

Intelligent Surveillance and Evaluation System in Courtroom Trials

Denny Anderson
University of Virginia
dra2zp@virginia.edu

Tamanna Hamid
University of Virginia
th4sf@virginia.edu

1 INTRODUCTION

Our country relies on a long established judicial system to determine the ramifications of questionable acts by members of society. Courtroom trials, law practices, and execution of sentences are common phenomena we rely on to address any misconduct and violence. Although the law is a system practiced almost to perfection, a single wrong decision made by a jury can have a catastrophic impact on the population as a whole. According to [1], there is an estimated 4.1% error rate for death-penalty cases, which accounts for only 0.1% of all prison sentences. The error for non-death-penalty cases, according to [2], is 2-5% due to false confessions, guilty pleas, flawed forensic evidence, and faulty eyewitness identification. An incredible reality lies in the fact that many tried and carried out cases are based on crimes that are not even committed in the first place. This means that as many as 100,000 people in the United States are found guilty and sentenced to prison but are innocent.

Such an error rate in verdicts shows that juries are not sufficiently capable to evaluate trials and make life-altering decisions about a person. Should there exist a complementary, scientific, facts-based system that will aid the jury in learning and understanding how the case is presented and acted out, the possibility of causing an error in judgment will decline to a great extent. This paper aims to address this problem by suggesting a likely solution using an intelligent system that determines these elements. We will discuss our intended surveillance system, possible working principle, evaluation methods and how it will change the course of law practice and in turn our judicial system.

2 IDEA

A typical court trial has an accumulation of people who are involved, directly or indirectly, in the case or the persons of interest. Apart from the suspect and plaintiff, the advocates, friends and family of each party are the ones mostly aware about the incidence in question. Although it is the duty of a law practitioner to use his skills to restore justice to a case, it is not uncommon for a lawyer to tweak evidence and statements to prove his guilty client as innocent. The same can be said about the people emotionally close to the persons of interest. Studying the reactions of these members of the court, it is possible to figure out lies and other unfair means being used in the trial for the suspect's benefit.

Our system focuses and analyzes the behavior of these members throughout the proceedings of the case. We aim to utilize facial recognition, gesture analysis, and voice modulation techniques to detect inconsistencies that can indicate possible lie detection and role play. With the help of a number of high-resolution video cameras located smartly around the courtroom, the system will track these behavioral traits, record and match them with proven traits to evaluate the integrity of the statements. Only the jury will be able to access this information during the trial. They can use the feedback provided by the system in combination with their personal judgment to present a fair, final verdict.

3 IMPORTANCE

The mission of our surveillance system is not only to help juries make better decisions in trials but help the judicial system overall. Cases and trials are quite subjective and the impact of setting a guilty person free due to our lack of recognizing behaviors in the trial reflects negatively on our society. The surveillance system is an effective method that can avoid such an occurrence. Since one of the primary persons of interest of this system is the lawyers of each party, it has the potential to recognize corrupt law practitioners and audit their eligibility to proceed in this profession in the future. Also, this system helps juries to overcome professional bias, since it focuses on the behaviors of the people instead of their physical appearance, race, and insignificant features like attractiveness that are unrelated to lie detection techniques.

4 IMPACT OF SUCCESS

If this proposed device works accurately, the projected impact will be the entire population. Whenever the legal system wrongly imprisons someone or frees a guilty person, everyone is affected. The consequences of a "not guilty" verdict on someone who actually committed the crime is very high. As a society, we rely on justice and fairness, and the population could start to feel that the judicial system is ineffective if too many wrong verdicts are made. Furthermore, allowing a criminal to go back into society without justice being served could cause other people to be adversely affected from that person's future criminal actions. Thus, reaching a true and correct verdict is extremely important.

In addition, a great amount of stress is placed on the jury. Twelve random peers must determine the fate of one person, which is a life-altering decision. Members of the jury are typically not experts and have a difficult time keeping track of all the evidence and everything that people have said during the long and drawn-out trial that may take up to a year of courtroom appearances. Also, the jury

is subject to implicit biases that they may not be aware of. For example, if the suspect reminds the jury of someone they liked or didn't like, it could influence whether they choose guilty or not guilty. So, this system is also expected to decrease the burden placed on juries in determining the outcome of a trial.

