The face of an imposter: Computer Vision for Deception Detection Research in Progress

Aaron C. Elkins
Imperial College London
a.elkins@imperial.ac.uk

Yijia Sun
Imperial College London
yijia.sun@imperial.ac.uk

Stefanos Zafeiriou
Imperial College London
s.zafeiriou@imperial.ac.uk

Maja Pantic
Imperial College London
m.pantic@imperial.ac.uk

Abstract

Using video analyzed from a novel deception experiment, this paper introduces computer vision research in progress that addresses two critical components to computational modeling of deceptive behavior: 1) individual nonverbal behavior differences, and 2) deceptive ground truth. Video interviews analyzed for this research were participants recruited as potential hooligans (extreme sports fans) who lied about support for their rival team. From these participants, we will process and extract features representing their faces that will be submitted to slow feature analysis. From this analysis we will identify each person's unique facial expression and behaviors, and look for systemic variation between truth and deception.

1. Introduction

Do liars behave differently than truthtellers? Theory and conventional wisdom suggest that they do. Not confined to theory alone, experimental research is replete with evidence that liars do in fact, speak, gesture, and behave differently [1]. One need only review the significant p-values and accounted for variance explained by deceptive behavior. Why then do we still have such difficulty translating significant findings into robust and accurate deception detection?

Translating statistical significance into computational models suitable for reliable deception detection is a challenging endeavor. Using video analyzed from a deception experiment, this paper introduces research in progress that addresses two critical components to computational modeling of deceptive behavior: 1) individual nonverbal behavior differences, and 2) deceptive ground truth.

2. Deception Detection

2.1. Nonverbal Behavior

One impediment to detection accuracy is overreliance on individual cue modalities, such as just the voice or facial expressions. While existing theories suggest liars will leak behaviors in response to the situational, cognitive, emotional, and strategic demands of the interaction [2], [3], there are no individual cues that universally reveal deceit.

The result is that individual behaviors occur concomitantly with deception frequently enough to discount chance observation, but infrequently enough to inhibit reliable deception detection. Some liars will reveal their deception through their language and voice, while others speak impeccably regardless. To address this problem, Derrick et al. [4] recommend modality fusion as a necessary step towards accurate deception detection—pooling cues to improve convergent validity, sensitivity, and reliability of detection.

2.2. Sensor Fusion

The intuition underlying fusion is that even though people do not exhibit nonverbal behavior identically in response to deception, they leak some form of behavior unintentionally. Someone practiced in composing their body posture or facial expression, may neglect their language or speech. By fusing multiple cues, it ensures that we aren't ignoring the appropriate cue for the specific person and situation.

Research into how best to reconcile multiple cues is still ongoing. For example, should we combine cues mathematically into an input vector prior to classification, or impose a voting model on independent classification models based on each cue? In addition to the scientific investigation still required, there are great practical constraints.

There is high complexity and expense added by each additional sensor. Most sensors require a prominent position directly center (0°) and in front the person being measured. In the best case, two or three sensors can achieve optimal placement and additional sensors must make compromises. In operational environments, we are many years from real-time multiple sensor fusion.

The same spirit that motivates sensor fusion can still be applied to individual sensors. This research introduces an approach to identifying nonverbal behavior differences unique to each person. Using computer vision analysis of the face and slow feature analysis, we identify each person's unique facial expression and behaviors, and look for systemic variation between truth and deception.

2.3. Slow Feature Analysis

Slow feature analysis (SFA) is an unsupervised learning algorithm for reducing an input vector or signal into a reduced set of features that vary the slowest [5]. Rooted in how human's perceive visual stimulus via the visual cortex, SFA identifies the features that vary the slowest overtime and discards quickly varying or changing features as noise.

For deception detection using computer vision features, SFA allows unique instances of facial behavior (onset, apex, offset) to be identified. Each person's video processed and submitted to SFA will result in different, but relevant facial expressions identified. For example, some people may make a Duchenne smile, while others a lip press.

SFA has only recently been employed in computer vision to provide online video segmentation of behavioral changes [6]. These segments of behavioral changes, unique to the person in the video, can then be submitted to a classification model. In this way, SFA allows us to abstract to significant facial behavior and look for systematic differences, regardless of specific cues that vary between persons.

3. Ground Truth

One challenge to predicting deception is ensuring that you have reliable ground truth. We can extract sophisticated features and identify significant behaviors, but we must be confident that we are training our computational models on actual deception and truthful instances. A majority of computational models are based on experimentally induced behavior to guarantee consistency and ground truth.

