

画像や動画を解析・生成する視覚情報処理の研究

高橋(裕) 研究室



高橋 裕樹
Hiroki TAKAHASHI

「百聞は一見に如かず」ということわざがあるように、人間の知覚の大部分を占める「視覚情報」には、情報を直感的かつ的確に伝えられるという特徴があります。コンピュータビジョン(CV)とは、コンピュータにこうした視覚知能を与える技術であり、これとは対照的に、コンピュータを使って情報を可視化する技術がコンピュータグラフィックス(CG)です。これらを包含する研究は視覚情報処理と呼ばれています。

機械学習による画像の領域分割

高橋裕樹准教授は、機械学習や深層学習、ニューラルネットワークといったさまざまな人工知能(AI)技術を使って、画像や動画の処理や解析、生成を行う視覚情報処理の研究に取り組んでいます。例えば、現実の世界にCG画像を合成し、実際に存在しないものを見えるようにする拡張現実感(AR)などの画像表示技術はこの分野の研究の一例です。反対に、現実の世界に存在する物体を隠蔽・消去する隠消現実感(DR)という技術もあります。

歩行を識別し個人を特定する

機械学習を用いた別の研究として、人間の歩行パターンである「歩容」を識別する技術も開発されています。防犯カメラの映像のような

一つは写真などから被写体を抜き出すような場合に使える、機械学習を使った画像の領域分割の研究です。通常、撮影時は被写体にフォーカスするため、背景はボケたような画像になります。高橋准教授は、被写体までの距離によってボケの度合いが異なることを利用して被写体を抽出する手法を考案しました。現状でも対象物体の7割程度が正確に抽出できています。また、ボケの状態と奥行きに関する情報があることから、今後精度がさらに上がれば、奥行き方向の情報なども取得できるかもしれません。

IP (Image Processing) / 機械学習

Focal Blur 領域分割
被写体までの距離による局所的なボケの違いを利用した被写体抽出
提案手法 [Shi14] [Chak10] [Zhuo11]
Ground Truth
1行: Blur map
2行: 初期分割
3行: 最終結果

視線インタフェース

歩容識別

異常行動検出

ウエハ不良パターン識別

機械学習を使った研究の概要

不特定多数の人が映る動画から、歩き方だけで個人をある程度特定できるそうです。歩容は顔や指紋などと同様に、個人ごとに異なるパターンを持つため、生体認証の一つとされています。「顔や指紋とは違って、多少遠くから撮影した映像でも認証できるため、利用のハードルは低いだろう」と高橋准教授は言います。服装の違いや、バッグなど携帯品の有無でも変わらぬ高い認識率を誇ります。似たような手法で、人が倒れ込んでいたり、人間同士がけんかしたりしているといった、通常とは異なる人の「異常行動」を検出する研究も進めています。そのほか、半導体ウエハの不良パターンを識別して検出し、半導

キーワード

画像処理、コンピュータビジョン(CV)、コンピュータグラフィックス(CG)、拡張現実感(AR)、機械学習

所属	大学院情報理工学研究所 情報学専攻
メンバー	高橋 裕樹 准教授
所属学会	電子情報通信学会、映像情報メディア学会、画像電子学会、芸術科学会
E-mail	rocky@inf.uec.ac.jp

歩容識別 [百瀬14]

- CASIA Gait Database 105名 6試行
- 60×90画素
- 特徴量: 4,644次元
 - 3-交差検定



図1 Gait Energy Image [Han06]

表1 背面歩容識別結果

	識別正答数(個)	識別率(%)
GEI [Han06]	605/630	96.0
接合背面歩容特徴	583/630	92.5
身体セグメント + 分割背面歩容特徴	615/630	97.6
分割背面歩容特徴	618/630	98.1

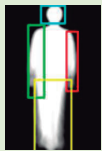
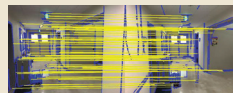
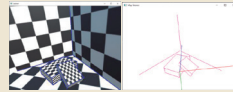


図2 局所GEI勾配

歩容識別の研究

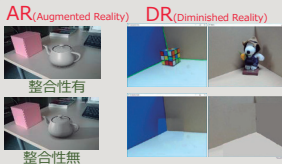
情報可視化のためのCV

線分を用いた3次元再構成



屋内ナビゲーション

区画線の可視化



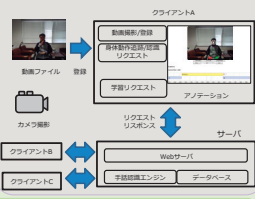
整合性有

整合性無

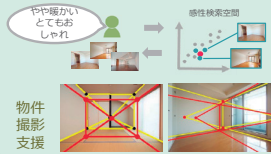
コンピュータビジョンの研究

視覚情報を用いた支援

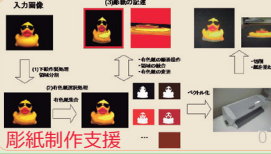
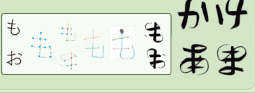
手話コミュニケーション支援



不動産支援



アンビグラム生成



視覚情報を使った多様な支援技術

体の生産における歩留まり向上につなげたり、視線を認識することによって、視線だけでタブレット端末などを簡単に操作することを目指した「視線インタフェース」を開発したりもしています。

動画から3次元構造を取得

一方、CVの研究では、壁面などに模様が少ない屋内の動画を撮影し、そこから線分を抜き出して床面や壁面などの情報を組み立てて空間を再構成することを試みています。再構成した画像に矢印などを図示すれば、例えば駅の構内や大学の建物内など、全地球測位

システム(GPS)の電波が届かない屋内のナビゲーションなどに利用できそうです。また、夜間の運転などで見にくくなる道路の区画線を可視化し、運転を補助する研究などもしています。

さらにAR技術では、現実の世界を撮影した画像にティポットなどのCG画像を合成するため、光源の方向を推定してこのCG画像に適切な影をつけることができました。高橋准教授によると、「物体の色を再現することに加えて、影をつけることでより現実感が増す」そうです。部屋の隅に置いた人形を画面上で消去する

DR技術もかなり精度が向上しています。

応用面では、視覚情報を使った多様な支援技術の研究に乗り出しています。例えば、手話によるコミュニケーション支援です。手話の様子を撮影してその手の動きを機械学習で解析し、単語に翻訳してテキストや音声として出力すれば、聴覚障害者と自然なコミュニケーションができるでしょう。

不動産検索や芸術支援も

不動産物件を探すための支援も行っています。これは駅からの距離や家賃などの数値的な情報だけ

でなく、「モダン」「明るい」「落ち着いた」といったような感性語のキーワードで物件を検索できるようにするという新しい試みです。物件の写真を魅力的に撮影するためのサポート手法なども提案しています。

このほか、芸術分野では、一つの文字を逆さにすると違う字として読める「アンビグラム」の制作を支援したりしています。アンビグラムでは、文字の構造を解析し、異なる文字との間に共通する部分を見つけて対応づけを行います。ひらがなの全文字を使うと

1081個(文字)のアンビグラムが作れる計算になりますが、そのなかで高橋准教授は、人間が作成可能な組み合わせの約45%であり、人間では作成が困難と思われる組み合わせの26%に相当する、418文字の印象的なアンビグラムを自動生成しています。膨紙アートでは、膨大な作業が必要な設計図の作製を補助することで、その後は制作者に作品作りに没頭してもらうことができるでしょう。このように、高橋准教授は視覚情報を使って多様な分野に応用の可能性を見いだしています。

【取材・文】藤木信穂