Toward a Human-Like Similarity Measure for Face Recognition

Behrooz Kamgar-Parsi
AI Center, ITD
Naval Research Laboratory
Washington, DC
202-404-7035
kamgar@aic.nrl.navy.mil

Collaborators:

Stephen Krawczyk, NRL and Michigan State University
Edgar Lawson, NRL and George Mason University
Roman Stanchak, NRL and University of Maryland

Behzad Kamgar-Parsi
Office of Naval Research

This research is another effort within the thrust of

FACE RECOGNITION GRAND CHALLENGE

Acknowledgements

- Jonathon Phillips, Army, DARPA, NIST
- Dave Murley, NSA
- ONR
- Naval Research Laboratory

Face Recognition: Human vs Machine Performance

O'Toole and Phillips compared the performance of several algorithms to those of humans

Why the Comparison?

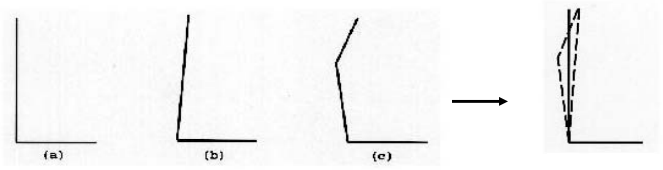
* FRGC approaches have advanced sufficiently to do so
* Nothing else to measure against

How good are humans in Face Recognition?

- How well does the human know the face in question?
- 4 of the 7 algorithms tested by O'Toole and Phillips outperformed humans (on faces unfamiliar to humans)
- **Humans are extremely powerful in recognizing faces that they see often (friends, colleagues, relatives, etc.)**
- The next milestone is to do as well as humans on this category
  - How could we reach this level of performance?
  - Mimic how humans do it
  - Mimic a human similarity measure instead of using a metric similarity measure

Why are metric similarity measures inadequate?

- They "describe" the **dissimilarity** of two visual patterns by a single number.
- Shapes are complex objects in high dimensional spaces.
- Describing their differences by single numbers amounts to projecting them into one-dimensional spaces.
- Hence, they will not remain separated, and get mixed.



Psychophysics Literature has consistently raised questions about using metric similarity measure for shape comparison

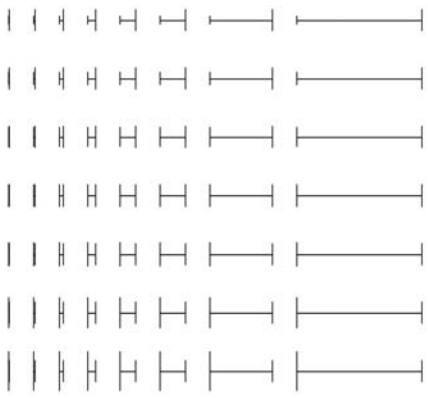
How to deploy a human-like similarity measure

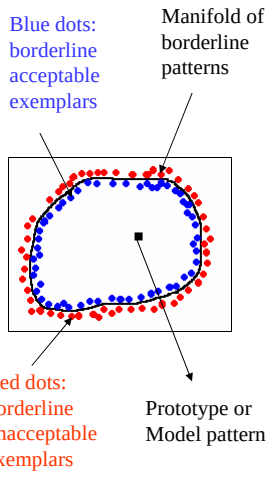
- It is not known how human similarity measure works
- Then, how could we deploy it?
- We propose:
  - Use the Human Eye/Intelligence as Teacher
  - Construct classification boundary in agreement with a human critic

The Eye as Teacher

- OBSERVATION: In the pattern space, according to a human critic, the manifold of borderline patterns encloses the prototype pattern and its visually acceptable matches.

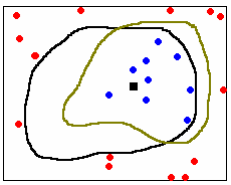
Acceptable instances of H are surrounded by patterns unacceptable as H



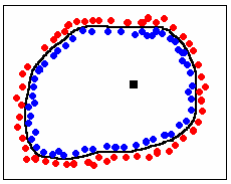


Inadequate real data

- **Typical data**
  - Sparse data
  - Not borderline/useful
  - Best classifier cannot make up for inadequate data
- **Desired data**
  - Dense borderline acceptable & unacceptable examples
  - **Must generate synthetic data**
  - Then, the classification boundary constructed by a neural network will be in good agreement with that of its teacher.



