CVPR 2013 TUTORIAL:
A CRASH COURSE ON VISUAL SALIENCY MODELING:
BEHAVIORAL FINDINGS AND COMPUTATIONAL MODELS

SALIENCY: APPLICATIONS IN VISION, IMAGE PROCESSING AND COMPUTER GRAPHICS

Neil Bruce

Department of Computer
University of Manitoba
Winnipeg, MB, Canada

Contact Information:

bruce@cs.umanitoba.ca
www.cs.umanitoba.ca/~bruce

UNIVERSITY OF MANITOBA

Overview

- Some applications in computer vision / image processing:
  - Mobile robots
  - Surveillance
  - Proto-object selection
  - Cropping/resizing/retargeting/montages
  - Compression/sampling
  - Graphics, rendering and art
  - Other
  - Beyond Salience

Applications

- Mobile robot navigation

Applications

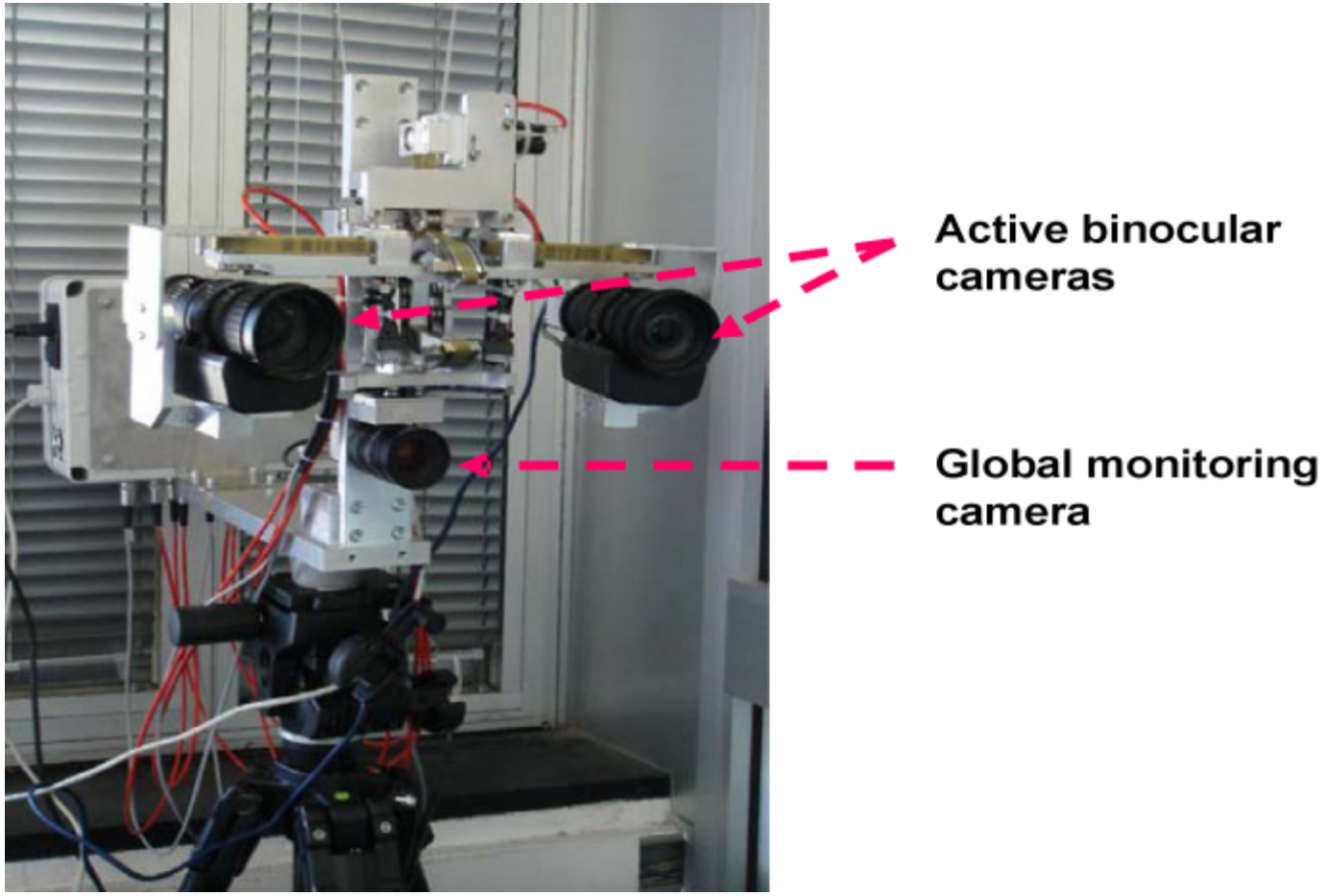
- Surveillance

SEARISE

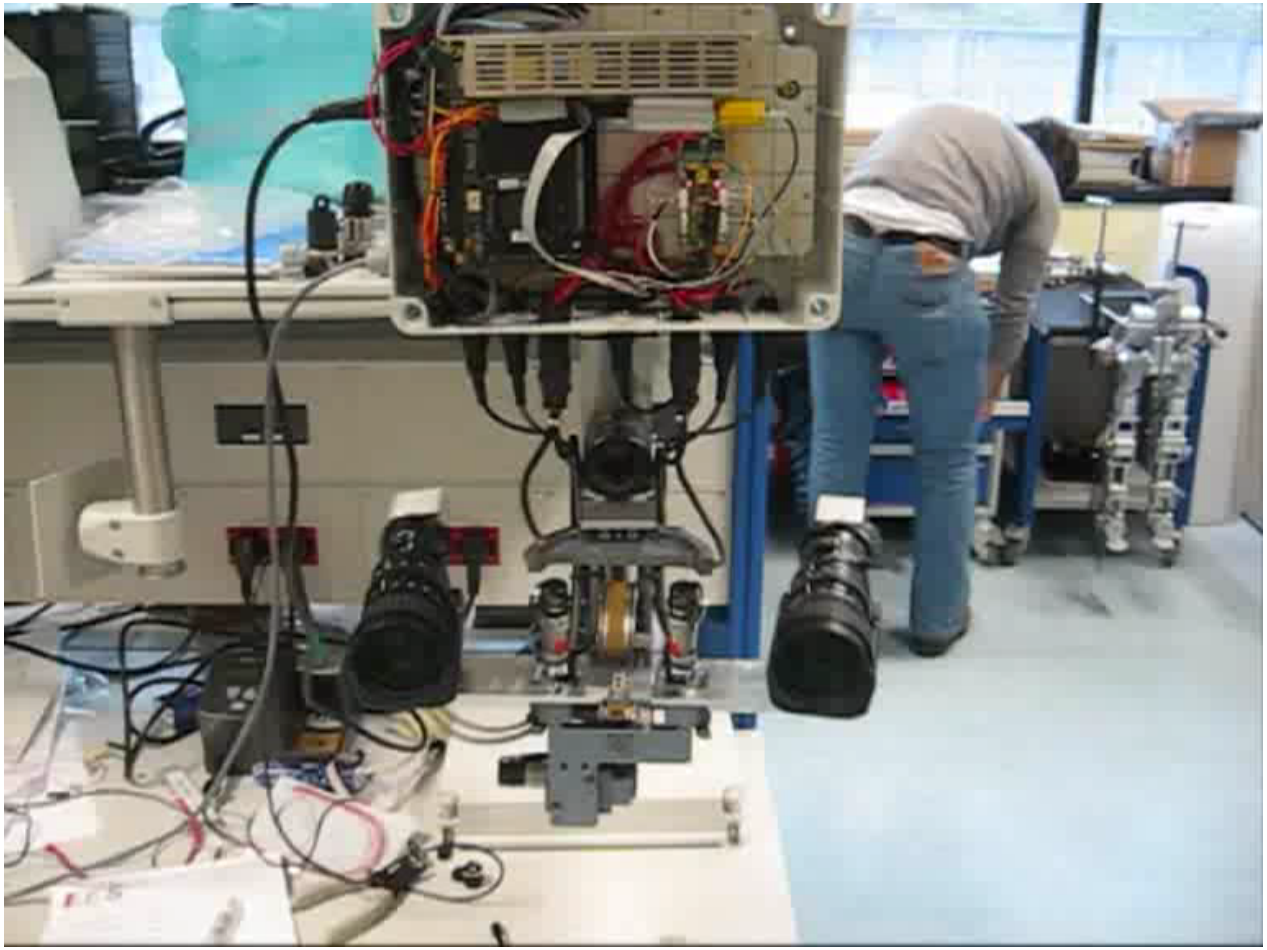
- Applying saliency to surveillance
- Smart Eyes: Attending and Recognizing Instances of Salient Events
- European Project including 7 University and Industrial Partners
- Wide angle low resolution camera – For detecting salient events
- Focused high resolution camera deployed based on salience for detailed focus and analysis

