

Human face recognition capabilities - psychological experiment

Mazanec Ján · Informačné technológie, Študentské práce

29.07.2009



This paper deals with the human face recognition abilities. A psychological experiment has been proposed and realized as a web questionnaire with 90 questions. The answers from 142 respondents were analyzed to confirm the hypothesis about human face recognition abilities for different recognition problems and the influence of different

face preprocessing.

The human ability to recognize faces has been studied by Aristotle or Darwin (Turk, 2001) and we still don't know exactly how this mechanism works by humans and after approx. 30 years of intensive research in the automatized face recognition we still cannot overcome the abilities of human (Sinha, et. al., 2006). Generally it is a pattern recognition problem, but face is too complex and changing object (due to the emotions, facial hair, face color change because of temperature or suntan, jewellery, age...).

Face is very important in communication (we use it for determining the sex, age, emotions, character...) and the ability of recognizing other people has been developed since the first evolution stages and can be observed also by animals. By human special brain cells has been found which are active by the "experts" - people who are in concrete domain very sensitive to details (Cottrell, 2004) - This phenomenon also supports the effect, that Asian people look to a European "all the same".

Face recognition is a general ability of humans and is independent of intellect (Lovell, 2002), but there exists "prosopagnosia" (face blindness) - disorder of face perception where the ability to recognize faces is impaired, while the ability to recognize other objects may be relatively intact.

It is important to say, that human use for recognizing other people also other visual features than face (clothes, gait, hair) and also nonvisual (smell, audition).

Even when human can perfectly recognize familiar persons, the ability to recognize identify unfamiliar persons is not perfect. Burton et. al. (1999) proved (with the security camera videos capturing the teachers), that students who knew these persons (have been attending their lessons) were more successful by recognizing the persons than students to which these persons were unfamiliar or even to policemen, by who we can assume that are used to recognize unknown persons.

Other experiments showed that even people who are in daily contact with lots of persons were unable recognize/reject peoples from the photos on the ID cards (Sinha, et. al., 2006; Anderson, 2001).

Results of these and similar experiments should bring us to think about how seriously could we take the eye-witnesses, what is the actual meaning of the security cameras and how much help the photographs on ID cards protect against its misuse. This is the place for automated face recognition systems.

I realized also a psychological experiment (in the form of questionnaire) to research the human face recognition abilities and here. This paper is divided following: In the second chapters are stated hypothesis I would like to confirm. In the third chapter is described the implementation and in the fourth chapter I analyze the results.

Hypothesis

I formulated these hypotheses:

1. *Human are not perfect in recognizing faces* - I would like to show, that the recognition rates of used algorithms are comparable to the human ability to recognize unfamiliar persons.
2. *The ability to correctly assign or refuse some person is relatively low* - I would like to show, that the role of a doorman, which is often mentioned as a possible application of automated face recognition system, is problematic for human.
3. *Human are not always certain by comparing two faces and often mistake* - I would like to show that photos on ID cards is not a good protection against its misuse and human are not always sure when comparing two faces.
4. *People recognise faces also using other features than face* - I would like to prove, that humans use also information from the surrounding of face (head shape, neck, beard, hair...) - not only from the face region (eyes, nose, mouth).

Implementation

The questionnaire has been implemented as a web application (free available at <http://mazanec.eu/face-rec>) using PHP with MySQL database for collecting the answers.

Respondent first read the instruction, filled basic data about himself and then continued to the questions. There was one question per page and should be answered by clicking on the pictures. At the end there was possibility to leave a note.

The questions were made of images taken from the FERET database (2001) with 3 types of preprocessing (see Fig. 1):

- *original* - no preprocessing, just downsized to 150px height.
- *“face”* preprocessing - face crop, downsized to 65x75px and then resized to 130x150px
- *“BIGface”* preprocessing - differs from “face” by leaving more surrounding of the face (the cropping ellipse is larger)

Using these 3 types of pictures I wanted to support hypothesis Nr. 4.



