



平成28年9月6日

信州大学
東北大学
科学技術振興機構（JST）
内閣府政策統括官（科学技術・イノベーション担当）

被災環境下での搜索・状況確認活動を支援する画像認識システム ImPACT タフ・ロボティクス・チャレンジによるロボットインテリジェンス・極限画像処理

内閣府総合科学技術・イノベーション会議が主導する革新的研究開発推進プログラム（ImPACT）タフ・ロボティクス・チャレンジ（プログラム・マネージャー：田所 諭）の一環として、信州大学 山崎 公俊（やまざき きみとし） 助教，アーノード・ソービ 研究員，東北大学 田所 諭（たどころ さとし） 教授のグループは、被災地での搜索活動などに用いる遠隔操作ロボットで撮影した画像から、周囲に存在するものの種別や探索対象の発見を支援するための画像認識システムの開発に成功しました。本システムは、ヘビ型ロボット、サイバー救助犬、ヒューマノイドロボットなど、災害環境で作業をおこなう様々なタイプのロボットへ搭載可能です。

本成果の一部は、2016年9月7～9日に開催される第34回日本ロボット学会学術講演会にて発表いたします。

ポイント

- 不定形物が乱雑に存在する環境で撮影された映像から、どこに何の物体が写っているか、表面状態はどうか、を自動で識別。
- 搜索対象（要救助者など）が存在しそうな領域の絞り込みに、効果を発揮。
- 事前学習のために必要とするデータ量が少なく、初めての現場でもすぐに効果を発揮。
- 操作者との対話を通して識別能力を向上・拡張していくことが可能。

<研究の背景と経緯>

熊本地震や阪神淡路大震災に代表される大規模地震災害では、倒壊した建物内に取り残された人の発見と救助が大きな課題です。信州大学 山崎助教，アーノード研究員，東北大学 田所教授を中心とするグループは、ImPACT タフ・ロボティクス・チャレンジの共同研究開発の一つのテーマとして、被災地での搜索活動を支援するための画像認識手法を研究しています。その一環として、災害現場ですぐに利用できる搜索活動補助のための画像認識システムの開発をすすめてきました。被災木造家屋を模擬した評価試験フィールドや、森林環境、熊本地震での倒壊家屋で撮影した画像データなどに対して検証をおこない、一定の性能が確認されました。

ImPACT タフ・ロボティクス・チャレンジで研究開発を進めるレスキューロボットは、1995年の阪神淡路大震災を契機に田所らのグループが提唱し、世界を牽引してきた研究分野です。田所教授は、災害ロボットの研究が世界的に高く評価され、IEEE Fellow（米国電子技術協会 IEEE の最高位メンバー資格）に昇格しています。また、山崎助教のグループは布などの不定形物の認識・操作に関する研究で成果を上げてきており、平成28年度文部科学大臣若手科学者賞などを受賞しています。本成果は、そこで培った技術を活用・発展させたものです。

<研究成果の概要>

本システムは、倒壊家屋などの被災現場を撮影した映像を解析し、その結果を操作者に提示することで、搜索活動を支援します。被災地での搜索活動などにおいては、瓦礫の中から人や物体を発見するために、ロボットやビデオスコープで撮影した映像の解析が重要になります。近年、人工知能技術の発達により画像認識分野では相当な進歩が見られますが、その一方で、映像から意味のある情報を得るための機能（画像認識器）を構築する作業に手間と時間がかかり、被災現場のような乱雑な環境への素早い適用が困難でした。これは災害対応において特に大きな障害になります。

私たちはこの課題に対し、映像から認識能力を得る処理（学習処理）を効率化することで、少ない事前情報から画像認識器を作り出す技術を開発しました。この技術により、初めての被災現場でもすぐに効果を発揮でき、映像解析や重要情報の発見といった作業への利用が期待できます。さらに本開発では、瓦礫を構成している材質を推定する技術や、それらの表面状態（濡れているのか、土で汚れているのかなど）を数値化する技術も開発しました。これらの技術により、崩れやすい場所や滑りやすい場所を発見しやすくなり、ロボットによる搜索活動における災害現場での崩落などの二次被害を防ぐことに役立つと考えられます。

