Towards Computational Proxemics: Inferring Social Relations from Interpersonal Distances

Marco Cristani
University of Verona (IT)
Istituto Italiano di Tecnologia, Genova (IT)
Email: marco.cristani@univr.it

Giulia Paggetti
University of Verona (IT)
Email: giulia.paggetti@univr.it

Alessandro Vinciarelli
University of Glasgow (UK)
Idiap Research Institute (CH)
Email: vincia@dcs.gla.ac.uk

Loris Bazzani
University of Verona (IT)
Email: loris.bazzani@univr.it

Gloria Menegaz
University of Verona (IT)
Email: gloria.menegaz@univr.it

Vittorio Murino
Istituto Italiano di Tecnologia, Genova (IT)
University of Verona (IT)
Email: vittorio.murino@iit.it

***Abstract*—This paper proposes a study corroborated by preliminary experiments on the inference of social relations based on the analysis of interpersonal distances, measured with onobtrusive computer vision techniques. The experiments have been performed over 13 individuals involved in casual standing conversations and the results show that people tend to get closer when their relation is more intimate. In other words, social and physical distances tend to match one another. In this respect, the results match the findings of proxemics, the discipline studying the social and affective meaning of space use and organization in social gatherings. The match between results and expectations of proxemics is observed also when changing one of the most important contextual factors in this type of scenarios, namely the amount of space available to the interactants.**

I. Introduction

Proxemics can be defined as the "[...] *the study of man's transactions as he perceives and uses intimate, personal, social and public space in various settings* [...]", quoting Hall [1], [2], the anthropologist who first introduced this term in 1966. In other words, proxemics investigates how people use and organize the space they share with others to communicate, typically outside conscious awareness, socially relevant information such as personality traits (e.g., dominant people tend to use more space than others in shared environments [3]), attitudes (e.g., people that discuss tend to seat in front of the other, whereas people that collaborate tend to seat side-by-side [4]), etc..

This paper focuses on one of the most important aspects of proxemics, namely the relationship between physical and social distance. In particular, the paper shows that interpersonal distance (measured automatically using computer vision techniques) provides physical evidence of the social distance between two individuals, i.e. of whether they are simply acquainted, friends, or involved in a romantic relationship. The proposed approach consists of two main stages: the first is the automatic measurement of interpersonal distances, the second is the automatic analysis of interpersonal distances in terms of proxemics and social relations (see Section IV for details).

The choice of distance as a social relation cue relies on one of the most basic and fundamental findings of proxemics: People tend to unconsciously organize the space around them in concentric zones corresponding to different degrees of intimacy [1], [2]. The size of the zones changes with a number of factors (culture, gender, physical constraints, etc.), but the resulting effect remains the same: the more two people are intimate, the closer they get. Furthermore, intimacy appears to correlate with distance more than with other important proxemic cues like, e.g., mutual orientation [5]. Hence, it is reasonable to expect that the distance accounts for the social relation between two people.

One of the main contributions of the paper is that the experiments consider an ecological scenario (standing conversations) where more than two people are involved. This represents a problem because in this case distances are not only determined by the degree of intimacy, but also by the need of ensuring that every person can participate in the interaction. This leads to the emergence of stable spatial arrangements, called F-formations (see Section II for more details) [6], that impose a constraint on interpersonal distances and need to be detected automatically. Furthermore, not all distances can be used because, in some cases, they are no longer determined by the degree of intimacy, but rather by geometric constraints. The approach proposed in this work is to consider only the distances between people adjacent in the F-formation (see Section V for more details) [6].

The other important contribution is that, in contrast with other works in the literature, the radii of the concentric zones corresponding to different degrees of intimacy are not imposed a-priori, but rather learned from the data using an unsupervised approach. This makes the technique robust with respect to the factors affecting proxemic behavior, like culture, gender, etc., as well as environmental boundaries. In particular, the experiments show how the organization into zones changes when decreasing the space at disposition of the subjects and how the unsupervised approach is robust to such an effect.

Standing conversations are an ideal scenario not only because they offer excellent examples of proxemic behavior, but also because they allow one to work at the crossroad between surveillance technologies, often applied to monitor the behavior of people in public spaces, and domains like Social Signal Processing that focus on automatic understanding of social behavior. This is expected to lead, on the long-term, to socially intelligent surveillance and monitoring technologies [7].

The rest of this paper is organized as follows. Section II introduces the main concepts of proxemics, and Section III provides a brief survey of the state-of-the-art in computational proxemics. Section IV presents the approach, and Section V reports the experiments and results. Finally, Section VI draws some conclusions.

II. Fundamentals of Proxemics

The wide spectrum of nonverbal behavioral cues displayed during social interactions (facial expressions, vocalizations, gestures, postures, etc.) is well known to convey information about social and affective aspects of human-human interaction (attitudes, personality, emotions, etc.) [8]. Proxemics has shown that the way people use, organize and share space during gatherings and encounters is a nonverbal cue and it conveys, like all other cues, social and affective meaning [9]. This section provides a short description of the main findings of the discipline, with particular attention to phenomena that can be observed in standing conversations, the scenario investigated in the experiments of this work.

