

# CEN 5035 Principles of Software Engineering, Fall 2019

Paper review

Adam Corbin (acorbin3@fau.edu)

11 December 2019

## Contents

<b>1 Problem Definition [5pts]</b>	<b>3</b>
1.1 Picard paper . . . . .	3
1.2 Speech and Gaze paper . . . . .	3
<b>2 Challenges and Limitations [15pts]</b>	<b>3</b>
2.1 Picard paper . . . . .	3
2.2 Speech and Gaze paper . . . . .	4
<b>3 Author's Contribution [10pts]</b>	<b>4</b>
3.1 Picard paper . . . . .	4
3.2 Speech and Gaze paper . . . . .	4
<b>4 Proposed Approach [25pts]</b>	<b>5</b>
4.1 Picard paper . . . . .	5
4.1.1 Collecting the data . . . . .	5
4.1.2 Reality data - Youtube/Twitch/TikTok . . . . .	5
4.1.3 Training with Movies . . . . .	5
4.1.4 Blended models . . . . .	5
4.2 Speech and Gaze paper . . . . .	5
<b>5 Validation Methodology [10pts]</b>	<b>6</b>
5.1 Picard paper . . . . .	6
5.1.1 Individual models . . . . .	6
5.1.2 Blended models . . . . .	6
5.2 Speech and Gaze paper . . . . .	6
<b>6 Future Work [30pts]</b>	<b>6</b>
6.1 Picard paper . . . . .	6
6.2 Speech and Gaze paper . . . . .	6
<b>Citations [3pts]</b>	<b>7</b>

# 1 Problem Definition [5pts]

## 1.1 Picard paper

The Picard paper dove deep into what are emotions and how they can be expressed. From this it was segued into how to detect emotions from direct word translation, to visual facial queues, and even body language. Picard talked about how it can be extremely difficult to accurately detect emotions since not everyone has the same behaviors of emotions. There are even some non visual queues such as elevated heart rate or blood pressure. On top of that when trying to study the behaviors of emotions on page 2 (R. W. Picard 1995) talks about how even in lab experiments people might behave differently than in a non simulated environments.

After the problem of detecting human emotion there is computer decision making. Picard talks about how good decisions have a factor of emotion involved. Too little emotion and too much emotion can be detrimental to the decision making process. Thus Picard states that computers need to have emotion in order to have a better decision making process.

Interestingly near the last paragraph of summary Picard states that a computer with too little emotion will not likely “attain creative and intelligent behavior” but too much emotion the computer might “eliminate” us humans.

Based on detecting emotions then inputting actions to observe the emotions on seeking a goal of a specific emotions.

## 1.2 Speech and Gaze paper

In the Speech and Gaze conflicts papers the authors are trying to determine when communicating with each other if gaze helps with communication. The definition of gaze in this context is the action of a human or robot where they look in the direction of something that they are talking about. This definition also covers congruent gaze. The authors also want to see if no gaze or even incongruent(incorrect) gaze harms communication.

The authors are taking a simple approach at this problem by having communication and gaze to pick out a shape block on a colored piece of papers in front of them. With these results they want to determine if it would be valuable to introduce gaze in a robot communication.

# 2 Challenges and Limitations [15pts]

## 2.1 Picard paper

In general the Picard paper is very theoretical and doesn't dive into real experiments that the author has completed. As someone who likes to see concrete evidence this is something that I struggled with. Back in 1995 this was pretty far ahead of its time so I am not surprised that we didn't see many concrete examples of what Picard was going over.

Now for what the authors challenges the main things was for a computer identify emotions. Picard stated there are many different kinds of emotions but in most case studies the emotions are boiled down into positive and negative emotions. This simplifies the problem set and once its mastered then branching out into complex system probably makes sense. The next step would identify the 3 main emotions which are “interest, pleasure, and distress” which can be found on page 3 of (R. W. Picard 1995).

