Proxemics measurement during social anxiety disorder therapy using a RGBD sensors network

Julien Leroy, François Rocca, and Bernard Gosselin

University Of Mons, TCTS Lab Mons, Belgium

Abstract. In this paper, we focus on the development of a new ecological methodology to study proxemics behaviors. We based our approach on a network of RGBD cameras, calibrated together. The use of this type of sensors lets us build a 3D multiview recording installation working in various natural settings. The skeleton tracking functionalities, provided by the multiple 3D data, are a useful tool to make proxemics observation and automatically code these non-verbal cues. Our goal is to propose a new approach to study proxemics behaviors of patients suffering from social anxiety disorder to improve observation capabilities of the therapist with an unobtrusive, ecological and precise measurement system.

1 Introduction

The study of your spatial behavior can reveal lot of information about you. Indeed, even if people feel free to move, many rules unconsciously shape their evolution, their use of space. Speed, direction, or trajectory are controlled and comply with various patterns constrained by social mechanisms [1]. Proxemics is a domain of research that investigates the way people use and organize physical space around them while interacting with others. E.T.Hall, an American anthropologist, introduced for the first time this concept in his studies about the human behavior in public space [2] and defined it as [...] *the study of man's transactions as he perceives and use intimate, personal, social and public space in various settings* [...]. Since this first development, much research has been conducted in the field of psychology and social science related to the concept of proxemics and one of its most important aspects: the interpersonal distances. The study of social spaces and proxemics can obviously lead to a better understanding of social mechanisms and context but also to developments in many areas such as robotics, human-computer interactions, teaching methodologies, etc. This paper focus on the development of a new methodology to study proxemics behaviors that try to solve some of the main issues raised by ethologists:

1. Unobtrusive means and ecological conditions
2. Easy setting in various environments
3. Precise measurements
4. Simplify and accelerate experiments

Our solution is based on a framework using multiple RGB-Depth(RGB-D) cameras. This choice, lets us create large 3D observation scene coupled with possibilities of users and skeleton tracking to extract non-verbal cues linked to proxemics behaviors. The main element of our system is a network of RGB-D cameras calibrated together to realize ecological recordings of the interactions. On these recordings, we apply skeleton tracking algorithm to automatically extract information about the behaviors that can be used by a psychologist to code the interaction like proposed by [3]. This research and developments take place in a long-term study of proxemics behaviors in social anxiety disorder (DSM-IV 300.23), also known as social phobia. Our goal, with this new methodology, is to create a computer vision tool to assist and help the psychologist in the treatment of patients suffering of this type of disorder. Firstly, by giving an interactive visualization of proxemics cues to the patient and his therapist, to help the analysis of the interaction by modeling interpersonal distances. Secondly, to precisely record the spatial behavior of patients through time and see if these proxemics information can be used as indicators of the patient's progress through his therapy.

This paper is organized as follows. Section 2 introduces the main concepts of proxemics. Section 3 presents a state-of-the-art of the methodologies and settings use to study proxemics behaviors. Section 4 details our system and how we tried to solve the main problems in ethological experiment. Section 5 reports some experiments on which we apply our methodology and section 6 draws some conclusions.

2 Proxemics fundamentals

Proxemics is the study of how man unconsciously organize, share, use physical space and the underlying meaning of these spatial behaviors both on social interactions and human psyche. The most important aspect of proxemics behaviors is probably the notion of interpersonal distances, these areas unconsciously established by people between them. The first part of this section will give some enlightenment about this social signal. The second part will be dedicated to detail some factors that can be used to caractherize proxemics interactions and we want to put in evidence with our methodology.

2.1 Interpersonal distances

People create unconscious territories around them, which define and determine the interactions they can have with other individuals. Those territories are like some invisible bubbles surrounding them and keeping them far from each other, unless space has some physical constraints (small room, crowded environment...). Interpersonal distances are a form of non-verbal communication between two or more persons defined by the social relationship they have. In a way, the measure of these distances is a clue that can tell us how people know each other. E.T.Hall has proposed a first model, based on the study of the spatial behavior of the

American middle class people, that divides the space around a person in four distinct regions:

1. Intimate distance (0 to 45 cm): a really close space with high probability of physical contact. It's a distance for touching, whispering or embracing someone. It indicates a close relationship like with lovers or children.
2. Personal distance (45 cm to 1.2 m): distance for interacting with relatives like family members or good friends. Unrequested penetration of this space will provoke discomfort, defensive postures and even avoidance behaviors.
3. Social distance (1.2 m to 3.5 m): distance for more formal or impersonal interactions. It's the distance you naturally pose when you meet stranger and establish a communication process with them.
4. Public distance (3.5 to infinity): distance for mass meeting, lecture hall or interactions with important personalities.

