An Approach for Intent Identification by Building on Deception Detection

Judee Burgoon, Mark Adkins, John Kruse, Matthew L. Jensen, Thomas Meservy,
Douglas P. Twitchell, Amit Deokar, Jay F. Nunamaker
Center for the Management of Information
University of Arizona
jburgoon@cmi.arizona.edu

Shan Lu, Gabriel Tsechpenakis, Dimitris N. Metaxas
Center of Computational Biomedicine Imaging and Modeling
Rutgers University
shanlu@cs.rutgers.edu

Robert E. Younger
Space and Naval Warfare Systems Center
younger@spawar.navy.mil

Abstract

Past research in deception detection at the University of Arizona has guided the investigation of intent detection. A theoretical foundation and model for the analysis of intent detection is proposed. Available test beds for intent analysis are discussed and two proof-of-concept studies exploring nonverbal communication within the context of deception detection and intent analysis are shared.

1. Introduction

Each nation has the obligation to safeguard its homeland against deception and infiltration by adversaries who may be planning hostile actions. One of the most daunting challenges of the 21st century is determining how to create and implement these safeguards. Achieving high information assurance is complicated not only due to the speed, complexity, volume, and global reach of communications and information exchange that current information technologies now afford, but also because of the fallibility of humans in detecting hostile intent. Those protecting the borders and public spaces are often handicapped by untimely and incomplete information, overwhelming flows of people and materiel, and the limits of human vigilance. The vulnerabilities of human agents are exacerbated by the very same technologies that enable the collection of massive amounts of information— information that must be sorted, analyzed, and synthesized.

The interactions and complex interdependencies of information systems and social systems make the problem even more difficult and challenging. We simply do not have the necessary means to specifically identify every dangerous individual around the world. Although automating the detection of intent is an appealing prospect, the complexity of detecting and countering hostile intentions defies a completely automated solution. A more promising approach is to integrate improved human efforts with automated tools for behavioral analysis, the end goal being a system that singles out individuals for further scrutiny in a manner that reduces false positives and false negatives.

In this paper, we present our current research efforts in the direction of developing automated tools to identify intent and deception. The paper is organized as follows: Section 2 discusses the relationship between intent, deception and behavior. Section 3 explains the model and the methodology we follow to identify intent based on suspicion level. Section 4 discusses available data sets for investigating intent and Section 5 addresses the current research based on the model and methodology. Finally, Section 6 discusses plans for future experiments and research work.

2. Deception, intent and behavior

Deception is a message knowingly transmitted with the intent to foster false beliefs or conclusions [1]. Over the past two and a half years, the Center for the Management of Information (CMI) at the University of Arizona has conducted over a dozen experiments to study

deception with over 2000 subjects [2-5]. These experiments have been instrumental in creating an understanding of the factors influencing deception, and have guided the building of automated tools for enhancing deception detection and the creation of training for security personnel [3, 6]. Furthermore, research into deception has led to the examination of the relationship between deception and intent. For purposes of our research about intent, we are concerned with those planning to commit or abet criminal or terrorist activity.

The connection between intent and suspicion is not direct. As shown in Figure 1, intent may be inferred by detecting the presence of deception in communication. The premise of inferring intent from deception rests on the idea that individuals with hostile intent will be deceptive about their intentions in order to avoid detection. If deception is present in the communication, deceptive behavioral cues will be displayed that may arouse suspicion.

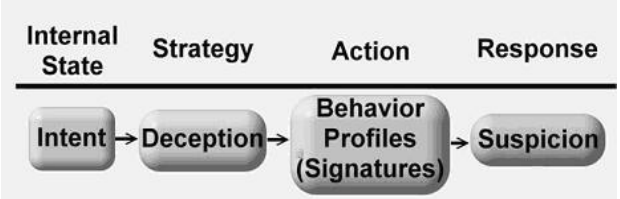

Figure 1. Role of deception in determining intent from behavior

While inferring intent from the presence of deception is a productive first step, correctly identifying intent is not as simple as merely searching for deceptive behavioral cues. Having a deceptive goal is only one of many internal states an individual may possess; other internal states that may indicate hostile intent include anger, tenseness, or fear as shown in Figure 2. The intent of a person, whether benign or hostile, is tied to his or her internal state and each internal state may have outward manifestations that take the form of observable behaviors. Behavioral cues may be indicative of any number of internal states, which may in turn point to different intents. For instance, hostile intent may be revealed in a typical angry or agitated state. However, it may also be displayed as excitement or happiness. While internal states may be manifest by any of a number of behaviors, a single behavior could indicate a number of emotions. Further, an internal state may be manifest by an unexpected behavior, as indicated by the dotted lines in Figure 2.

