Saliency Maps

Paul Bienkowski

Seminar “Brain Modelling”

2014-07-14

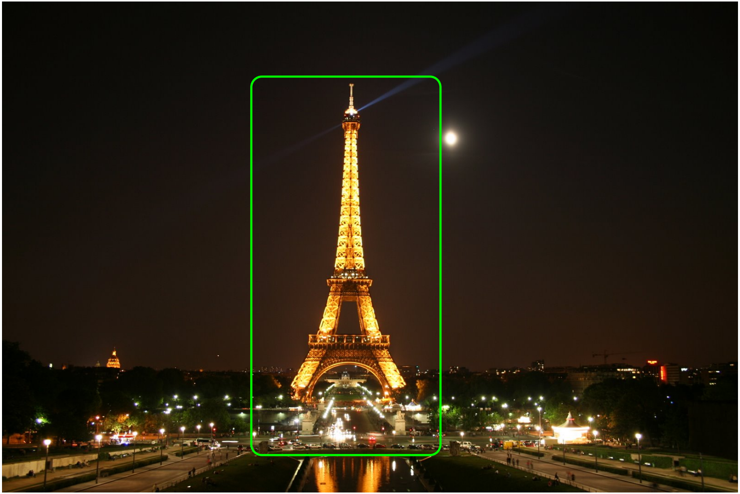
Outline

Introduction

Introduction

Introduction

Introduction



Introduction

Definition

salient

most noticeable or important

Introduction

Definition

salient

most noticeable or important

Definition

map

topographically arranged representation of information in a multidimensional space

Introduction

Definition

salient

most noticeable or important

Definition

map

topographically arranged representation of information in a multidimensional space

Definition

saliency map

A 2D map representing visual saliency of a corresponding visual scene.

Motivation

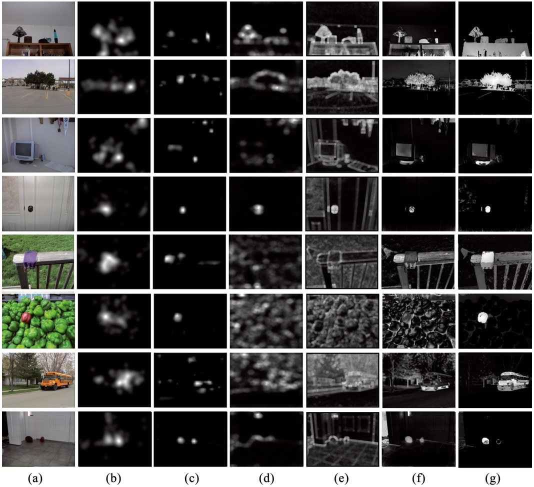
Biological

- plan eye movement
- see things "from the corner of the eye"

Technical

- pre-select important areas from images for image processing
- performance gain in computer vision applications
- object detection (computer graphics application)

Example Maps



Prequisites

Contrast

- *Color intensity contrast* is the difference in color perception
- distance within *Lab color space*
  $\rightarrow$ models perception better than RGB
- requires reference color
  - global average $\rightarrow$ global contrast
  - local average function $\rightarrow$ spatial contrast

Contrast

- *Color intensity contrast* is the difference in color perception
- distance within *Lab color space*
  $\rightarrow$ models perception better than RGB
- requires reference color
  - global average $\rightarrow$ global contrast
  - local average function $\rightarrow$ spatial contrast

Types of contrast

- color
- intensity
- motion
- orientation
- depth

differences

Ventral/dorsal pathway

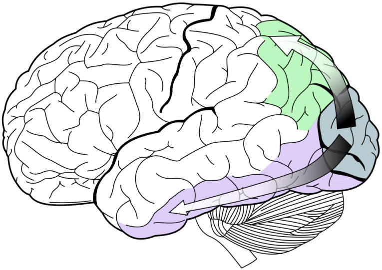


Figure: Ventral and dorsal stream

What?	Where?
Object recognition / identification	Spatial processing
Temporal cortex	Parietal cortex
Ventral stream	Dorsal stream