利用方法

- **環境認識支援**：ロボットに搭載したカメラから送られてくる映像は、それら特有の視点や狭い視野で得られることが多く、現場を正しく把握することは容易ではありません。本システムでは、映りこんでいる物体の種類などを同定し強調表示することで、操作者の環境認識の支援が可能です。また、すでに現場に詳しい人から環境認識に関する指示を受けて、環境認識能力を短時間で獲得することも可能です。
- **搜索支援**：特定の対象の搜索をおこなう場合、（救助犬に匂いを嗅がせるように）その対象が映っている画像、または類似のものをシステムに事前に学習させることで、対象がカメラ映像に映ったときにアラートを鳴らしたり、記録を取ったりすることが可能です。それにより、遠隔操作中に対象の見逃しを防止する目的や、録画映像を通して重要情報を発見する目的への利用が可能です。また、崩れやすい場所・滑りやすい場所を判断する作業、遠隔操作者が継続的には映像を閲覧できないような搜索状況での利用も有効と考えられます。

本成果は、以下の3つの研究開発成果から構成されます。

(1) 現場に素早く対応できる認識システム

図1に、開発した画像認識システムの利用例を示します。ヘビ型ロボットの先端に取り付けたカメラから撮影をおこない、その映像に対して本技術により処理をおこなったものです。

左上に表示される入力映像に対し、ひとまとまりとみなせる領域を検知し、右上画像のように検知領域を色付けて強調します。中段に示したのはユーザーインターフェースです。右上画像の検知領域が意味のあるものと判断できる場合、操作者はその領域にカテゴリ^{注1)}名を付けます。この作業により、本システムは対象領域を学習し、同じカテゴリに所属する可能性が高いと判断される画像領域を同じ色で強調するようになります。操作者は必要に応じて領域を修正することもできます。このようなカテゴリ名の付与および誤認識の修正作業を通して、現場の映像から迅速に認識能力を獲得させることができます。認識能力が十分に得られた後は、認識のみのモードでシステムを動作させることで、搜索活動・状況確認活動を続けることが可能です。

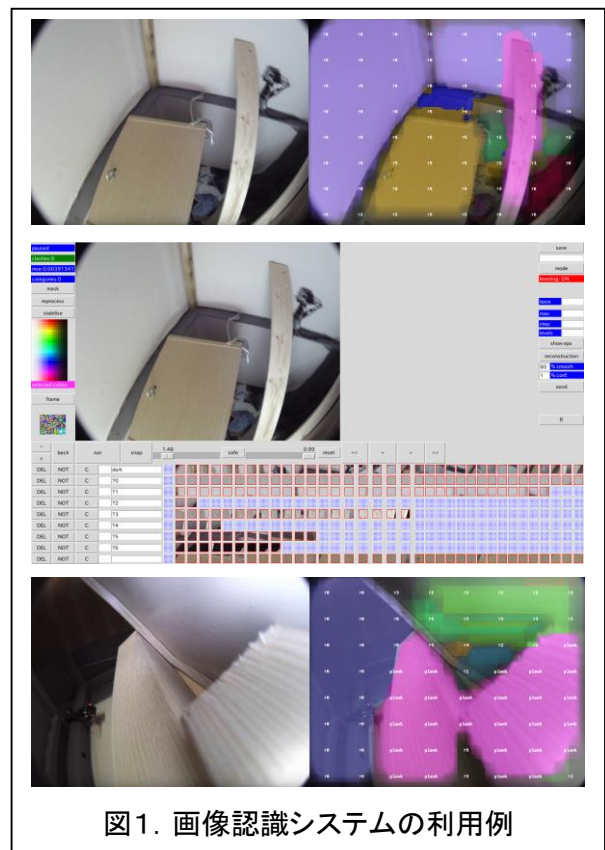


図1. 画像認識システムの利用例

(2) 乱雑なテクスチャ状況下での環境認識

図2は、森林環境で活動訓練をしているサイバー救助犬にカメラを搭載し、そこから得られた映像に対する認識処理の結果を示しています。木や草が乱雑に生えていたり、枯れ葉が堆積している森林環境の環境認識をおこなうには、乱雑なテクスチャ（模様）から有意なパターンを見つけ、種別をおこなうことが有効です。