SEARISE: Data

SEARISE

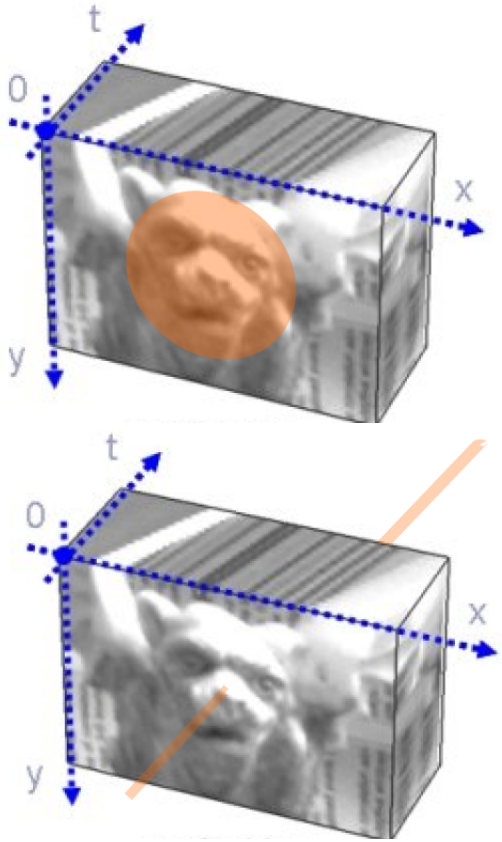


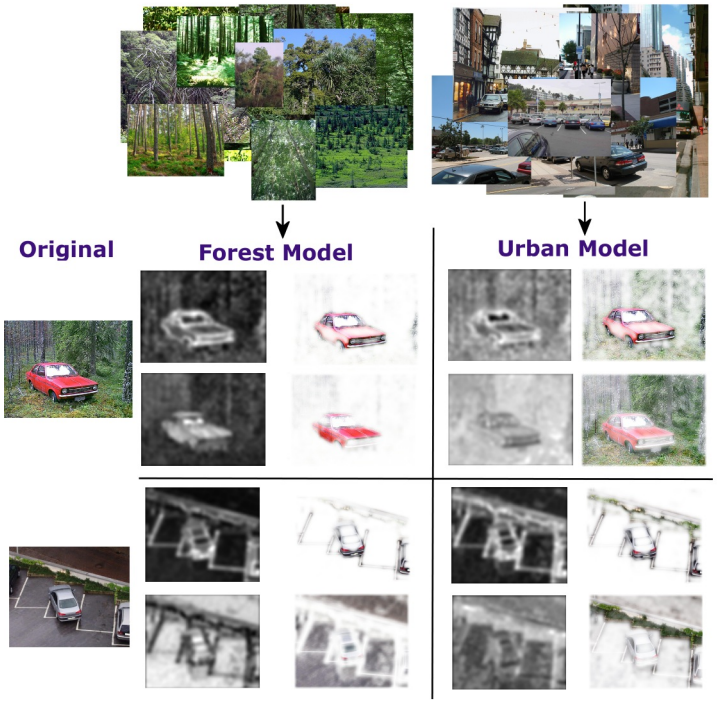
SEARISE



SEARISE: Data

- Region of support need not be spatial

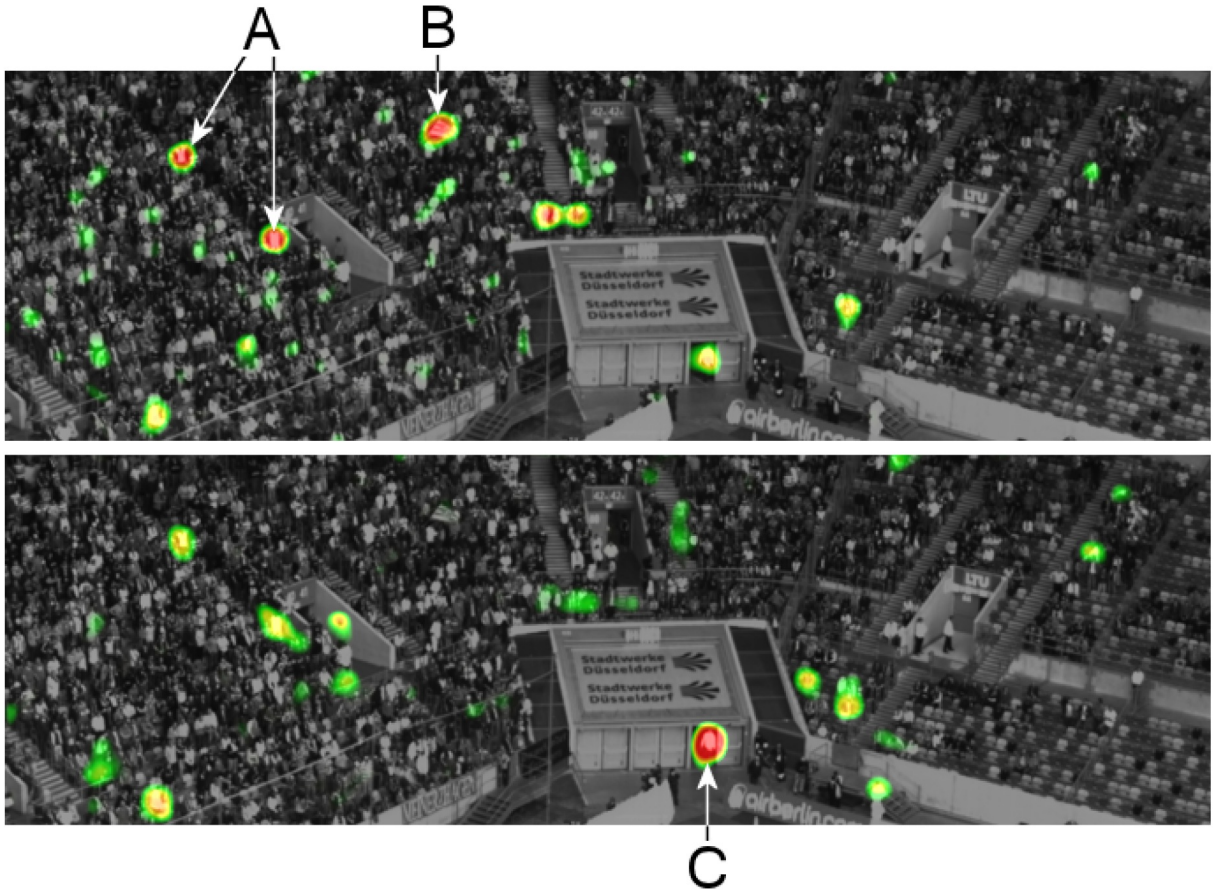




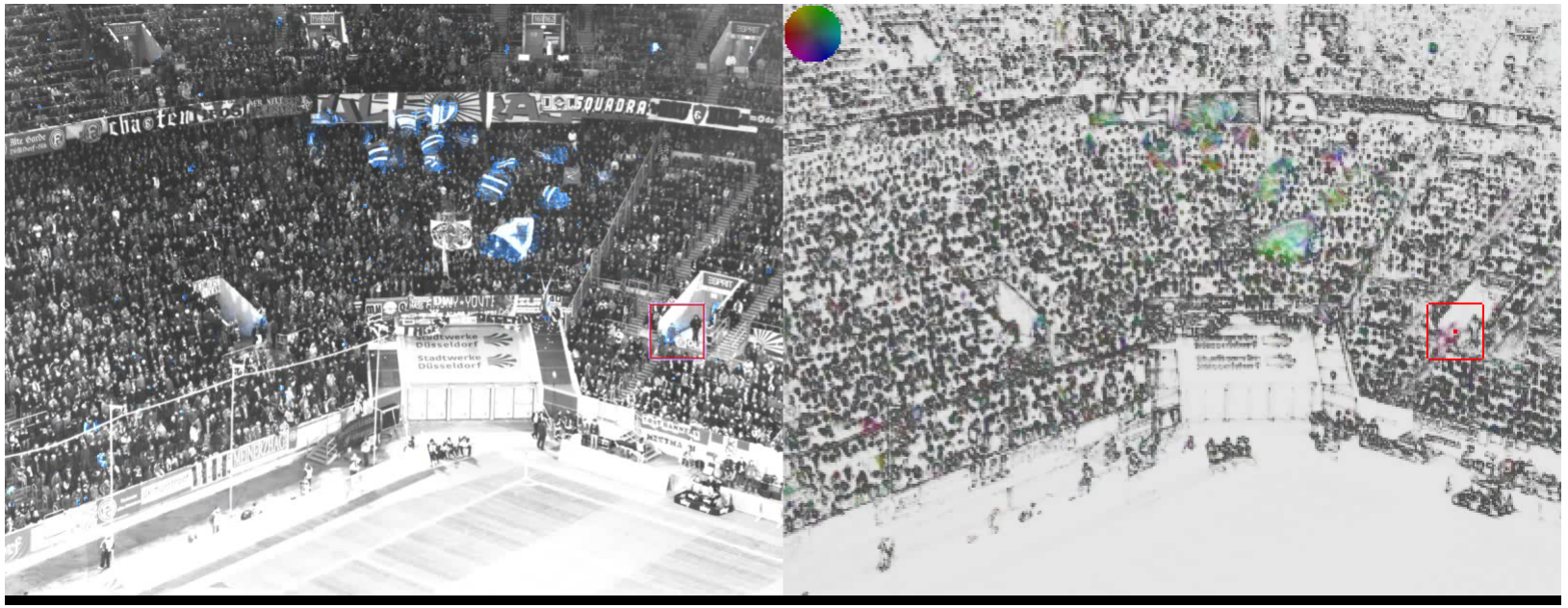
(Bruce and Kornprobst, 2009)

Salience in SEARISE

Some single frame examples



