

マシンビジョンに新たな次元をもたらす ディープラーニング

アンドリュー・ウィルソン

ディープラーニング、人工知能、コグニティブシステムは新しいものではないが、マシンビジョンシステムへの適用は今始まったばかりである。

マシンビジョンにおける革命として一部では推進されている、コンピュータプログラミングを必要とすることなくオブジェクトを解析および分類するシステムの機能を表すものとして現在、多数の用語が用いられている。人工知能(AI: Artificial Intelligence)とディープラーニングは、そのような概念を推進するために用いられる多数の専門用語のうちの2つである。

しかし、その誇張的な表現の根底で、そうした概念を支える基本的技術は、よりシンプルに説明することができる。例えば、従来のマシンビジョンシステムでは、部品上のバーコードを読み取り、寸法を判断し、欠陥を検査することが必要になる場合がある。その

ようなシステムを構築するために、通常は、データマトリックスコードの判読に適用できる標準ツールを備える市販のソフトウェアや、GUI(グラフィカルユーザーインターフェース)を用いて部品寸法を測定するキャリパー・ツールセットなどが利用される。

そうして測定した部品を、あらかじめ定められた条件に適合するか否かによって、正常か異常かに分類することができる。このような測定方法とは異なり、いわゆる「ディープラーニング」ツールは、画像分類器の範疇に区分するのがより適切である。バーコードデータの読み取り専用のソフトウェアとは異なり、それらは例えば、画像の中のオブジェクトが存在するか、正常か、

異常かを判別するように設計されている。したがってそうしたツールは、補完的な役割を果たすものである。

「ニューラルネットワークなどのディープラーニングツールは、他のマシンビジョン技術を補完するものである。例えば、そのようなニューラルネットワークによって、データマトリックスコードが画像内に存在する確率を判定することができる」と、加マトロックスイメージング社(Matrox Imaging)の研究者およびイノベーション担当ディレクターを務めるアルノー・リナ氏(Arnaud Lina)は述べている。「しかし、その判読にはやはり、従来のバーコードアルゴリズムが用いられる」(リナ氏)。

エレガントな分類器

ニューラルネットワークをベースとするツールは、部品の存在の有無や、画像内のオブジェクトが正常か異常かの判定に使われることが多い。こうしたツールは、k近傍法(k-Nearest Neighbor: k-NN)や決定木といった、画像分類器として知られるアルゴリズムの範疇に属する。さまざまな種類の分類器を示した図については、ジェイソン・ブラウンリー(Jason Brownlee)著「A Tour of Machine Learning Algorithms」(機械学習アルゴリズムの紹介)(2013年11月、<http://bit.ly/VSD-TMLA>)を参照してほしい。

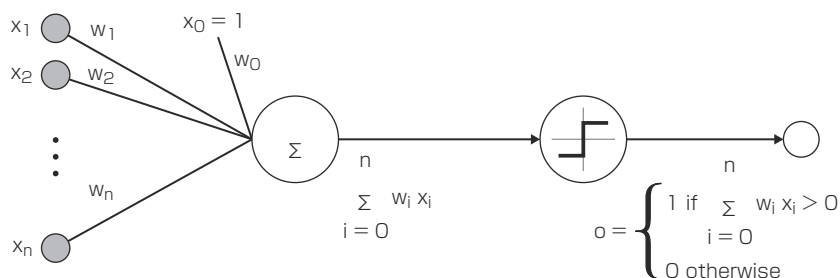


図1 フランク・ローゼンブラット氏によって1957年に発明されたパーセプトロンは、バイナリ入力集合を受け取り、各入力に連続値加重を乗じ、その加重入力の合計値をしきい値判定することによって、脳の神経細胞をモデル化する。(提供: アンドレイ・クレンコフ氏)

そうした分類器の多くが、マシンビジョンの用途に適用可能である。独エムブイテック・ソフトウェア社(MVTec Software)などはすでに、トレーニング済みのニューラルネットワーク、サポートベクターマシン(SVM: Support Vector Machine)、混合ガウスモデル(GMM: Gaussian Mixture Model)、k近傍法といった分類器を、「HALCON」ソフトウェアパッケージの中で提供している。ここで、ディープラーニングネットワークをゼロからトレーニングして、有効な認識結果を達成するには、1つのエラークラスにつき数十万枚ものサンプル画像が必要であることに注意しなければならない。しかし、HALCONには広範なディープラーニングネットワークが統合されているので、比較的少ないサンプル画像だけでトレーニングが可能で、時間と費用を大きく節約することができる。