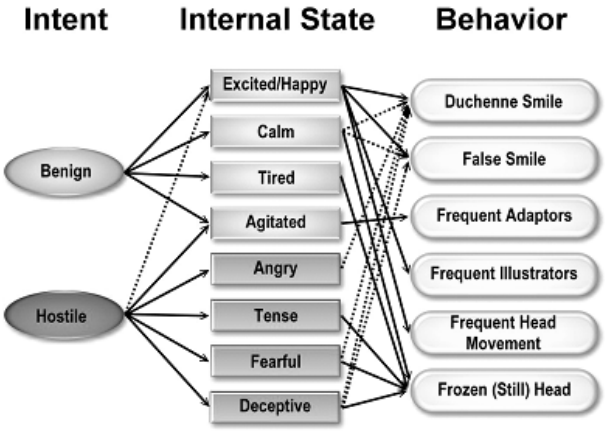

Figure 2. Relationship between intent, internal state and behavior

The complex relationships between behaviors and internal states render the task of identifying intent particularly challenging. Our current research focuses on understanding these mappings and leveraging experience in the area of deception detection to produce methods which identify the true intent of a person. With established methods in place, the creation of a tool to aid in intent detection becomes possible.

3. Theoretical foundation

Several theories and models offer useful perspectives on the linkage between intent and overt behavioral manifestations that elicit trust or suspicion. Three theories that are especially germane—interpersonal deception theory, expectancy violations theory, and signal detection theory—are integrated to produce a model of suspicious and trust-eliciting verbal and nonverbal communication. Additionally, a theory-guided taxonomy is useful for clustering verbal and nonverbal behaviors into appropriate profiles of suspicious and non-suspicious behavior in order to identify those who have the highest probability of hostile intent.

Interpersonal deception theory (IDT) is the key theory used for mapping behavioral cues into profiles [7]. IDT depicts the process-oriented nature of interpersonal deception and the multiplicity of pre-interactional, interactional, and outcome factors that are thought to influence it. Among its relevant precepts is the assumption that deception is a strategic activity subject to a variety of tactics for evading detection. It also recognizes the influence of receiver behaviors on sender displays, and it views deception as a dynamic and iterative process, a game of moves and countermoves that enable senders to make ongoing adaptations that further hamper detection. Consequently, a theory of suspicious

and trust-eliciting behavior must take into account a variety of moderator variables, each of which may spawn a different behavioral profile.

Expectancy violations theory (EVT) is concerned with what nonverbal and verbal behavior patterns are considered normal or expected, what behaviors constitute violations of expectations, and what consequences violations create [8]. Its proponents contend that specific behavioral cues are less diagnostic than whether a sender's behavior conforms to or violates expected behavioral patterns and that receivers are more likely to attune to such violations. In other words, it is more useful to classify communication according to whether it includes behavioral anomalies, deviations from a baseline, or discrepancies among indicators. Behavioral patterns which include deviations and anomalies are predicted to influence receiver judgments of credibility and deceit. EVT distinguishes between positive and negative violations. Positive violations may actually foster perceived trustworthiness and credibility, whereas negative violations should foster suspicion. Expectancy violations theory is thus relevant to the process of comparing the behavioral profiles against the expected norms.

The process of interpreting different verbal and nonverbal cues and clustering them together in the form of behavioral profiles to contrast with the expected behavioral profiles is non-trivial and challenging, considering the large variation in the behaviors of different human beings. The key segments of this dynamic process are the characteristics of actors, features of transmission channels, features of messages, and the information exchange process itself.

Finally, a threshold for deriving the output of suspicion or trust is based on signal detection theory (SDT). Developed by Green and Swets [9], SDT defines two sets of probabilities in a signal detection test, in which two possible stimuli types must be discriminated. In the context of intent identification, the two possible stimuli types are hostile and benign intent. If the actual intent is hostile and the output judgment is suspicion, the trial is a "hit." If the actual intent is benign and the output is judged suspicion, it is a "false alarm." If the actual intent is hostile but the judgment is one of trust, it is a "miss." Finally, if the actual intent is trustworthy and the judgment is one of trust, it is a correct decision as shown in Table 1.

According to SDT, the output of such a binary test is based on the value of a decision variable, which in the context of intent identification is the suspicion level. The threshold value of the decision variable is called the criterion. For humans, the selection of a criterion is not only related to the value of actual stimuli but also related to their psychological characteristics. In other words, the criterion is a function of perceived stimuli, which, in the context of intent detection, are the behavioral profile deviations. The SDT calculation methods described in [10] can be used to study the distribution of the values of the suspicion level variable across the behavioral profile deviations to determine the appropriate criterion for the final decision making.