Selection, Tracking and IOR



Selection, Tracking and IOR

High resolution focus of attention

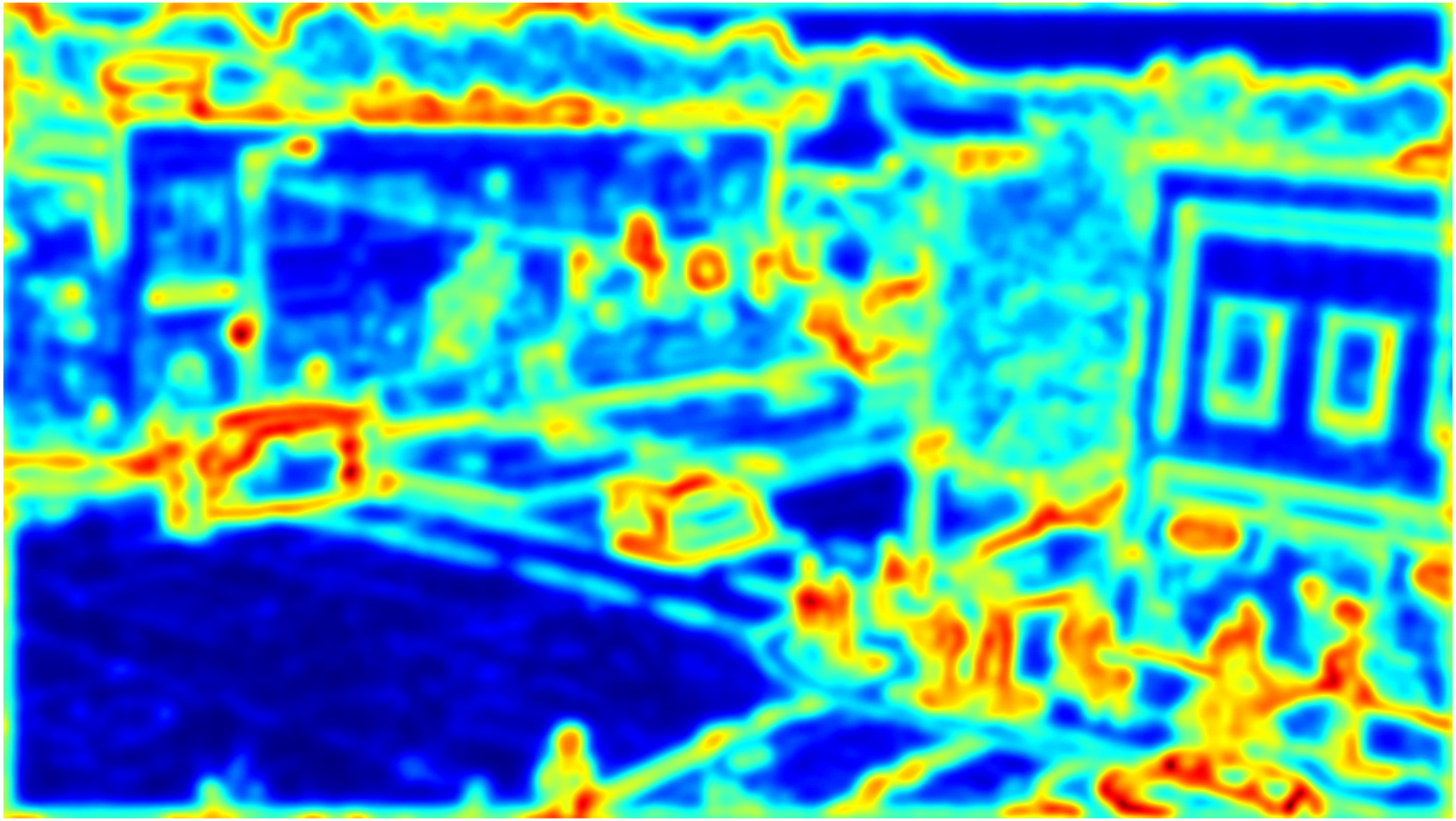
Active vision, and multi-resolution

- Wide view, zoom view
  - Focal processing vs. focal sensing
  - Allows for benefits to be had at the acquisition stage
- Other specific domains may benefit from this paradigm…

