

知っておきたいキーワード

サリエンシーディテクション

中澤 篤志[†]

[†]大阪大学 サイバーメディアセンター

"Saliency Detection" by Atsushi Nakazawa (Cybermedia Center, Osaka University, Osaka)

キーワード：画像処理，サリエンシー，サリエンシーマップ，注視点

まえがき

静止画像やビデオから、人が興味を引く領域を抽出する手法が注目されています。この技術はサリエンシーディテクション (Saliency Detection) と呼ばれ、ビデオの解析やビデオ圧縮などさまざまな分野に応用できるため、近年盛んに研究が行われています。

サリエンシーディテクション技術自

体は、画像処理の分野では比較的古くから行われたものですが、近年、このように盛り上がりを見せている背景として、HDDビデオレコーダやYouTubeなどのインターネット動画が身近になってきていることが大きいと思われる。

このような動画の中から人の興味を引く領域の検出ができれば、ビデオからユーザーが見たい部分だけ取出す、ビ

デオコンテンツの認識を行う、自動要約の作成を行うといったビデオの自動編集技術に応用が可能です。また、「人が興味を引く領域」がわかれば、CM等の映像作成において、視聴者がどこに注目するのかを、映像のみから判断することができます。

このようにサリエンシーディテクションは、さまざまな分野に応用が可能です。

サリエンシーマップ

サリエンシーマップ (Saliency Map) は、入力画像に対応するサリエンシーの強さを表したマップのことで、画像処理により求める手法が数多く提案されています。サリエンシーマップは、画像領域に対する「人の興味」の度合いを表したもので、エッジや色などの画像特徴を組合せることで求められます。図1にサリエンシーマップの例を示します。

KochとUllmanは、人間の脳内の神経回路からのアナロジーから、画像の局所的な特徴を組合せ、

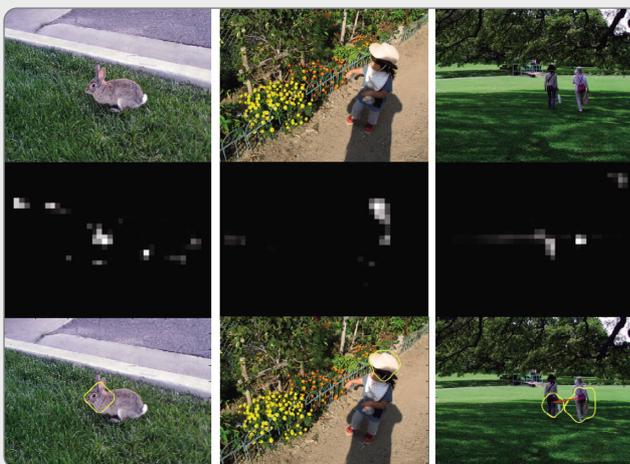


図1 サリエンシーマップの例

(上段：入力画像，中段：サリエンシーマップ，下段：サリエンシーマップから得られた注目領域(黄色枠))

画像内のある点がどの程度周辺と異なっているかを評価し、一つのサリエンシーマップを作るというアイデアを提案しました¹⁾。具体的な実装では、画像の色、輝度、エッジ方向と動きを統合して求められています(図2)。これらのモデルを改良した手法もさまざま提案されています^{2) 3)}。このツールキットは作者のホームページよりダウンロードすることができます⁴⁾。

また、近年では、奥行き情報を用いGPUでサリエンシーマップの高速計算を行う手法も提案されています⁵⁾。

WaltherとKochの手法を実装し配布されているSaliencyToolboxを用いて、サリエンシーマップを計算した例および実行画面を図2、図3に示します^{6) 7)}。入力画像(図1上段)に対してサリエンシーマップ(図1中段)が計算

されています。結果として人の主観に近い点が注目領域となっていることがわかります(図1下段)。また、実行画面(図3)をみるとこのサリエンシーマップは、色(ColorCM)、輝度(IntensitiesCM)、エッジ方向(OrientationsCM)等の情報を融合して計算されていることがわかります。

近年ではこれらの低レベル(画像のシグナルレベル)の画像解析からの情報だけではなく、さまざまな情報を統合的に組合せた方法を提案しています。画像認識の国際会議CVPR2010では、Early and Biologically-inspired Visionというセッションが設けられ、新たなサリエンシーマップの研究が発表されています。

