、およびコンポーネントの中心部と最も近いクラスタの中心との距離の平均をFigure 3に示す。クラスタの中心と接続部との距離はIntra-modal conditionで有意に小さく、クラスタの中心とコンポーネントの中心部との距離はInter-modal conditionで有意に小さかった。

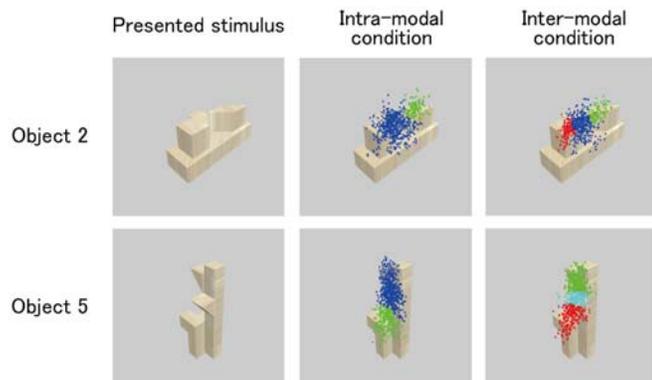


Figure 2. Estimated fixation clusters. Each cluster was presented by different color; there was no color correlation across pictures. Very small clusters were excluded (less than 3% of fixations).

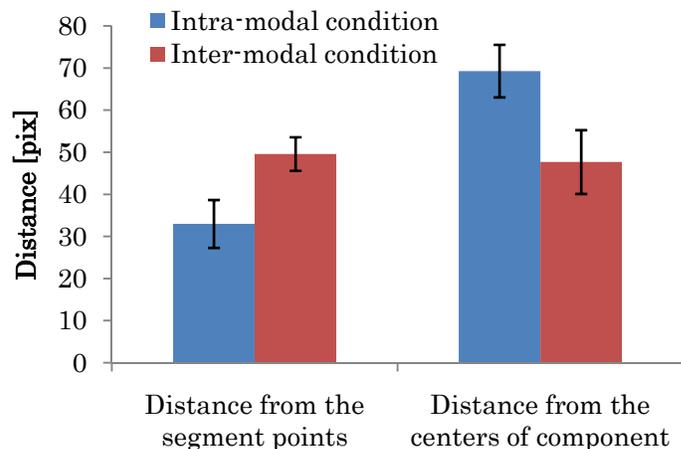


Figure 3. The mean distance between the center of cluster and the segment points or the centers of component.

考察

本研究では、モダリティ内とモダリティ間の物体認識において学習時に注視される物体部位が異なることを明らかにするために、新奇な三次元物体を学習中の眼球運動を計測した。その結果、Figure 1に見られるように眼球運動が2つの条件で有意に異なった。

さらに、それぞれの条件でどのような学習戦略が取られているのかを明らかにするために、どのような情報を含んだ領域がより頻繁に注視されたのかを調べた。その結果、Intra-modal conditionでは物体のコンポーネントの接続部がより頻繁に注視された。コンポーネントの接続部は多くのエッジや頂点などが集中しており、視覚的な情報量が多いと推定される。しかし、エッジや頂点の集まりは回転によって遮蔽されやすく、形や向きなども変化しやすい。このため、これらの情報を基に作成された表象は視点の変化の影響を受けやすくなると考えられる。

一方で、Inter-modal conditionではコンポーネントの中心部がより頻繁に注視されていることが明らかになった。各コンポーネントは複数のエッジや頂点の組み合わせによって決定されるGeonである。Geonは回転によって成績が低下しにくく、偶然の見え(accidental view)以外では再認が容易であることが知られている。Inter-modal conditionでは、特定の領域に集中した特徴ではなく、各コンポーネントを形成する特徴に注目していた可能性が示唆される。

本研究で見られた再認モダリティによって変化した眼球運動は、モダリティ内とモダリティ間の物体認識が異なる学習戦略によって行われていることを示唆している。

引用文献

Lacey, S., Peters, A., & Sathian, K. 2007 Cross-modal object recognition is viewpoint-independent. *PLoS ONE*, 2(9), e890.

上田祥行・齋木潤 2007 視覚と触覚を用いた物体認識における視点独立性 基礎心理学研究, 26(1), 11-19.