Using nonverbal video-features of players as an indicator of a social bond between deceiving players in a game of Werewolves.

Coen Dekker coen.dekker7@gmail.com,
Dirrik Emmen dirrik@gmail.com,
Natalia Rios Rivera natalia.rios.rivera@gmail.com,
Sabrina Verhage sabrinaverhage@gmail.com

17 November 2014

Delft University of Technology MMI Department, Leiden University Media Technology Department, The Netherlands

Under supervision of:
Joost Broekens, http://www.joostbroekens.com/

Abstract
Collaborative deceiving roles in a multiparty role playing game may be identified by analyzing the behavior of the collaborating subjects. This paper explores an automated nonverbal affective system for detecting collaborative deceiving behavior. Assuming people attempt to control behavior so as to maintain deception the system is focused on analyzing features of gross body movements of a deceiver listening to its 'partner in crime'. The game of Werewolves was played to test this assumption. The recorded games were analysed and converted to a workable dataset. This paper discusses all the steps of the research including the techniques used for doing the experiments, working with raw data towards a dataset, and analysing that dataset. It is shown that this data can be used and interpreted in several directions.

Introduction
In everyday life people usually understand any situation, position or movement of nonverbal behavior. The nonverbal behavior that a receiver signifies is conceptualized by Ekman [8] as coding; "...the difference of correspondence between the act and its meaning". The information provided from nonverbal behavior bring people insights to be aware of certain roles and identities of the actors in the environment. However, deceptive cues require more attention to be recognized by the observer [9].
It has been proven that people act different when they are lying than when they are telling the truth. In 1981, Zuckerman, DePaulo and Rosenthal published the first exhaustive meta-analysis of cue to deception in which they reported the differences occurred during deceptive communications compared with truthful ones [21]. This gives rise to the question if people also act different when they are listening to lies in which they are involved.
The cues that people construe to code nonverbal behavior are generally visual. The encoding process primarily covers the interpretation of proxemics [13] which is the use

of space within interpersonal communications, chronometric [19] which is the time involved in nonverbal communication and facial expressions, gestures and postures analysis [8]. Zuckerman considers hands as illustrators which bring movements that are fully linked to speech [21]. The head and hands provide insights on the nature of the emotion while body provide information about the intensity of the emotion [6]. Additionally, head and hands have been proved as cues of actual indicators of nonverbal deception [16]. Consequently, head and hands movements are considered to provide cues on deceptive behavior in multiparty role playing games.
Different approaches of nonverbal deceptive behavior have been defined [20]. The *emotional approach* proposes that a liar may feel fear, guilty or excitement and the emotional state will provoke concrete behaviors (e.g. guilt evokes gaze aversion). The *cognitive approach* proposes that performing a lie requires a more cognitive complex task. Finally, the *attempting control approach* suggests that a liar is likely to control his behavior in order to look normal. An example of this conception is that liars think that movements performed during their lie can make them appear more suspicions and therefore, they tend to avoid movements that are not essential for their speech. Consequently, liars movements occur in an unusual degree of rigid and inhibition body movement performance during speech [20].

This paper is focused on multiparty role playing games, specifically the game of Werewolves. Using the *attempting control approach* it relies on the claim that communication is not a unidirectional activity in which the sender creates a deceitful message and the receiver is absolutely passive. In contrast, it buttress the understanding that deception is a dynamic communication event where deceptive patterns can be found either in the sender or the receiver and in which the dynamics are explained from the principles of the Interpersonal Deception Theory (IDT) [1]. IDT reveals how individuals perceive deception in conscious and unconscious levels by means of interaction. This theory analyzes the process during communication, in which the sent message is influenced by the behavior and reactions of the recipient, and vice versa. The theory is concerned with the recipients possible awareness of the senders deceptive role.
Furthermore, this paper is focused on detecting deceptive roles in videos by analyzing gross movements. The analysis of body posture and facial expression are discarded since they require heavy computational processing and also, facial expressions are not always in a fixed position in front of the camera range. This project is different from other nonverbal approaches since the detection of deceptive roles is based on listeners data analysis rather than on speaker data analysis.