Models and algorithms

Koch et al. 1987

Shifts in selective visual attention: towards the underlying neural circuitry

- How to handle large amounts of sensory data?
- $\rightarrow$ only process parts of them: SELECTIVE ATTENTION

Koch et al. 1987

Shifts in selective visual attention: towards the underlying neural circuitry

- How to handle large amounts of sensory data?
- $\rightarrow$ only process parts of them: SELECTIVE ATTENTION

- decomposition into early representation of basic features
- retinotopical feature maps
- combined into one general saliency map

Koch et al. 1987

Shifts in selective visual attention: towards the underlying neural circuitry

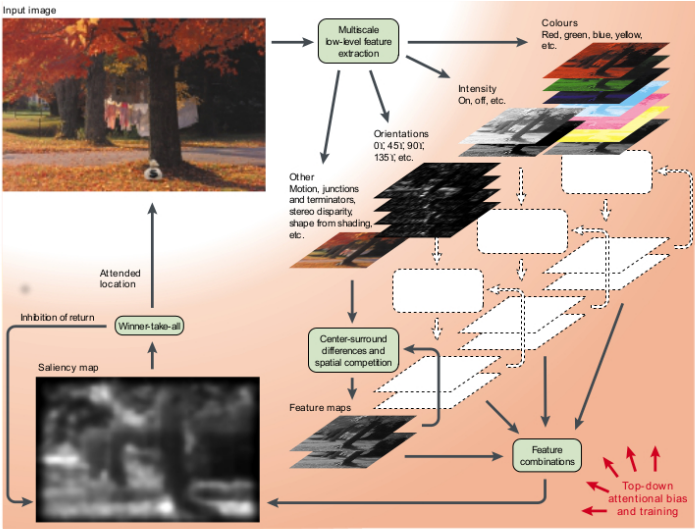
- How to handle large amounts of sensory data?
- $\rightarrow$ only process parts of them: SELECTIVE ATTENTION

- decomposition into early representation of basic features
- retinotopical feature maps
- combined into one general saliency map

"A 'saliency map', that is, an explicit two-dimensional topographical map that encodes stimulus conspicuity, or saliency, at every location in the visual scene."[2]

Koch et al. 1987

The saliency map



Koch et al. 1987

The saliency map

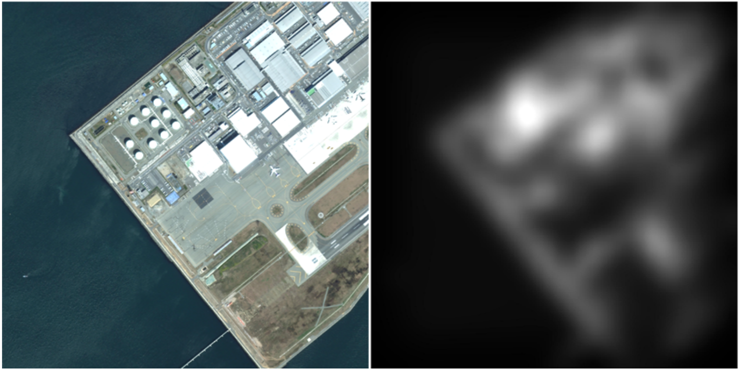
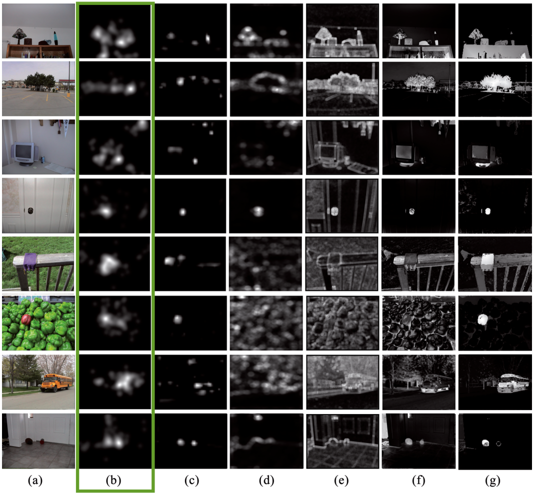