Halfblue,
Wikimedia Commons

Attention for proto-object selection

Attention for proto-object selection

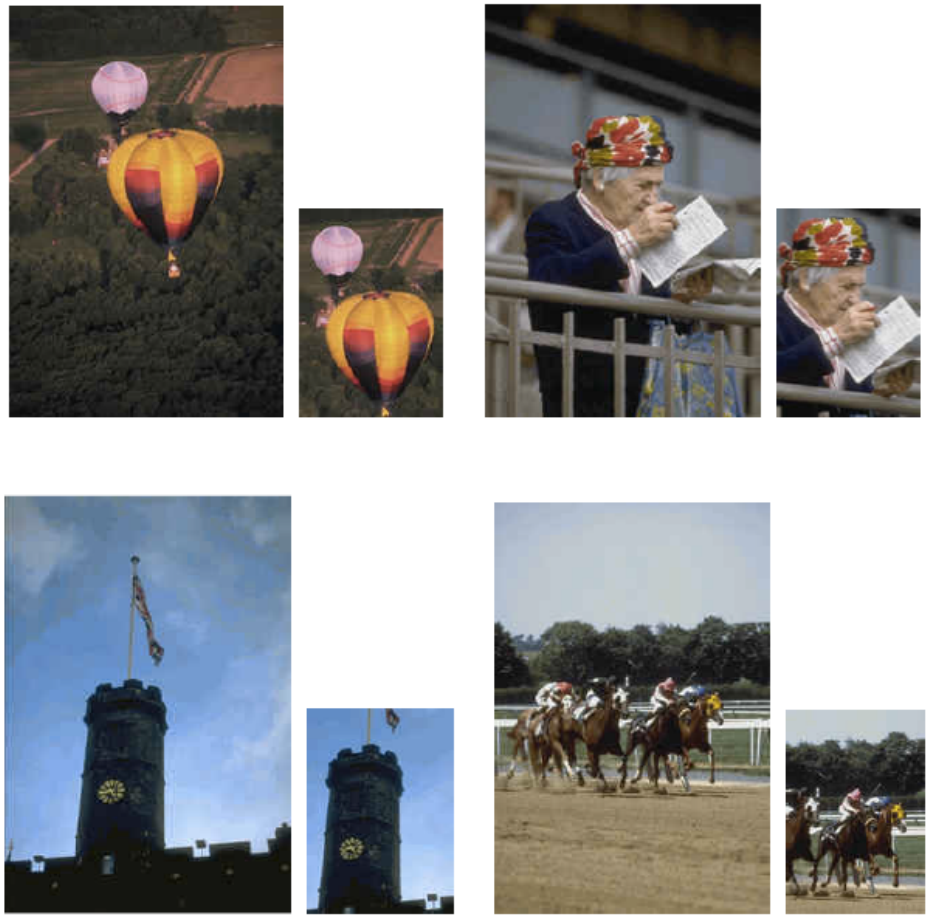


Attention for proto-object selection

Attention for proto-object selection

IMAGES: CROPPING, RESIZING AND RE-TARGETING

Auto-cropping



Steniford, ICVS 2007

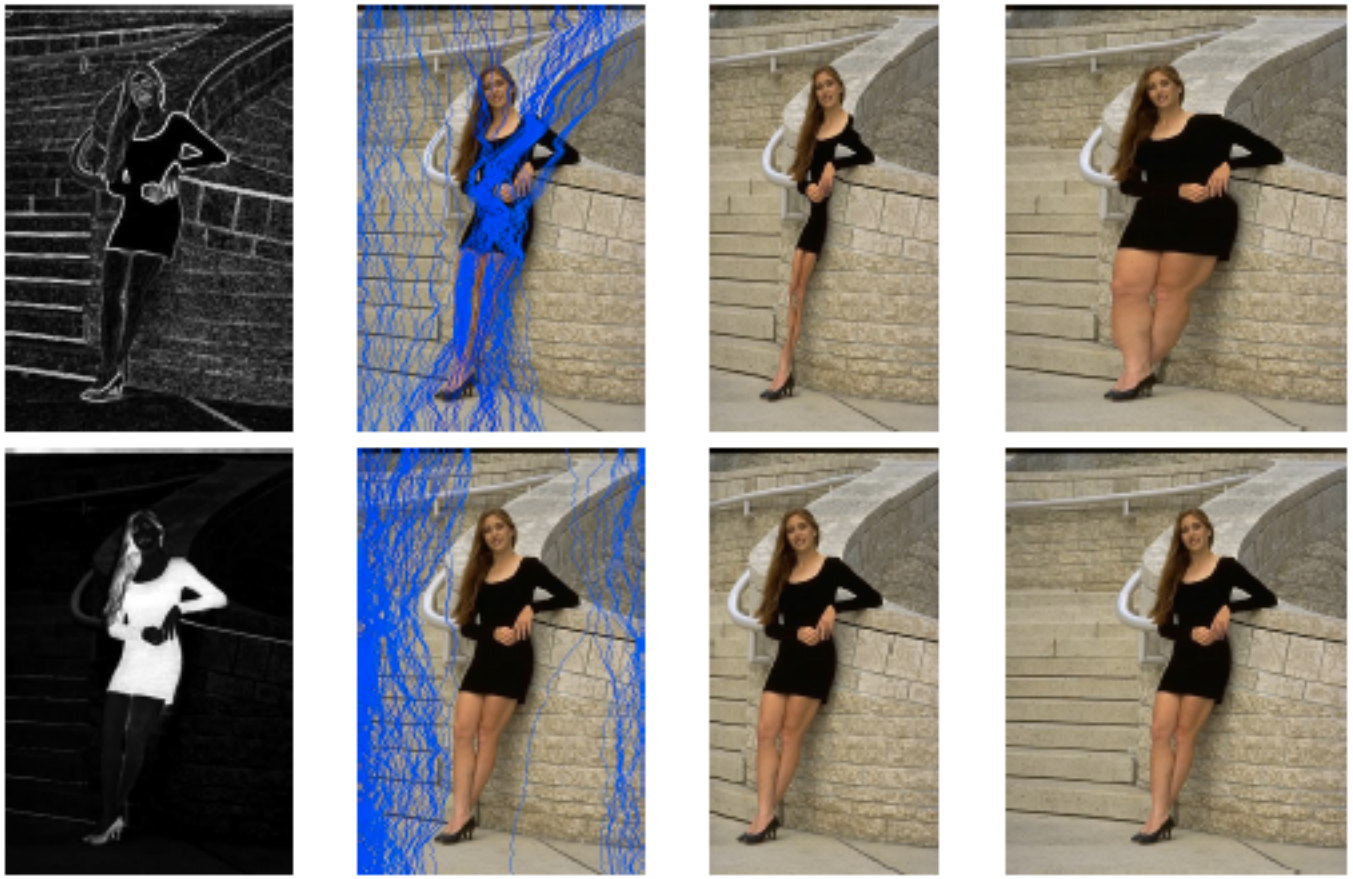
Automatic Thumbnailing

Marchesotti et al., ICCV 2009

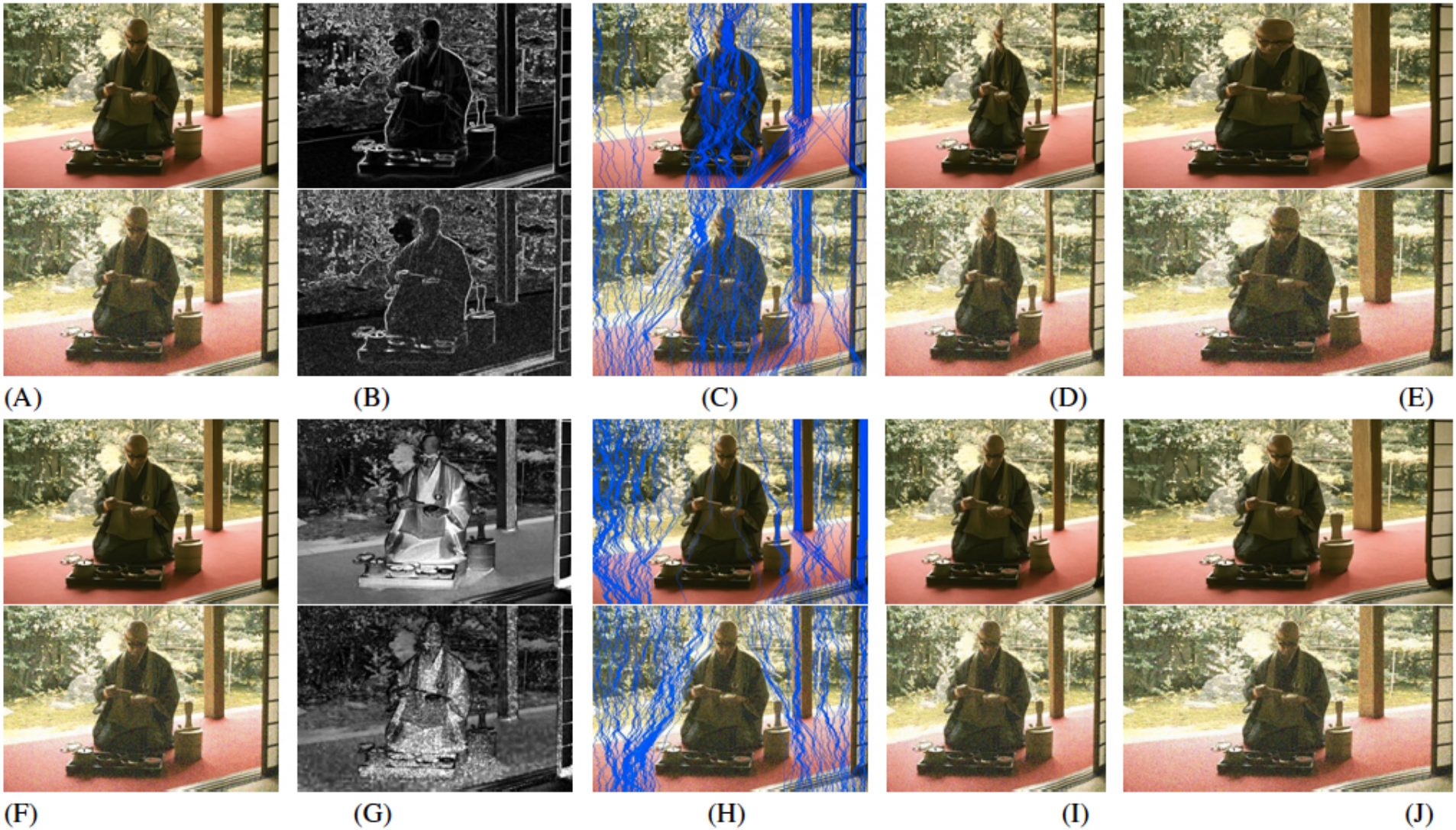
Content aware resizing

- Content aware resizing (Achanta + Susstrunk, 2009)