Table 1. Possible judgments from SDT

		Judgment	
		Suspicion	**Trust**
Actual Intent	**Hostile**	"Hit"	"Miss"
	Benign	False Alarm	Correct Decision

We have integrated these multidisciplinary theories and models into a single systemic framework that guides our experimental work and tool development. The model is shown in Figure 3. It is a decision model for judging how trustworthy an individual is on a trust-suspicion spectrum, based on demonstrated behavioral cues. The actual intent of the individual can be considered as input for the model, which is demonstrated in the form of the behavioral cues, either verbal or nonverbal, which include kinesic, proxemic-haptic, chronemic, linguistic, content, meta-content, and paralinguistic cues. These behaviors are affected by the individual's emotions as well as the context in which the communication takes place. Moreover, the behaviors are also influenced by the interaction of sender and receiver actions, cognitions and their mutual influence. These micro-level behavioral cues can be clustered into macro-level profiles and located along conceptual dimensions of arousal, power, pleasantness, and intensity as shown in Figure 3.

The four-dimensional macro-level profile portrays an individual's emotional and cognitive status, as a combination of active/passive, dominant/submissive, tense/relaxed, and pleasant/unpleasant, based on observed behaviors. This observed macro-level profile of the sender can then be compared with two normal or expected profiles stored in a repository.

First, the macro-level profile is compared with the individual's historical macro-level profiles across multiple episodes within a given context. For instance, if Victor is usually dominant in conversation with others, and this dominance is usually associated with such behaviors as speaking loudly, then an archive of Victor's typical behavior patterns will include dominance and loudness as part of the individual-level set of expectations. When individual-level profiles are not

available, only the second set of expectations will be utilized. This second set of expectations is comprised of a general profile of normal behavior across people within the same scenario. For example, when questioned face-to-face, guilty suspects may show a combination of verbal brevity, vocal tension, and over control of movement. The result is that such individuals typically look more tense, unpleasant, aroused, and submissive than those with benign intent.