1.1 Related work

Research regarding deception techniques has evolved over time. The first insights were psychology experiments that provided cues on deceit based on observers' skill to detect deception [5, 7, 14]. Then, multiple studies have provided evidence that deceivers act different from truth tellers and differences include the lack of head movement [2].
Meservy [15] proposed a video-taped interaction system to obtain accurate indicators of nonverbal deception based principles of the Interpersonal Deception Theory (IDT) [1]. The system gathers data from video input and also considers inter-relational features

from IDT. The system focused on head and hands cues of deception. The research approach categorizes nonverbal behavior from the video frames input and is classified into patterns through innovative methods. The classification is as follows: the raw data is split into discreet units, the general metrics are extracted from discreet units, the features are extracted and inferred from the metrics and finally, selected features are used to classify each unit. In this sense the system performs a really sophisticated classification outside the scope of this project. Additionally, a method called "blob analysis" (Figure 1 and 2) was used to position the hands and head and from there calculate the center point, define the lengths and angle of axis and to determine the meaningful features.

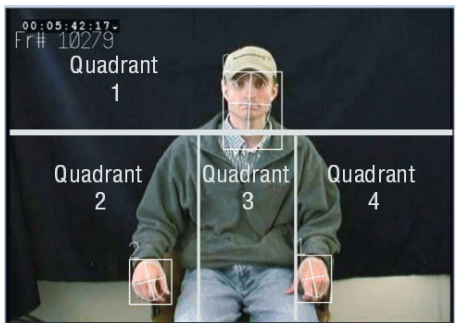

Figure 1: sample frame in which hands and face are identified [15].

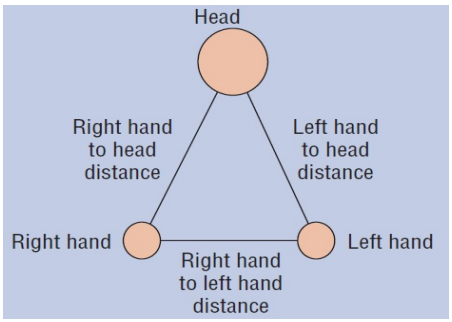

Figure 2: distances between blobs [15].

In 2010 Hung [11] looked at dominance in multiparty conversations using the multiparty role-playing game 'Werewolves'. The research focused on audio signals from the participants speech. The features analysed were the speaking length, the number of turns, number of successful interruptions, number of times being interrupted, and the number of 'speaking first' to predict deceiving roles. Even though the research did not analyse body behavior, it did provide insights on how to examine natural interventions in audio-visual deceptive behavior during the Werewolves game. Additionally Hung created a Werewolf database, the 'Idiap Wolf Corpus', available for acquisition online [12]. Later, a thesis student Raiman [18] coached by Hung, developed an automated nonverbal audio-visual system to extract deceptive features from motion and speech during a multiparty conversation. For this he used Hung's 'Idiap Wolf Corpus' Werewolf database [12]. The system looked at gross body movement of the players to predict deceiving roles. Inspired by Meservy, Raiman divided the body into multiple areas. Raiman however tracked the head and not the hands, he proposed a complementary division of the screen where the difference between 'Body Left' (BL) and 'Body Middle Left' (BML) indicates closed and open posture (Figure 3).

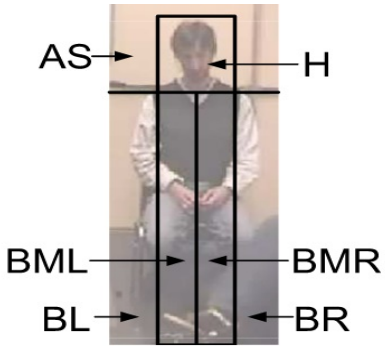

Figure 3: Regions of Raiman's body division [18]

Another work that is related to the research in some extend is the embodied agent proposed by Heylen [10]. The listening behavior was implemented in a semi-autonomous embodied chat-robot that was part of the Sensitive Artificial Listener. The interesting argument that concerns to the Werewolf research is that it addresses attention in listening as a multi modal phenomenon and it faces the difficulties of automatically extracting subtle patters in nonverbal behaviors displayed by listeners in a multi-party conversation.