Gofermanらは、心理的な知見に基づき、従来の低レベル情報(画像のコ

ントラストや色情報)、画像領域全体に対する各領域の特異度、画像特徴の構成ルール、顔などの高レベル情報の4種類を組合せて認識を行う方法を提案しています⁸⁾。具体的には、画像を多くのパッチに分け、その類似度を画像全体から判定することにより、特異な領域を判定したり、背景パッチは画像全体の離れた位置にも分散して分布することを利用しています。

Wangらは、人の注視領域は最もinformativeな点であるという画像エントロピーに基づく方法を考案し、Site Entropy Rate (SER) と呼ばれる指標を開発しています⁹⁾。これを使って、画像やビデオの中から注目領域を検出する研究を発表しています。

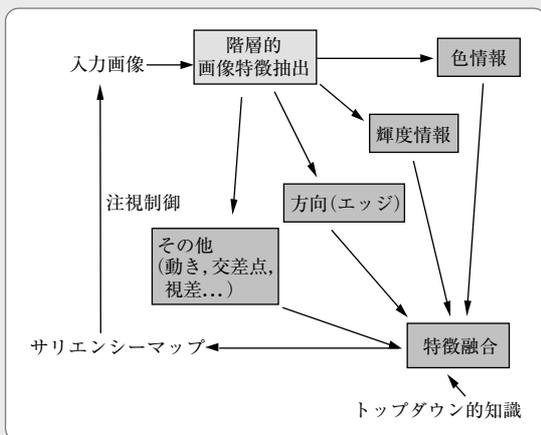


図2 KochとUllmanの手法によるサリエンシーマップ検出法(SaliencyToolbox⁷⁾による)

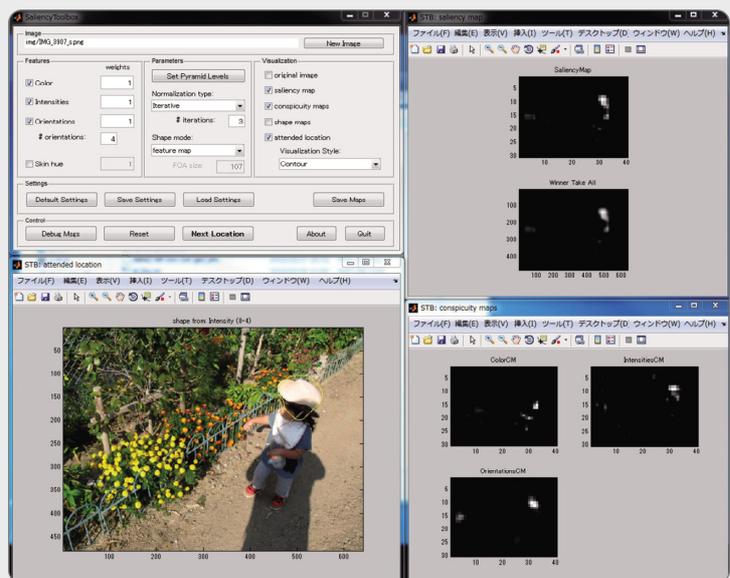


図3 SaliencyToolboxの実行例

視線検出技術との連携

サリエンシーは、人の画像に対する注意度を表現したものであることから、視線検出技術と深い関係があります。ParkhurstやUnderwoodらの研

究では、画像をユーザに提示すると同時に視線も記録し、画像から検出されたとのサリエンシーの情報に対してユーザが良く引きつけられるかを検出しています^{10) 11)}。

逆に画像の特徴と注視点との対応が

得られれば、画像のみから視線方向を推測することが可能になります。菅野らは、入力動画画像から得られたサリエンシーマップと視線の関係を学習して、画像のみから視線を推定する方法を提案しています¹²⁾。

応 用

サイレンシー情報は、人が画像中で注意を引く点を検出する手法であるため、多くの利用可能性が考えられ、実際にも応用する研究が行われています。

前述した研究⁵⁾では、GPUによる画像のレンダリングを、サリエンシーマップを用いて選択的に行う手法が提案されています。サリエンシーマップの

値の高い部分は、人がよく注視していることを意味するため、この部分を重点的に高精度でレンダリングする等の制御を行っています。

Donoserらは、サリエンシーに基づいた画像のセグメンテーションを行う方法を提案しています¹³⁾。この手法では、画像のセグメンテーションの指標として、サリエンシーマップに基づく結果を最適化させるような手法を選択

