

【ロボット・AI 技術分野】

仮訳

## 学習するカメラ (英国)

2020 年 10 月 13 日

被写体をその場で学習して認識するカメラを開発した  
ブリストル大学とマンチェスター大学の共同研究は、  
インテリジェントなカメラを一步前進させるかもしれない。

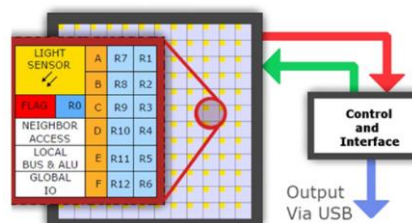
既存のシステムの感知・処理能力には課題があることは、ロボット工学・人工知能(AI)の研究者の間では周知されている。今でも視覚情報処理システムは、画像記録用に設計されたデジタルカメラのようなセンサーと、ビデオゲームの画像処理の高速化用に設計された画像処理装置(GPU)のようなコンピューティングデバイスを組み合わせているのだ。

つまり、センサーとプロセッサ間で視覚情報を記録・送信した後、AIシステムが被写体を認識する、ということだ。しかし、視界には、例えば自動運転車が走行する際の街路樹の葉の詳細など、タスクとは無関係なものも多く含まれることがよくある。現行では、これら全ての情報をセンサーが細部詳細まで捉え、タスクとは無関係なデータがシステムを詰まらせるので、エネルギーと処理時間を浪費する。インテリジェントな機械の効率性を実現するには、別のアプローチが必要だ。

ブリストル大学とマンチェスター大学は共同研究で2つの論文を発表し、センシングと学習を組み合わせたAIシステム用の新しいカメラの作製を提示した。

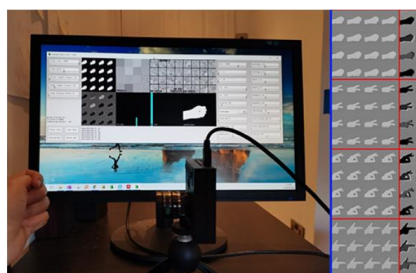


SCAMP-5d 視覚情報処理システム  
画像クレジット: *The University of Manchester, 2020.*



SCAMP-5d のハードウェアアーキテクチャ。  
256x256PPA 配列のピクセルプロセッサを内蔵。各ピクセルプロセッサには光センサー、ローカルメモリレジスタ、その他機能コンポーネントが含まれる。画像クレジット: *The University of Manchester, 2020.*

「効率的な知覚システムを作製するためには、私たちがこれまで行ってきた方法を超え、限界を押し上げる必要があります。」と、ブリストル大学のロボット工学・コンピュータービジョン・モバイルシステム専門の教授で、研究室主宰者(PI)のWalterio Mayol-Cuevas氏は言う。



SCAMP-5D ビジョンシステム上の畳み込みニューラルネットワーク(CNN)。手のジェスチャーを 8,200 フレーム/秒で分類。

「自然のシステムが視界を処理する方法から、着想を得ることができます。私たちはすべてを認識しているわけではありません。眼と脳が協働して物体を認識しており、場合によっては無関係な物体を低減する脳の働きを、眼がサポートしているのです。」

このことは、画像を検出した時点で直接ハエのような物体を確定する、カエルの眼の検出機能が実証している。

ブリストル大学のDr Laurie Bose 氏が一方を、Yanan Liu 氏がもう一方を率いた両論文は、この目標に向け、2つの改良点を明らかにした。視覚認識を可能にするAIアルゴリズムの一つである畳み込みニューラルネットワーク(CNN)を、直接像平面で実行したのだ。研究チームが開発したCNNは、これらの画像の記録や、処理パイプラインへの送信が不要で、毎秒何千回のフレーム分類が可能だ。研究者たちは、手書きの数字、手のジェスチャーやプランクトンまでも分類する実証も検討した。

本研究は、カメラの目の前の物体の種類や、起こっている事象等の高レベルな情報をシステム全体に伝える視覚システムを有する、インテリジェントなAIカメラがある未来を提案する。新アプローチは画像記録が不要なため、システムははるかに効率的で安定する。

本研究は、マンチェスター大学で回路・システム・PIが専門の教授、Piotr Dudek氏と、彼の研究チームが開発したSCAMPアーキテクチャにより実現した。SCAMPとは、ピクセルプロセッサアレイ(PPA)と称する、カメラプロセッサチップだ。PPAは、超並列処理を実行するために、相互に通信するプロセッサが各ピクセルに埋め込まれている。これは、CNNと視覚アルゴリズムには理想的だ。

「センシング、プロセッシングとメモリをピクセルレベルで統合することで、高性能で低レイテンシのシステムのみならず、低電力で高効率のハードウェアの可能性が見込めます。」と、Dudek教授は言う。

「SCAMPデバイスは、既存のカメラセンサーと同様の実装面積で実行でき、画像を捉えたその場で汎用的な超並列処理を実行する性能を備えています。」

ブリストル大学の飛行力学上級講師で、本プロジェクトメンバーのTom Richardson博士は、SCAMPアーキテクチャを軽量のドローンと統合している。

「これらのカメラの素晴らしいところは、新しい機械学習能力のみならず、作動速度と軽量の構成です。」と、同氏は言う。

「これらは、文字通りその場で学習して飛行でき、高速で高機敏性の航空プラットフォームには最適です!」

本研究には、英国工学・物理科学研究会議(EPSRC)が資金を提供しており、AIシステム的设计時には、存在する様々な疑問を投げかけることが重要であることを示した。また、カメラのような、しばしば当たり前とされているものには、より効率的なインテリジェント・マシンの実現目標に向けた改善が可能で、また改善すべきである。

## 論文

'Fully embedding fast convolutional networks on pixel processor arrays' by Laurie Bose, Jianing Chen, Stephen J. Carey, Piotr Dudek and Walterio Mayol-Cuevas presented at the European Conference on Computer Vision (ECCV) 2020

'High-speed Light-weight CNN Inference via strided convolutions on a pixel processor array' by Yanan Liu, Laurie Bose, Jianing Chen, Stephen J. Carey, Piotr Dudek, Walterio Mayol-Cuevas presented at the British Machine Vision Conference (BMVC) 2020

翻訳：NEDO（担当 技術戦略研究センター）

出典：本資料は、ブリストル大学の以下の記事を翻訳したものである。

“Cameras that can learn”

(<https://www.bristol.ac.uk/news/2020/october/scamp.html>)

(Reprinted with permission of the University of Bristol.)