独ステマ・イメージング社(Stemmer Imaging)なども、「Common Vision Blox (CVB) Manto」ソフトウェアにおいてSVMを採用している。

異なるグループのフィーチャー値が著しく重なる場合に、SVMを使用すれば、多次元のフィーチャー空間を生成して異なる欠陥グループを隔離することができる。蘭PRシス・デザイン社(PR Sys Design)の「perClass」ソフトウェアには、そうした分類器が多数提供されている。同ソフトウェアには、k近傍法、ニューラルネットワーク、ランダムフォレスト、SVMなど、さまざまな分類器が含まれている。この画像解析ソフトウェアは「MATLAB」ベースのツールボックスで、開発者はデータをインタラクティブに操作し、データの中から画像分類に最良のフィーチャーを選択し、多数の種類の分類器

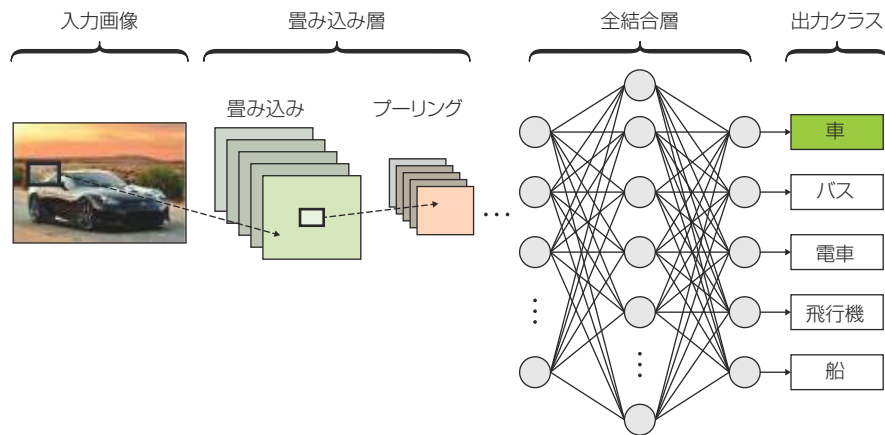


図2 畳み込み演算子がエッジなどのフィーチャーの検出に用いられるのと同様に、CNNでは、畳み込み層がフィーチャー抽出に用いられる。畳み込み画像処理では、ガウシアンぼかしやメディアンフィルタなどの画像フィルタによって、この処理が行われる。一方、CNNのアーキテクチャは、網膜出力によってエッジ検出などのフィーチャー抽出を行う、人間の視覚系を模倣する。(画像提供:米マサチューセッツ工科大)

をトレーニングし、その性能を最適化することができる(Vision Systems Design誌2015年2月号の記事「Machine learning leverages image classification techniques」(画像分類手法を活用する機械学習)、<http://bit.ly/VSD-MLL>を参照)。

脳の模倣

画像内のオブジェクトの識別に、多数の異なる分類器が使用できるのと同様に、画像分類機能の実行に採用できるニューラルネットワークの種類は、数え切れないほど多い。そうしたニューラルネットワークは、人間の視覚系と脳で用いられる生体神経網の模倣を試みるものである。その最も単純なものが、フランク・ローゼンブラット氏(Frank Rosenblatt)によって1957年に発明された「パーセプトロン」(Perceptron)である(図1)。

パーセプトロンは、バイナリ入力集合(隣接神経細胞)を受け取り、各入力に連続値加重(各隣接神経細胞に対するシナプス強度)を乗じ、その加重入力の合計値をしきい値と比較して、十分に大きければ“1”、それ以外の場合は“0”を出力することにより、脳の神経細胞をモデル化する。これは、生体神経細胞の発火するか否かの動作を模倣している(アンドレイ・クレンコフ氏(Andrey Kurenkov)著「A 'Brief' History of Neural Nets and Deep Learning」(ニューラルネットとディープラーニングの「短い[簡単な]歴史」)、<http://bit.ly/VSD-NNDL>を参照)。

現在、より洗練されたネットワークでは、多層の「ニューロン」(神経細胞)を使用して、画像分類処理が行われている。これらの「隠れた層」により、ニューラルネットワークに提示された画像データ内のフィーチャーが検出さ

