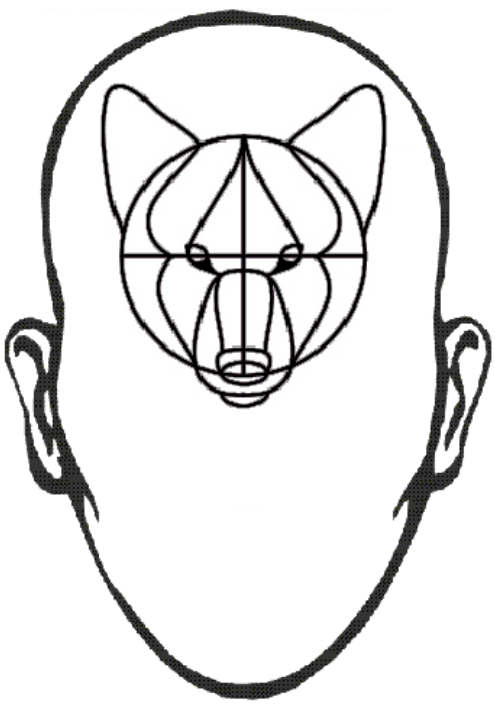

How well can we perceive animal faces compared to human faces?

DT2350: Human Perception for Information Technology

KTH, *2014-10-19*

Jesper Norberg jenor@kth.se

Vladimir Grozman grozman@kth.se

Supervisor: Roberto Bresin

Abstract

In this project we set out to determine whether us humans' facial recognition system extends to other animals' faces as well. If this was the case we reasoned, then the speed for identifying the faces of said animals in an image should be comparable to the speed for identifying the faces of humans in images of similar setting. To measure this, we created a testing software in which we measured test participants' speed in identifying human faces and animal faces. Our results show that there is a significant difference in identification speed between the faces of humans and the faces of other animals, with human faces being perceived much faster.

Contents

1 Introduction

1.1 Background

During a lecture in this course, "Introduction to vision, including perception of objects and scenes", the perception of faces was discussed. Compared to other animals, we humans have a very advanced facial recognition system, capable of recognizing even blurred faces, a feat computers find very difficult to achieve. We have neurons responding specifically to faces, and even areas such as the fusiform face area, where these specialised neurons are clustered.[1] Facial recognition also happens incredibly fast compared to recognition of other objects. These types of observations strengthen the idea that faces are on a level of their own in our recognition system. Some research even suggests that we are born with information regarding the structure of faces.[2] We then proposed a question to the lecturer; could our special perception of faces be extended to other animals' faces as well? The teacher wasn't sure, and so we decided to check the existing research. Though we found that there have been software implementations of animal facial recognition[3], as well as several studies of various animals' ability to recognize human faces, we were unable to find any research regarding whether our specialized facial recognition applies to other animals' faces as well. We thought this would be an interesting question to pursue.

1.2 Goal

In this project, we aimed to gain insight in the question of whether the human specialization in facial recognition extends to other animals' faces as well. If this was the case then we reasoned that the speed for identifying the faces of said animals in an image should be comparable to the speed for identifying the faces of humans in images of similar setting. We decided to limit the scope of the study to only three other animals; Bolivian Titi monkeys, Green Tree snakes, and German Shepherd dogs. We decided on these three animals based on the amount of suitable images we were able to acquire, as well as on each of them being typical examples of their respective common groupings; monkeys, snakes and dogs. To further narrow the scope of the study, we decided to convert the images to grey-scale. For test participants we decided to use our fellow students at KTH.

1.3 Acknowledgements

We worked together in the development phase of the project. Vladimir Grozman then wrote the testing program, and Jesper Norberg found and edited the images used in the project, as well as conducted the tests and performed the statistical analysis.

2 Method

2.1 The images

Figure 1: Example of large-face image, here with a titi monkey.

To establish how us humans' ability to identify human faces compares to our ability to identify other animals' faces, we needed to test this ability on several humans, ie. conduct a study. In order to create this study we needed to decide on what core attributes we wanted to measure. Since the scope of this project did not allow for larger scale studies, we had to limit the test in order to provide qualitative data despite having a small testing group. The only variable we chose to measure was as such identification time; how long it took the test subject to correctly identify the location of the face when displayed a picture containing an animal. When designing facial recognition software, the two core attributes are time and correctness,[4] so we thought this to be a fair way to evaluate our facial recognition capabilities. Next came deciding on what pictures to use and how said pictures should be presented to the test participants in order to optimize unbiased and valid results. We decided that since the scope of the study was so small, we should narrow it down to just a few species, and only use grey-scale, to reduce complexity. We decided on three species beside humans; Bolivian Titi monkeys (Callicebus donacophilus), Green Tree snakes (Dendrelaphis punctulata), and Domestic dogs (Canis lupus familiaris), specifically the breed German Shepherd. The species/breeds were decided based on the amount of suitable images we were able to acquire, as well on each of them being typical examples of their commonly used groupings, ie. monkeys, snakes and dogs respectively.

