HUMAN ROBOT INTERACTIONS : HEART, EYE GAZE, HEAD POSE

Enes ŞANLI (Middle East Technical University, Ankara, Turkey)

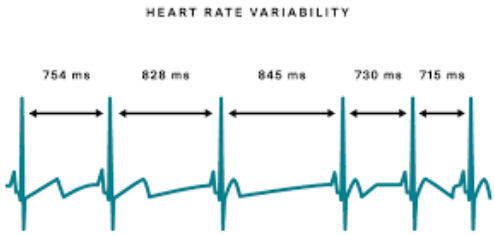

Figure 1: Heart rate variability

The aim of this research was to find out whether and how heart, eye gaze and head pose could be used to develop HRI. You can click on the link to see how some of the codes work in real life: Link

1 Introduction

Human-Robot Interaction (HRI) is a field of study dedicated to understanding, designing, and evaluating robotic systems for use by or with humans. Interaction, by definition, requires communication between robots and humans.[1] In this research, human emotion's effects on heart, eye gaze and head pose are investigated. Upon observing these emotions, it is tried to enhance the interaction between humans and robots. This is done by setting the robot's responses according to human emotions.

2 Heart

Researchers investigated the effects of emotions on heart according to two variables. These are Heart Rate Variability (HRV) and Heart Rate (HR). Heart rate variability is where the amount of time between your heartbeats fluctuates slightly. Heart rate is the number of times someone's heart beats per minute (bpm). After the researchers had done investigations on four different emotions that are amusement, fear, anger and neutral, they found out that HRV and HR show different results. When in HR, being amused can be detected significantly, the others seem to stay still with not much of a difference. However, when HRV is used to detect the levels of emotions, it is observed that even though neutral is not significantly different, the other three feelings are detected easily.[2]

2.1 Observing Heart by Camera

Researchers tried to detect HRV and HR by using a normal camera, such as a phone's or a computer. This article investigates two different ways that are used to achieve this aim.

1. The Change in the Color of Face
2. Head Movements

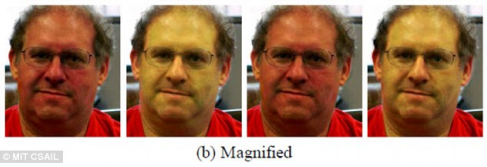

Figure 2: Color Change

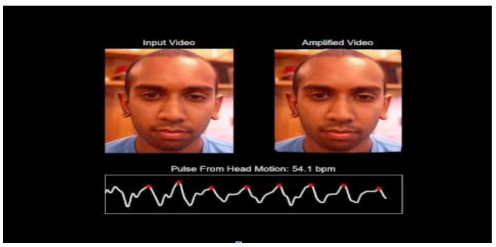

Figure 3: Head Movements

In both of these ways, Eulerian Video Magnification was used. Eulerian video magnification is a set of simple and robust algorithms that can reveal and analyze tiny motions. It is a new type of microscope, not made of optics, but of software taking an ordinary video as input and producing one in which the temporal changes are larger.[3]

2.1.1 The Change in the Color of Face

In this method, the changings in the color of the face caused by the circulation of the blood were observed. Each changing circle in color accounts for one beat and as a result of this HR and HRV could be measured. However, for this to happen, the person who was being observed had to stay as still as possible.

2.1.2 Head Movements

In this method, the head movement that is caused by the force of blood that was pumped by the heart was observed. Thanks to Eulerian Video Magnification, the observability of this movement was made easier. The benefit of this according to the first method is that even if a person has a mask on or if we see the back of someone's head, we can see the results nonetheless. On the other hand, the movements done by the observed person affect the results more than the first method.

2.2 Codes

The codes tried for heartbeat detection via the camera are as follows. But most of these codes did not work correctly when tried.

1. https://github.com/irfan798/head-pulse-track
2. https://github.com/Anirudh0707/Pulse_Detection_From_Videos
3. https://github.com/thearn/webcam-pulse-detector
4. https://github.com/prouast/heartbeat
5. https://github.com/habom2310/Heart-rate-measurement-using-cam
6. https://github.com/xliucs/MTTS-CAN

There are applications that work as well. However, regardless of the situation, the head of the examined person should stay as steady as possible.

3 Eye Gaze and Head Pose

By examining the eye gaze, blink rate and head poses of people it is possible to understand what they pay attention to and how they feel. But in this research, the fundamental aim is to find out what people pay attention to as understanding how they feel is harder. So, we will try to understand what object people are looking at. While doing this, we will compare head pose and eye gaze. Because the direction at which people are looking does not always show what they exactly look at compared to eye.

3.1 Detecting Eye Gaze and Head Pose

In order to find the direction of the eye, researchers used different algorithms such as CNN-based models.[4] However, as the eye shapes of people are different, there are angular errors in both eye and head directions. In addition to that, as the distance between the camera and people increases; it gets harder to detect the eye and the head which results in the enhancement in the angular error.

3.2 Codes

While trying to find out the accuracy of the codes they trained and tested, they used different databases. In this research, we tried to use the codes that has the least angular error rate even though they are in different databases. As some of the codes are too old or the environment which they were tried are not like ours we encountered some problems. The codes are as the following:

1. https://github.com/NVlabs/few_shot_gaze
2. https://github.com/Ahmednull/L2CS-Net
3. https://github.com/thohemp/6DRepNet
4. https://github.com/Tobias-Fischer/rt_gene
5. https://www.youtube.com/watch?v=-TVUwH1PgBs

4 Conclusion

Among the non-verbal-ques that we investigated, heart is not suitable for our research for now because of the problems mentioned above in detecting heart rate and the few number of researches in this area. At the same time, eye gaze and head poses are not good enough to understand people's emotions because of the error rate we have. Nonetheless, it is possible to understand the objects people look at and to what they pay attention by considering their eye gaze and head poses. So it seems possible to advance in this research.

References

[1] A. Sameer, "Human robot interaction (hri)," 04 2016.

[2] Y. Wu, R. Gu, Q. Yang, and Y.-J. Luo, "How do amusement, anger and fear influence heart rate and heart rate variability?," *Frontiers in Neuroscience*, vol. 13, p. 1131, 10 2019.

[3] H. Shahadi, H. Albatat, Z. Jabbar Al-allaq, A. Thahab, I. Haider, and Shahadi, "Eulerian video magnification: a review," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, pp. 799–811, 05 2020.

[4] A. Abdelrahman, T. Hempel, A. Khalifa, and A. Al-Hamadi, "L2cs-net: Fine-grained gaze estimation in unconstrained environments," 03 2022.

Enes Şanlı [sanli.enes@metu.edu.tr] Computer Engineering student at Middle East Technical University.