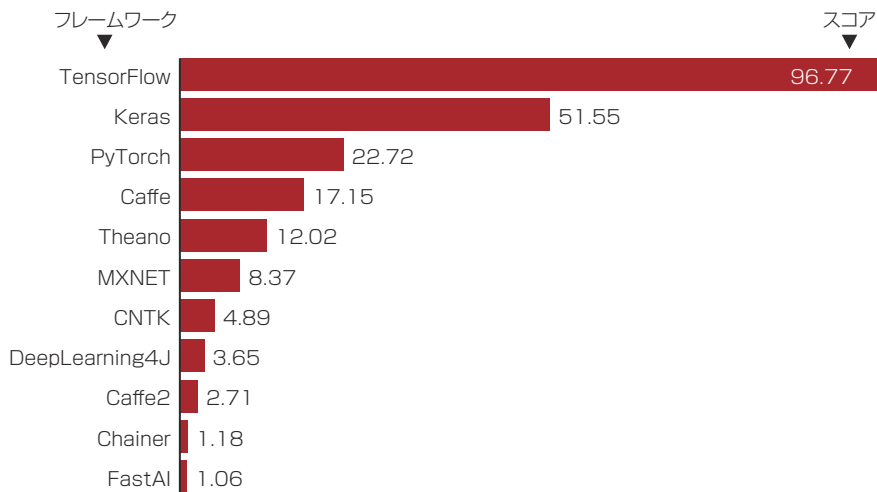


図3 Kerasは、11のデータソースを基に、グーグル社の機械学習ライブラリ「TensorFlow」が、最も広く採用されている機械学習ソフトウェアであるとしている。(提供:Keras)

れる。単純なパーセプトロンから、ディープフィードフォワード(Deep Feed Forward:DFE)や畳み込みニューラルネットワーク(CNN:Convolutional Neural Network)にいたるまでの、無数のニューラルネットワークアーキテクチャを、手書き解析や音声認識などの用途に利用することができる(アンドリュー・チルコフ氏(Andrew Tch)著「The mostly complete chart of Neural Networks, explained」(ニューラルネットワークのほぼ完全なチャートを解説)、<http://bit.ly/VSD-CCNN>を参照)。

マシンビジョンシステムに搭載されるニューラルネットワークとして最も広く採用されているのは、おそらくCNNである。それにはいくつかの理由がある。まず、CNNのアーキテクチャは、人間の視覚野とパターン認識のメカニズムをより忠実に模倣するように設計されている。これが可能なのは、CNNが、異なる機能を実行する複数のニューロン段で構成されるためであ

る(図2)。

簡略化されたこの図では、説明上、畳み込み層とプーリング層が分けて示されている。実際には、それらはCNN全体の一部である。畳み込み演算子がエッジなどのフィーチャーの検出に用いられるのと同様に、CNNでは、畳み込み層がフィーチャー抽出に用いられる。

畳み込み画像処理では、この処理を行うガウシアンぼかしやメディアンフィルタなどの画像フィルタを、FPGA(Field Programmable Gate Array)にオフロードすることができる。

一方、CNNのアーキテクチャは、網膜出力によってエッジ検出などのフィーチャー抽出を行う、人間の視覚系(HVS:Human Visual System)を模倣したものとなっている。CNNでは、この畳み込みがフィーチャー抽出を行い、入力画像のフィーチャーを表す役割を担う。これらの畳み込み層は、それぞれ受容野を持つ各ニューロンが、一連の加重を介して1つ前の層のニュー

ロンに接続される、フィーチャーマップを構成する。加重はトレーニング可能で、適用されるフィルタの種類を決定する。

フィーチャーを抽出した後は、計算速度を上げるために、いわゆる「プーリング層」によって画像データ表現の空間サイズが縮小される。続いてこの画像データは、さらなるデータ処理のためにネットワークの最終層に供給される(ワシム・ラワット氏(Waseem Rawat)とゾンホォイ・ワン氏(Zenghui Wang)共著「Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review」(画像分類のためのディープ畳み込みニューラルネットワーク:包括的概観)、Neural Computation, Volume 29, Issue 9, 2017年9月、<http://bit.ly/VSD-DCNN>を参照)。