Figure 2: Example of small-face image, here with a human.

Figure 3: Example of no-face image, here with a green tree snake.

To create some regularity, we used three types of image settings; large face (Figure 1), small face (Figure 2), and no face (Figure 3). For the large face images, the face took up about 6% of the image. For the small face images, the face took up about 1.5% of the image. On the images with no face at all, at least part of the body of the respective animal could be seen, but the face

was not in the image. We thought that the different image types would make the test participants more focused on the task at hand, as well as provide some diversity in the collected data. We spread out the positions of the faces in the images, so that it was roughly random where in the image the face would be. This was done so that the test participants would not be able to decide where to place their gaze beforehand.

All the images were adjusted to the format 16:9, which is the standard on many screens and laptops, including the one used in the test. This enabled us to present the images in fullscreen without the images being stretched or cropped. In total, 60 images were created, 5 per animal type and category. Our images primarily came from Google Images[5] and Flickr[6]. All image editing was done in Adobe Photoshop CS4[7].

2.2 The testing software

Besides the images, we needed some software in order to conduct the test. We wrote a customized software for running the tests as well as measuring the relevant parameters in Javascript and PHP. We specified the face area for each image in the software, which could then identify if the face had been correctly identified. Besides reading and interpreting input, the program was also able to shuffle the images for each test. In order to record the test participants data, we created a local MySQL server using XAMPP[8], in which we recorded all the test data in SQL. This data could then be extracted to Excel where we made our statistical analysis. The time taken was recorded in milliseconds.

2.3 The test

For the test, we decided to use a laptop with touch screen in order to make the interaction with the images more intuitive; the test participants could simply "point" in order to click on the screen. There were two different types of behaviours expected from the test participants. The first was if there was a face in the image. In this case they should click (press) where on the screen the face was. The X/Y coordinates of the click were recorded and matched against the predefined face area (see 2.2) for correctness. The second type of behaviour was when there was no face in the image. The test participants were then supposed to press the SPACE key on the laptop keyboard. All the 60 images were displayed in each test, in a random order.

When the test program started it displayed instructions for the test. In retrospect, we should probably have added a warning for snakes, just in case a test participant would have suffered from ophiophobia. Luckily we had no incidents in this project though. After reading the instructions the test participants were to click the screen to begin the test. A timer would then appear in the middle, drawing the test participants gaze for two seconds, after which a random image would appear. Depending on the image type, the test participant would then either click the face or press the SPACE key if there was no face in the image. Regardless of reaction or correctness, any input would hide the image,

with the timer appearing again. Our goal with this timer was to make the test participants relax their gaze in the middle of the screen before each image, to facilitate optimal testing conditions. The test took roughly 3 minutes, and was over when all 60 images had been processed by the test participant.

For test participants we used our fellow KTH students. In total 10 students participated, all completing the test in its entirety. Due to our budget no compensations were handed out for participation.

3 Results

From the results we removed all incorrect answers, and also all the extreme values where time taken was over 4 seconds. In total these removals amounted to less than 5% of the test data. On the remaining data we did a basic data analysis (Figure 4, Figure 5), as well as performed ANOVA (Analysis of Variance) between relevant categories. We used two different alpha values for our ANOVA; $\alpha = 0.05$ which we consider to show a slight difference, and $\alpha = 0.001$ which denotes a significant difference.

For images with faces, we found significant differences between humans and all other animals. The implication is very clear when you look at Figure 4. It seems we are much better at recognizing humans than the other animals present in the study. There is also a slight difference between dogs and snakes, as well as between monkeys and snakes. We found no difference between dogs and monkeys however.

For images without faces, we found no differences at all, which can be hinted in Figure 5.

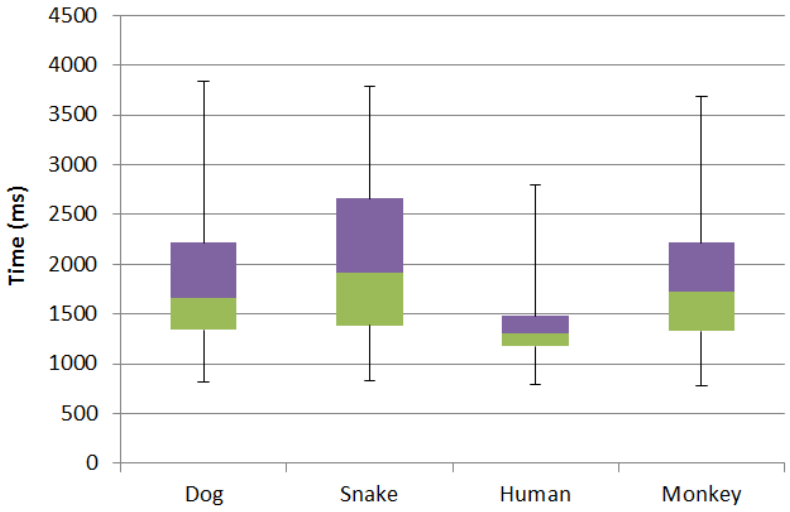

Figure 4: Box plot of the correctly identified images with faces.