Figure: Sample Koch/Itti saliency map

Example Maps



Koch et al. 1987

Visual search

To search a visual scene, the saliency map and two mechanisms are required:

Koch et al. 1987

Visual search

To search a visual scene, the saliency map and two mechanisms are required:

Winner-take-all

- maximum value and location are determined
- using complex pyramid-shaped neural network
- directs the next saccade

Koch et al. 1987

Visual search

To search a visual scene, the saliency map and two mechanisms are required:

Winner-take-all

- maximum value and location are determined
- using complex pyramid-shaped neural network
- directs the next saccade

Inhibition of return

- concept to prevent going back to seen locations
- simple to implement using inhibitory synapses

Koch et al. 1987

Visual search

Special task: Search for a person in a group of people

- ventral stream detects faces
- feed face locations into saliency map (dorsal stream)
- search salient (face) locations for matching face

Itti/Koch et al. 2001

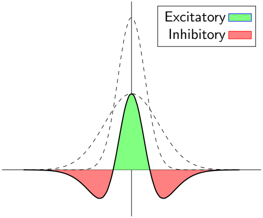
Computational Modelling of visual Attention

- Itti refines Koch's model
- describes early stages in detail
- uses simple localized filters
    - center-sourround
    - gabor kernels

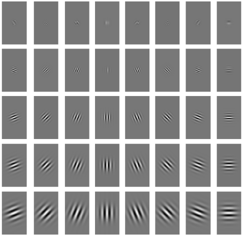
Itti/Koch et al. 2001

Computational Modelling of visual Attention

- Itti refines Koch's model
- describes early stages in detail
- uses simple localized filters
  - center-sourround
  - gabor kernels



Center-sourround / DIFFERENCE OF GAUSSIANS



Gabor kernels

Perazzi et al. 2012

Contrast Based Filtering for Salient Region Detection

- fast and reliable *application-oriented* algorithm
- salient *region* detection
- contour detail stays intact

Perazzi et al. 2012

Contrast Based Filtering for Salient Region Detection

1. segment input image using SLIC
   - *Simple Linear Iterative Clustering*
   - (edge-preserving) K-means clustering algorithms
   - geodesic image distance in Lab color space
   - segments are called *elements*

Perazzi et al. 2012

Contrast Based Filtering for Salient Region Detection

1. segment input image using SLIC
   - *Simple Linear Iterative Clustering*
   - (edge-preserving) K-means clustering algorithms
   - geodesic image distance in Lab color space
   - segments are called *elements*
2. analyze segment uniqueness
   - rarity of the element color
   - compared to all other segments
   - weighted by the distance between them
   - Gaussian weight function for fast localized uniqueness

Perazzi et al. 2012

Contrast Based Filtering for Salient Region Detection

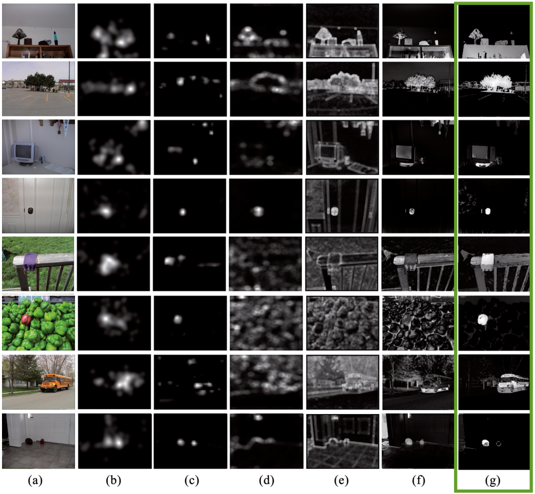
1. segment input image using SLIC
   - *Simple Linear Iterative Clustering*
   - (edge-preserving) K-means clustering algorithms
   - geodesic image distance in Lab color space
   - segments are called *elements*
2. analyze segment uniqueness
   - rarity of the element color
   - compared to all other segments
   - weighted by the distance between them
   - Gaussian weight function for fast localized uniqueness
3. analyze segment distribution
   - occurence of a color elsewhere in the image
   - low variance $\rightarrow$ compact object $\rightarrow$ more salient