Typical

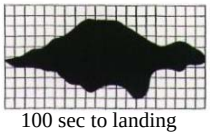


Desired

Automatic Generation and Labeling of Borderline Data

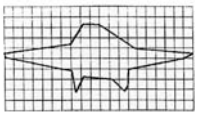
(A previous problem: Detection of Approaching Aircraft)

15 sec to landing



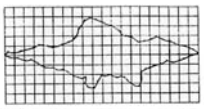
100 sec to landing

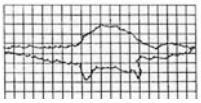
Aircraft prototype →



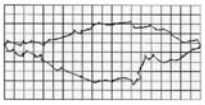
A human expert decides how deformed borderline examples must be

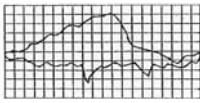
Borderline acceptable →





Borderline unacceptable →

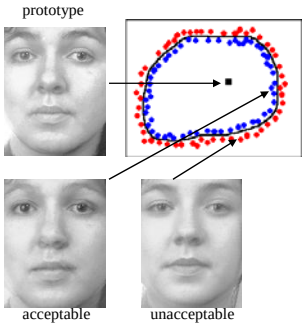




- ''Aircraft Detection: A Case Study in Using Human Similarity Measure,'' Kamgar-Parsi, Behrooz, Kamgar-Parsi, Behzad, Jain, A.K., Dayhoff, J., IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 23, pp. 1404-1414, December 2001.

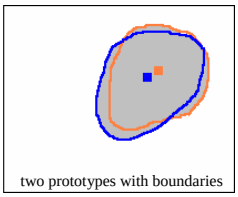
Guided Deformation for Face Recognition

- Deform the image without leaving the space of human faces
- Morph the image of the target person toward images of other people to produce 2 large sets of borderline exemplars:
  - with reduced resemblance to the imaged person, yet recognizable by humans as that person (positive exemplars)
  - look-alikes: 'others' who look almost like the imaged person (negative exemplars)

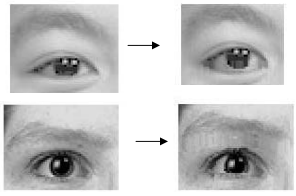


A single image is not enough to represent a face
More than one prototype is needed

- If not enough photos are available, carefully synthesize more prototype images
- For example:
  - Develop a set of intra-person variation operators to generate more prototypes (natural variations and appearances of a given face)



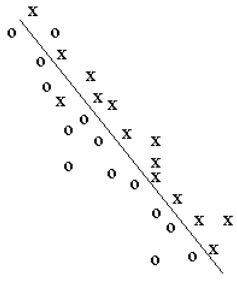
"Synthetic Eyes," Kamgar-Parsi, Behrooz, Kamgar-Parsi, Behzad, Jain, A. K., Proc. 4th International Conference on Audio- and Video-Based Biometric Person Authentication, pp. 412-420, June 2003



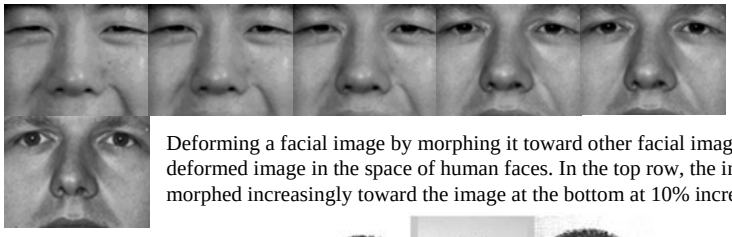
Concluding remarks

- Humans do not rank face/images as a precursor to recognition. Upon encounter, humans simply either recognize the face or reject it as unfamiliar.
- Like humans, our system does not rank images, as it is capable of rejecting a face which is not of interest by simply examining that face, and recognizing those on which it is trained.
- Error rate goes down as the number of images used to morph increases.
- We have a live system. Large-scale tests are waiting the approval of human subject institutional review board.

Not all the generated data has to be accurate



Boundary created by the classifier, with most of the generated exemplars created where they are intended.



Deforming a facial image by morphing it toward other facial images keeps the deformed image in the space of human faces. In the top row, the image at the left is morphed increasingly toward the image at the bottom at 10% increments.