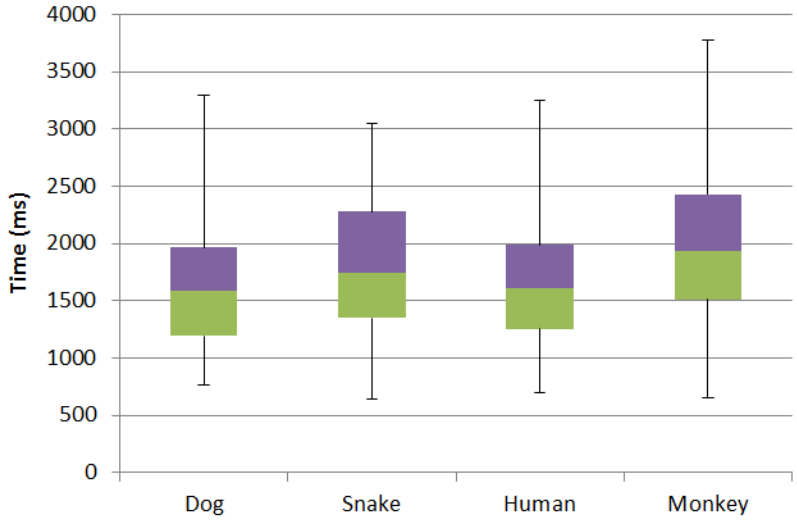

Figure 5: Box plot of the correctly identified images without faces.

4 Discussion

Though the test data was small, each student generated 60 data points, which gave us more than enough data for reliable results. Our results clearly show that there is a significant difference in speed between our ability to recognize human faces and other animals' faces. Given that our choices of images, animals and testing method were valid, we can consider our initial question answered.

Our results also indicate that we are about as good at recognizing dogs and monkeys, whilst slightly worse at identifying snakes. We hypothesize that this has much to do with mammals bearing a far closer resemblence to us than reptiles do. This is in line with existing science, as previous studies show that face recognition involves computing distinctive features and comparing them to the average of previously observed exemplars.[9]

One interesting aspect of our results is the fact that there from our data seems to be little if any difference in how fast we dismiss there being a face in an image depending on the animal body present in it. This suggests that we have no specialized functions for exclusion, merely for identification.

There is always room for more research, and suggestions for further studies include doing a more ambitious study on how we humans respond to various animals' faces. Things that we left out but are of value include eye movement, neural activity, and colors. Previous research suggests that familiarity plays a huge role.[10] As such, a dog owner would be significantly better at identifying dogs' faces than someone not used to animals. This also to some extent explains the rather vast spread in our results for the various animals, whilst the first three

quartiles for humans are very compact. The fact that all the test participants were social people used to meeting people may have played a huge role in this result.

5 Conclusion

It seems that we humans are significantly better at recognizing the faces of other humans than the faces of other species. It seems we are also, to a lesser extent, better at recognizing the faces of other mammals rather than the faces of reptiles. This is likely because the former are closer to human faces. We also seem to be equally good at ruling out there being a face in an image regardless of species. This suggests that our facial specialization does not extend to exclusion, only inclusion.

We managed to answer our question, and have hopefully contributed to a very interesting field of research. As such we consider the project a great success.

6 References

[1] N. Kanwisher, *Domain specificity in face perception*, Nature, 3, 759 (2000)

[2] E. Bruce Goldstein, *Sensation and perception*, 8th ed, chapter 5 (2010)

[3] PetRecognition. http://www.petrecognition.com/ (2014-10-19)

[4] P. Sinha, B. Balas, et al., *Face recognition by Humans*, IEEE, 94, 1948 (2006)

[5] Google Images. https://www.google.se/imghp (2014-10-19)

[6] Flickr. https://www.flickr.com/ (2014-10-19)

[7] Photoshop. http://www.adobe.com/products/photoshop.html (2014-10-19)

[8] XAMPP. https://www.apachefriends.org/index.html (2014-10-19)

[9] G. Rhodes, I.G. McLean, Perception 19, 773 (1990)

[10] I. Gauthier, et al. Nature Neuroscience 3, 191 (2000).