Fig. 1: Example of original image, image after "face" preprocessing and after "BIGface" preprocessing

There were 3 types of questions:

- Type 1 - Assign 1 face to one of 4 faces (supporting hypothesis 1) - see Fig. 2.
- Type 2 - Compare 2 faces - in 50% of the questions it were the same persons, in the rest 50% it were not. Possible answers were "yes", "rather yes", "rather no" and "no" (supporting hypothesis 3) - see Fig. 3.
- Type 3 - Assign 1 face to one of 3 faces or reject - in 50% of questions the right answer was one of the faces, in the rest 50% it was the "none from offered" (supporting hypothesis 2) - see Fig. 4.



Fig. 2: Example of question type 1



Fig. 3: Example of question type 2



Fig. 4: Example of question type 3

There were 9 combinations of type of preprocessing and type of question – for each I made 10 questions – so 90 question pages together.

According to my previous experiments with PCA and LDA for recognizing faces (Mazanec, 2008) I picked out these pictures:

- for question of type 1 I chosen the mostly incorrect recognized subjects and for the 3 possible answers I picked up the pictures, which were mostly incorrectly recognized
- for questions of type 2 I picked up mostly incorrectly recognized couples of images
- questions of type 3 I made up from most incorrectly recognized images and 2 (3) other mostly incorrectly recognized with them

The questions were sorted following: first questions with images without preprocessing, then images with “face” preprocessing and last with “BIGface” preprocessing. First there were 10 questions of type 1, than 10 questions of type 2 and last 10 question of type 3 for every image type. Expected time for the whole questionnaire was 10 minutes – which I find acceptable for the respondents. At the end, the respondent has been informed about his total percentage result as a motivation factor.

Results

I asked my family members and friends to fill the questionnaire and to pass it to their friends – so the answers were from quite a small circle of people and there have been only a few answers which could be considered for untrustworthy or random.

The questionnaire was online approximately one week and totally approx 150 answers were collected. After elimination of

- incomplete answers
- answers with suspicious low score (under 50%, when the average was above 80%)
- answers where the respondent filled, that he just “quickly clicked through”
- multiple answers from the same respondent (based on the IP, time or note)

There were 142 complete answers: 61 from women and 81 from men with average age 25.8 years. Only 4.23% respondents mentioned they have already seen the FERET database – so for most it were unknown faces.

The average score was 82.15% of correct answers. Graph of incorrect answers is on

the Fig. 5 and the results for every category of images in the Table 1. The average time to fill the questionnaire was 11minutes and the average time for 1 question was 7seconds.

Preprocessing	Type of question			FAR	FRR
	type1	type2	type3		
original	98,24%	88,52%	87,82%	5,07%	7,11%
“face”	79,01%	69,86%	64,86%	21,20%	13,94
“BIGface”	89,58%	84,93%	81,06%	12,46%	6,48%

Table 1: Summary of results

From the results we can deduce these conclusions:

- Hypothesis 1 has been accepted - human cannot perfectly recognize unknown faces - 100% success has been achieved only by 3 questions from the first 10 questions (images without preprocessing, question type 1)
- By comparing the results with different type of preprocessing the hypothesis 4 has been accepted - humans recognize faces not only according to the basic face region (ellipse including eyes, mouth and nose). Best result has been achieved with images without preprocessing and the worse with the “face” preprocessing. Also some respondents said the “face” images were most difficult to recognize.
- After analyzing the questions of type 2, where the respondent could select by comparing 2 faces also the uncertain answers (“rather yes”, “rather no”), the hypothesis 3 is accepted - people are not always sure and mistake often. Interesting is, that there were more uncertain answers by those questions, where the overall score was worse (see Fig. 6)
- Hypothesis 2 is accepted when analyzing the results of questions of type 3 (see Tab. 1), where the computed FAR (false accept rate) and FRR (false refuse rate) even for original images preprocessing are above 5% (present-day system achieve scores under 1% (Phillips, et. al., 2007). By these questions the results are generally worst - we can state that this type of problem is the most difficult for people - but also the most common - e.g. by confirming someone’s identity from an ID card.
- Analyzing the relation between the time needed to answer a question and the average correctness of the answers, could be stated that people need more time to answer more problematic questions (where the recognition results were worse)