5 RELATED WORK

Polygraphs are rarely admissible in court, despite being 90% accurate [4]. Current polygraphs are made of several components, including two pneumographs that measure respiratory rate, a blood pressure cuff to measure blood pressure and heart rate, and galvanometers that measure galvanic skin resistance (sweat on the fingertips), and very few polygraphs are able to record basic arm and leg movements [4]. Also, polygraphs must be performed while the suspect is sitting still in a chair, and they must only respond "yes" or "no" to each question asked. In addition, the type of question being asked as well as its phrasing are crucial to a polygraph's accuracy. For example, some questions can be worded in such a way that requires more explanation than just a simple yes/no response, and the suspect can interpret a question differently than what was originally intended [5].

There are many more indicators that detectives use to figure out if someone is lying, which current lie detectors are unable to detect. Behavioral pauses or delays in speaking, a verbal and non-verbal disconnect (e.g. someone saying yes while shaking their head no), hiding or covering the mouth or eyes, clearing the throat or swallowing, touching the face, and grooming gestures (e.g. fixing a piece of clothing) are some indicators that a person may be lying [6]. Some other indicators include the person's eyes darting back and forth, rapid blinking, closing the eyes for more than one second at a time, looking to the right (for right-handed people), face touching, pursed lips, wiping excessive sweat off the face and neck, blushing, and shaking their head would normally not be considered when a person is undergoing a standard polygraph examination [7]. Other common indicators include a person changing their head position quickly, standing very still, repeating words and phrases, providing too much information, touching or covering their mouth, covering vulnerable body parts like the throat, shuffling their feet, difficulty speaking due to dry mouth, staring without blinking too much, and frequently pointing with their fingers [8]. All these cues for lie detection are not considered during standard polygraphs. Our system would monitor the above signs in order to give a more accurate estimation of whether or not someone may be lying.

Tselia Data Lab developed a camera algorithm that detects lies based on facial signals, but it takes several calibration questions to achieve its maximum 75% accuracy rate [9, 10]. Thermal imaging has also been shown to be 70% accurate in determining when a person is lying. The thermal camera is focused on a person's skin around the eye (because the skin is thinner there), and the camera can detect a warming in the face of just 0.1°C [11].

TruthTek is a system that uses facial, gesture, ocular, speech stress, voice stress, environmental, thermal/infrared, and physiological

sensors to provide over 99% accuracy in emotion and lie detection [12]. However, this sensor must be placed in very close proximity to the person's face, but this will most likely not be feasible in a courtroom, where we plan on using a few sensors several meters away from the person. These sensors also only focus on the person's face, rather than their gestures like hand, arm, feet, and eye positions and movements.

6 GENERAL SYSTEM PREREQUISITES

This system has the possibility to revolutionize the law and order system, but it comes with certain prerequisites and tests that should be carried out. One of them is to get members of the jury familiar with the surveillance system and how it determines inconsistencies. This can be achieved quite simply. When a certain person is tested for eligibility of jury duty, one of the test modules can be based on our surveillance system. Candidates will learn about the tracking mechanism and what different gestures mean. They will be given a thorough training about lie detection and body language. This will enable the prospective members to understand the system's method of operation, how it is perceiving the actions it records, etc. This additional module is advantageous to all because it will give the potential jury members the ability to recognize these traits themselves in the court as well. However, the members will have the freedom to make their own judgements about each individual in real time. Also, this system needs to be a covert tool in the courtroom so that potential suspects are not aware of exactly how it works. We emphasize on its confidentiality so deeply because it is absolutely essential that the knowledge of its existence do not reach any member of the court other than the judge and the jury. Otherwise potential hackers and anti-law organizations may attempt to alter the system for their benefit, disrupting the flow of the program and therefore the court proceedings and consequent outcomes from the system. This is a possibility filled with dire repercussions that we aim to absolutely avoid.

7 DESIGN FOR SOLUTION

A potential design for our system would include approximately three sensors placed strategically throughout the courtroom. Some possible placements might include one in the jury's point of view, so the sensor can "see" what the jury sees; one in the judge's point of view facing the defendant, plaintiff, lawyers, and audience members; and one in the audience's point of view facing the witness stand. These locations are meant to capture as many people as possible in order to pick up on their gestures and facial movements. The sensors will transmit the data to a central source location, where the algorithm and software can analyze the data and determine a confidence interval showing the probability of the truth or falsity of the given statements.