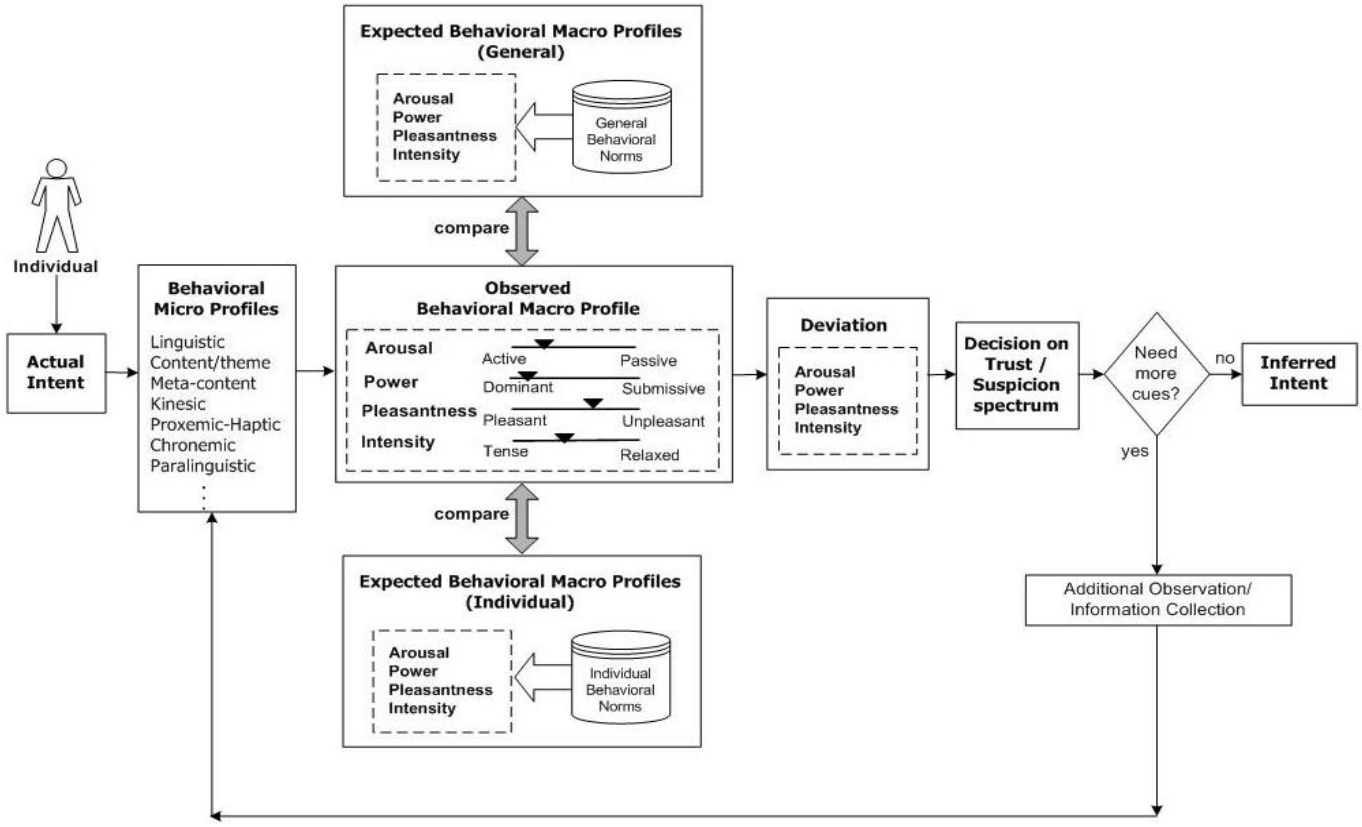

Figure 3. Decision model for trust and suspicion

The deviation between the observed behavioral profile and the expected individual and group profiles, in either positive or negative directions, indicates a suspicion level. By setting a proper threshold on this deviation measure and given a certain context, the automated tool will thus be able to indicate the probability that the sender is suspicious or trustworthy.

4. Past experiments

In order to analyze the behavioral profile of a sender, a rich data set is needed. The research team has compiled data from its past and current research projects to form test beds which are good examples of human behavior. The available data sets are briefly described below along with the experiment designs to illustrate their richness and potential for this research.

4.1 Behavioral analysis interviews

Because experimentally generated data often lack the high motivation and severity of consequences found in real-world circumstances, we have obtained videotapes of criminal suspects being subjected to pre-polygraph interviews in which a standard protocol, the Behavioral Analysis Interview, is used. The interviewees in each video were found guilty based on confession or by independent corroborating evidence. This test bed enables controlling for the degree of nervousness common for anyone subjected to criminal interviews from the behavior patterns uniquely associated with criminal conduct. We believe that the criminals' behavior will be a valid surrogate for hostile intent or suspicious behavior evidenced in other contexts.

4.2 Mock theft experiment

The Mock Theft Experiment [4, 5] was designed at the University of Arizona to reveal cues that can be used in detecting deception. In this experiment, some participants played the role of a thief while others where simply present during the theft. All participants were subsequently interviewed by untrained and trained interviewers via text chat, audio conferencing, or face-to-face interaction. A companion observer study includes third-party assessments of senders' trustworthiness and thus can serve as a second form of independent verification of whether the interviewees' language, content, and nonverbal behavior appeared trustworthy or suspicious.

4.3 Airport scenarios

Four actors were hired to assist in performing a proof-of-concept study to determine the feasibility of identifying behavioral states from gestures and body movement. They participated in scenarios designed to simulate airport screening procedures. The scenarios included seated interaction, standing interaction, queuing, and locomotion. Within each scenario, each actor was asked to demonstrate three states: relaxed, agitated, and over-controlled.

4.4 Machine learning training set

Seven CMI employees created a set of videos used to train machine learning tools to identify human gestures and movement. Twenty gestures involving the fingers, hands, arms, trunk and head were repeated 10-12 times by each participant. These gestures were recorded and then manually coded. Although the gestures were the same across participants the size and speed of the gesture varied.

5. Current research

In an effort to test the concept of automatically identifying nonverbal cues that arouse suspicion, two proof-of-concept studies were conducted at the University of Arizona. These studies focused on analyzing nonverbal cues found in body movement and hand gestures.

According to Ekman [11], human gestures can be divided into three categories: emblems, illustrators, and adaptors. Emblems are hand gestures that have a consistent meaning to everyone within a cultural group. An example emblem is a head nod used to mean "yes." Illustrator gestures are movements that accompany speech. They can emphasize what is being said or can diagram a difficult concept in space. They usually have no meaning independent of speech, for example demonstrating a right turn when giving oral directions. Finally, manipulators or adaptors are gestures that fulfill some physical or psychological need. Examples of adaptors include scratching, preening, picking and fidgeting.

Previous research has suggested that liars gesture differently than truth tellers do [11-13]. A common perception among those who try to identify deception is that an increased number of adaptor gestures such as fidgeting and foot tapping are indicators of deception. Numerous studies have shown that adaptor gestures may not be linked with deception [13].

However a decrease in the number of illustrator gestures has been linked with deception [11-13] Increased cognitive load has been suggested as a possible reason for the decrease in illustrator gestures. The deceiver may be more concerned with fabricating a realistic lie than with punctuating that lie with hand gestures.