それを実現するため、本グループでは、以下に示す2段階の認識処理を開発しました。図2上段は、1段階目の処理結果を示しています。入力画像を等間隔に区切り、その区切り（グリッドと呼びます）ごとに種別（犬、地面、木々、木の幹など）をおこなうことで、画像をおおまかに分割します。その後、2段階目をおこないます。各グリッド内部の画素ひとつひとつに対して、種別（木の幹、葉、草、空など）をおこないます。図2下段に二つの処理結果を示します。左側にある原画像の風景が、右側で色分けしてあるように細かく分割され、森林のような乱雑な環境下での環境認識が可能になりました。

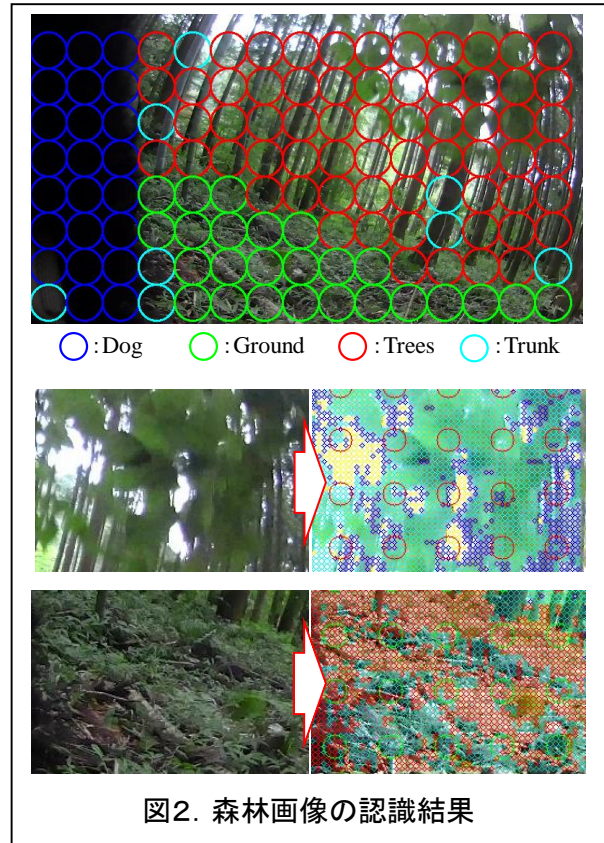


図2. 森林画像の認識結果

(3) 瓦礫構成物の表面状態の記述と種別

被災した建物内部の映像の解析をおこなう場合、材質などの種別に加え、汚れた場所・濡れた場所の検出や、汚れ・濡れの程度の推定を適切におこなうことで、現場に入って作業を行う人々の安全性向上につながります。しかし、建材や内装に利用する材質は、特徴的な模様が少なく、撮影画像からこれらの情報を得ることは容易ではありませんでした。

本研究開発では、この課題の解決に向けて画像から有意なテクスチャ情報を得るための方法を開発しました。図3上段に示すように、水や土をつけた建材等の画像から、材質等や表面状態の違いを適切に表現できる手法を開発しました。図3下段は、入力画像をグリッドに分け、テクスチャの違いによって各グリッドを色分けしたものです。表面状態に係らずブロック、木板、背景紙がテクスチャ情報に基づいて分類することが可能になりました。

