

A Thesis for the Degree of Ph.D. in Engineering

Speed-up of  
Biologically Inspired  
Image Categorization for  
Information Retrieval and  
Object Localization

February 2014

Graduate School of Science and Technology  
Keio University

Takuya Minagawa

# 主 論 文 要 旨

報告番号	甲 乙 第	号	氏 名	皆川 卓也
主論文題目： Speed-up of Biologically Inspired Image Categorization for Information Retrieval and Object Localization (生物学的知見に基づいた画像分類手法を情報検索と物体位置検出に応用するための高速化)				
(内容の要旨)				
<p>撮影された画像中に「何」が写っているのかを判断することは、人間にとっては比較的易しいタスクであっても機械にとってはまだまだ難しい課題である。霊長類が物体を「何」であるか認識する場合、特に物体提示後 100 ミリ秒から 130 ミリ秒で行われる「瞬間的な認識」に関しては、視覚皮質の腹側経路におけるフィードフォワードパスで行われていることがわかっている。この腹側経路におけるフィードフォワードパスを計算アルゴリズムとして近似したものとして HMAX モデルが知られているが、脳とコンピュータのアーキテクチャが大きく異なるため、この HMAX をアプリケーション向けに使用しようとした時、その処理速度が大きな問題となる。</p> <p>そこで本論文では、HMAX 特徴を実用的なアプリケーション構築のために高速化することを目指す。</p> <p>HMAX はもともと物体カテゴリ認識向けに構築されたモデルのため、まずは同タスク向けに高速化を行い、その応用例として情報検索サーバーを構築する。高速化にあたり、そのボトルネックである S2 層の処理負荷を軽減することを中心に、特徴量の簡略化、処理領域の削減、計算する特徴数の削減などを組み合わせることで、計算時間の削減を行う。</p> <p>更に HMAX 特徴を物体検出という、より高度なタスクへ適用する。この場合画像上の各位置でカテゴリ認識を行うことになり、更に計算の負荷が大きくなる。そこで、HMAX が形状との類似度を算出する特徴であることに着目し、画像中に存在する物体の持つ形状との類似度を手がかりに Coarse-to-Fine に物体の候補領域を絞り込む手法を提案する。これにより、Sliding Window を用いて HMAX を使用した際に発生する冗長性を大幅に削減することができる。</p> <p>提案手法の有効性を示すため、カテゴリ認識では Caltech101 や屋外シーン画像データに対して、物体検出では UIUC Car データセットや FDDB データセットに対し、その処理速度及び認識精度を定量的に検証した。それにより認識率の低減はわずかに抑えながら、カテゴリ認識ではおよそ 37 倍、検出では 250 倍以上の大幅な処理速度の改善を行うことができたことを確認した。</p>				

## SUMMARY OF Ph.D. DISSERTATION

School School of Science for Open and Environmental Systems	Student Identification Number	SURNAME, First name MINAGAWA, Takuya
<p>Title</p> <p>Speed-up of Biologically Inspired Image Categorization for Information Retrieval and Object Localization</p>		
<p>Abstract</p> <p>It is relatively easy task to understand “what” in an image not for a machine but for a human. Primates can recognize category of an image contents very rapidly: between 100 and 130 msec. A feedforward path of ventral stream in the visual cortex is known to be activated for this “rapid categorization” task. HMAX model were invented to approximate the activity of the feedforward path of ventral stream, however it is so slow to compute because of the different architecture between machines and brain.</p> <p>This thesis aims to speed up HMAX feature computation for applications.</p> <p>HMAX was originally applied to object categorization tasks, thus we first try to reduce the processing cost for that task and to build an image based information retrieval system. Mainly S2 layer that is the most time consuming process in HMAX is treated to speed up by simplified features, elimination of processing area, and reduction of the number of feature patches.</p> <p>Then more difficult application that is object localization is addressed. This task is more time consuming because object categorization must be executed on every sliding window. Because HMAX feature is extracted by computing similarities between an image and image patches, our approach searches the similar regions to the shape of target object as the candidates by coarse-to-fine. This approach reduces the redundancy of HMAX and sliding window enormously.</p> <p>The object categorization method was evaluated by Caltech-101 and scene image datasets, and the object localization method was evaluated by UIUC car and Fddb datasets. They showed that our categorization method is about 37 times faster than original HMAX, and our localization is 250 time faster or more than sliding window approach, with little reduction of recognition rates.</p>		