Building on past research concerning illustrator gestures, the two proof-of-concept studies explore methods of automatically identifying and classifying gestures for the purpose of creating profiles of normal behavior and investigating what might be considered deviations. The first study investigates machine learning techniques to automatically identify gestures from a video segment. The automatically identified gestures are important kinesic indicators in determining the arousal and intensity dimensions of the four-dimensional behavioral profile described in Section 3. The second study investigates a new method of gesture classification in order to differentiate between illustrator and adaptor gestures. Illustrator and adaptor gestures provide additional information that can be included in the individual's behavioral profile and will possibly provide insight into deception and hostile intent. The results of both proof-of-concept studies are illustrated below.

5.1 Blob analysis

Central to the recognition of nonverbal signals including individual gestures in video is the ability to recognize and track body parts such as the head and hands. The Computational Biomedicine Imaging and Modeling Center (CBIM) at Rutgers University refined a method called "blob analysis" which provides this foundation [14-16]. MIT's PFinder is another tool based on blob analysis [17].

Using color analysis, eigenspace-based shape segmentation, and Kalman filters, the software is able to track the position, size, and angle of different body parts with great accuracy. Figure 4 shows a single frame of video which has been subjected to blob analysis. The ellipses in the figure represent the body parts' position, size, and angle.

Figure 4. Blobs capture the location of the head and hand

Blob analysis extracts hand and face regions using the color distribution from an image sequence. A three-dimensional look-up-table (3-D LUT) is prepared to set the color distribution of the face and hands. This 3-D LUT is created in advance using skin color samples. After extracting the hand and face regions from an image sequence, the system computes elliptical "blobs" identifying candidates for the face and hands. The 3-D LUT may incorrectly identify candidate regions which are similar to skin color, however these candidates are disregarded through fine segmentation and comparing the subspaces of the face and hand candidates. Thus, the most face-like and hand-like regions in a video sequence are identified. From the blobs, the left hand, right hand and face can be tracked continuously.

From each blob a number of measurements are recorded for each frame in an image sequence. As demonstrated in Figure 5, the center of the blob is captured as x and y coordinates. These coordinates are based on the pixels contained in each frame. Further, the lengths of the major and minor axes of the ellipse are recorded in pixels. Finally, the angle of the major axis of the blob is recorded. Table 2 contains an example data stream based on a single blob.

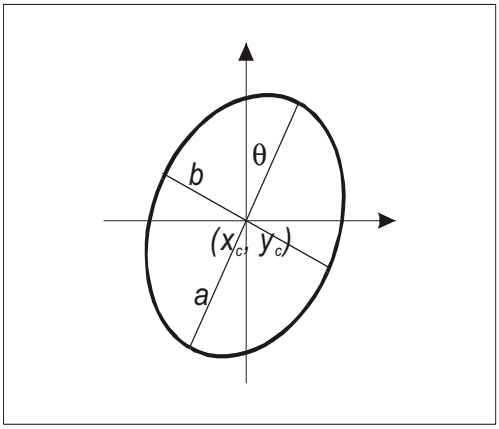

Figure 5. An example of a blob that surrounds the hands and head

Table 2. Measurements from the head blob. Similar measurements are collected from the hand blobs

Frame	X	Y	Major Axis Length	Minor Axis Length	Major Axis Angle
1	322	149	90	63	0.48
2	322	149	90	63	0.48
3	323	149	88	64	0.46
4	323	148	89	65	0.49

From positions and movements of the hands and face we can make further inferences about the trunk and relations to other people and objects. This allows the identification of gestures, posture and other body expressions.

5.2 Study 1: Neural networks used in time series analysis of gestures

In analyzing body movement, one approach is to include the element of time in the analysis of movement. Neural networks have been used frequently over the last few decades for static pattern recognition and pattern recognition in time. Two prevalent methods for analyzing time series data using neural networks are Time Delay Neural Networks (TDNN) and Recurrent Neural Networks (RNN) [18]. TDNNs are structured in much the same way as standard neural networks with an input layer representing each of the inputs, a hidden layer fully connected to the input layer, and an output layer fully connected to the hidden layer. Instead of inputting a single input vector for each time step, TDNNs use a sliding window of some number of consecutive input vectors. This allows the TDNN to recognize patterns that

are approximately the size of the sliding window. RNNs, on the other hand, are not restricted to a single window size. RNNs have some form of feedback in the network itself. In a common architecture introduced by Elman [19], all of the output nodes feed back into the hidden nodes with a delay of one time step. This technique gives the network a memory of arbitrary length.

5.2.1 Methodology. To train our neural networks, the Machine Learning Training Set described above was used. The videos were manually segmented and labeled and the gestures in the video were subjected to blob analysis. From the blob analysis the x and y coordinates of the center, major and minor axes lengths and angle of major axis were collected from each blob for every frame in the video clip. The data were normalized according to minimum and maximum values in each column giving each value in each frame a value between 1 and 0. The normalization allowed training of the neural network and comparability across individuals.