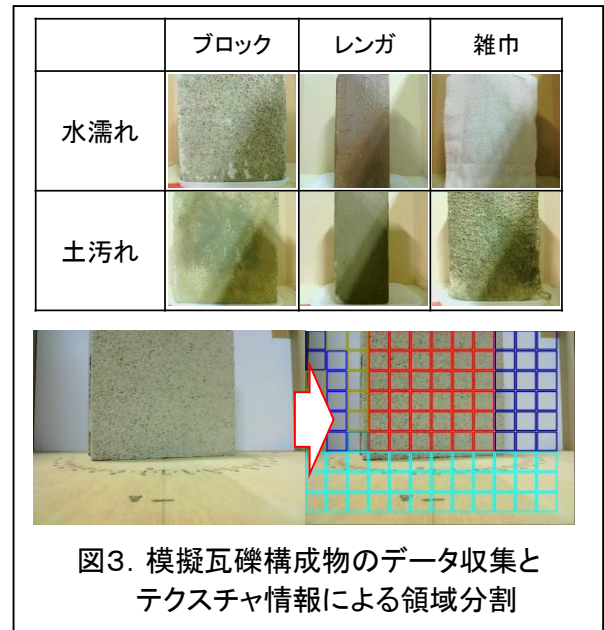


図3. 模擬瓦礫構成物のデータ収集とテクスチャ情報による領域分割

<技術について>

上で紹介した3つの研究成果について技術的な説明をおこないます。

(1) 現場に素早く対応するための認識システム

技術的課題は以下の通りです。

- カテゴリ情報が与えられている訓練データから認識器を構築する手法（教師あり学習）は、高い認識性能が見込める一方で、大量の訓練データを必要とします。そのため、どのようなカテゴリが存在するのかを事前を知ることが難しい被災現場では、十分な訓練データを集める余裕がなく、搜索活動に迅速に対応することが困難です。一部の訓練データのカテゴリが未知でも利用できる手法（半教師あり学習）では、訓練データのハードルは緩和されますが、データの分割処理に時間がかかるため、映像の実時間処理が困難でした。

- 従来の教師あり学習・半教師あり学習では、認識器を構築している最中に新たなカテゴリを追加したり、複数のカテゴリを一つに融合、あるいは一つのカテゴリを複数に分割することが困難でした。

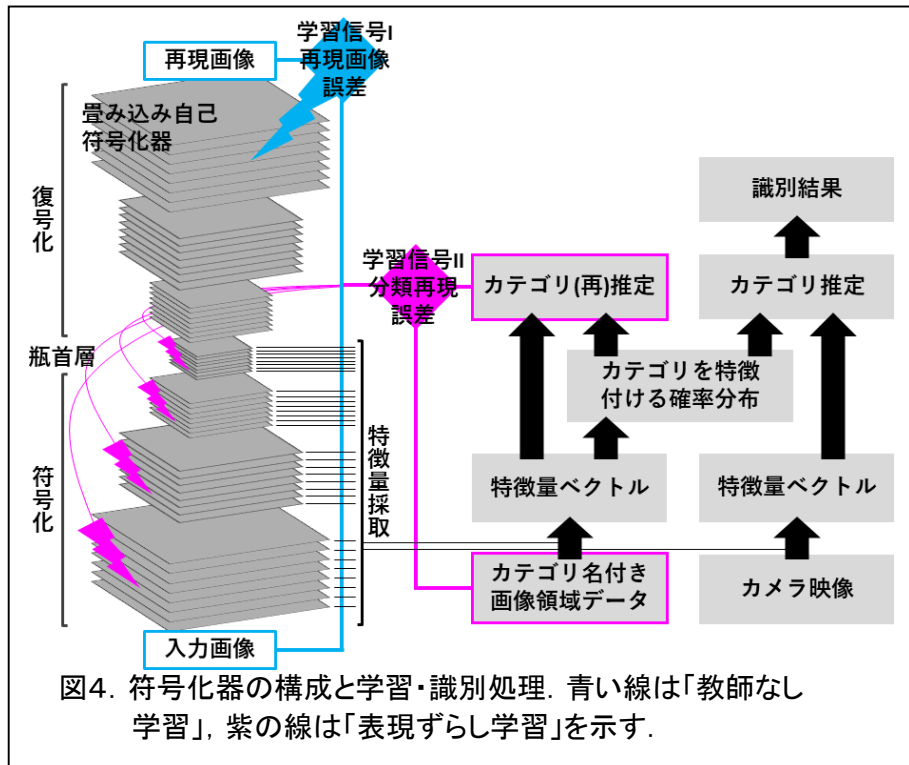
上記の課題に対し、本研究開発では、検索すべき環境の訓練データが手に入れにくいという前提の下で、操作者からの少ない指示に基づき、指示されたカテゴリを見分ける能力をできるだけ早く獲得する画像認識システムの開発をおこないました。

開発したシステムの核となるのは、畳み込み自己符号化器^{注2)}（以下「符号化器」と呼ばれるニューラルネットワークです。符号化器の役割は、入力画像を分割して得た小さな画像領域から、カテゴリ間の区別が明確になるような「特徴ベクトル」を生成することです。画像領域の特徴ベクトルが、操作者からカテゴリ名が付与された画像領域の特徴ベクトルと類似する場合は、その画像領域には同じカテゴリ名を付与します。このような仕組みのもとで高い認識性能を得るための鍵となるのは、画像領域から生成される特徴ベクトルの「類似性」をうまく操作することで、システムの識別結果を操作者の感覚に沿わせる部分です。