Deception is often characterized by high stakes or consequence environments that induce high arousal or stress in liars. Because of this, experimentally induced or sanctioned lying is often criticized as unrealistic or unrepresentative. This is debatable, as there is much more to deception than simply stress, such as cognitive effort, memory, emotions, and behavioral control. However, one aspect of deception that remains important is motivation.

3.1. Motivation

DePaulo and Kirkendol [7] introduced their motivation impairment effect on their self-presentational perspective [8] of deception. They predict that increased motivation moderates self-regulation causing redoubling of effort and awareness of behavior that causes additional leakage of deception cues. Conversely, motivation has been found to actually improve verbal and nonverbal performance and behavior [9].

Despite this apparent contradiction, motivation remains an important moderator to deceptive behavior. However, we have very few mechanisms to increase the true motivation for experimental participants to lie. The most common motivation is to reward or punish participants monetarily.

For this research, we sought to increase the motivation of participants through the intrinsic content of their lies, rather than external factors or operant stimulus.

4. Hooligans and Imposters

Elkins, Derrick, and Gariup [10] introduced an experimental paradigm that focuses on imposter lie behavior. This paradigm focuses on a specific lie scenario or situation in which some participants lie about their identity. The important takeaway from this research is that lying about your identity produces different behavioral cues than other types of lies.

For example, a person telling the truth to biographical questions may take longer (increased cognitive effort) than a liar who is reciting canned or memorized responses. In other scenarios, you might expect the opposite for liars who are fabricating messages on the spot. It is critical that any computational model not conflate deception scenarios.

Based on the imposter scenario, an experimental paradigm designed to increase intrinsic motivation was design and conducted—requiring high and low football fans to lie about their favorite football team.

4.1. Hooligans

Hooligans are sports fans with extreme interest and identification with their favorite team. Hooligans in sporting events are so affected by the outcome of their team that they are more likely to commit acts of violence against rival fans. To increase the intrinsic motivation of the lie behavior, potential hooligans were recruited to lie about their favorite team (i.e., claim support for their rival team) in an imposter experimental scenario.

5. Method

5.1. Participants

Ninety participants from the Netherlands were recruited to complete a mock screening scenario for a football match. Before arriving on the day of the experiment the participants completed pre-surveys to collect demographic, favorite/rival football teams, and level of hooliganism.

5.2. Hooliganism Measure

An important aspect of this experiment is increasing the intrinsic motivation of the lie behavior. To accomplish this we first needed to identify participants who hold the strongest and weakest support for their teams. To this end, we adapted and validated a series of questions that measured football fan level, importance of winning, and proclivity towards violent behavior. These factors are directly related to the violent and extreme positions held by football hooligans.

From the sample, the average fan level (7 point scale) of a participant was 3.13 (SD=1.86) with an average reported propensity for sports violence was 1.76 (SD=1.45). The majority of participants reported they would not engage in any violence. Because of this, the football fan level was the primary measurement of hooliganism for this study, as there was virtually no difference between participants and their reported violence.

Next we identified each of the participant's favorite and rival teams. The top teams listed as favorite were Ajax (N=36) and Feyenoord (N=16). The top teams listed as a rival were Ajax (N=34) and Feyenoord (N=35). Because these were the most prevalent teams, they were selected as the match for the mock football game. All participants were asked to support their actual favorite or rival (imposter) team.

5.3. Condition Assignment

Based on the reported hooliganism, a stratified random sampling was incorporated were those that reported both high and low levels of hooliganism were assigned to be imposters in the scenario. There were 30 participants assigned to the imposter group for this scenario. Imposter participants were contacted separately and given instructions to falsely claim to support the team that was actually their rival.

6. Procedure

Upon arrival, participants were given a ticket for the football match and instructions to complete a screening interview with the AVATAR [11], [12], an automated embodied conversational interviewer (ECA). Regardless of condition, participants were instructed to appear honest and attempt to pass through the screening checkpoint. For imposters, this meant lying to the AVATAR by claiming they supported the team that was actually their rival.

Figure 1 below illustrates a typical automated interview with a participant. Participants were shown different images relevant to the interview as well as an ECA interviewer who asked the questions.

Figure 1: AVATAR Interview with Participant

After completing the AVATAR interview, participants then spoke with human border guard screeners, and completed a post-survey.

6.1. AVATAR and Sensors

The advantage of using the AVATAR interviewer is that all screening questions are asked identically to increase between-subject comparability and all sensors measuring behavior were automatically segmented by questions. The sensors measuring behavior during this interview were an eye tracker, high definition video camera, and microphone.

For this research, the video recordings captured during the interview were analyzed and processed for the computer vision analysis.

7. Current Progress

Since the completion of the experiment the participant videos have been preprocessed. This stage includes extracting the image frames from the video files (30 frames per second, 1280x720 resolution). For each of these images, the face and facial landmarks were tracked using Active Orientation [13] and Appearance-based [14] tracking models. Figure 2 below illustrates a participant face with a facial landmark model fit to their face.