5.2.2 Results. We constructed a TDNN with a sliding window of 30 frames or one second. A visual inspection of the results shows that the network was only able to recognize one of the tested gestures with consistent accuracy. However, the RNN we constructed did much better. Likely because of its ability to recognize patterns of variable length, it was able to recognize most of the gestures. Although most gestures were recognized by the RNN, both networks occasionally confused similar gestures such as "right hand in" and "right hand touching face." There are two possible ways to mitigate such confusion: a larger training set or more optimized network architecture could be employed. While the former requires video work, the latter has been shown to be a difficult problem. Currently, only heuristics exist for determining optimal neural network architectures.

These preliminary results indicate that gesture recognition using RNNs may be a promising method for identifying gestures from subjects on video. TDNNs, on the other hand did not perform as well as RNNs, although they may be useful for limited gestures where the length of the gesture is known.

5.3 Study 2: Quadrants used in gesture analysis and deception detection

While identification of gestures and body movement may be possible through time series neural networks, a large and representative training set is required for accurate results. Obtaining, manually coding, and preparing such a training set is time consuming and expensive. Further, as one moves from one context to another, a training set loses its representative nature. A neural network trained to identify a set of gestures in one condition may not be able to identify gestures in another condition.

As mentioned in Study 1, the neural network struggles with identifying gestures that are similar in nature. Two similar gestures such as "Right hand in" and "Right hand touching face" are often confused. The difference between defined gestures such as "Right hand in" and "Right hand touching face" is of less importance than the classification of each gesture into the categories: illustrator or adaptor gestures.

Automatically classifying gestures as illustrators or adaptors is extremely difficult with blob analysis. The data from each blob is limited to a center point, height, width, and angle of the main axis. With such limited data a hand scratching the chest would look very similar to an illustrator gesture displayed in front of the body. However self-adaptors are usually confined to movement near the body, while illustrator gestures can be expressed near the body or away from the body.

This difference between illustrator gestures and adaptor gestures has led to the idea of dividing up the space on a video frame and tracking the hands as they enter, remain within, and exit a certain area. Of most importance is when the hands are close to body and when they are extended and away from the body.

5.3.1 Using quadrants for identifying movement. In order to determine when the hands are close to the body and when they are extended, we use the blob data to make approximations about the location of the body. The head is the anchor point for identifying the body. Once the head is identified using blob analysis, the rest of the body's location is inferred by using the center point, height, and width of the head blob to create quadrants as shown in Figure 6.

The formulas for determining the four quadrants are show in Table 3 where x and y are the center point, h is the height and w is the width of a blob. All measurements are done in pixels where the first pixel in the upper left corner is 1, 1 and values increase to the right along the x axis and down the y axis.

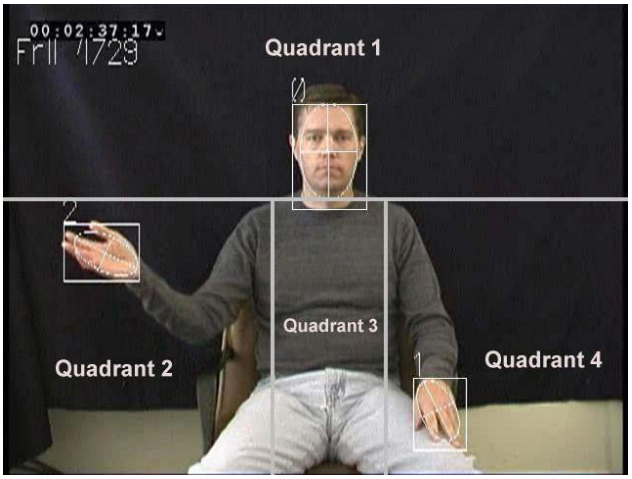

Figure 6. Quadrants calculated from the head blob

In Figure 6, the right hand can be classified in quadrant 2 because the y value of the center point of the blob identifying the right hand is greater than the y value of the center point of the head blob minus one half of the height of the head blob and the x value of the center point of the right hand blob is greater than the x value of the head blob minus the width of the head blob.

Table 3. Calculations of quadrants based on head blob data

Quadrant Number	Boundaries of Quadrants
Quadrant 1	$y < y_{\text{head blob}} - h_{\text{head blob}} / 2$
Quadrant 2	$y < y_{\text{head blob}} - h_{\text{head blob}} / 2$ AND $x < x_{\text{head blob}} - w_{\text{head blob}}$
Quadrant 3	$y < y_{\text{head blob}} - h_{\text{head blob}} / 2$ AND $x > x_{\text{head blob}} - w_{\text{head blob}}$ AND $x < x_{\text{head blob}} + w_{\text{head blob}}$
Quadrant 4	$y < y_{\text{head blob}} - h_{\text{head blob}} / 2$ AND $x > x_{\text{head blob}} + w_{\text{head blob}}$

The calculation of quadrants and tracking the position of the hand blobs as they move between quadrants uses a set of simple equations which are quickly executed. Using the data from the head blob to create quadrants to classify the position of hands is portable across various distances. Since bodily proportions remain consistent with distance the quadrants would not be affected if the camera were near or further away. Further, the quadrants are calculated for each frame individually so if a person takes a step to one side, the quadrants will adjust according to the location of the head.