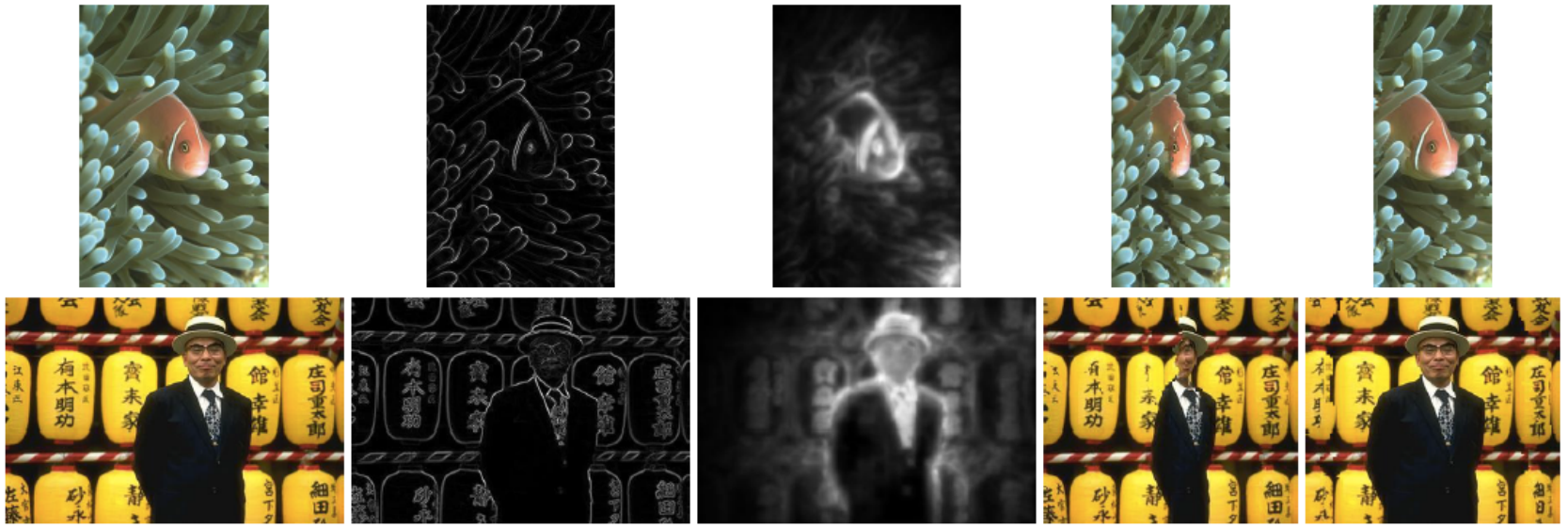
Seam carving



Seam carving



Re-targeting



(Goferman, Zelnik-Manor, Tal, 2010, 2012)

Photo montages

- Data summarization (Goferman, Zelnik-Manor, Tal, 2010, 2012)

Photo montages

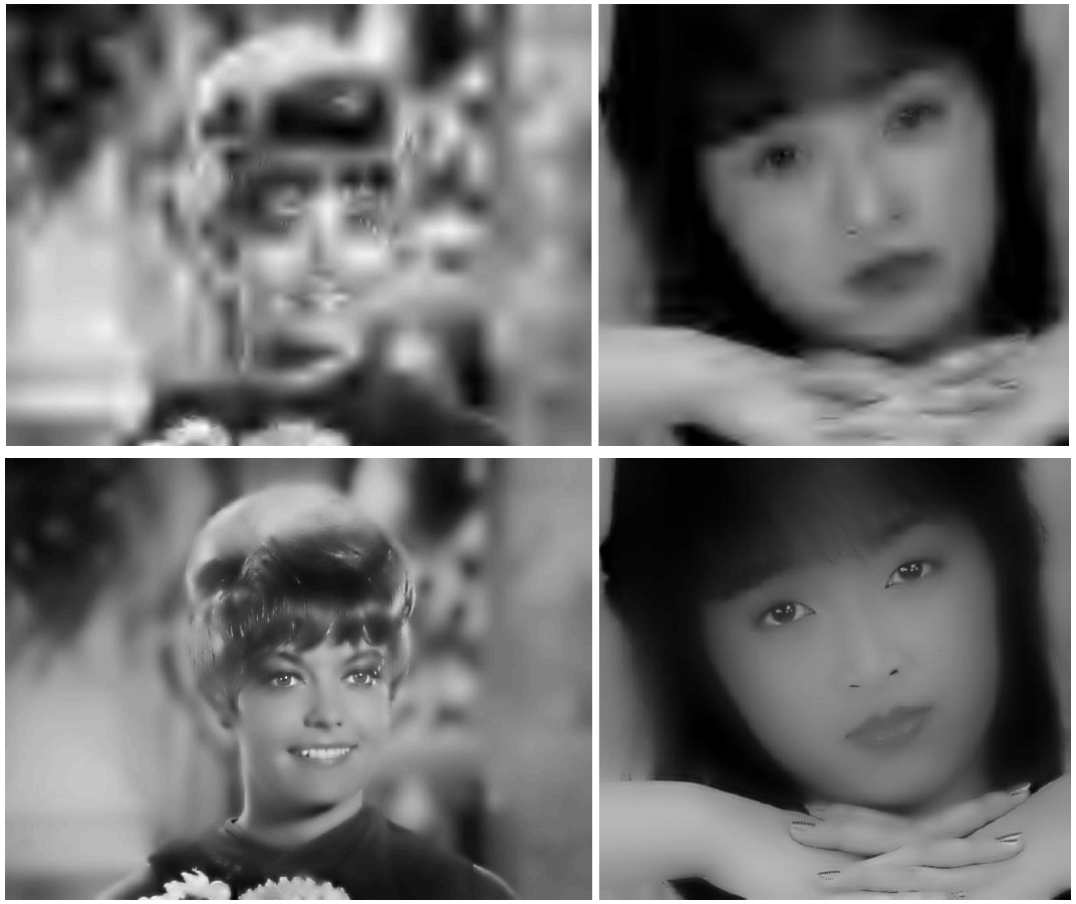
(a) Mosaic Collage (b) Google's Picasa Collage (c) Picture Collage