プログラミングは不要

ニューラルネットワークでは、他の手法で必要となる従来型のプログラミングやFPGAのプログラミングが不要である。CNNに対するオブジェクト認識のトレーニングは、3つの異なる方法で行うことができる。1つめは教師あり学習で、画像に手動でラベル付けを行い、そのデータを使用することにより、未知の画像の分類を行う出力モデルを生成するものである。2つめは教師なし学習で、画像データのタグ付けが必要な教師あり学習に対し、こちらは、画像データ内のパターンを検出して、複数の未知の画像からモデルを構築することができる。続いてそれらの結果を解釈することにより、さらに正確な教師あり分類器を開発することができる。

このような教師なし学習システムは、往々にして不正確であるため、ラ

表1 多数のマシンビジョン企業によって現在、マシンビジョンの用途に特化したソフトウェアの開発にオープンソースのツールキットが利用されている製品が、提供されている。

機械学習ソフトウェア		
企業	製品	特長
アダプティブ・ビジョン ポーランド www.adaptive-vision.com	「Adaptive Vision Studio」用の ディープラーニングアドオン	ニューラルネットワークに基づく 教師ありと教師なしモードで使用可能
コグネックス 米国 www.cognex.com	「VisionPro ViDi」	ViDi red は、異常や外観上の欠陥の検出に使用される ViDi green は、オブジェクトやシーン全体の分類に使用される ViDi blue は、画像内の1つまたは複数のフィーチャの検出や特定に使用される
シス・システムズ 米国 www.cyth.com	「Neural Vision」	ニューラルネットワーク - 教師あり学習モード
ディービオ (Deevio) ドイツ www.deevio.ai	「Deevio」	ニューラルネットワーク - 教師あり学習モード
フリアーシステムズ カナダ www.flir.com/mv	「FireFly」カメラシリーズ	ディープニューラルネットワークとビジョンの用途向けに Neural Compute Engine を採用 FireFly カメラに搭載 「モビディウス」の項を参照のこと
IDS イメージングデベロップ メントシステムズ ドイツ https://en.ids-imaging.com	「NXT」カメラ 「Vision App Software」	Movidius ニューラルネットワークアクセラレータ搭載の推論カメラ 開発者は、トレーニング済みのニューラルネットワークをオブジェクト認識と分類用に カメラに読み込み可能 「モビディウス」の項を参照のこと
インスペクト (Inspekto) イスラエル https://inspekto.com	「Plug & Inspect」ソフトウェア	センサパラメータの自己設定、生産ライン変更に対する自己適応、検出技術の自己調整
ラオンピープル (Laon People) 韓国 www.laonpeople.com	「NAVI v.2」ディープラーニング ソフトウェア	教師あり学習と教師なし学習 欠陥有無の検査とオブジェクト分類に使用
マトロックスイメージング カナダ www.matrox.com/imaging	「MIL 10 Processing Pack 3」	教師あり学習に基づく、量み込みニューラルネットワーク (CNN) 技術
ムーンビジョン (The MoonVision) オーストリア www.moonvision.io	「The MoonVision」	教師あり学習
モビディウス (Movidius、インテル傘下) 米国 www.movidius.com	「Myriad X」プロセッサ	専用の Neural Compute Engine ビジョンおよびニューラルネットワークライブラリをソースコードとして提供 DIY型カメラキットを https://bit.ly/2fxvFar にて提供
パークラス (perClass BV) オランダ www.perclass.com	Matlab ベースの「perClass Toolbox」	認識アルゴリズムのトレーニング用
エムブイテック・ソフトウェア ドイツ www.mvtec.com	「HALCON」 「MERLIC」	教師あり学習に基づく、量み込みニューラルネットワーク (CNN) 技術
Sualab 韓国 www.sualab.com	「SuaKIT」	セグメント化、分類、検出 教師あり学習と教師なし学習
ザイリックス (Xilinx) 米国 www.xilinx.com	ACAP (Adaptive Compute Acceleration Platform)	画像認識処理の高速化に使用可能

ベル付きデータを取り入れることによってトレーニングが必要な画像枚数を減らす、半教師あり学習を利用するのが有効かもしれない。これが3つめの方法である。

「トレーニング方法の種類に関しては、当社のツールのほとんどは、教師あり学習に基づいている」と、ポーランドのアダプティブ・ビジョン社 (Adaptive Vision) でゼネラルマネー