However, using the center point, width, and height of the head blob to approximate the position of the body carries some risks. Measurements can be incorrectly altered by head movement such as head tilting and looking side to side. Measurements can also be influenced by camera positioning. If the camera is at an angle, the boundaries of the quadrants will be altered due to the orientation of the head in relation to the camera. Defining hand position by quadrants is most reliable when the person is directly oriented to the camera.

5.3.2 Quadrants in deception detection. To explore the use of blob analysis and quadrants in the identification of deception, a set of 28 clips were collected from the Mock Theft Experiment. The clips contained answers to interview questions where the respondent is either truthful or deceptive. Seventeen responses were truthful and 11 responses were deceptive. Each clip was segmented to contain the answer to only one question. The answer lengths varied from 4 seconds to 41 seconds, with a mean length of 24.5 seconds. The angle of the camera was slightly to the right of the interviewer so interviewee does not have a direct orientation to the camera. However, the camera angle was consistent for all interviews, thus allowing for comparability.

The following measurements were taken for each hand blob: total time in each quadrant, normalized time in each quadrant based on clip length, total distance the center point moved, normalized distance the center point moved, total time the hand blobs were together, and normalized time the hand blobs were together.

The following measurements were taken for the head blob: total distance the center point moved, normalized distance the center point moved, total time the head blob and a hand blob were together, and normalized time the head blob and a hand blob were together.

5.3.3 Results. Analysis of the measurements between truthful responses and deceptive responses indicates the normalized amount of time the deceivers' left hand was in quadrant 2 and quadrant 3 was greater than truth tellers' (quadrant 2 $p < 0.05$, quadrant 3 $p < 0.05$) Further the normalized amount of deceivers' head movement was less than truth tellers' ($p < 0.05$). While not significant, another interesting finding is that the normalized amount of time deceivers kept their hands together was greater than truth tellers ($p < 0.10$).

The amount of time deceivers' left hands remained in quadrant 2 (the quadrant covering the body) and the amount of time the deceivers' hands were together support the idea that deceivers gesture less than truth tellers and that such a difference can be automatically identified. The difference in head movement supports other studies analyzing deceivers' movements and may be a worthwhile measurement to include in the behavioral profile [13]. The difference in the amount of time

deceivers' left hands remained in quadrant 3 was unexpected and may have been influenced by the angle of the camera.

Clearly, automatically creating a kinesic behavioral profile from gestures as required by the decision model in Figure 3 is very difficult. While these proof-of-concept studies are simplistic in their approach, they do show that such an approach may be possible. In order to gain more acceptable results a larger datasets must be analyzed and camera angles must be carefully controlled.

6. Future steps

Future efforts to expand our understanding and ability to detect hostile intentions will include developing a more comprehensive data set for establishing hostile and benign behavior and creating methods for combining multiple cues in a more robust model of intent. By considering multiple cues from different communication channels and by using realistic data, our accuracy in detecting deception and hostile intent may increase.

6.1 Border crossing experiment

In collaboration with U.S. Customs & Border Protection (CBP), the Center for the Management of Information at the University of Arizona is developing the capability to conduct studies and establish baselines for specific behaviors. These studies will be extremely rich in behavioral cues and will provide another ecologically valid test bed where subjects should have high motivation due to serious possible consequences.

With data from contextually valid sources such as interviews with people who have high motivation to deceive, one could investigate cues that officers use to determine probability of hostile intent. Rich data sets may also be available for training and testing machine learning tools in applicable settings where contextual constraints such as lighting, space and equipment issues are present. Finally, one could validate past research in an actual setting where deception and interaction with people harboring hostile intent occurs each day.

6.2 Fusion of multiple cues

The development of a fusion engine to combine the nonverbal cues with previously identified text-based and audio indicators will strengthen reliability in intent detection. Developing a system that combines multiple data sources to detect intent is an extreme challenge but we expect it will have high payoff. The data streams from each type of indicator of intent will be fed into the fusion engine and the engine will merge the probabilities of hostile intent based on the weight of reliability for each method.

The fusion engine should also be adaptable to various conditions and contexts where one indicator may be more reliable than another or where a data stream may be unavailable. The result of fusing multiple indicators together should be a single indicator of hostile intent combined with a confidence level. An observer may then use the intent indicator and confidence level when forming opinions about hostile intent.