Each individual sensor will contain a high-resolution, wide range video camera, a microphone and thermal camera to record several people at once and to detect variances in a person's speech, gesture, and body temperature. Some additional sensors may likely be mounted in an inconspicuous location such as the ceiling. The intention is that the sensors will be more passive in picking up data

so that people don't have to communicate directly to the sensor, it doesn't cause any distractions, and the way trials are done can remain nearly the same. Since the sensors would be several meters away from the persons of interest, they will likely not be able to pick up on very subtle cues like pupil dilation. However, they should be able to pick up on slight changes in body temperature as well as body gestures and facial expressions to determine emotion. The central software will consist of several functions running concurrently, analyzing the data it receives from the various sensors. The results calculated must be conveyed very clearly in a user-friendly form so that the jury can make conclusions without having to analyze raw data. The main system should also be able to compile all the statements given by the different people and divide the statements up between truthful and deceptive with a probability associated with each to aid the jury in making a final verdict.

8 POSSIBLE TECHNOLOGICAL IMPLEMENTATIONS

For the purpose of this surveillance system, there are a number of available technologies that we can employ for each of its aspects. In that regard, we divide our system into five operations:

- figure detection and gesture analysis
- facial recognition and expression assessment
- body temperature monitoring
- voice modulation
- speech to evidence in text analysis

Each operation will be run by individual programs (sub-programs) within the central software, where each will provide its own evaluation to the members of the jury at the end of the trial. A detailed proposal of the methodology of each operation is listed below.

8.1 Figure Detection and Gesture Analysis

A potential detection system for this sub-program is Face++ [13]. This is a very comprehensive system that features body detection, outlining and even gesture recognition algorithms. Integrating Face++ to our central system, we will develop the sub-program to track how a person is moving around in the courtroom, measure changes in stance, arm and body movements, from which indicators of irregular behavior can be derived. For example, the program can detect when a person is fidgeting or looking sideways, which are signs of nervousness. Another similar technology we are considering is the FORTH 3D Human Body Tracker [14]. This provides a more reliable solution to the gesture analysis issue as it offers to determine the orientation and articulation of human body skeleton from a visual sample with no pre-set markers, for example, an image from a video camera. Apart from these, a more basic system named Phantom [15] can also be employed for this cause. This system is quite dated, requiring very basic operation requirements. For example, it is able to detect human bodies from a monochromatic image, which greatly reduces the cost for the surveillance cameras. However, it is being considered for its mathematically appropriate working principle and implementation of this dated technology is secondary to the more evolved version

of Face++ and FORTH for the sake of greater flexibility in our system implementation.

8.2 Facial Recognition and Expression Assessment

This operation is an extension of the previous one and requires more extensive use of the well-known detection technologies. Apart from body and gesture detection, Face++ is a likely resource for facial recognition. Another suitable system is the Amazon Rekognition, Amazon's new AI image processing tool [16]. Both these systems are able to provide quite accurate assessments of facial expressions and can be utilized to develop an algorithm tailored to our needs. For example, the high-resolution cameras can detect the changes in eye movements and blinking patterns, which will be analyzed by the program for possible indications of deceptive behavior.

8.3 Speech to Evidence Text Analysis

To deploy this operation, we develop a full-fledged algorithm with no external technology involved. This includes a speech to text conversion algorithm and database matching functions, which will be achieved by natural language processing technology. So, the audio feed collected from microphones in the courtroom will be changed to text form, and keywords from this will be searched on the database of written evidence presented to the court. A key aspect of this operation will be the use of deep learning for the system to analyze written evidence and relate data gathered from them to the audio data. For initial deployment, we can expect this operation to be slightly time-consuming than the rest of the operations because of its intensive processing. However, the analyzed data should be available to members of the jury during deliberation.

8.4 Voice Modulation

Voice stress analysis measures slight inaudible or almost inaudible fluctuations in the human voice called "micro-tremors" [17]. A series of tremor fluctuation patterns indicate the speaker is under a state of anxiety, looking for words to convince his/her audience. For the purpose of this operation, we will develop an algorithm to identify these micro-tremors from the audio feedback. One key notion in this operation is the recognition of individual voices from several human sources. By initially identifying a character in the conversation based on the sound pitch, tone and frequency, the algorithm will assign a designated decibel range for each voice. Data in form of negative or positive deviation from this range will be recorded by the system each time the source attempts to lie. One of the key advantages of this operation is that unlike video feed, this will require data from only particular areas in the courtroom. So, if we focus on the audio feed received from the microphone installed near the witness box, it is possible to obtain most of the conversation between the lawyer, suspect and witnesses.