2 Research question

Research by Hung [11] and Raiman [18] suggests that participants in a game of Werewolves move and talk different depending on the role they are playing. In 2012 however, Hung mentions [4] having found evidence that werewolves also act different when another werewolf is leading the conversation (Figure 4). Hung's hypothesis is that "...in multiparty settings, liars who collaborate to dupe others behave differently in their nonverbal behavior, in when they choose to join a conversation, and even in how they behave while their "partners in crime" are speaking." [4]

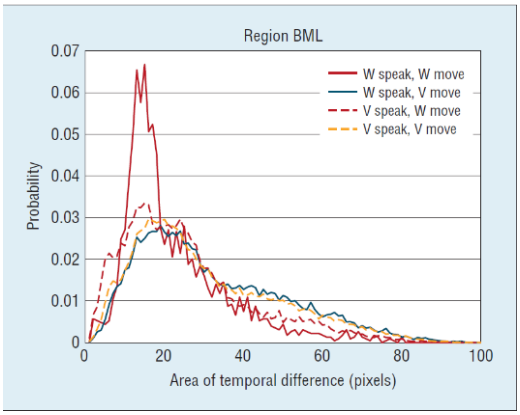

Figure 4 [4]: Hung's graph showing that werewolves move different when another werewolf is leading the conversation

The graph in Figure 4 shows a clear peak in the probability of low visual activity, suggesting the behavior of one deceptive speaker influences the behavior of the other. Indeed when one werewolf is talking the other has to pay attention; they are both collaborating in a deceiving scheme and want to win the game. Thus can, by analysing players movements, be suggested that the behavior of one deceptive speaker influences the behavior of the other:

"Can the nonverbal video features of players be used as an indicator of a social bond between deceiving players in a multi-party conversation?"

2.1 Hypotheses

Multiple researches have provided insights on deceptive nonverbal behavior while a liar speaks [5, 7, 14, 20, 21] and other researches claim on the importance of extracting broad features during the communication process to accomplish a multidimensional research [1, 2, 3]. Hung's suggestion that the behavior of one deceptive speaker influences the behavior of the other is supported by the earlier mentioned *attempting control approach* of nonverbal behavior proposed by Vrij [20]. This implies that liars tend to control their behavior so they do not appear suspicious to their listener. Liars can feel guilt or fear and so their emotional state makes them inhibit their movements when they speak. Additionally, evidence has provided insights that determine that liars move their head less than truth tellers [2]. One can say that the same conclusions are likely to happen when the liar is playing the role of receiver in a conversation since the emotion will not immediately finish but continue overtime.

Moreover, Burgoon [3] studied nonverbal behavior dynamics based on the IDT analysis and provides evidence and understanding of deception and honest behavior in valuable ways. The transmit signals are metaphorized as a game of movements by deceiver and deceived and is addressed in a meaningful approach for the current research. The game metaphor implies the strategic behavior of the deceiver is performed in a multifunctional, multidimensional, multimodal and multivariate approach during communication. The assumption that receivers are not passive but instead active message-processors from which a deception profile can be also analyzed is a center point in the current proposal. Additionally, the above mentioned study provides evidence on reciprocal patterns of deception so that, deceivers expressed more submissive movement patterns than truthful ones [3]. Therefore, the study conclusions imply that deceivers tend to reduce and control their information behavior.

Furthermore, Hung suggests that people sharing a lie within a group, move less when one of the liars is speaking [11]. With these findings in mind it is expected that the deceiving players in this research will actually move less in relation to the other players, and specifically the head should move less.

3 Method

In order to study the hypotheses four different cases during a Werewolf game were looked at. These four cases are inspired by Hung [4] and are as followed:

- The movements of a Werewolf when another Werewolf is talking.
- The movements of a Villager when another Villager is talking.
- The movements of a Werewolf when a Villager is talking.
- The movements of a Villager when a Werewolf is talking.

These cases were studied through an experiment in which eight people are playing a game of Werewolves. Although Hung created the 'Idiap Wolf Corpus' database [12], the decision was made to do a new experiment. This because Hung's video material was of low quality.

3.1 Materials

For this research one Sennheiser wireless microphone was used per player and one for the narrator. Besides this four GoPro Hero 3+'s were used. Each GoPro Hero 3+ captured two players. The microphones were synchronized to a central recording device while the camera's were set up separately.

