My 10-Year Research Vision: Building the Prosopagnosic Computer

FG 2015

Matthew Turk

Computer Science Department and

Media Arts and Technology Program

University of California, Santa Barbara

http://www.cs.ucsb.edu/~mturk

Theme of the talk

- Invited to talk about:
  - New and under-explored problems
  - Promising new approaches emerging in related fields
  - New ways of evaluation, datasets, benchmarking, etc.
  - New applications
  - New interdisciplinary research opportunities
  - Future role of professional organizations in the field

Theme of the talk

- Invited to talk about:
  - New and under-explored problems
  - Promising new approaches emerging in related fields
  - New ways of evaluation, datasets, benchmarking, etc.
  - New applications
  - New interdisciplinary research opportunities
  - Future role of professional organizations in the field

Theme of the talk

- Invited to talk about:
  - New and under-explored problems
  - Promising new approaches emerging in related fields
  - New ways of evaluation, datasets, benchmarking, etc.
  - New applications
  - New interdisciplinary research opportunities
  - Future role of professional organizations in the field

- What I'll actually talk about:
  - **Old** things and the **past**!

Prosopagnosia (a.k.a. “Face blindness”)

Greek: *prosopon* = face, *agnosia* = not knowing

Prosopagnosia (a.k.a. “Face blindness”)

Greek: *prosopon* = face, *agnosia* = not knowing

An impaired ability to recognize faces, while other aspects of visual processing and intellectual functioning remain largely intact.

Somewhat analogous to being “tone deaf” or “color blind”

Prosopagnosia (a.k.a. “Face blindness”)

Greek: *prosopon* = face, *agnosia* = not knowing

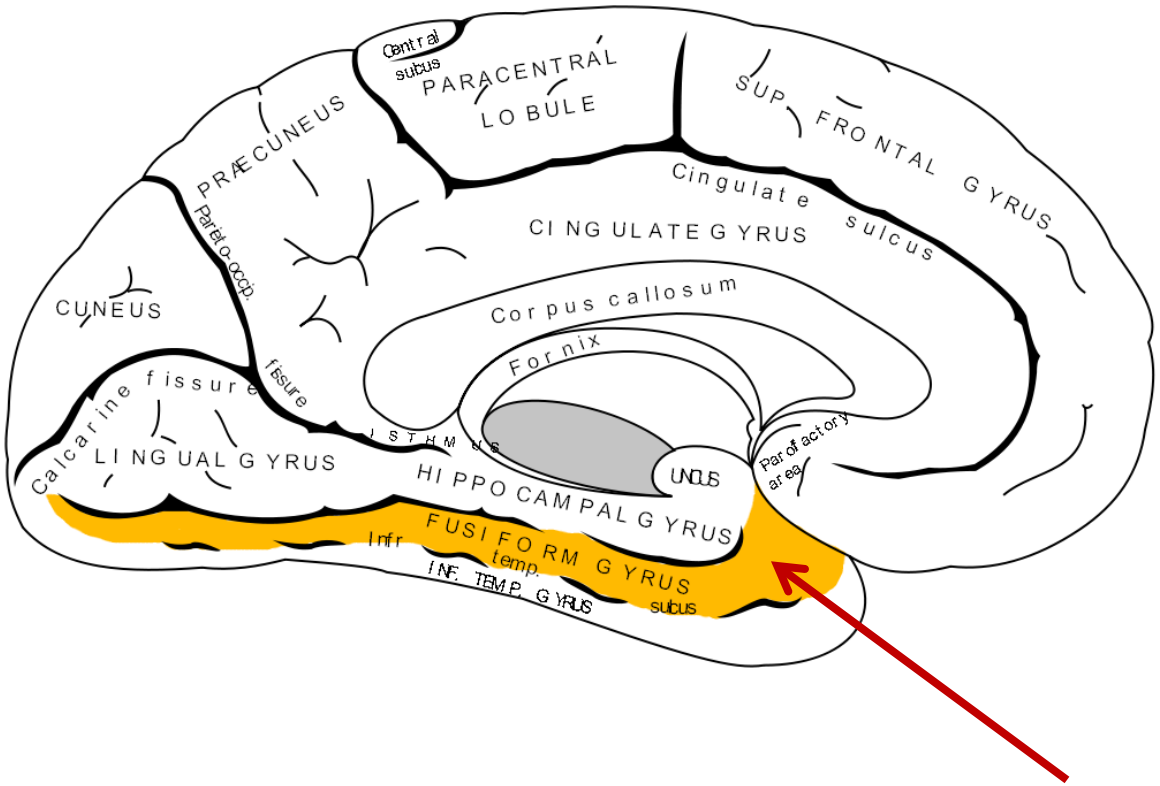
An impaired ability to recognize faces, while other aspects of visual processing and intellectual functioning remain largely intact.