7. Conclusion

To alleviate some of the problems associated with high information assurance, the development of a tool to assist humans in judging intent is proposed. The creation of a tool is possible through adherence to a theoretically-based model and use of representative data sets. The proposed model for identifying hostile intent is a combination of three theories from various disciplines and it will be used in a contextually rich test bed. The approach in the model has been initially validated in the proof-of-concept studies and although the proof-of-concept studies presented here are a small first step, our approach shows promise in addressing the challenge of intent determination.

8. References

[1] D. Buller and J. Burgoon, "Interpersonal deception theory," *Communication Theory*, vol. 6, pp. 203-242, 1996.

[2] L. Zhou, D. Twitchell, T. Qin, J. K. Burgoon, and J. F. Nunamaker, Jr., "An exploratory study into deception detection in text-based computer-mediated communication.," presented at 36th Hawai'i International Conference on System Sciences, Big Island, HI, 2003.

[3] L. Zhou, J. K. Burgoon, D. Twitchell, and J. F. Nunamaker, Jr., "Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communication," *Group Decision and Negotiation*, In Press.

[4] J. K. Burgoon, J. P. Blair, and E. Moyer, "Effects of Communication Modality on Arousal, Cognitive Complexity, Behavioral Control and Deception Detection During Deceptive Episodes," presented at Annual Meeting of the National Communication Association, Miami Beach, Florida, 2003.

[5] J. K. Burgoon, J. P. Blair, T. Qin, and J. F. Nunamaker, "Detecting Deception Through Linguistic Analysis," presented at NSF/NIJ

Symposium on Intelligence and Security Informatics, 2003.
[6] J. George, D. P. Biros, J. K. Burgoon, and J. Nunamaker, "Training Professionals to Detect Deception," presented at NSF/NIJ Symposium on "Intelligence and Security Informatics", Tucson, AZ, 2003.
[7] J. K. Burgoon, D. B. Buller, L. K. Guerrero, and W. A. Afifi, "Interpersonal deception: XII, information management dimensions underlying deceptive and truthful messages," *Communication Monographs*, vol. 63, pp. 50-69, 1996.
[8] J. K. Burgoon, "A communication model of personal space violations: Explication and an initial test," *Human Communication Research*, vol. 4, pp. 129-142, 1978.
[9] a. S. J. A. Green D. M., *Signal Detection Theory and Psychophysics*: New York: Wiley, 1966.
[10] H. Stanislaw, Todorov, N, "Calculation of signal detection theory measures," *Behavior Research Methods, Instruments, & Computers*, vol. 31, pp. 137-149, 1999.
[11] P. Ekman, *Telling Lies: Clues to Deceit in the Marketplace, Politics and Marriage*. New York, New York: W. W. Norton & Company, 1985.
[12] A. Vrij, K. Edward, K. P. Roberts, and R. Bull, "Detecting Deceit via Analysis of Verbal and Nonverbal Behavior," *Journal of Nonverbal Behavior*, vol. 24, pp. 239 - 263, 2000.
[13] A. Vrij, *Detecting Lies and Deceit: The Psychology of Lying and the Implications for Professional Practice*. West Sussex: John Wily & Sons Ltd, 2000.
[14] K. Imagawa, S. Lu, and S. Igi, "Color-Based Hands Tracking System for Sign Language Recognition," presented at Proceedings of 3rd International Conference on Automatic Face and Gesture Recognition, 1998.
[15] S. Lu, D. Metaxas, D. Samaras, and J. Oliensis, "Using Multiple Cues for Hand Tracking and Model Refinement," presented at IEEE CVPR 2003, Madison, Wisconsin, 2003.
[16] S. Lu, G. Tsechpenakis, D. N. Metaxas, M. L. Jensen, and J. Kruse, "Blob Analysis of the Head and Hands: A Method for Deception Detection," presented at Hawaii International Conference on System Science (HICSS'05), Hawaii, 2005.
[17] C. R. Wren, A. J. Azarbayejani, T. Darrell, and A. P. Pentland, "Pfinder: Real-Time Tracking of the Human Body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 780-785, 1997.
[18] Y. Bengio, *Neural Networks for Speech and Pattern Recognition*. London, United Kingdom: International Thompson Computer Press, 1995.
[19] J. Elman, "Finding Structure in Time," *Cognitive Science*, vol. 14, pp. 179-221, 1990.

9. Acknowledgements

Portions of this research were supported by funding from the U. S. Air Force Office of Scientific Research under the U. S. Department of Defense University Research Initiative (Grant #F49620-01-1-0394) and Department of Homeland Security - Science and Technology Directorate under cooperative agreement NBC2030003. The views, opinions, and/or findings in this report are those of the authors and should not be construed as an official U.S. Government position, policy, or decision.