In order to extract the movement of a player a system was developed by using 'Max/MSP/Jitter' to analyse the difference in pixels between frames. First the system located the body using 'cv.jit'; computer vision for Jitter [17]. Considering head and hands provide insights on the nature of the emotion while body provide information about the intensity of the emotion [21]. The decision was made to divide the body into multiple segments similar to Raiman [18]. The cv.jit.faces function enabled us to locate the face. Whenever the face was not detected it took the last position of the face(e.g. when looking to the side). Once the face was located the body was divided in four segments, namely the 'head', 'body', 'outside-head' and 'outside-body' as displayed in Figure 4. The average width of the extracted body was set to three times head width. The average height of the extracted body was set to two times head height. The space outside of the extracted head and body positions were named 'outside-head' and 'outside-body'. Movements captured in the outside areas relate to a more open posture while the 'head' and 'body' movements relate to a more closed posture [18].

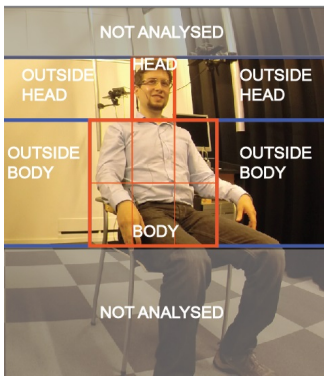

Figure 4: Division of body segments

For extracting body motion, frame subtraction was used, also known as temporal difference. This was done by feeding the video's in 30 frames per second, in grey-scale, to the system which subtracted consecutive frames to see a change in pixels. This value was treshholded and translated to a binary value, with a value of '1' indicating the pixel has changed or 0 if the pixel did not changed. The system analyzed each frame per video per player. The amount of pixels that changed each frame were translated to percentages of the total pixels in the body segment.
Besides this the same system analyzed the audio of the player simultaneously, the audio was synchronized with the video beforehand. Since there are unique audio samples for each participant the decision was made to relate sound intensity to what sound is closest to the microphone; in this case the participant self. If a collection of high intensity peaks is longer than 1000ms the assumption is made that the participant is speaking. The output of this analysis was represented in a binary format; for each frame a value of '1' indicates this participant was speaking. The output was checked and verified by manually labeling a game.

3.2 Experiment

In this chapter the procedure of the invitation, experimenting and the Werewolf game is discussed.

3.2.1 Participants and Procedure

With four unique groups of eight people seven games of Werewolves were played. The decision was made to play several rounds with the participants because the roles were distributed randomly every round which changes the game completely.
The participants were gathered at the faculty of Electrical Engineering, Mathematics and Computer Science from the TU Delft. The participants volunteered to participate as players and they were asked to sign a consent form where they agreed to be recorded during the experiment. The experiments were conducted at the INSYGHTlab at the same faculty. Before each game the participants were numbered, this way they could refer to each other by number so they didn't have to remember names. For each game the narrator explained participants the rules and once all agree on having understood the rules and felt comfortable then the game started and the equipment commenced gathering the data.

3.2.2 Experiment Setup

The setup consisted of a central table around which the eight players were seated as seen in Figure 5. Each side of the table had a GoPro camera mounted on top. The setup was enclosed by black curtains to ensure a more stable scene. All players were hooked up with a microphone and were in sight of one of the four GoPro camera's. The narrator was also hooked up with a microphone but out of sight of the GoPro camera's, narrating from behind the curtain.
Each game resulted in four video files, two players each, and nine audio files per game, one for each player and one for the narrator. Each video file was cut in two to create a unique video file per participant.

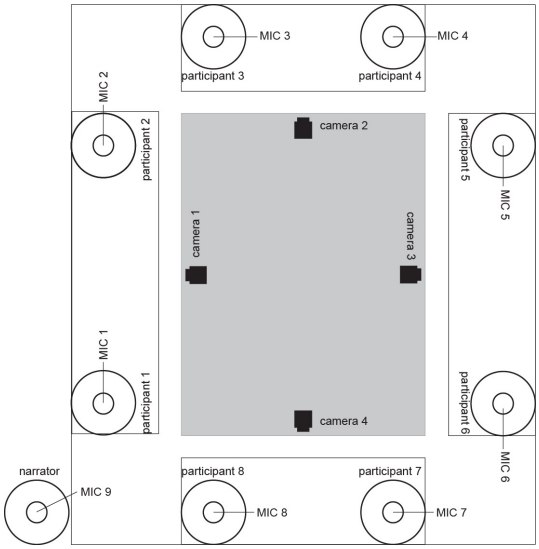

Figure 5. Setup used for the Werewolves experiment.