Somewhat analogous to being “tone deaf” or “color blind”

May be congenital (developmental) or acquired (resulting from injury or degeneration)

Affects ~2% of the population

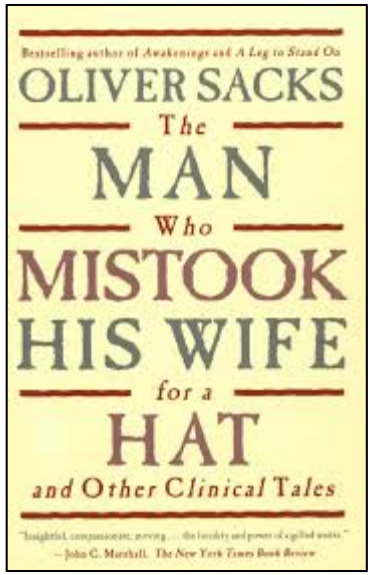
Prosopagnosia (a.k.a. “Face blindness”)

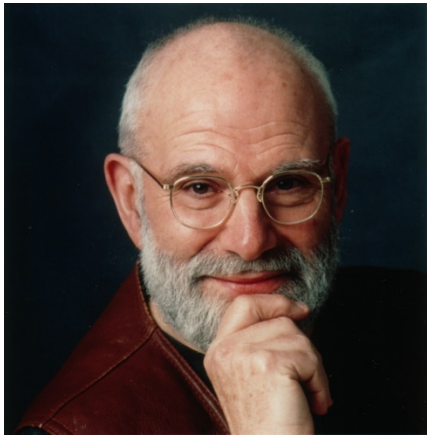


Typically associated with damage in the *fusiform gyrus* (a.k.a. the *fusiform face area*), in the temporal and occipital lobes

Prosopagnosia (a.k.a. “Face blindness”)

1985

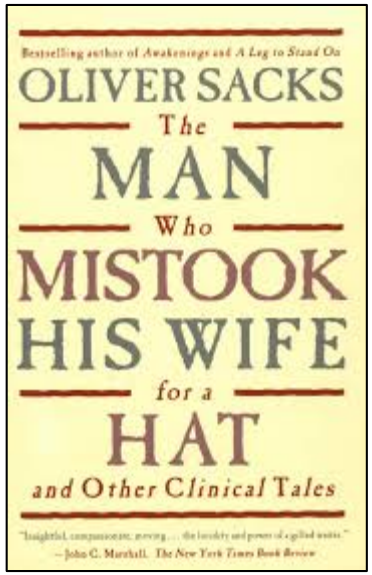


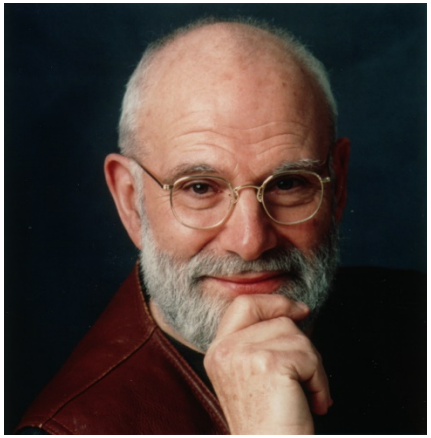


Oliver Sacks

Prosopagnosia (a.k.a. “Face blindness”)

1985





Oliver Sacks

Brad Duchaine

Prosopagnosia (a.k.a. “Face blindness”)

- Prosopagnosia is characterized by an inability to recognize people (incl. close friends, relatives, and self from face only)
- Relatively normal eyesight
- Perception of facial features, face detection, and face/nonface discrimination may not be impaired
- Often accompanied by impairments in recognition of facial expression, sex, attractiveness, age, race, etc.
- Prosopagnosics are usually able to discern identity via other channels such as voice, gait, hair, or clothing
- Causes severe social problems

Prosopagnosia (a.k.a. “Face blindness”)

Bill Choisser

- Recognized his developmental prosopagnosia in his late 40s
- Very intelligent, MIT student, lawyer, engineer
- Reports no other visual difficulties
- Recognizes people using hair, facial hair, and clothes
- Online book about prosopagnosia (2002): www.choisser.com/faceblind

Prosopagnosia (a.k.a. “Face blindness”)

- For decades, neuroscience researchers have debated two alternative views of face recognition:
  - **Face-specific view:** Face recognition is carried out by dedicated, face-specific mechanisms
  - **Expertise view:** Face recognition leverages more general mechanisms that are also used in general visual recognition

- Prosopagnosic patients have been a rich source of investigation for this question!