8.5 Body Temperature Monitoring

Thermographic cameras and sensors are an efficient method of detecting face and limb temperature. When a person lies, his facial temperature drops due to increased stress on the brain, which is evidently constructing falsified statements. However, temperature around the nose and inner corner of the eyes tend to rise in these situations. This is called the Pinocchio Effect [18]. For the purpose of the proposed system, we will develop an algorithm that detects abnormal temperature rises in particular areas of the body, record these deviations as positive or negative data and map them against the variables time and individuals. These will then be analyzed by the system and enable it to detect liars and deceptive actions. Like the speech to text conversion and analysis operation, this operation would be a little time consuming, as it requires data processing and matching as well. However, results from the operation are likely to be available when the jury members are deliberating.

9 TECHNICAL PREREQUISITES

In order to deploy this system, it is essential to have access to necessary hardware to improve the quality of data collected and assessed. So, we need to equip our system with high-resolution video cameras. Video feeds with a minimum resolution of 1080x1920 (commonly termed as full HD) are needed for the proper detection of human bodies and movement tracking. This feed will be transmitted to its corresponding sub-program in the central system, where images of the feed can be taken for further analysis. Apart from this, thermal sensors and good quality microphones are required to pick up body temperature data and noiseless audio feed respectively.

The system also requires an extensive database server and communication system. For the preliminary stage of this system, a wireless communication system is not required, so high amounts of data about a case can be easily transmitted to the processor. The scale of the server, however, is hardly negotiable due to the large amount of data to be stored and processed.

10 DEPLOYMENT AND TESTING

Due to confidentiality purposes, a typical survey or test cannot be conducted when this system is deployed. However, for initial implementation we can use this system in smaller courtrooms where there is less evidence and the possibility of lying is limited like a case in the civil court. By installing sensor nodes in at least three primary locations, we can assess the outcomes of individual sub-programs and overall outcome of the system.

REFERENCES

- [1] Hughes, Virginia. "How Many People Are Wrongly Convicted? Researchers Do the Math." National Geographic. April 28, 2014. Accessed September 23, 2017.
- [2] Ferner, Matt. "A Record Number Of People Were Exonerated In 2015 For Crimes They Didn't Commit." HuffPost. February 3, 2016. Accessed September 23, 2017.
- [3] Terris, Callie K. "Jury Bias: Myth and Reality." July 22, 2016. Accessed September 23, 2017.
- [4] McGinnis, Patrick B. "How a Polygraph Works." How It Works. October 12, 2009. Accessed September 23, 2017.
- [5] "Is a Polygraph Test Admissible as Evidence." HG.org. Accessed September 23, 2017.
- [6] Nguyen, Vi-An. "Former CIA Officers Share 6 Ways to Tell If Someone's Lying." Parade. July 25, 2013. Accessed September 23, 2017.
- [7] Gillett, Rachel, and Samantha Lee. "How to tell someone's lying to you by watching their face." Business Insider. January 30, 2016. Accessed September 23, 2017.
- [8] Giang, Vivian, and Jacquelyn Smith. "11 Signs Someone Is Lying To You." Business Insider. April 30, 2014. Accessed September 23, 2017.
- [9] Grigonis, Hillary. "Could an AI Camera Detect Lies Better than a Polygraph? One Tech Firm Thinks So." Digital Trends. October 25, 2016. Accessed September 23, 2017.
- [10] "AI Camera Might One Day Detect Lies Better Than a Polygraph." Nvidia Developer. November 4, 2016. Accessed September 23, 2017.
- [11] Thornhill, Ted. "Feeling flush? How hot-faced liars could be caught out by thermal-imaging cameras." Psych Central. September 13, 2011. Accessed September 23, 2017.
- [12] "TruthTek: A revolutionary lie detection system." Indiegogo. Accessed September 23, 2017.
- [13] "Body Detection." Face++. Accessed October 24, 2017.
- [14] Michel, et al. "Apparatuses, Methods and Systems for Recovering a 3-Dimensional Skeletal Model of the Human Body." FPO. March 24, 2016. Accessed October 24, 2017.
- [15] Leskovec, Jurij. "Detection of Human Bodies using Computer Analysis of a Sequence of Stereo Images." University of Stanford. May 27, 1999. Accessed October 24, 2017.
- [16] "Amazon Rekognition." AWS. Accessed October 24, 2017.
- [17] "About Liar Liar." Source Forge. Accessed October 24, 2017.
- [18] University of Granada. "'Pinocchio effect' confirmed: When you lie, your nose temperature rises." ScienceDaily. December 3, 2012. Accessed October 24, 2017.