People update, control and adjust these spaces continuously. Obviously social spaces are much more complex social mechanisms and depend on many parameters. They continuously evolve and adapt to people circumstances. They should be seen as dynamic and elastic territories varying with lot of parameters like: culture, sex, age, gender, size, social position, relationship or physical appearance. Social sciences have already explored a lot this subject and showed the importance of this concept for explaining how people behave but also in some medical case like with schizophrenic patients [4]. Many behavioral experiments have shown the importance and impact that these distances could have on our actions. An example is the unauthorized penetration of these territories that will cause a feeling of discomfort for people. This can lead to an aggressive response of the subject who may feel oppressed by the presence of the intruder. In the case of social anxious patients, this kind of behavior is often accentuated. The main feature that appear about their proxemics behavior is a tendancy to have a bigger personal space, a larger confort zone that keeps them away from social interraction but also generates a hight level of anxiety when people invade their territory. An other observed defensive behavior is the body orientation where they try to not stay in front of the others. Situation understood like a miss of implication in the social interaction.

2.2 Visual cues to code proxemics interactions

E.T.Hall proposed a system to code proxemic behaviors [3] like a function of height indicators. In our methodology, we were only interested by some visual indicators that we wanted to highlight and measure:

1. Postural identifer. Minimal information about the position of a subject: prone, sitting or standing.
2. Sociofugal-sociopetal orientation (SPF axis). It represents the spatial position, orientation that pushes or pulls people to interact together. It's a function of the bodies' orientation, shoulders position between each person. Hall proposed to code this on a height positions compass. Two subjects

face to face would be maximum sociopetality and back to back, maximum sociofugality with intermediate possibilities.
3. Vision. Here is the gaze orientation that is use like an factor of involvement. Hall propose three zones base of the anatomy of the retina: fovea (direct gaze, 12 degrees horizontal visual angle), macula (20 degrees), peripheral vision (up to 180 degrees).

3 Proxemics measurement systems: state-of-the-art

Until recently, measurements on proxemics behaviors were performed and coded manually. Some resort to the use of video recordings using special rooms with for example a floor grid [5]. This method lets them measure, a posteriori, interpersonal distances and code proxemics behaviors but often with bias due to unnatural settings and the conditions of the experimentation. More recently, an approach using virtual reality appeared [6]. This method gives good results because people tend to behave the same way in virtual world and in the real world [7]. As it is based on projection and an heavy equipment is often installed on the patient, it is not obvious that they can be considered ecological measurement. To our knowledge, few computer vision systems have been designed to help solve these issues. [8] and [9] proposed two solutions using cameras in a bird-eye setting. This is an interesting solution but difficult to apply in various environment like a classroom or a meeting room.

4 Our approach for analysing proxemics interactions based on multiple RGB-D sensors

Our system is divided into recording units, each consisting of a computer, connected to one or more RGB-D cameras. To obtain a 3D multi-view system, Figure 1, that lets us observe behavioral scenes, each camera must be calibrated in space but also in time to avoid delay when we merge the different recordings. The spatial calibration can be performed in various ways: by manually manipulating point clouds or by using a SLAM approach [10] that allows us to know at any moment the position of the sensor.

The cameras network operates with a master-slave architecture on a LAN network, Figure 2. Each recording unit is a slave waiting UDP command from an distant master. The master is used to send a startup message on the network to initialize the recording. During the recording, a log file is created containing the timestamp of each saved frame. With this method, records can be resynchronized if it's necessary.

4.1 Data extraction and analysis

On each record is applied a skeleton tracking algorithm [11]. Then, knowing the camera positions, all the skeletons are remapped in the same world coordinate

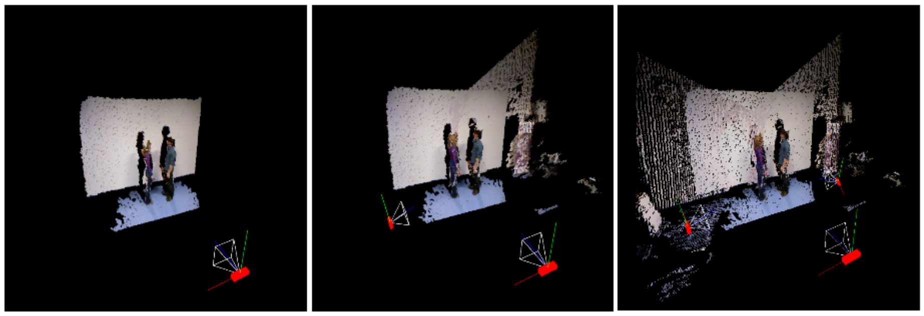

Fig. 1. Example of 3D multiview of the same proxemic scene with one, two and three calibrated sensors.