Figure 2: Untracked and Tracked Participant Face

The purpose of fitting the model is two-fold first the face is identified within each image, and second, the specific regions of the face within the face are demarcated. After the faces were identified, the face part of the images was extracted from each frame and normalized to make all facial images comparable.

8. Next Steps

With all of the facial images extracted and normalized, the next step for this research is to extract the images circumscribed by the eye and mouth landmarks. These eye and mouth images will then be submitted to feature extraction to represent the eyes and mouth for each frame. The features that will be extracted for each image include intensity values, Gabor filters, Local Binary Patterns, and gradient-based descriptors.

The extracted features will then be submitted to SFA to reduce the features to the slowest varying features, which represent person-dependent behavioral actions. After which, we will experiment with machine learning and classification methods to classify deceptive speakers with and without the covariate of motivation (high vs. low fan level).

9. Conclusion

This paper introduces research in progress on the application of computer vision for automated deception detection using a novel imposter scenario and person-specific computer vision feature reduction. While additional work is still needed, the approached discussed could contribute to addressing between person nonverbal behavior variability that confounds existing single modality based deception detection methods.

10. Acknowledgements

This research was made possible through collaboration with Frontex, Research and Development Unit, and the National Center for Border Security and Immigration (BORDERS) at the University of Arizona.

11. References

[1] B. M. DePaulo, J. J. Lindsay, B. E. Malone, L. Muhlenbruck, K. Charlton, and H. Cooper, “Cues to deception.,” *Psychological Bulletin*, vol. 129, no. 1, pp. 74–118, 2003.

[2] P. Ekman and W. V Friesen, “Nonverbal leakage and clues to deception,” *Psychiatry*, vol. 32, no. 1, pp. 88–106, 1969.

[3] D. B. Buller and J. K. Burgoon, “Interpersonal Deception Theory,” *Communication Theory*, vol. 6, no. 3, pp. 203–242, Aug. 1996.

[4] D. C. Derrick, A. C. Elkins, J. K. Burgoon, J. F. Nunamaker Jr, and D. D. Zeng, “Border Security Credibility Assessments via Heterogeneous Sensor Fusion,” *IEEE Intelligent Systems*, vol. 25, no. May/June, pp. 41–49, 2010.

[5] L. Wiskott and T. J. Sejnowski, “Slow feature analysis: unsupervised learning of invariances.,” *Neural computation*, vol. 14, no. 4, pp. 715–70, Apr. 2002.

[6] S. Liwicki, S. Zafeiriou, and M. Pantic, “Incremental Slow Feature Analysis with Indefinite Kernel for Online Temporal Video Segmentation,” in *11th Asian Conference on Computer Vision*, 2012.

[7] B. M. DePaulo and S. E. Kirkendol, “The motivational impairment effect in the communication of deception,” in *Credibility Assessment*, J. Yuille, Ed. Deurne, Belgium: Kluwer, 1989, p. 1996.

[8] B. M. DePaulo, “Nonverbal behavior and self-presentation.,” *Psychological bulletin*, vol. 111, no. 2, pp. 203–43, Mar. 1992.

[9] J. Burgoon and K. Floyd, “Testing for the motivation impairment effect during deceptive and truthful interaction,” *Western Journal of Communication (includes …*, no. September 2012, pp. 243–267, 2000.

[10] A. C. Elkins, D. C. Derrick, and M. Gariup, “The Voice and Eye Gaze Behavior of an Imposter: Automated Interviewing and Detection for Rapid Screening at the Border,” in *Conference of the European Chapter of the Association for Computational Linguistics*, 2012.

[11] A. C. Elkins, D. C. Derrick, J. K. Burgoon, and J. F. Nunamaker Jr, “Predicting Users’ Perceived Trust in Embodied Conversational Agents Using Vocal Dynamics,” in *Forty-Fifth Annual Hawaii International Conference on System Sciences*, 2012.

[12] J. F. Nunamaker, D. C. Derrick, A. C. Elkins, J. K. Burgoon, and M. Patton, “A System Model for Human Interactions with Intelligent, Embodied Conversational Agents,” *Journal of Management Information Systems*, 2011.

[13] G. Tzimiropoulos, J. Alabort-i-medina, and S. Zafeiriou, “Generic Active Appearance Models Revisited,” pp. 1–14.

[14] S. Liwicki, S. Zafeiriou, G. Tzimiropoulos, and M. Pantic, “Fast and robust appearance-based tracking,” *Face and Gesture 2011*, pp. 507–513, Mar. 2011.