ジャーを務めるミハウ・ザーディボン氏 (Michał Czardybon) は述べている。「手動でアルゴリズムを設計する必要がないので、これが最も効果的だ。ただし、ラベル付きのトレーニングデー

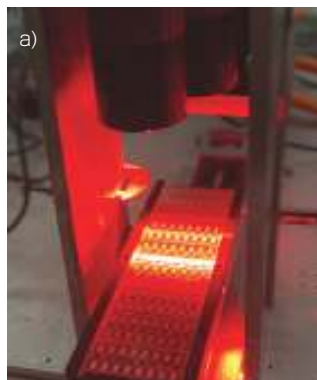


図4 シス・システムズ社のNeural Visionソフトウェアは、(a)半導体のリードフレーム検査、(b)ロボットのピック&プレイス・システム、(c)医療装置検査など、幅広い分野で利用されている。(写真提供：シス・システムズ社)

タを用意する必要がある。当社のツールの1つである『Anomaly Detection』では、半教師あり学習が採用されている。オブジェクトの複数の『正常』画像を開発者に供給してもらい、それらから逸脱しているかどうかだけを認識するように、ネットワークをトレーニングする」(ザーディボン氏)。

ザーディボン氏は、「このようなソフトウェアを教師なしモードで動作させたいと考える人々に対しては、当社の技術はまだ準備が整っていない。しかし、少数のラベル付きサンプルに、数百件のラベルなしサンプルを組み合わせることによって精度を高める、半教師あり学習について、当社はさらに研究を進めている」と述べた。

このようなネットワークを構築するために、市販とオープンソースの両方の開発ツールが利用できる。ニューラルネットワークベースのアプリケーション

の開発に利用できる市販ツールとしては、英ウルフラム社(Wolfram)の「Version 11」、米マスワークス社(Math Works)の「Deep Learning Toolbox」、スペインのアルテルニクス社(Artelnics)の「Neural Designer」などがある。その処理を行うオープンソースコードを提供する企業もある(「Deep Learning Software」(ディープラーニングソフトウェア)、<http://bit.ly/VSD-DLS>と、「Comparison of deep learning software」(ディープラーニングソフトウェアの比較)、<http://bit.ly/VSD-CDLS>を参照)。

オープンソースのソリューション

米インテル社(Intel)、米IBM社、米エヌビディア社(NVIDIA)、米マイクロソフト社(Microsoft)、米グーグル社(Google)といった大手企業のすべてが、機械学習アルゴリズムを開発するためのオープンソースコードを提

供している。

興味深いことに、Pythonで記述された高レベルのニューラルネットワークAPIを提供するKeras(<https://keras.io>)の開発元は、11のデータソースに基づく、ディープラーニングフレームワークの比較チャートを示している。それによると、グーグル社の機械学習ライブラリ「TensorFlow」が、最も広く採用されている機械学習ソフトウェアであるという(図3)。

Kerasの開発者であるフランソワ・ショレット氏(François Chollet)によると、Kerasは、高レベルのインターフェースを提供するが、低レベルの演算は処理しないという。それらの演算は、グーグル社のTensorFlowか、加モントリオール大(Université de Montréal)のMILA(Montréal Institute for Learning Algorithms)で開発されたオープンソースプロジェクト「Theano」(<http://deeplearning.net/software/theano>)のオープンソースコードによって実行される。

開発者はこのようなツールを利用して、ビジョンベースのアプリケーションを実装している。例えばインテル社は、蘭フィリップス社(Philips)と提携して、GPUを使わずにX線やCTスキャンに関するディープラーニング推論を行うシステムにおいて、「OpenVINO」ツールキットの利用を実証した。汎用CPUを使ってそのような解析を実行すると、数時間から数日を要する可能性がある。

既存のCPUアーキテクチャの性能を拡張するために、フィリップス社はOpenVINOツールキットとディープラーニング手法を採用した。GPUを使用する必要性を取り除くことにより、標準的なCPUでディープラーニングを行う場合と比べて、速度は200倍近く向

上している。しかしながら、このようなツールが利用できるのは、医用イメージングの分野だけではない。インテル社のシニア技術コンサルティングソフトウェアエンジニアであるジェフ・マクアリスタ氏(Jeff McAllister)は、こちらのビデオ(<http://bit.ly/VSD-OV>)の中で、Open VINO ツールキットをマシンビジョンの用途で使用方法を紹介している。