本成果である符号化器では、以下の二つの学習プロセスを並行して実行することで、上記の実現を可能にしました。

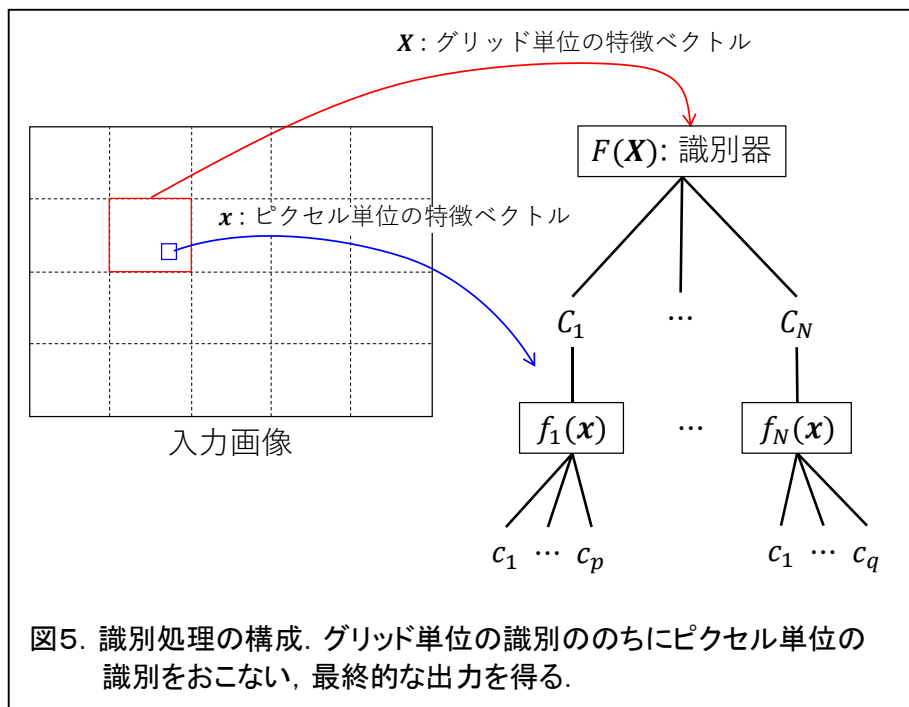
- A) カテゴリ情報が与えられていない訓練データから有意な法則を抽出する手法（教師なし学習）を用いて、入力画像の圧縮・展開を学習させます。この学習処理により、入力映像をうまく特徴付ける特徴表現が獲得されます。符号化器は多層構造を持ち、画像データが層を通るにつれて、より一般化された特徴表現になります。その全ての層を通して得られた特徴ベクトルを利用することにより、細かいテクスチャに特徴付けられるカテゴリも、大まかな形に特徴付けられるカテゴリも、認識できるようになります。以上に述べた教師なし学習処理は随時おこなうものですが、この処理にはどの画像領域がどのカテゴリに所属するかの事前情報は不要です。また、あらかじめ用意されたさまざまな映像（例えば同じカメラで撮られた映像や多様な被災現場の映像など）から、事前に高性能の符号化器を得ておくことも可能です。
- B) 今回新たに開発した「表現ずらし」という学習法を用います。「表現ずらし」では、カテゴリ名が付与された画像領域から特徴ベクトルを得たのち、カテゴリの「類似性」に基づいてその特徴ベクトルを少しだけ操作します。まず、各カテゴリに所属する全画像領域の特徴ベクトルから、各カテゴリを特徴付ける多変量確率分布を計算します。その後、すでに付与されているカテゴリ名とは別に、特徴ベクトルとカテゴリの確率分布だけを利用して、もとの画像領域にカテゴリ名を自動で付与します。ここで、操作者から付与されたカテゴリ名とシステムが自動で付与したカテゴリ名が合致しない画像領域に関しては、符号化器がそのカテゴリを区別する特徴を掴んでいないと判断します。その場合、誤ったカテゴリ名が付与された画像領域の特徴ベクトルと、実際そのカテゴリに所属する画像領域の特徴ベクトルの距離が増えるように、符号化器の更新処理を行います。この新技術により、操作者から指示されたカテゴリを見分ける識別性能を獲得していきます。