frame. In the case where a user would have several skeletons due to overlapping areas, they can be merge by measuring the distance between similar joints. With the resulting multiview skeleton 3D data, Figure 3, we can automatically process some proxemics information:

1. Interpersonal distances: they can be measure classicaly with the distance between users' centroids or head-to-head distances.
2. Postural identifier can be obtain with geometric constrains on the skeleton. Measuring angles produced by the legs can detect a sitting position. The angle formed by the head-torso line and the floor plane provides an estimate of the inclination of the upper body. This can also be used for the kinesthetic factors.
3. Sociofugal-socialpetal axis can be compute using the 3D shoulders' positions projected on the floor plane. It's the angle between the shoulders' axis of two interacting people.
4. Vision factor. To our knowledge, it is not actually possible to extract the gaze direction with RGB-D camera in our operating conditions. So we try to at least extract the 3D head orientation that should remain a good indicator. For this if the user is not too far away from the sensor, we use a 3D head pose estimation algorithm [12] to follow the head movements.

5 Experiment and discussion

We tested our system in the context of behavioral group therapy of patients suffering from social anxiety disorder. The goal was to use it to film exercises already integrated into therapy and to discuss what additional possibilities of measurement and analysis our system could provide. The system consisted of two recording units. Two cameras were sufficient to cover the entire area where took place the therapy sessions. These group therapy sessions are dedicated to the understanding of non-verbal communication and the importance it can have

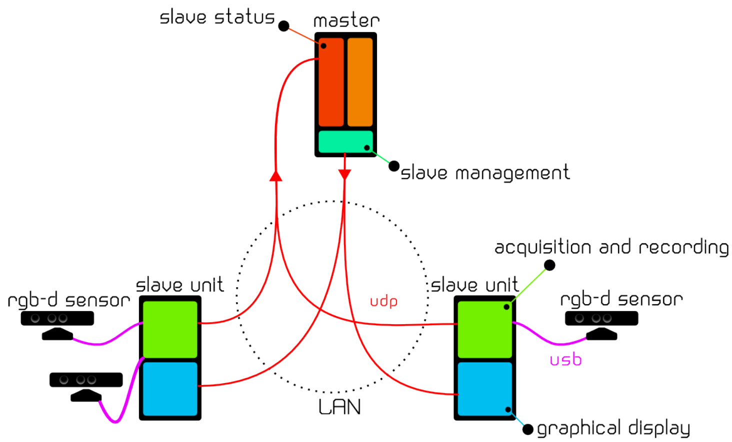

Fig. 2. Network infrastructure. One or two sensors can be connected to a slave unit. Record commands are send by a master unit.

in social interactions. They aim to provoke awareness in patients about this form of language and the problems that can affect it. Typically with such patients, it is the situations of discomfort and escape behaviors due to the proximity of others that are addressed. The recorded sessions dealt with territoriality, management of personal space and the physicality of the patient during dyadic interaction. Two exercises are performed by the therapist:

– the meeting. He asked two patients to simulate a meeting, typically a handshake, followed by a brief discussion of a few minutes. The aim is to observe the evolution over time of the spatial behavior and gestures as well as the involvement in the discussion.

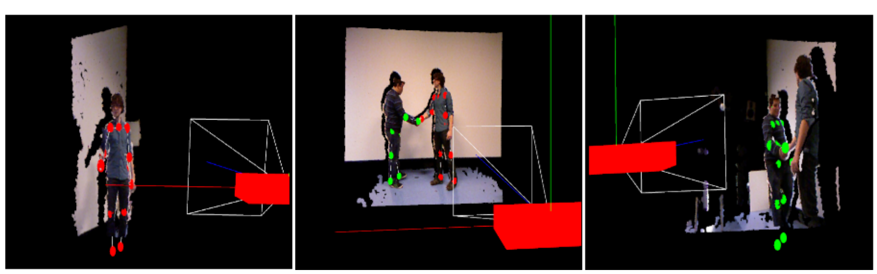

Fig. 3. Skeleton tracking applied on each view. These data are use to measure proxemics information like interpersonal distances or SPF orientation.