Wang et al. CVPR 2006

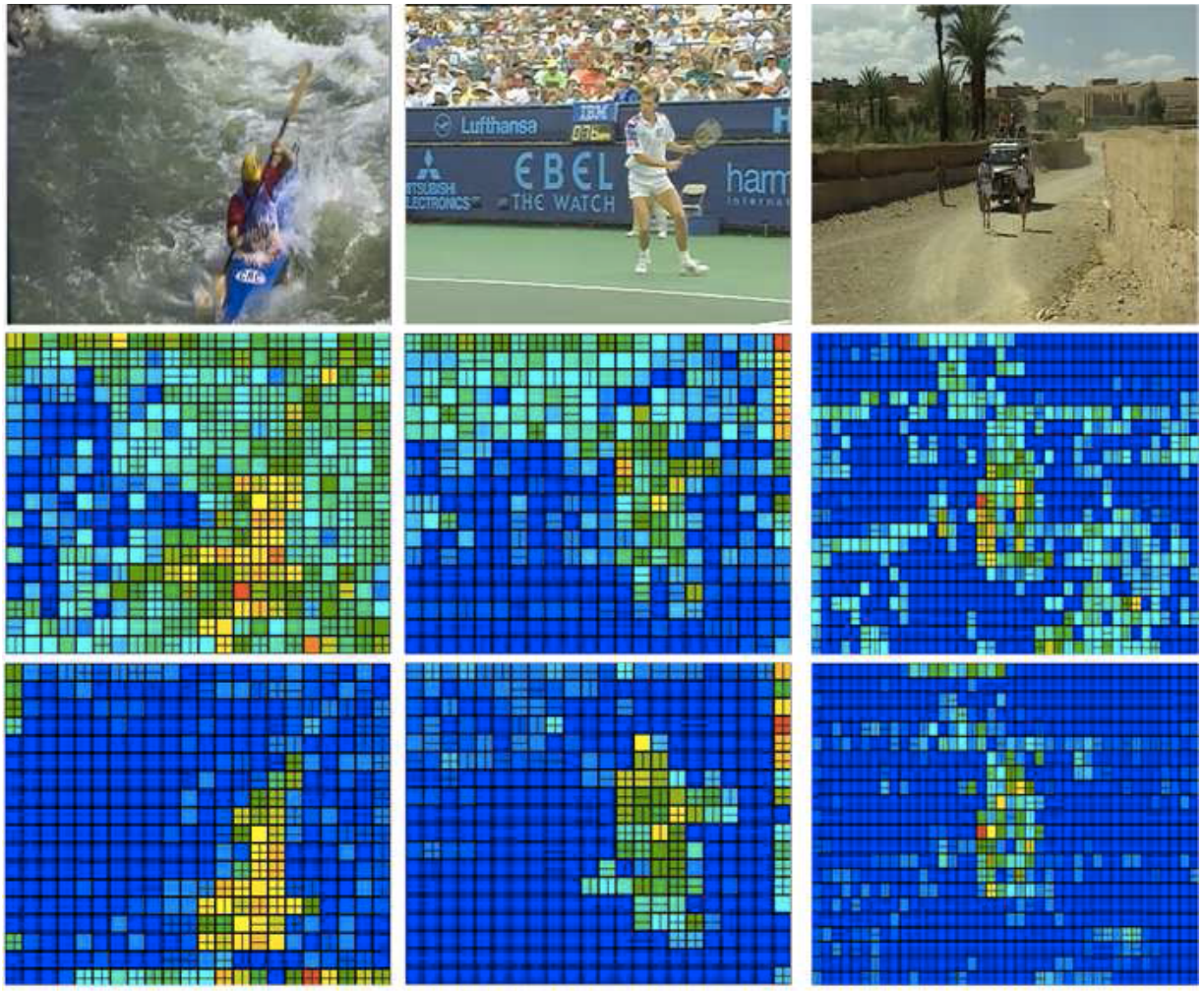
PERCEPTUALLY MOTIVATED COMPRESSION, SAMPLING

Why is this important?

- Compression and quality assessment



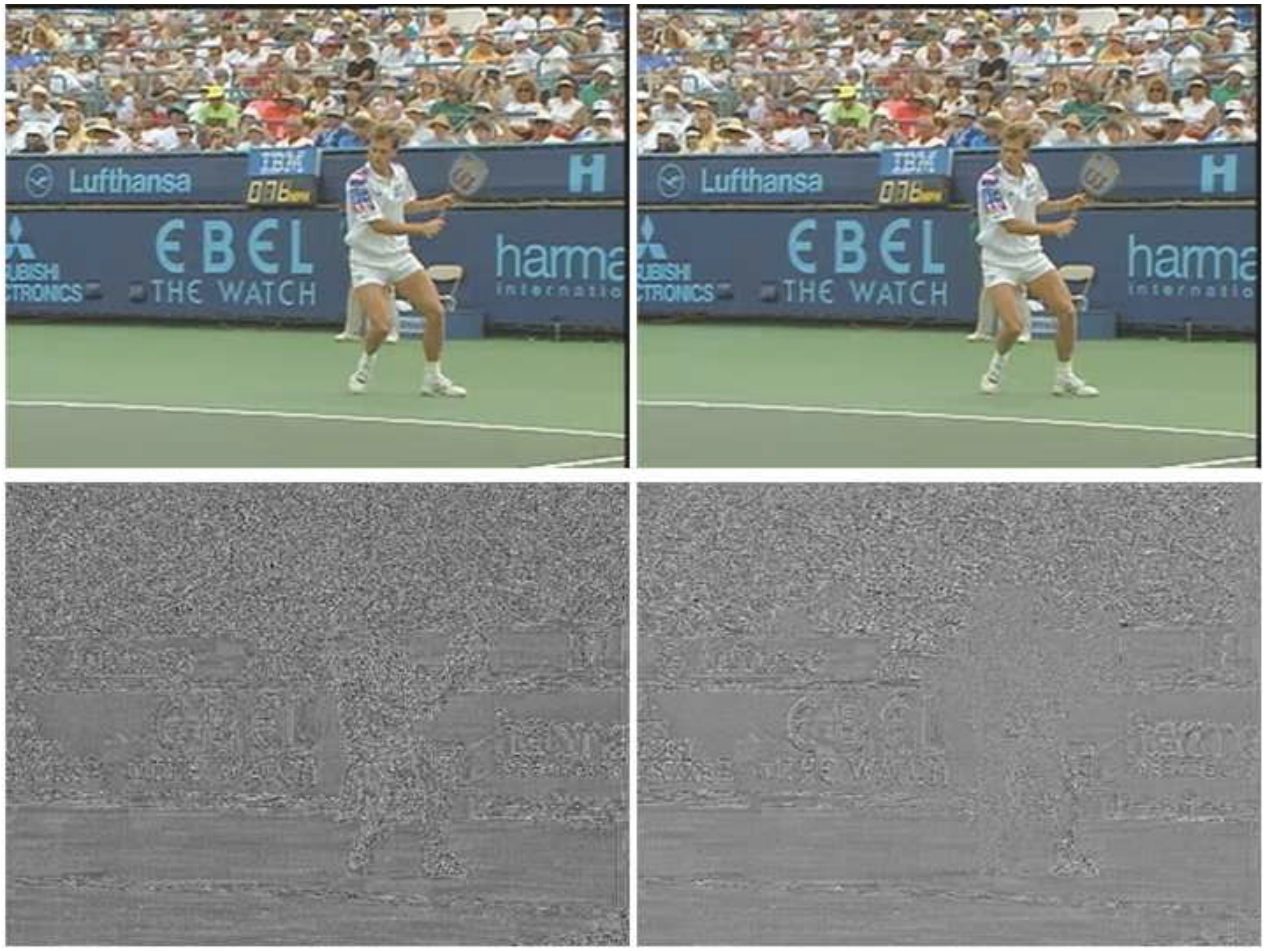
Video Compression



Le Meur, Le Callet, Barba, 2005 (unpublished)

See also:
Le Meur, Castellan, Le Callet, Barba, ICIP 2006

Video Compression



Spatiotemporal compression

Guo and Zhang, IEEE TIP, 2010

(See also Dhavale + Itti, Signal processing and applications, 2003)

Frame-rate Up-conversion

(a) Original

(b) MMVP

(c) Method of [4]