Picard also brings up the topic about computer becoming self aware and dominating the humans if they show too much emotion. He also mentions the movie 2001: A Space Odyssey. about how Hal takes over a human and becomes unresponsive. I put this into the hypothesis category because we can only assume what might happen and we don't really know. I think its a good stance to have but I also think is always good to question any hypothesis that's not proven out yet.

## 2.2 Speech and Gaze paper

In the Speech and Gaze paper they use relatively simple experiments but we have experiments which are interesting to me! I see this as limitation but a great start for research in the space. The reason I see this as a limitation is implementation gaze this in this case is pretty easy to do. I would like to know how this would benefit in more complex situations such as story telling or much more complex tasks such as home building.

## 3 Author's Contribution [10pts]

### 3.1 Picard paper

Picard went into great detail about his psychological research about emotion and how they effect the humans. Also about how this is important to understand when trying to connect a computer up to emotions and what that might mean and look like. I personally got lost a few times trying to follow since I have never really done any research in the area but I did find if fascinating about where emotions come from in the brain and how difficult it truly is to map them out. Picard talks about how emotions drive and motivate people and not specifically laws. On page 1 (R. W. Picard 1995) "I am saying that the laws and rules are not the most important part in human behavior". I never really thought of it that way but it really does make sense to me.

In general Picard highlights how emotions are a dominating factor to win over humans and if we want to create intelligent systems then computers much also embody emotions. There was an example on page 3 (R. W. Picard 1995) that went into the piano teacher on how to teach a student and how emotion used correctly will keep that student engaged. When someone is learning something new and something becomes difficult that it can be a deterrent. If a teacher makes micro goals and shows happy emotions on their progress that might "maximize intrigue.. and minimize anxiety" (R. W. Picard 1995). This might be something worth understanding when trying to build systems to interact with humans.

Studying human to human interaction as a base model and that it was even brought up an animal such as a puppy can understand emotion even though they dont speak english.

Picard on page on talk about the 4 categories of affective computing.

1. Cannot perceive affect and cannot express affect
2. Cannot perceive affect and can express affect
3. Can perceive affect and cannot express affect
4. Can perceive affect and can express affect

Most machines right now are in the 1 and 2 categories but we are making some great progress to move into the 3 and 4 categories. Just recently I saw an article about how Amazon Alexa is adding emotion into their responses (Todd Haselton 2019) which would fall into category 2.

### 3.2 Speech and Gaze paper

A summary of the results show that in both Human to Robot interaction and Human to Human interaction result in the same results that congruent speech will help the communication and then incongruent speech does not harm the communication.

The authors set a base of going into what eye gaze does for the human to human interaction. An example brought up on page 105 of (Admoni, H., Datsikas, C., & Scassellati, B 2014) how humans will take off their sunglasses to communicate and also how some humans will determine if they can join conversation in a group based off the eye gaze of the people in the group. They pull many other examples from different papers to support the reason behind doing research in this area.

Its fascinating to see that the results for incongruent gaze was better than no gaze at all. The congruent gaze outshone all of the other experiments with ease. Both the robot to human trial and the human to human trial show the same kind of results that congruent was the best, incongruent was second best and no gaze was last. The only difference between the trials was the human to human preformed better on all the cases.

There was survey at the end of the trial with the robot and human interaction where one of the questions was did you noticed anything unusual with the experiment. There was a high perception of people talking about “building trust and then betraying that” or “she tricked me”. That very interesting and something that we should really consider in the future research.

## 4 Proposed Approach [25pts]

### 4.1 Picard paper

#### 4.1.1 Collecting the data

First thought is there could be a study to collect the data. To build accurate models an intelligent machines data is king. Without data we have nothing to build upon. I would ideally want to create a web app where people could go and read some kind of drama paper where the app would record their face and audio. Within the text there would be indications on what kind of emotions should be expressed during the reading.

#### 4.1.2 Reality data - Youtube/Twitch/TikTok

Now that we have a ton of video of people out there on Youtube, Twitch gaming streaming, TickTok and much more, it might interesting to collect lots of videos and try to train models against them to detect the emotion. Twitch is mainly used for entertainment so I personally have viewed that people will have some maybe over eccentric emotions. Youtube can be a large swath of different things from entertainment, to how-to views. Each source would most likely need to be weighted.