図4は符号化器の構成と学習プロセスを示しています。入力映像に対し、学習処理B)で述べた確率分布を用いて各領域の所属カテゴリを推定・表示します。学習処理B)の効果により、所属カテゴリの推定は単純になり、実時間処理が可能です。更に、二つ学習処理を一つの効率の良い更新処理にまとめることにより、学習しながらでも数fpsの実時間処理を可能としました（GPUを使用した場合）。



(2) 乱雑なテクスチャ状況下での環境認識

技術的課題は基本的に(1)と同じですが、本技術は特に、地面や樹木などで見られる乱雑なテクスチャへの対応を重視しています。処理の基本的な構成は図5に示すとおり、2種類の識別器から構成されます。ひとつは、グリッド単位で切り出した画像領域から特徴ベクトルを求め、それを入力として識別をおこなうもの(図中では $F(X)$ と表現)です。もう一つは、ピクセル単位で求めた特徴ベクトルを入力とするもの(図中では $f(x)$ と表現)です。グリッド単位の識別器は一つのみ用意されるのに対し、ピクセル単位の識別器は、グリッド単位識別でのカテゴリの数だけ用意されます。 $f_i(x)$ の出力は、グリッドの識別結果に応じて定められます。例えばグリッド単位の識別で地面と判断されると、ピクセル単位では草、落ち葉などといった限られたカテゴリのみが識別の対象となります。



本方式で重要なのは、グリッド単位、ピクセル単位でどのような特徴量を用いるかです。森林画像では再現性のある形状情報を得ることが難しいため、形状依存の少ない特徴量が必要となります。本技術では、テクスチャ情報と色情報の組み合わせを識別に利用することとし、テクスチャに関しては、さまざまなエッジ検出器の集合（フィルタバンク）を利用します。そして、フィルタバンクを利用して得られた反応値の列から頻度ヒストグラムを算出することで、形状不変性を有する特徴ベクトルを生成します。

フィルタバンクを利用する場合の課題のひとつは処理時間がかかることですが、この課題を解決する手法も開発しました。処理時間がかかる理由は、多数のフィルタカーネルを用意し、それぞれで畳み込み積分計算をおこなう必要があるためです。一方で、本研究の目的である被災現場での適用を考えると、小型軽量で持ち運びが容易な計算機を用いて、遠隔移動体の使用者がストレスなく見られるフレームレートで処理できることが求められます。そこで、本技術では、畳み込み積分の部分をボックスフィルタを用いて近似しました。これにより、GPU等のハードウェアマルチスレッド環境を利用しなくても、グリッド単位の識別を6~7fpsでおこなうことが可能になりました。

本識別処理におけるそのほかの特徴として、条件付き確率場（CRF）を導入しています。一つ一つのグリッドを個別に識別するのではなく、周囲のグリッドとの関連性を考慮することで識別率を向上させています。

（3）瓦礫構成物の表面状態の記述と種別

建材や内装および室内に配置されることが多い材質は、森林の場合と同様に決まった形状を持っていないため、表面性状や木目などによる見た目の違いから識別などをおこなう必要があります。本研究開発では、細かなテクスチャの違いを記述するための識別手法（テクスチャ解析手法）に着目し、フィルタバンク2種と色情報の組み合わせにより、瓦礫を構成する材質の識別を可能にしました。評価に利用した材質は、コンクリートブロック、MDF材、木板など10種類程度です。ただし、半透明の物体（ビニール袋）、光沢の強い物体（アルミ板）などの識別は未だ難しく、今後の課題と言えます。

また、本グループでは、材質表面の水や泥による汚れ度合いを数値化する技術も開発しました。材質表面が汚れることにより見た目が変化しても、各材質が持つもとのテクスチャが激しく欠損しないかぎり、材質を推定することが可能になります。

<今後の展開>

開発した技術は、事前に訓練データを用意しづらいタスクへの対応に適しています。また、倒壊家屋内部などのような雑多な環境でも機能するよう工夫をこらしているため、従来手法では十分な認識性能が得にくい対象にも利用できる可能性があります。災害対応だけでなく、インフラ点検、農業、林業、水産業などへの展開も考えていきます。