(d) Proposed

Jacobson and Nguyen, ICASSP, 2011

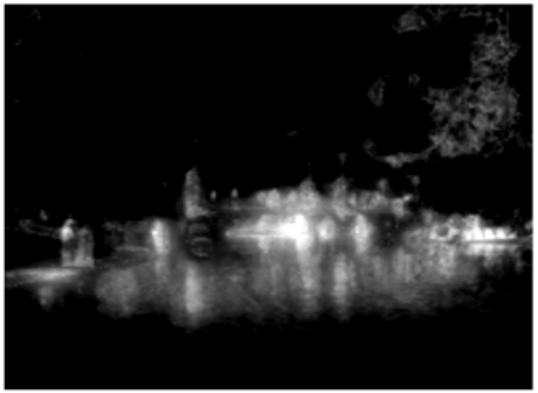
GRAPHICS, RENDERING AND ART

Graphics, Rendering, Art

- Rendering

Graphics, Rendering, Art

- Artistic Effects



Graphics, Rendering, Art

Original Various rendering effects

Margolin, Zelnik-Mayor, Tal
CGI 2012

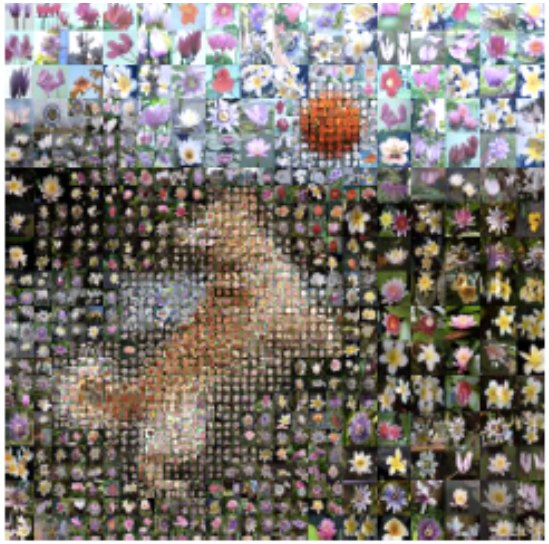
Graphics, Rendering, Art

Input

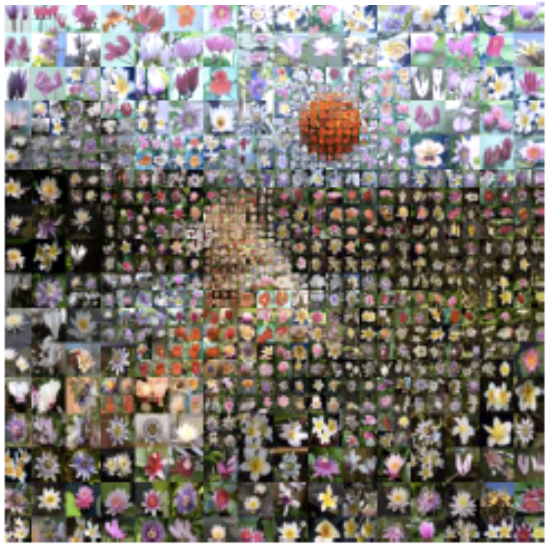
Our

[9]

Input



Our

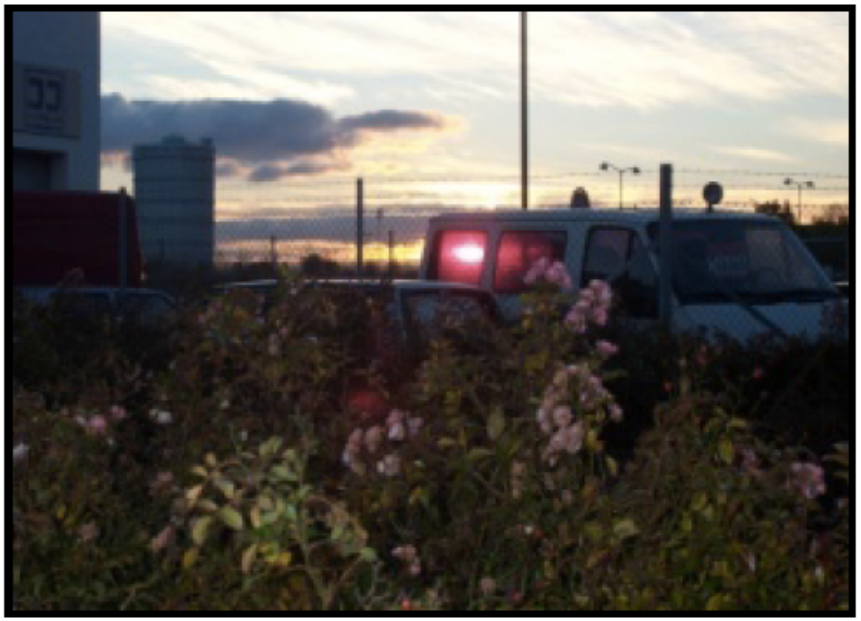


[9]

MISCELLANEOUS

Feature selection

e.g. Aesthetic class prediction



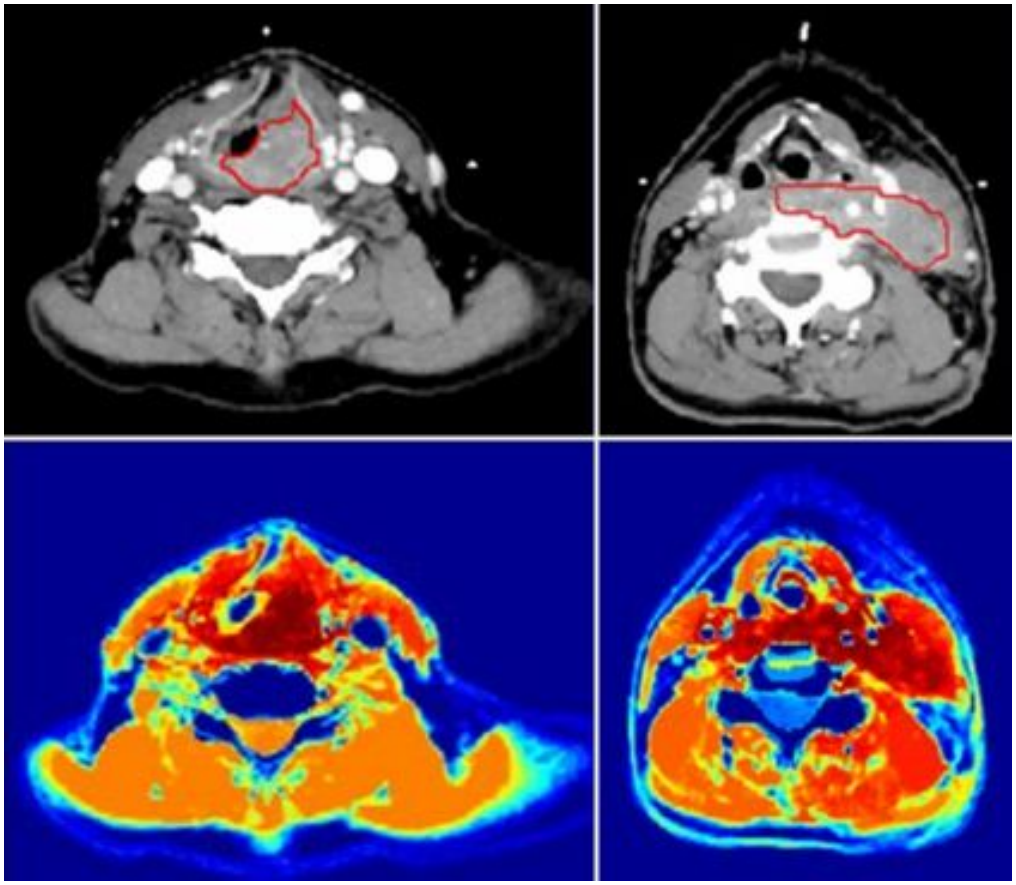
Wong, Low, ICIP 2009

Other examples:
Object recognition
Memorability
Crowd analysis
etc.