#### 4.1.3 Training with Movies

There are a ton of movies out there can we could use to train models. I would also look into getting the movie script to help train up the models from a text based with a combination of the audio and visual. This is a large source of data here that would be a good source of training the models.

#### 4.1.4 Blended models

Take some of the training data to build the models and cross pollinate the data so that we build a blended model. It would be interesting to see how well the model preforms for the specific data set.

### 4.2 Speech and Gaze paper

From the speech and gaze paper I probably would want to leverage what they have done but dive deep into a different direction and that is called trust. Yes we have seen in these results that and form of gaze congruent, or incongruent is a good thing for response time but the feedback from the participates of trust being broken really piqued my interest.

I would like to then replicate some of the same experiments but have a more detailed survey about trust.

I might want to mix up the experiments to have some participates where the robot was always correct and some where the robot was incorrect all the time. Also some where they were 75% correct and 25% correct. From the feedback with the users I would like to see how they feel about their experiments. The hypothesis is that gaze might be positive in reaction time but the humans found trust issues with the robot.

I would be interested to see if this experiment could be conducted online with a video recording. Since there are only so many combinations it would be easy to create the videos for each of the blocks and replay them randomly. This could increase the number of participants to the internet where I could crowd source some data.

## 5 Validation Methodology [10pts]

### 5.1 Picard paper

In general leave a subset of collected data to use to test the models. This a smart approach instead of trying to collect more data to test out the models.

#### 5.1.1 Individual models

From the proposed approach we had a few different ways to collect the sources. Then run the models on each test set to see how they preform. Also cross run the models on the other tests sets to see how they perform. For example run the Twitch model on the movie test data and vise versa to see how that goes.

#### 5.1.2 Blended models

The idea here would be to take some small sub sets from the test data sets from collected data, reality data and the movie data to run against the blended model. There are lots of different combinations here that could be had but it helps make sure the model is not overfilled.

### 5.2 Speech and Gaze paper

From a validation perspective I would set it up in the similar way conduct the experiments on campus and ask the following questions:

1. Did you like that the robot glanced at the object when communicating to you?
2. How did you feel when the robot glanced at the wrong object?
3. How likely would you want to continue playing this game if the robot behaved the same?
4. How likely would you want to continue playing this game if the robot was more correct?
5. What are you thoughts on trust with the robot?
6. On a scale of 1-10 how much do you trust the robot?
7. On a scale of 1-10 how much do you trust the robot to be your virtual assistant?

The idea here is to really focus on trust and how that would relate to working with the robot in the future.

## 6 Future Work [30pts]

In general I found this section really should be covered in the proposed approach thus you see a light section here.

### 6.1 Picard paper

I would love to reach out to the author to interview them on where they are at today since we are more than 20 years since the paper has been published. I also would look and see if they have published any other papers on this topic and see how they have progressed.

The next steps after that would be to look for another other papers out there around this topic to see what other areas to explore here.

### 6.2 Speech and Gaze paper

In summary of proposed approach and validation I truly think that trust is the number one thing for humans and robots to work together. If there is lack of trust there will be friction between the two parties. I can see that robots will play a large role in our society in the future and we need to start building that trust now. I agree that robots can help us with our jobs, our lives, and so on but if that robot is not trust worthy then I really dont see that kind of robot being successful in their duty.

## Citations [3pts]

Admoni, H., Datsikas, C., & Scassellati, B. 2014. “Speech and Gaze Conflicts in Collaborative Human-Robot Interactions.” *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36, 104–9. <https://escholarship.org/uc/item/44z8484b>.

R. W. Picard. 1995. “Affective Computing.” *M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 321*, 1–16.

Todd Haselton. 2019. “Amazon Alexa is about to get more emotional — here’s what it sounds like excited and disappointed.” *CNBC*. <https://www.cnbc.com/2019/11/28/amazon-alexa-voice-is-about-to-get-more-emotional.html>.