することで、シーン全体で良いセグメンテーションが可能になると主張しています。

その他、画像の要約法であるSeam Curvingと同様の方法を、サリエンシーを導入することで実現する方法¹⁴⁾や、サムネイル画像をサリエンシーを用いて自動作成する研究¹⁵⁾など、さまざまな応用研究もなされています。

む す び

サリエンシーディテクション技術は比較的古くから研究が行われている研究ですが、近年ではまた多くの研究が出てきました。これは、インターネット動画などのメディア理解へのニーズや、人間の特性を理解したメディア処理へのニーズが高まっていることに起因していると考えられ、理論面・応用面にわたり、今後より活発な研究や利活用が行われる技術であると考えられます。

(2010年10月4日受付)

参 考 文 献

- 1) C. Koch and S. Ullman: "Shifts in selective visual attention: towards the underlying neural circuitry", *Human Neurobiology* 4:219-227 (1985)
- 2) L. Itti & C. Koch: "Computational Modeling of Visual Attention", *Nature Reviews Neuroscience* 2, 3, pp.194-203 (2001)
- 3) H. Yee, S.N. Pattanaik and D.P. Greenberg: "Spatio-temporal sensitivity and visual attention for efficient rendering of dynamic environments", *ACM Transactions on Graphics*, 20, 1, pp.39-65 (Jan. 2001)
- 4) iLab Neuromorphic Vision C++ Toolkit, <http://ilab.usc.edu/toolkit/downloads.shtml>
- 5) P. Longhurst and K. Debattista and A. Chalmers: "A GPU based Saliency Map for High-Fidelity Selective Rendering", *AFRIGRAPH '06: Proceedings of the 4th international conference on Computer graphics, virtual reality, visualisation and interaction in Africa*, pp.21-29 (2006)
- 6) D. Walther and C. Koch: "Modeling attention to salient proto-objects", *Neural Networks* 19, 1395-1407 (2006)
- 7) Saliency Toolbox: <http://www.saliencytoolbox.net/index.html>
- 8) S. Goferman, L. Zelnik-Manor and A. Tal: "Context-aware saliency detection", in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2010)
- 9) W. Wang, Y. Wang: "Measuring Visual Saliency by Site Entropy Rate", in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2010)
- 10) D. Parkhurst and K. Law and E. Niebur: "Modelling the role of saliency in the allocation of visual selective attention", *Vision Research*, 42, 1, pp.107-123 (2002)
- 11) G. Underwood, T. Foulsham, E. van Loon, L. Humphreys and J. Bloyce: "Eye movements during scene inspection: A test of the saliency map hypothesis", *European Journal of Cognitive Psychology*, 18, 3, pp.321-342 (2006)
- 12) Y. Sugano, Y. Matsushita, Y. Sato: "Calibration-free gaze sensing Using saliency maps", in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2010)
- 13) M. Donoser, M. Urschler, M. Hirzer, H. Bischof: "Saliency Driven Total Variation Segmentation", in *IEEE Conference on Computer Vision (ICCV)* (2009)
- 14) L. Shi, J. Wang, L. Xu, H. Lu, C. Xu: "Context Saliency Based Image Summarization", *IEEE international conference on Multimedia and Expo* (2009)
- 15) L. Marchesotti, C. Cifarelli and G. Csurka: "A framework for visual saliency detection with applications to image thumbnailing", in *IEEE Conference on Computer Vision (ICCV)* (2009)



なかざわ あつし
中澤 篤志 1999年、大阪大学大学院基礎工学研究科博士前期課程修了。2001年、同大学大学院博士後期課程修了。同年、科学技術振興事業団研究員(東京大学生産技術研究所)。2003年、大阪大学サイバーメディアセンター講師。2007年~2008年、ジョージア工科大学GVU Center客員研究員。2010年10月より、JSTさきがけ研究員(兼任)。画像計測、コンピュータグラフィックス、ロボティクス人体動作解析および生成の研究に従事。博士(工学)。