Appropriateness depends on specific decision being made

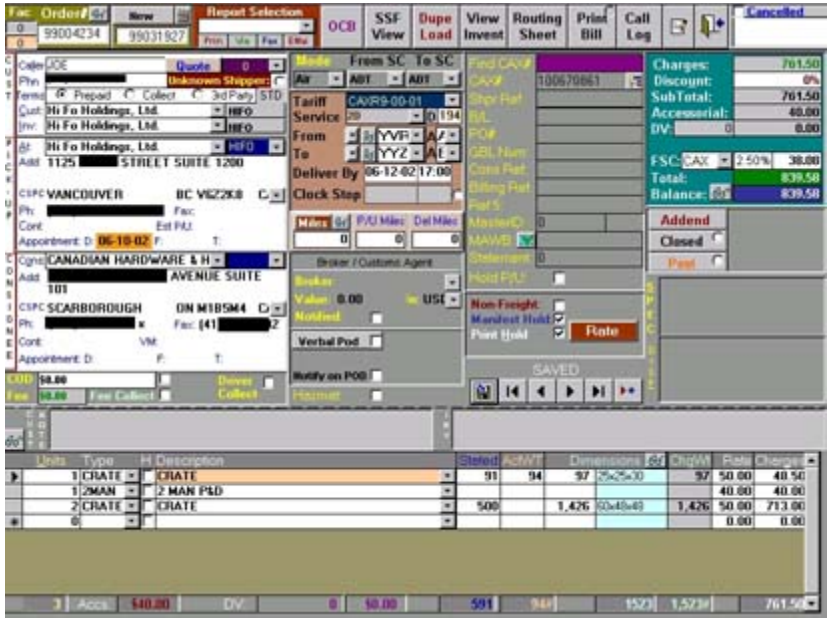
Other possibilities

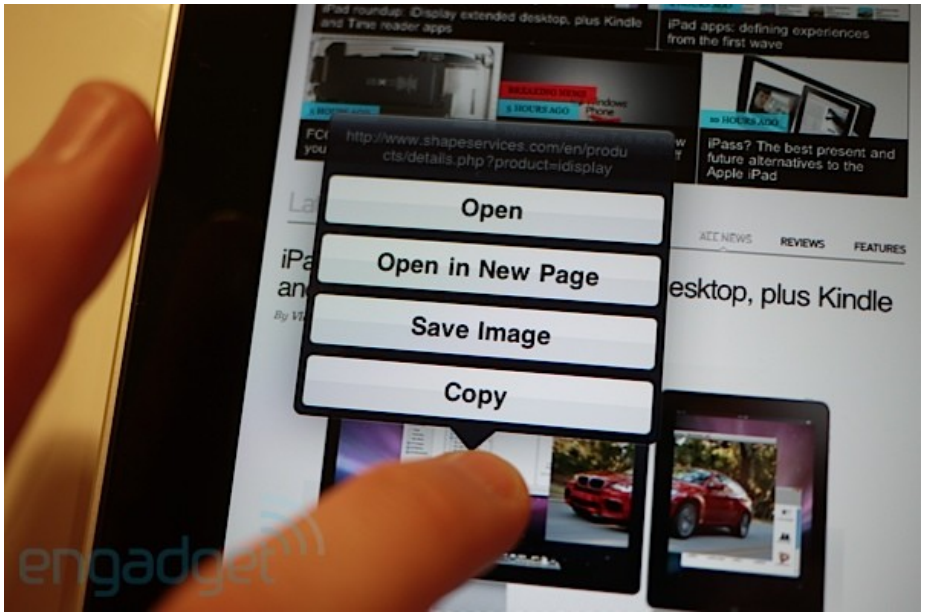
- Pathology detection



Other possibilities

- Interface design / HCI





Other possibilities

(Really) Big Data

Other possibilities

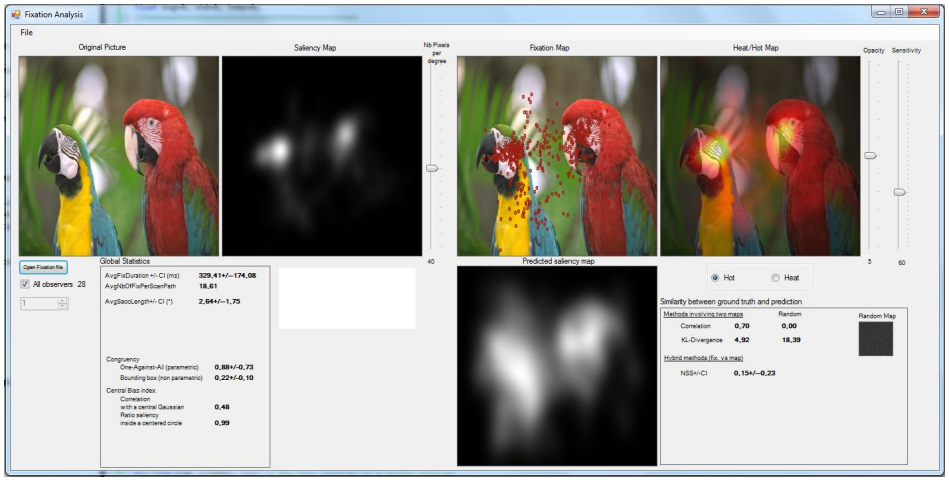
- Real world scenarios

Examining Eye Movements

- Fixation analysis software:

  LeMeur and Baccino
  Behavior Research Methods (BRM) 2012

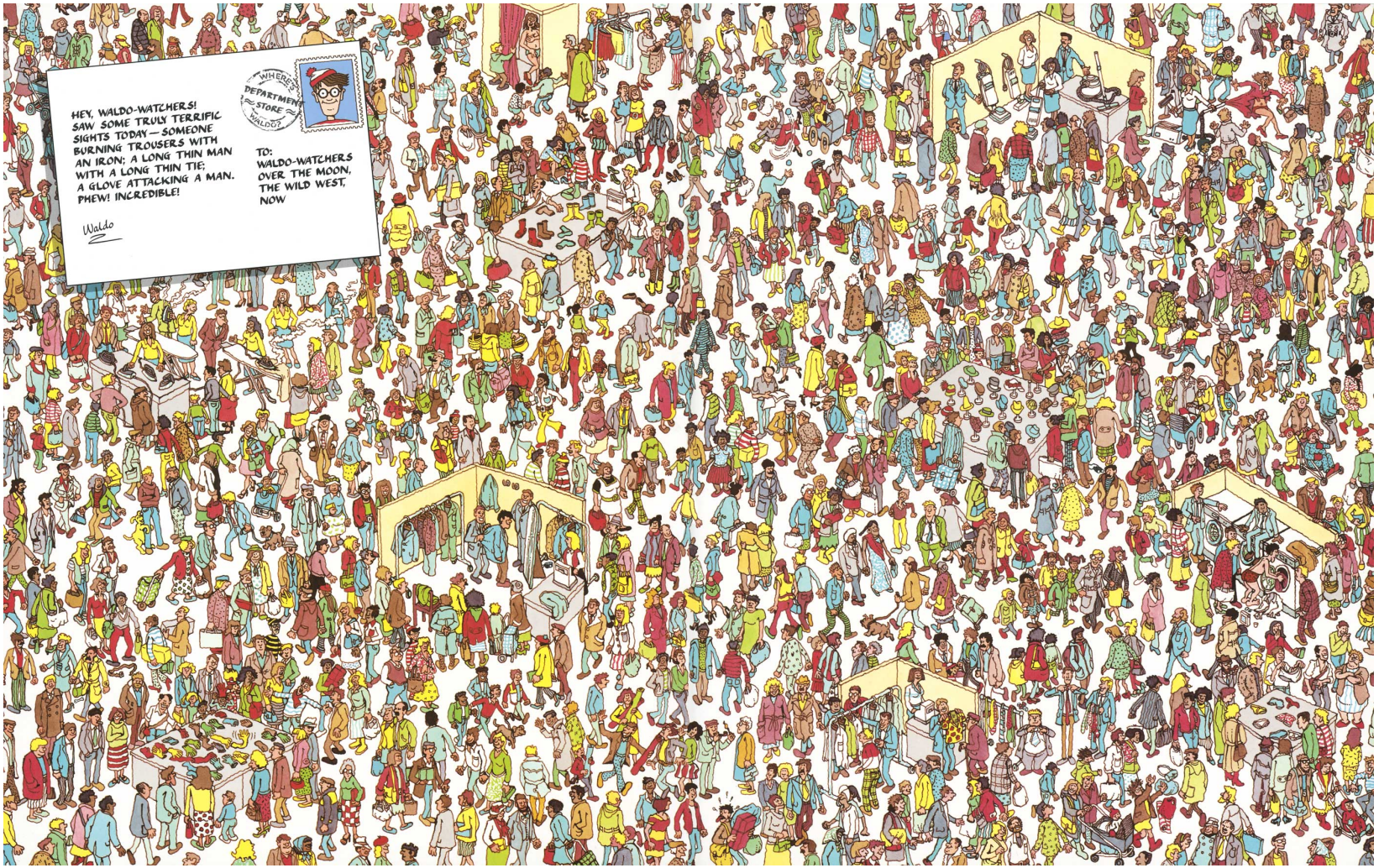
  http://people.irisa.fr/Olivier.Le_Meur/publi/2012_BRM/index2.html



- See Ali Borji's page for additional evaluation software – which he will discuss!

BEYOND SALIENCY?

Beyond saliency



What is attention?

- Some elements of visual processing are evidently **focal**
- Computational foundations:

  Number of image locations (P)
  Number of object/event prototypes in memory (N)
  Number of measurements made at each image location (M)
- Theorem 1:
  Unbounded (Bottom-Up) Visual Matching is NP-Complete, with time complexity an exponential function of P - Worst-case is $O(N2^{PM})$
- Theorem 2: Bounded (Task-Directed) Visual Match has time complexity linear in P - Worst-case is $O(NP2^{M})$
- Addition of attentive, task-specific guidance is sufficient to reduce this complexity to linear or better
- Basic theorems, proved in (Tsotsos 1989, 1990a, 1995)
  See also (Tsotsos and Bruce, 2008)

Summary and conclusions

Summary and Conclusions

- Many possible applications across various branches of computer science from computer vision to image and video processing, HCI, robotics and human behavior
- Right choice for role of salience (and algorithm) may be application dependent
- Salience can be useful for many applications, but doesn't solve everything – think about complexity too!