3.2.3. The Werewolf game

The Werewolf game is a role playing game for large groups and "is a game of accusations, lying, second-guessing, assassination and mob hysteria" [4]. This game was chosen on the basis of previous research by Hung [4][11][18]. The decision was made to simplify the game by removing special roles like the seer. This changes the game but does not change the behavior of the werewolves. It also allowed for new players to get easier into the game. The game now only has villagers and werewolves, which gives us a clear focus to research on.

The game proceeds as follows: players are villagers in a village. The village is terrorized by cursed villagers who turn into werewolves at night to eat one villager each night.

The game starts with all players getting a role card from the narrator at random, this card determines a player's role for this game. A player can either be: A villager (V) - finding out who the werewolves are or a werewolf (W) - deceive your fellow villagers and kill the town one by one.

Night phase: It is nighttime in the village, the narrator will ask all players to close their eyes and be quiet. Some villagers will turn into a werewolves and the narrator will ask them to wake up. The werewolves see each other and secretly 'discuss' which villager they will kill. They have to be unanimous and can only kill one villager. When a target is chosen the narrator will ask them to close their eyes again. The werewolves will turn into normal humans again. From here the players move on to the Day phase.

Day phase: It is daytime and the narrator will tell who was killed and ask everyone else to open their eyes. The killed villager turns his/her card but cannot speak anymore this game. For the remaining villagers the debate starts, who is a werewolf?! The debate can go on freely, until the narrator will ask everyone to move on to the voting.

It is time to vote! All villagers quiet down and raise one hand. The narrator counts down from 3 and on 0 everyone should point at the one they want to lynch this round. The

villager with the most votes is killed. Vote will be redone if there is a tie. The killed villager turns his/her card but cannot speak anymore this game. The remaining players move on to the next Night phase. The game continues this loop until all villagers or werewolves are killed. The narrator will inform the players when this is the case.

4 Results

In total the system processed 3 games creating 24 data files with 1 data row per frame of the player video resulting in approximate 30 000 rows per file. First the audio columns of each player was added in each file, describing per file which player was talking during a certain frame. After this the rows were reduced based on playing time of the player. From this two new data files were created per player based on the role they played, for each case we only used data if a single player was talking.

For a villager (V):
The movements of a V when another V is talking. *(VV)*
The movements of a V when a W is talking. *(VW)*

For a werewolf (W):
The movements of a W when another W is talking. *(WW)*
The movements of a W when a V is talking. *(WV)*

This resulted in 16 data files per game, creating 48 files in total. These data files were used to create only four data files by appending all VV tables with each other, all VW, all WW, and WV, creating four data files, containing all three games. Each file contained the movement values for head, outside-head, body and outside body.

For each of the four cases the first quartile, median and third quartile were 0.0.

For CW head μ=0.5509, outside head μ=0.03293, body μ=0.2556 and outside body μ=0.03878.
For CC head μ=0.7087, outside head μ=0.04037, body μ=0.3141 and outside body μ=0.05865.
For WC head μ=0.709, outside head μ=0.03996, body μ=0.1514 and outside body μ=0.04391.
For WW head μ=0.6992, outside head μ=0.01369, body μ=0.1421 and outside body μ=0.008191.
The maximum values varied between the four cases are shown in table 1. Density lines per case are shown in picture 6, categorized on body part.

	head	outside head	body	outside body
CW	52.5150	16.6470	40.5830	18.05900
CC	52.7940	17.00800	50.7470	33.78300
WC	55.096	16.28300	46.8170	14.80100
WW	40.5830	2.88700	32.5440	4.814000

Table 1. Maximum movement values per case.