– th search of the personal space. This exercise is dedicated to the search of his comfort zone to interact with someone. He asked a patient to stand still back to a wall (to limit the escape behavior), then a second participant must move forward and find the most comfortable to begin a discussion, then the roles are reversed distance. The distances of the two patients are compared and lead a discussion about their feelings during the experiment.

The first observation made on the videos is the inability of the practitioner to detect and manage all behavioral signs shown. Indeed, by the time constraints, he is obliged to perform the exercises simultaneously with several groups of patients. It is therefore difficult for him to fully concentrate on the interactions and raise the indices as we can do with the video. The second observation is that patients tend to distort the observations by "cheating" during the exercises. Six different patients were registered, four men and two women, for a total of 8 different dyadic groups. For each group, three exercises were filmed. The sessions were recorded twice: first at the beginning of therapy and the second at the end (4 months with a weekly meeting). The objective of this double standard was to see if it was possible to measure behavioral changes in patients and the impact of therapy could have on them. We present below an example studied with our system. The most interesting patient, in a proxemics perspective, was a man of more than forty years (H1) presented by the therapist as having little presence and suffering from a lack of recognition towards others. The patient describes himself as "invisible". Its spatial behavior has the distinction of being invasive, without taking into account the behavioral response of others. For the therapist, it is the result of an affirmation need and an excessive difficulty in reading the nonverbal behavior of other participants. In our signals, it is denoted by:

– an personal space invasion of the other which also causes several escape
– almost no gesture

At the end of therapy, his behavior has changed, his interpersonal distance increased (11%), approaching the average distance, causing less discomfort with his partner. For other patients, the behavior is more classic with a tendency to approach the others, sticking to the model, as patients are much more familiar. These encouraging results suggest that it should be possible to quantify the impact that therapy could have in a proxemics view point.

6 Conclusion

In this paper, we propose a new methodology to study proxemics behaviors. The use of multiple RGBD cameras and the underlying tracking possibilities lets the experiment be natural, unobstrusive and precise. This should lead to good ecological data. The accuracy of our methodology makes it easy to identify patients with spatial behavior disorders but also to raise events that can escape the therapist. The initial results seem to show that it is possible to measure the impact of therapy on the patient behavior.

Fig. 4. Patient H1 (white), a meeting simulation. It is interesting in this case to observe the proximity of the two subjects, despite their stature and the little knowledge they have of each other. This proximity will result in H2 (red) several escape movements.

References

1. Murino, V., Cristani, M., Vinciarelli, A.: Socially intelligent surveillance and monitoring: analysing social dimensions of physical space. In: Science. (2010) 51–58
2. Hall, E.T.: The Hidden Dimension. Volume 6. Doubleday (1966)
3. Hall, E.T.: A System for Notation of Proxemic Behavior. American Anthropologist **65** (1963) 1003–1026
4. Nechamkin, Y., Salganik, I., Modai, I., Ponizovsky, A.M.: Interpersonal distance in schizophrenic patients: relationship to negative syndrome. The International journal of social psychiatry **49**(3) (2003) 166–174
5. Ickinger, W.J.: A study on simulation of proxemic behavior. Methodology (1994) 1–17
6. Wieser, M.J., Pauli, P., Grosseibl, M., Molzow, I., Mühlberger, A.: Virtual social interactions in social anxiety–the impact of sex, gaze, and interpersonal distance. Cyberpsychology behavior and social networking **13**(5) (2010) 547–554
7. Friedman, D., Steed, A., Slater, M.: Spatial Social Behavior in Second Life. Intelligent virtual agents **4722**(1 997) (2007) 252–263
8. Amaoka, T., Laga, H.: Personal Space Modeling for Human-Computer Interaction. ComputingICEC 2009 **5709** (2010) 60–72
9. Cristani, M., Paggetti, G., Vinciarelli, A., Bazzani, L., Menegaz, G., Murino, V.: Towards computational proxemics: inferring social relations from interpersonal distances. In: Context. (2011) 290–297
10. Henry, P., Krainin, M., Herbst, E., Ren, X., Fox, D.: RGB-D Mapping : Using Depth Cameras for Dense 3D Modeling of Indoor Environments. Work **1**(c) (2010) 9–10
11. OpenNI organization: OpenNI Programmer Guide. (November 2010) Last viewed 17-03-2012 12:22.
12. Fanelli, G., Weise, T., Gall, J., Gool, L.V.: Real time head pose estimation from consumer depth cameras. In: Proceedings of the 33rd international conference on Pattern recognition. DAGM'11, Berlin, Heidelberg, Springer-Verlag (2011) 101–110