Identifying faces over large changes in illumination: Human vs. Machine Performance

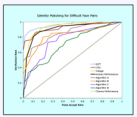
A. J. O'Toole[1], P. J. Phillips[2], J. Dunlop[1], X. An[1], S. Weimer[1] & V. Natu[1]

[1] The University of Texas at Dallas, [2] National Institute of Standards & Technology

Contact: otoole@utdallas.edu

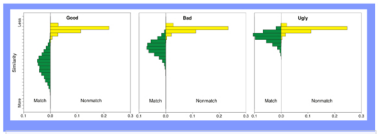
BACKGROUND

- human vs. machine face recognition
  - international face recognition algorithm competitions
    - Face Recognition Vendor Test (FRVT 2006) (Phillips et al., 2010)
    - Multiple Biometric Evaluation (MBE 2009-2011)
  - top performing algorithms in FRGC and FRVT 06
    - better than humans at identity matching in frontal images
      - controlled and uncontrolled illumination (O'Toole et al., 2007)

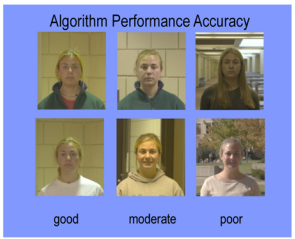


O'Toole et al. (2007)

- Is face recognition from frontal images a solved problem?
  - **No**, recent algorithm test - broad performance range
    - indoor ambient and outdoor natural illumination?
      - *"The Good, Bad, and Ugly Challenge Problem"* (Phillips et al., 2011)



- range of challenge
  - identity pairs constant - only the images change



GOAL

- *When humans and machines disagree, which decision should be considered more reliable ?*

ALGORITHM

- FRVT 2006
  - **fusion of 3 top performing algorithms**
    - algorithms match identity in pairs of images
    - similarity score matrix, where s_{ij}
      - *computed similarity between the i^{th} and j^{th} image*
    - algorithm fusion
      - fused similarity score S_f over algorithms *(k = 1:3)*
      - $S_f = \sum_k (s_k - \text{median}_k)/\text{MAD}_k$
        - *MAD_k mean absolute deviation for the k^{th} algorithm*
  - **stimuli**
    - set 1 - 1085 images of 457 individuals
    - set 2 - 1085 images of same 457 individuals
    - ***similarity matrix*** = all possible pairs of set 1 and 2
      - 3,297 matched identity pairs
      - 1,173,928 non-matched pairs
    - ***constraints***
      - no duplicated images in set 1 and set 2
      - number of images of each person in set 1 and set 2 are equal
      - matched image pairs always taken on different days
    - ***partitioning of scores***
      - similarity scores for matched pairs of each individual ranked
        - good partition - top 1/3
        - moderate partition - middle 1/3
        - poor partition - bottom 1/3

EXPERIMENTS

Exp. 1 - Good, Moderate, Poor

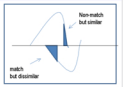
- *conditions*
  - good
  - moderate
  - poor
- *stimuli (per condition)*
  - 40 matched
  - 40 non-matched identity pairs
- *procedure*
  - 2 sec exposures, unlimited time for response
- subjects (*n* = 21)

Exp. 2A & 2B - Moderate & Poor

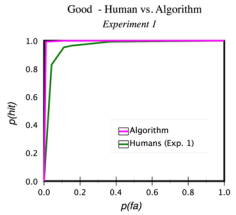
- Exp. 2A - **moderate pairs** (*n* = 23)
- Exp. 2B - **poor pairs** (*n* = 30)
  - *stimuli (per experiment)*
    - 120 matched
    - 120 non-matched

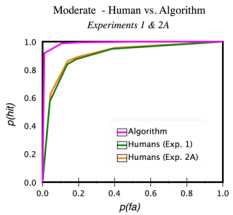
Exp. 3 - Worst Case Scenario

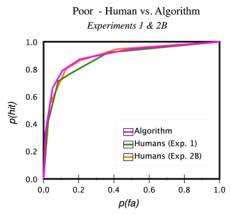
- algorithm performance 100% **INCORRECT**
- similarity score
  - match pairs << non-match pairs
- stimuli
  - 50 matched & 50 non-matched
- participants (*n* = 17)

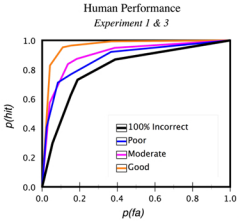


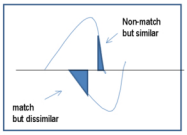
RESULTS

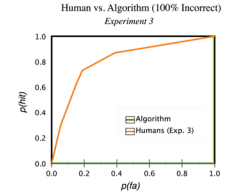


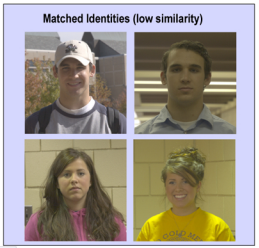


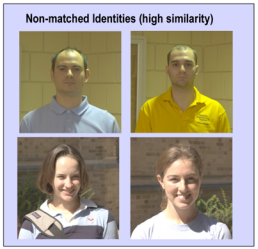












RESULTS SUMMARY

- algorithms more accurate than humans "on average"
- algorithm advantage diminishes as challenge level increases
- **conclusion** - humans still better in the most difficult cases
- ***observation***
  - algorithms make both trivial and non-trivial errors
  - humans only make non-trivial errors

REFERENCES

- Phillips, P. J., Beveridge, J. R. Draper, B.A.Givens, G. O'Toole, A.J., Bolme, D.S. et al. (2011). An Introduction to the good, the bad, and the ugly face recognition challenge problem. *Proceedings of the 9th IEEE Automatic Face and Gesture Recognition Workshop.*
- Phillips, P.J., Scruggs, W. T., O'Toole, A.J., Flynn, P. J., Bowyer, K. W., Schott, C. L., Sharpe, M. (2010). FRVT 2006 and ICE 2006 large-scale experimental results. *IEEE: Transactions on Pattern Analysis and Machine Intelligence, 32*(5), 831-846.
- Phillips, P.J., Flynn, P. J., Beveridge, R., Scruggs, W. T., O'Toole, A.J., Bolme, D., Bowyer, K. W., Givens, G. H., Lui, Y.M., Sahibzada, H. Scallan, J. Weimer, S. (2009). Overview of the Multiple Biometrics Grand Challenge. *Proceedings of the 3rd International Conference on Biometrics*, 2009.
- O'Toole, A. J., Phillips, P. J., Jiang, F., Ayyad, J., Pénard, N., & Abdi, H. (2007). Face recognition algorithms surpass humans matching faces across changes in illumination. *IEEE: Transactions on Pattern Analysis and Machine Intelligence, 29*(9), 1642-1646.

ACKNOWLEDGEMENT

Work funded by the Technical Support Working Group of the DOD