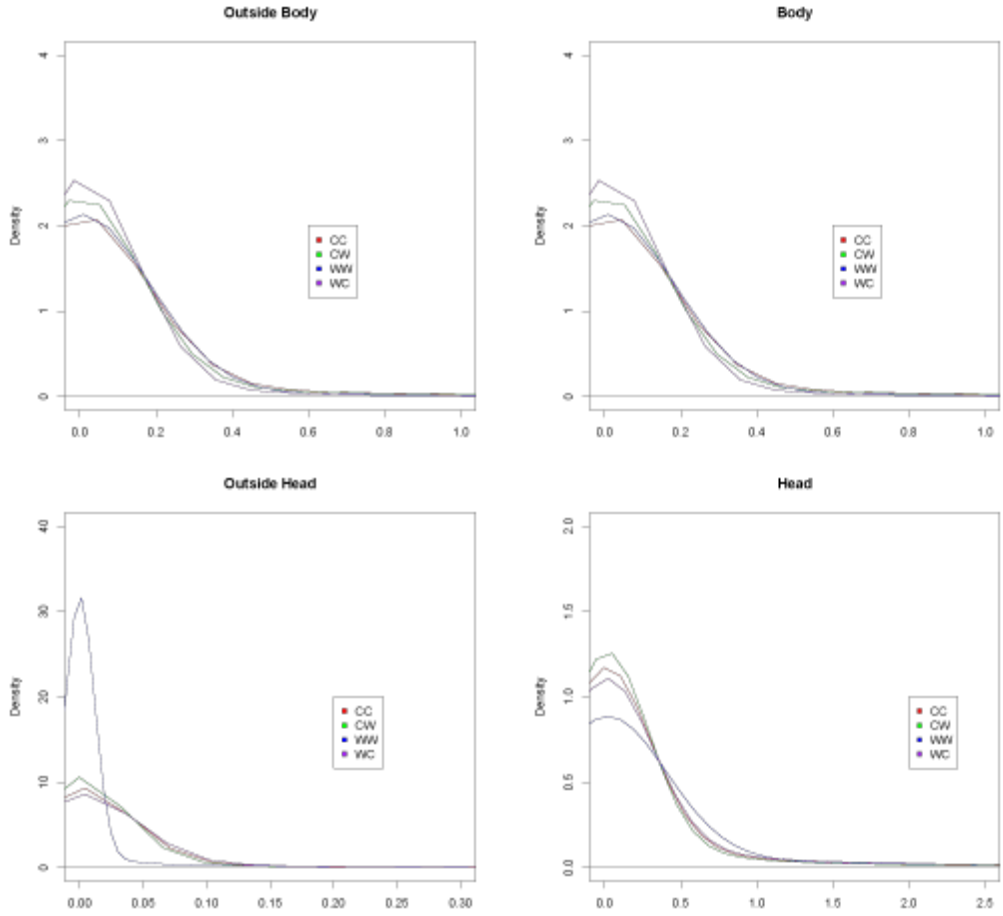

Figure 6. Density lines per case for movement values.

4.1 Discussion

An interesting note to take from the results is that the maximum movement values for a Werewolf listening to another Werewolf is remarkably less than the other three cases. Which could suggest less movement. Yet, when looking at the density lines it shows less lower values for body, outside-body and head whenever a Werewolf is listening to another Werewolf compared to the other 3 cases which could suggest *more* movement. This could be explained by that a werewolf might actually have less explicit and expressive movements, decreasing the maximum values, but more implicit movement increasing the overall movement.

Something to take into account is that since only three games are examined, it should be considered that head movements are influenced by looking at other players. The players next to the player leading the conversation the most are turning their head the most in order to look at him/her while speaking.

It also needs to be pointed out that the research did not use seven unique groups. The groups of 8 people therefore played two or three games, which means that they know each other a little bit better during the second and third game. Which could influence the way the game is played.

Lastly the test setup was simply structured, with the 4 GoPros it allowed to record almost 360 degrees. This wide angle also resulted in that the narrator could not move around freely during the day part of the game without disrupting the video footage. Despite hiding behind a curtain it might have occurred that the curtain moved a bit in the moment of hiding after a night part of the game.