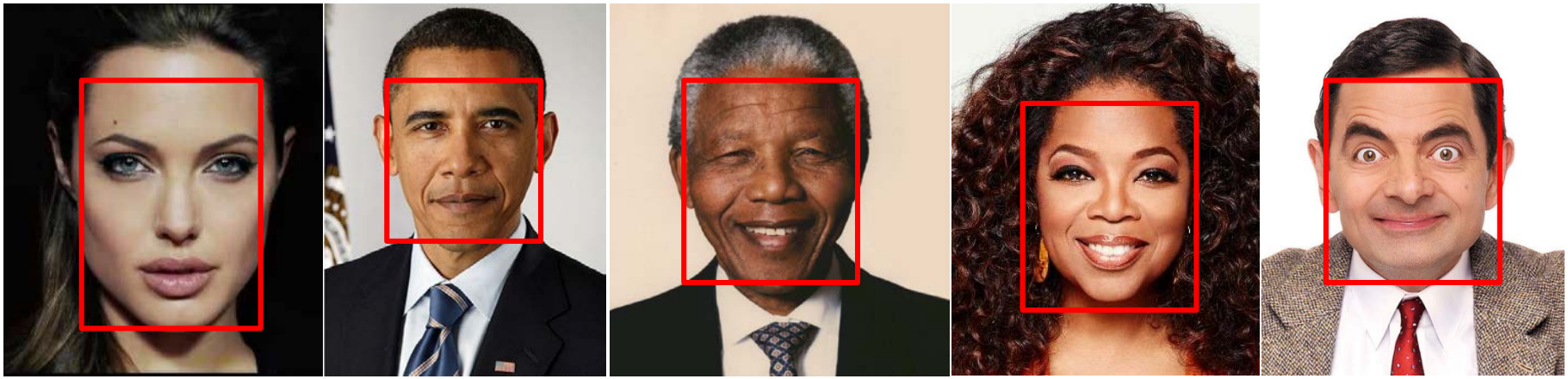
Prosopagnosia (a.k.a. “Face blindness”)

- The study of prosopagnosia, although it is a very specific dysfunction, involves exploring a wide range of neural, developmental, cognitive, and behavioral phenomena

- It has informed models of how the visual system works and thus enabled researchers to pose questions and new experiments intended to illuminate how biological systems achieve face and object recognition

Building the prosopagnosic computer

Building the prosopagnosic computer



? ? ? ? ?

This is not what I mean.

Building the prosopagnosic computer

Can we build visual recognition systems (including face and gesture recognition) that:

1. Work more like biological recognition systems
   - Multiple recognition strategies
   - Long-term learning
2. Help to inform those doing visual neuroscience and visual psychophysics, and be informed by them:
   - To refine models of perception
   - To determine experiments to run

Building the prosopagnosic computer

Can we build visual recognition systems (including face and gesture recognition) that:

1. Work more like biological recognition systems
   - Multiple recognition strategies
   - Long-term learning
2. Help to inform those doing visual neuroscience and visual psychophysics, and be informed by them:
   - To refine models of perception
   - To determine experiments to run

More generally, can the FG (computer vision) community engage more deeply with the biological vision community so that there are significant bidirectional contributions?

- Real synergy between the communities that benefits both!