Perazzi et al. 2012

Contrast Based Filtering for Salient Region Detection

1. segment input image using SLIC
   - *Simple Linear Iterative Clustering*
   - (edge-preserving) K-means clustering algorithms
   - geodesic image distance in Lab color space
   - segments are called *elements*
2. analyze segment uniqueness
   - rarity of the element color
   - compared to all other segments
   - weighted by the distance between them
   - Gaussian weight function for fast localized uniqueness
3. analyze segment distribution
   - occurence of a color elsewhere in the image
   - low variance $\rightarrow$ compact object $\rightarrow$ more salient
4. combine segment uniqueness and distribution and apply to original pixels

Example Maps



Hou et al. 2007

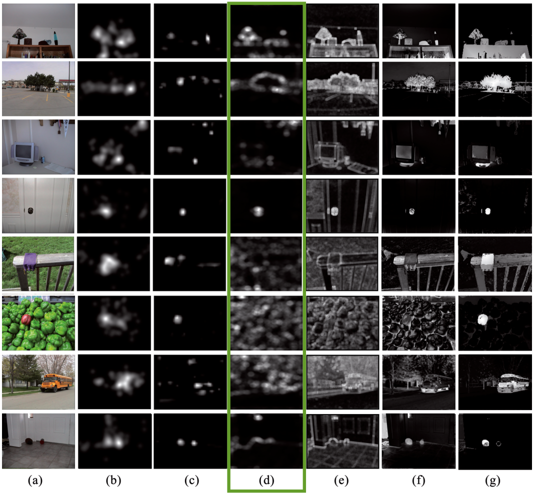
- in natural scene images, spatial frequency analysis (Fourier transform) yields a typical average pattern
- can be used to determine *untypical* frequencies
  $\rightarrow$ not average image data but interesting object
- reverse lookup for saliency map generation

Hou et al. 2007

- in natural scene images, spatial frequency analysis (Fourier transform) yields a typical average pattern
- can be used to determine *untypical* frequencies
  $\rightarrow$ not average image data but interesting object
- reverse lookup for saliency map generation

- independent of low-level features
- filter background noise from data
- flexible system
- *not neuro-biologically plausible*

Example Maps



Algorithm features

- Blurriness
  - caused by center-sourround filters or downsampling for performance
  - no problem in finding single POI[1] (winner-take-all)
  - impedes object detection

[1] point of interest

Algorithm features

- Blurriness
  - caused by center-sourround filters or downsampling for performance
  - no problem in finding single POI[1] (winner-take-all)
  - impedes object detection
- sort POI by intensity
  - opposite of boundary detection (binary decision)
  - required for visual search

[1] point of interest

Algorithm features

- Blurriness
  - caused by center-sourround filters or downsampling for performance
  - no problem in finding single POI[1] (winner-take-all)
  - impedes object detection
- sort POI by intensity
  - opposite of boundary detection (binary decision)
  - required for visual search
- iterative analysis
  - approximate quickly, refine over time
  - useful in robotics

[1] point of interest

Algorithm features

- Blurriness
  - caused by center-sourround filters or downsampling for performance
  - no problem in finding single POI[1] (winner-take-all)
  - impedes object detection
- sort POI by intensity
  - opposite of boundary detection (binary decision)
  - required for visual search
- iterative analysis
  - approximate quickly, refine over time
  - useful in robotics
- spatio-temporal analysis
  - analyze video data, movement detection

[1] point of interest

Discussion

Comparison

Approach	Itti/Koch	Perazzi	Hou
Model type	neural	computational	computational
Object detection	impossible	great (with contours)	good
Sortable by intensity	good / inhibition of return	fair	fair
Iterative analysis	simple	not possible	n/a
Movement detection	use 3-dim. input space	possible	difficult

Table: Comparison between approaches and their features and capabilities

Neuro-biological relevance

- Koch/Itti
  - based on neuro-biological findings
  - easily implementable using feedforward neural networks
  - combinable with winner-take-all/inhibition of return for visual search

Neuro-biological relevance

- Koch/Itti
  - based on neuro-biological findings
  - easily implementable using feedforward neural networks
  - combinable with winner-take-all/inhibition of return for visual search
- Perazzi **and** Hou
  - computational models
  - no similar findings in biological research
  - no obvious neural network implementations
  - perform tasks associated with both ventral *and* dorsal stream

Neuro-biological relevance

- Koch/Itti
  - based on neuro-biological findings
  - easily implementable using feedforward neural networks
  - combinable with winner-take-all/inhibition of return for visual search
- Perazzi **and** Hou
  - computational models
  - no similar findings in biological research
  - no obvious neural network implementations
  - perform tasks associated with both ventral *and* dorsal stream

Looking for a neuro-biological model of visual attention?