マシンビジョンに向けて

こうしたあらゆる開発ツールが提供されている状況を考えると、多数のマシンビジョン企業によって提供されている製品において、マシンビジョンの用途に特化したソフトウェアの開発にこれらのツールキットが利用されるとしても不思議ではない(表1)。どのツールを利用しているかや、自社製品においてそれ以外にも分類ツールを利用しているかについて語ることは控えつつも、これらの企業は、マシンビジョンにおけるニューラルネットワークの可能性を評価する手段を、開発者に提供していることになる。

そのようなソフトウェアパッケージがどのように利用されているかを示す、数多くの事例が現時点で存在する。例えば、中国ユーマイ・ダイカスティング社(YuMei Die Casting)は、部品の鋳造時に欠陥を検出するために、インテル社のコンピュータビジョンとディープラーニングソフトウェアに基づくシステムにおいて、インテル社の「Joint Edge Computing Platform」を利用している。

そうしたアプローチの能力を認識した米コグネックス社(Cognex)は2017年、スイスのViDiシステムズ社(ViDi Systems)とそのViDiソフトウェアスイートを買収した。同ソフトウェアは

最近、加AISテクノロジーズグループ社(AIS Technologies Group)によって、シリンダ内径検査システムに実装されている(Vision Systems Design誌2018年3月号の記事「Smart vision system ensures 100% cylinder bore inspection」(スマートビジョンシステムで、100%のシリンダ内径検査を保証)、<http://bit.ly/VSD-3DBI>を参照)。

また、独ディー・ファイン社(d-fine)は、独サイデナデル・マシネンバウ社(Seidenader Maschinenbau)向けに、医療製品の検査用システムにおけるCNNの性能をテストするためのシステムを開発した。このような手法の有効性を従来のマシンビジョン手法と比較するための、結果解析が現在行われている(「Medical Products Inspected with Deep Neural Networks」(ディープニューラルネットワークによる医療製品の検査)、<http://bit.ly/VSD-MPDL>を参照)。

米シス・システムズ社(Cyth Systems)は、このような技術を提唱して数多くの製品を開発しており、「Neural Vision」という独自のAIベースのソフトウェアを提供している。このソフトウェアは、半導体のリードフレーム検査(<http://bit.ly/VSD-NVFI>)、医療装置検査(<http://bit.ly/VSD-CNV>)、ロボットのピック&プレイス・システム(<http://bit.ly/VSD-CPAI>)など、幅広い分野で利用されている(図4)。

カメラへの搭載

このようなソフトウェアを採用するシステムの多くがPCベースだが、カ

メラベンダーも、この処理を実行できるスマートカメラを提供するという機会に着目し始めている。例えば、独IDSイメージングデベロップメントシステムズ社(IDS Imaging Development Systems)と加フリアーシステムズ社(FLIR Systems)である。

IDS社の「NXT」カメラは、特別に開発されたAIビジョンアプリケーションを利用して、あらかじめトレーニングされた人工ニューラルネットワークの読み込みで使用できるほか、FPGAベースのAIアクセラレーションによって推論回数の増加が図られている。同カメラはVISION 2018において、AIベースのオブジェクト認識システムとして出展された。同じ展示会で披露されたフリアー社の「FireFly」には、インテル社のビジョン処理ユニット(VPU: Vision Processing Unit)である「Movidius Myriad 2」が搭載されており、トレーニング済みのニューラルネットワークを、その上に直接読み込むことができる。

こうした発表に伴って、スマートカメラは、従来のマシンビジョンアルゴリズムとニューラルネットワークベースの分類器を実行するようになり、開発者には使用するアルゴリズムの柔軟性がもたらされる。これを推進するのは、グーグル社、インテル社、マイクロソフト社などの企業が提供するオープンソースコードになる見込みで、開発者はそれらを活用することにより、商用のイメージングソフトウェアをこれまでよりもはるかに迅速に市場に投入できるようになる。

.....
Vision Systems Design2019年4月号より転載。

著者紹介

アンドリュー・ウィルソン(1956~2019)は、Vision Systems Design誌の創刊者兼編集者。画像処理、マシンビジョン、コンピュータサイエンス分野の専門家であり、数多くの技術記事を書いた。

LFWJ