Building the prosopagnosic computer

- For decades, vision researchers have been investigating mechanisms of human face recognition, e.g.:
  - Are faces special? Does face processing have a different neural substrate and processing approach from other visual recognition?
  - *Configural* vs. *holistic* processing (feature-driven vs. whole stimulus integration)
- Evidence from many sources:
  - Neurophysiological case studies
  - Psychophysics
  - Single-cell recording
  - Neuroimaging

Building the prosopagnosic computer

- For decades, vision researchers have been investigating mechanisms of human face recognition, e.g.:
  - Are faces special? Does face processing have a different neural substrate and processing approach from other visual recognition?
  - *Configural* vs. *holistic* processing (feature-driven vs. whole stimulus integration)
- Evidence from many sources:
  - Neurophysiological case studies
  - Psychophysics
  - Single-cell recording
  - Neuroimaging
  - Computational models of face recognition

Building the prosopagnosic computer

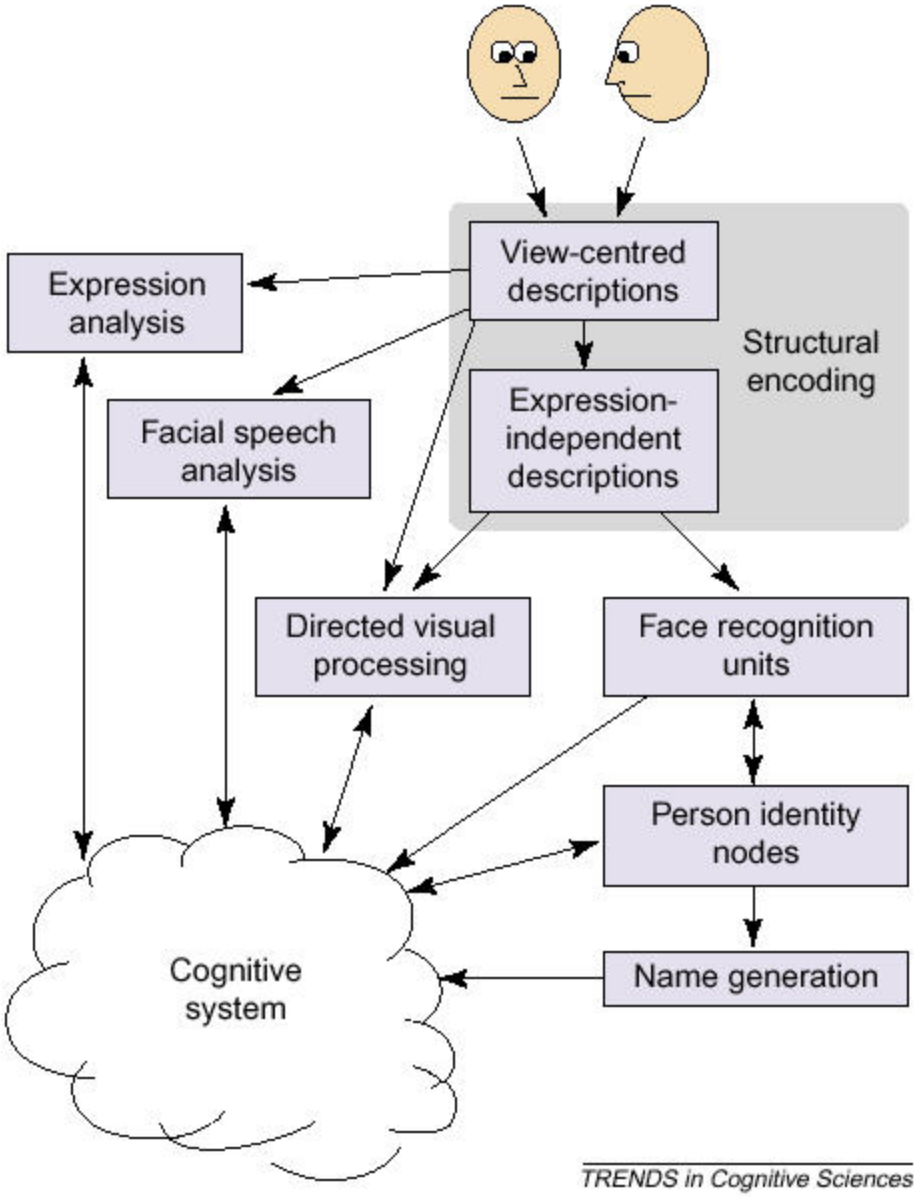
Move beyond the paradigm of

for each particular subproblem

Optimizing individual components of a complex system doesn't lead to overall system optimization!