Neuro-biological relevance

- Koch/Itti
  - based on neuro-biological findings
  - easily implementable using feedforward neural networks
  - combinable with winner-take-all/inhibition of return for visual search
- Perazzi **and** Hou
  - computational models
  - no similar findings in biological research
  - no obvious neural network implementations
  - perform tasks associated with both ventral *and* dorsal stream

Looking for a neuro-biological model of visual attention?
$\Rightarrow$ Koch/Itti!

Conclusion

Conclusion

- many different saliency detection methods (named three)

Conclusion

- many different saliency detection methods (named three)
- either **neural** or **computational** models

Conclusion

- many different saliency detection methods (named three)
- either **neural** or **computational** models
- Koch/Itti is state-of-the-art neuro-biological model for saliency

Conclusion

- many different saliency detection methods (named three)
- either **neural** or **computational** models
- Koch/Itti is state-of-the-art neuro-biological model for saliency
- other algorithms include object detection etc, more task specific

Conclusion

- many different saliency detection methods (named three)
- either **neural** or **computational** models
- Koch/Itti is state-of-the-art neuro-biological model for saliency
- other algorithms include object detection etc, more task specific
- broad area of applications in science and industry

Appendix

Bibliography

Xiaodi Hou and Liqing Zhang.
Saliency detection: A spectral residual approach.
In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

Laurent Itti and Christof Koch.
Computational modelling of visual attention.
Nature reviews neuroscience, 2(3):194–203, 2001.

R. M. Klein and J. Ivanoff.
Inhibition of return.
3(10):3650, 2008.
revision #90806.

Raymond M Klein.
Inhibition of return.
Trends in cognitive sciences, 4(4):138–147, 2000.

Christof Koch and Shimon Ullman.
Shifts in selective visual attention: towards the underlying neural circuitry.
In *Matters of Intelligence*, pages 115–141. Springer, 1987.

Bibliography

Norbert Krüger, Peter Janssen, Sinan Kalkan, Markus Lappe, Ales Leonardis, Justus Piater, Antonio Jose Rodriguez-Sanchez, and Laurenz Wiskott.
Deep hierarchies in the primate visual cortex: What can we learn for computer vision?
Pattern Analysis and Machine Intelligence, IEEE Transactions on, 35(8):1847–1871, 2013.

Federico Perazzi, Philipp Krahenbuhl, Yael Pritch, and Alexander Hornung.
Saliency filters: Contrast based filtering for salient region detection.
In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 733–740. IEEE, 2012.

Pamela Reinagel and Anthony M Zador.
Natural scene statistics at the centre of gaze.
Network: Computation in Neural Systems, 10(4):341–350, 1999.

Image sources

3 http://www.npebd.com/wp-content/uploads/2014/06/Eiffel-Tower.jpg

1 http://upload.wikimedia.org/wikipedia/commons/thumb/f/fb/Ventral-dorsal_streams.svg/1024px-Ventral-dorsal_streams.svg.png

2 http://cogpsy.info/wp-content/uploads/2012/10/Itti-Koch-2001-Computational-modeling-of-visual-attention.jpg

6 http://opticalengineering.spiedigitallibrary.org/data/Journals/OPTICE/23430/057008_1_9.png

3 http://opticalengineering.spiedigitallibrary.org/data/Journals/OPTICE/24206/OE_51_2_026201_f006.png

4b http://www.intechopen.com/source/html/16592/media/image74.png

16 http://stanford.edu/~philkr/imgs/sf_teaser.png

Questions?

Questions?

Thanks for your attention!