本成果は、以下の事業・研究プロジェクトによって得られました。

- 内閣府 革新的研究開発推進プログラム（ImPACT）
<http://www.jst.go.jp/impact/>
プログラム・マネージャー：田所 諭
研究開発プログラム：タフ・ロボティクス・チャレンジ
研究開発課題：災害環境探索のための補助視覚の研究開発
研究開発責任者：山崎 公俊
研究期間：平成27年度～平成28年度

この研究開発課題では、被災環境で使用するさまざまなタイプのロボットに搭載するための極限画像処理技術の開発に取り組んでいます。

■田所 諭 I m P A C Tプログラム・マネージャーのコメント ■



I m P A C Tタフ・ロボティクス・チャレンジは、災害の予防・緊急対応・復旧、人命救助、人道貢献のためのロボットに必要不可欠な、「タフで、へこたれない」さまざまな技術を創りだし、防災における社会的イノベーションとともに、新事業創出による産業的イノベーションを興すことを目的とし、プロジェクト研究開発を推進しています。

瓦礫の中から人や物体を発見するためには、ロボットやビデオスコープが撮影した映像の解析が重要です。本研究開発は、人工知能技術の適用と、映像からの学習を効率化する技術の開発により、瓦礫のような複雑なシーンからどこに何が写っているか、また、材質や表面状態などを認識できる技術を創り出したものであり、まさにタフな画像識別を可能にした成果です。今後、さらなる性能向上と実用化を進めることによって、地震災害の瓦礫等における人命救助の飛躍的向上を図るとともに、設備診断などの産業分野に対しても、広く波及効果が期待されます。

<用語解説>

注1) カテゴリ：倒壊家屋で撮影した画像には、壁、机、本、衣類などのさまざまな物体が乱雑に存在していると考えられます。本資料における「カテゴリ」とは、上述したような物体の種類のことを意味します。

注2) 畳み込み自己符号化器：ニューラルネットワークによる次元圧縮のアルゴリズムの一種です。入力層と出力層へ与えるデータは同じものとし、それらの中間には、入力・出力データよりも少ない次元の層を設定します。そして、出力データが入力データと同じになるように、畳み込みネットワークの重みを調整します。図2では複数の中間層を設けており、もっとも次元の少ない層を瓶首層（ボトルネックレイヤ）と呼んでいます。

<<お問い合わせ先>>

<画像認識システムに関すること>

山崎 公俊（やまざき きみとし）

信州大学 学術研究院工学系 助教

〒380-8553 長野県長野市若里4-17-1 機械システム工学科南棟4階

Tel: 026-269-5155

E-mail: kyamazaki@shinshu-u.ac.jp

Solvi Arnold（ソービ・アーノード）

信州大学 工学部機械システム工学科 研究員

〒380-8553 長野県長野市若里4-17-1 機械システム工学科南棟4階

Tel: 026-269-5159

E-mail: s_arnold@shinshu-u.ac.jp

田所 諭（たどころ さとし）

東北大学大学院情報科学研究科 教授

〒980-8579 宮城県仙台市青葉区荒巻字青葉6-6-01

TEL: 022-795-7025

E-mail: staff@rm.is.tohoku.ac.jp

< I m P A C Tの事業に関すること>

内閣府 革新的研究開発推進プログラム担当室

〒100-8914 東京都千代田区永田町1-6-1

Tel: 03-6257-1339

E-mail: g.sentan.pro@cao.go.jp

<IMPACTプログラム内容およびPMIに関すること>

科学技術振興機構 革新的研究開発推進室

〒102-0076 東京都千代田区五番町7 K's 五番町

Tel: 03-6380-9012 Fax: 03-6380-8263

E-mail: impact@jst.go.jp

<報道担当>

信州大学 総務部総務課広報室

〒390-8621 長野県松本市旭 3-1-1

Tel: 0263-37-3056 Fax: 0263-37-2182

E-mail: kouhoushitsu@gm.shinshu-u.ac.jp

科学技術振興機構 広報課

〒102-8666 東京都千代田区四番町5番地3

Tel: 03-5214-8404 Fax: 03-5214-8432

E-mail: jstkoho@jst.go.jp