E.g., face processing is not a set of isolated pattern recognition tasks:

- Face/feature/head detection, tracking, pose estimation
- Face recognition
- Facial expression analysis



Bruce and Young (1986) cognitive model of face processing

Methodological challenges for FGR: #1

Methodological challenges for FGR: #1

Multimodality

Methodological challenges for FGR: #1

Multimodality

- An increased emphasis on multimodal approaches to detection, recognition, and analysis of humans
  - See Thursday's keynote by Louis-Philippe Morency
- Integration with audio/voice, context, other sensing modalities and information channels
- Do people use multiple methods or strategies for body tracking/modeling, gesture recognition, and activity analysis (as we do with recognizing identity)?

Methodological challenges for FGR: #1

Multimodality

- An increased emphasis on multimodal approaches to detection, recognition, and analysis of humans
  - See Thursday's keynote by Louis-Philippe Morency
- Integration with audio/voice, context, other sensing modalities and information channels
- Do people use multiple methods or strategies for body tracking/modeling, gesture recognition, and activity analysis (as we do with recognizing identity)?

Methodological challenges for FGR: #2

A renewed engagement with the biological vision community

- Neuroscience
- Psychophysics
- Computational vision

Methodological challenges for FGR: #2

A renewed engagement with the biological vision community

- Neuroscience
- Psychophysics
- Computational vision

Modeling, experimentation, and computation done in cooperation, aiming both to understand biological systems and to advance computational systems

Methodological challenges for FGR: #2

A renewed engagement with the biological vision community

- Neuroscience
- Psychophysics
- Computational vision

Modeling, experimentation, and computation done in cooperation, aiming both to understand biological systems and to advance computational systems

Methodological challenges for FGR: #3

A renewed engagement with the AI community

Methodological challenges for FGR: #3

A renewed engagement with the AI community

- Face and gesture recognition in the broader contexts of vision and cognition
  - Not just machine learning and neural networks
- FGR in the context of goal-directed and reactive behaviors

Methodological challenges for FGR: #3

A renewed engagement with the AI community

- Face and gesture recognition in the broader contexts of vision and cognition
  - Not just machine learning and neural networks
- FGR in the context of goal-directed and reactive behaviors
- For example:
  - Takeo Kanade has talked about computer vision's "divorce" from AI (as a result of "getting married too early") and the resulting "signal to symbol mismatch"
  - Marvin Minsky: "Early on, computer vision hit up against the problem of knowledge representation, and there are still no good representations for knowledge in vision systems."

Methodological challenges for FGR: #3

A renewed engagement with the AI community

- Face and gesture recognition in the broader contexts of vision and cognition
  - Not just machine learning and neural networks
- FGR in the context of goal-directed and reactive behaviors
- For example:
  - Takeo Kanade has talked about computer vision's "divorce" from AI (as a result of "getting married too early") and the resulting "signal to symbol mismatch"
  - Marvin Minsky: "Early on, computer vision hit up against the problem of knowledge representation, and there are still no good representations for knowledge in vision systems."

Methodological challenges for FGR: #4

Proper use of the term robustness

Methodological challenges for FGR: #4

Proper use of the term robustness

- *Robustness: a characteristic describing the ability of a model, test or system to effectively perform while its variables or assumptions are altered.*

Methodological challenges for FGR: #4

Proper use of the term robustness

- *Robustness: a characteristic describing the ability of a model, test or system to effectively perform while its variables or assumptions are altered.*
- A claim of robustness requires formal (theoretical or experimental) consideration of performance in the face of variation or uncertainty in input variables, assumptions, environmental conditions, etc.

Methodological challenges for FGR: #4

Proper use of the term robustness

- *Robustness: a characteristic describing the ability of a model, test or system to effectively perform while its variables or assumptions are altered.*
- A claim of robustness requires formal (theoretical or experimental) consideration of performance in the face of variation or uncertainty in input variables, assumptions, environmental conditions, etc.
- It does not mean:
  - Low error rate
  - Works well on the data set
  - Works well on multiple data sets
  - Uses a robust statistic or measure

Implications

- Go beyond single “best” method and leverage multimodality – collaborate!
- Learn the neuroscience and perceptual psychology
  - And teach them computational thinking
- Realign CV and AI in building more complex, advanced perception systems

Implications

- Go beyond single “best” method and leverage multimodality – collaborate!
- Learn the neuroscience and perceptual psychology
  - And teach them computational thinking
- Realign CV and AI in building more complex, advanced perception systems
- None of this is meant to downplay or replace current approaches in FGR, but to provide the appropriate context for long-term success
  - To promote the science of FGR as well as the engineering of it

Questions?

http://www.cs.ucsb.edu/~mturk

http://ilab.cs.ucsb.edu