5 Conclusion and further research

In this research it is shown how a system can gather and analyse data from a multiparty set-up. Multiple steps are given on how to handle and process the data. It is shown that this data can be used and can be interpreted in several directions. For future research it might be interesting to only look at explicit movements by tracking the hands.

Reflecting back at the research question if nonverbal video features of players can be used as an indicator of a social bond between deceiving players in a multi-party conversation - this research gives directions to possible indicators for deceiving roles in a game of Werewolf, such as less explicit movement and more implicit movement in relation to villagers. Yet more data should emphasize and prove this in a significant way. In addition the research provides potential novelty insights on inter-relational nonverbal features for projects that concern on the listener behavior during a conversation. Additionally, it has potential applications for questioning purposes in airports, events and so on. Further research might prove that the way people behave while listening to a question or an argument of a party-member can provide indicators of deceptive behavior. Ultimately leading to an affective system recognizing deceiving people in a conversation.

References

1. Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. Communication theory, 6(3), 203-242.
2. Buller, D. B., Burgoon, J. K., White, C. H., & Ebesu, A. S. (1994). Interpersonal Deception VII Behavioral Profiles of Falsification, Equivocation, and Concealment. Journal of language and social psychology, 13(4), 366-395.
3. Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. Journal of Nonverbal Behavior, 18(2), 155-184.
4. Chittaranjan, G., & Hung, H. (2010, March). Are you a werewolf? detecting deceptive roles and outcomes in a conversational role-playing game. InAcoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on (pp. 5334-5337). IEEE.
5. DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. Psychological bulletin, 129(1), 74.
6. Ekman, P., & Friesen, W. V. (1967). Head and body cues in the judgment of emotion: A reformulation. Perceptual and motor skills, 24(3), 711-724.
7. Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. Langley Porter Neuropsychiatric inst San Francisco ca nonverbal behavior research project.
8. Ekman, P., & Friesen, W. V. (1981). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. Nonverbal communication, interaction, and gesture, 57-106.
9. Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar.Psychological Science, 10(3), 263-266.
10. Heylen, D., Nijholt, A., & Poel, M. (2007). Generating nonverbal signals for a sensitive artificial listener. In Verbal and Nonverbal Communication Behaviours(pp. 264-274). Springer Berlin Heidelberg.
11. Hung, H., & Chittaranjan, G. (2010, October). The idiap wolf corpus: exploring group behaviour in a competitive role-playing game. In Proceedings of the international conference on Multimedia (pp. 879-882). ACM.
12. Idiap Research Institute, (2010) Practical description of the data, http://www.idiap.ch/dataset/wolf
13. Jones, S. E. (1971). A comparative proxemics analysis of dyadic interaction in selected subcultures of New York City. The Journal of Social Psychology, 84(1), 35-44.
14. Kraut, R. E. (1978). Verbal and nonverbal cues in the perception of lying.Journal of personality and social psychology, 36(4), 380.
15. Meservy, T. O., Jensen, M. L., Kruse, W. J., Burgoon, J. K., & Nunamaker Jr, J. F. (2008). Automatic extraction of deceptive behavioral cues from video. InTerrorism Informatics (pp. 495-516). Springer US.
16. Meservy, T. O., Jensen, M.L., Kruse, J., Burgoon, J.K., Nunamaker, J.F., Jr., Twitchell, D.P., Tsechpenakis, G., Metaxas, D.N.(2005). Deception detection through automatic, unobtrusive analysis of nonverbal behavior. Intelligent Systems, IEEE, 20(5), 36-43.
17. Pelletier, J.M., (2010), Computer Vision for Jitter: collection of Max/MSP/Jitter tools for computer vision applications. http://jmpelletier.com/cvjit/

18. Raiman, N., Hung, H., & Englebienne, G. (2011, November). Move, and I will tell you who you are: detecting deceptive roles in low-quality data. InProceedings of the 13th international conference on multimodal interfaces (pp. 201-204). ACM.
19. Roelofs, A. (2008). Attention to spoken word planning: Chronometric and neuroimaging evidence. Language and Linguistics Compass, 2(3), 389-405.
20. Vrij, A., & Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. Journal of Nonverbal Behavior, 20(1), 65-80.
21. Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. Advances in experimental social psychology, 14, 1-59.