It needs to be said, that the percentage scores just imply the conclusions and is not possible to compare it with results from an automated face recognition system as first intended to by setting the hypothesis 1, because just the most problematic images were chosen for the questionnaire. Generally I doubt is possible to create conditions when we could objectively compare face recognition rates of human and computer.

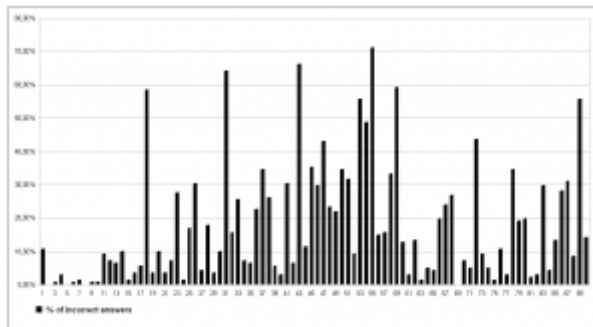


Fig. 5: Percent of incorrect answers for every question

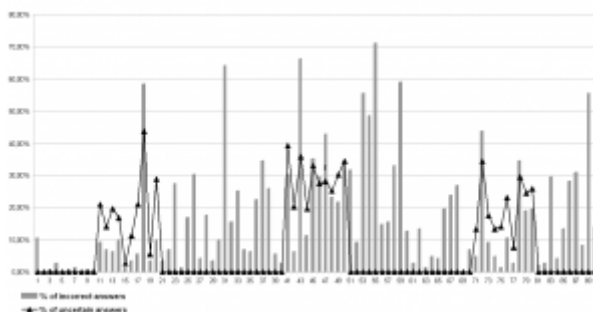


Fig. 6: Percents of uncertain answers ("rather yes", "rather no") with the graph of correct incorrect answers.

Conclusions

In this experiment I wanted to prove, that the human face recognition capabilities are not perfect and show that there is place for automated face recognition systems (e.g. automated doorman). Secondly, that humans recognize face not only according to area of eyes, mouth and nose, but they use more information also from surrounding of the face - this needs to be taken into account by preprocessing of images for face recognition systems.

References

1. Anderson, R., Security Engineering. WILEY, 2001.
2. Burton, A. M., S. Wilson, M. Cowan, and V. Bruce (1999), Face recognition in poorquality video: Evidence from security surveillance., Psychological Science, vol. 10, pp. 243-248
3. Cottrell, G.W. (2004), What can computational models tell us about face processing?, Lecture, "Introduction to Cognitive Science" course, Cognitive Science Department, UC San Diego, USA
4. FERET Database, <http://www.itl.nist.gov/iad/humanid/feret/>, NIST, 2001.
5. Lovell, G. (2002), Face Recognition, Tutorial Handouts, "Cognitive Psychology" Course, University of Stirling, UK
6. Mazanec, J. (2008) Face Recognition in Biometrics Based on PCA and SVM Methods, Diploma Thesis, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Bratislava
7. Phillips, P. J., W. T. Scruggs, A. J. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott and M. Sharpe, (2007) FRVT 2006 and ICE 2006 Large-Scale Results
8. Sinha, P., B. Balas, Y. Ostrovsky, and R. Russel (2006), Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About, Proceedings of

-
- the IEEE, Vol. 94, No. 11,
9. Turk, M. (2001), A Random Walk through Eigenspace, IEICE TRANS. INF. & SYST., vol. E84-D, pp. 1586-1595.
-

This paper was published at 11th Conference of Doctoral Students, ELITECH '09.