From a social point of view, two aspects of proxemic behavior appear to be particularly important, namely interpersonal distances and spatial arrangement of interactants. The rest of this section focuses on both aspects, including the most important factors that influence them.

A. Interpersonal Distances

Interpersonal distances have been the subject of the earliest investigations on proxemics and one of the main and seminal findings is that people tend to organize the space around them in terms of four concentric zones associated to different degrees of intimacy:

- *Intimate Zone*: distances for unmistakable involvement with another body (lover or close friend). This zone is typically forbidden to other non-intimate persons, except in those situations where intrusion cannot be avoided (e.g. in elevators).
- *Casual-Personal Zone*: distances established when interacting with familiar people, such as colleagues or friends. This zone is suitable for having personal conversations without feeling hassled. It also reflects mutual sympathy.
- *Socio-Consultive Zone*: distances for formal and impersonal relationships. In this zone, body contact is not possible anymore. It is typical for business conversations, consultation with professionals (lawyers, doctors, officers, etc.) or seller-customer interactions.
- *Public zone*: distances for non-personal interaction with others. It is a zone typical for teachers, speakers in front of a large audience, theater actors or interpersonal interactions in presence of some physical barrier.

In the case of Northern Americans, the four zones above correspond to the following ranges: less than 45 cm (intimate), between 45 and 120 cm (casual-personal), between 120 and 200 cm (socio-consultive), and beyond 200 cm (public). While the actual distances characterizing the zones depend on a large number of factors (e.g., culture, gender, physical constraints, etc.), the partition of the space into concentric areas seems to be common to all situations.

B. Spatial Arrangement: The F-Formations

The spatial arrangement during social interactions addresses two main needs: The first is to give all persons involved the possibility of participating, the second is to separate the group of interactants from other individuals (if any). The result are the *F-formations*, stable patterns that people tend to form during social interactions (including in particular standing conversations): "*an F-formation arises whenever two or more people sustain a spatial and orientational relationship in which the space between them is one to which they have equal, direct, and exclusive access*" [6].

In practice, an F-formation is the proper organization of three social spaces (see Figure 1): O-space, P-space and R-space. The O-space (the most important component of an F-formation) is a convex empty space surrounded by the people involved in a social interaction, every participant looks inward into it, and no external people are allowed in this region. The P-space is a narrow stripe that surrounds the O-space and that contains the bodies of the interactants, the R-space is the area beyond the P-space. There can be different F-formations:

- *Vis-à-vis*: An F-formation in which the absolute value of the angle between participants is approximately 180^o, and both participants share an O-space.
- *L-shape*: An F-formation in which the absolute value of the angle between participants is approximately 90^o, and both participants share an O-space.
- *Side-by-side*: An F-formation in which the absolute value of the angle between participants is approximately 0^o, and both participants share an O-space.
- *Circle*: An F-formation where people is organized in a circle, so that the configuration between adjacent participants can be considered as a hybrid between a L-shape and a Side-by-side F-formation.

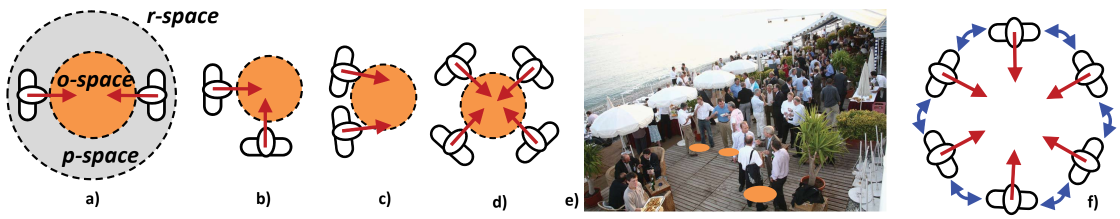


Figure 1. F-formations: a-d) The component spaces of an F-formation: vis-a-vis, L, side-by-side, and circular F-formations, respectively. O-spaces are drawn in orange. e) An example of cocktail-party scene where some o-spaces are superimposed in orange. f) The proposed scheme of evaluation of pairwise interpersonal distances in an F-formation.

The same contextual factors that influence the concentric zones described above, affect F-formations as well.

C. Context Effects on Proxemics

Proxemic behavior is affected by a large number of factors and culture seems to be one of the most important ones, especially when it comes to the size of the four concentric zones described above. In particular, cultures seem to distribute along a continuum ranging from "contact" (when the size of the areas is smaller) to "non-contact" (when the size of the areas is larger) [2]. Further evidence in this sense is proposed in [10], where people from "contact" cultures are shown to approach one another more than the others, and in [11], where the culture effect has been shown to depend on whether one considers shape of territory, size, central tendencies of encroachment, or encroachment variances (the observations were conducted on beaches). In the same vein, interpersonal distances seem to be affected by ethnicity: e.g., black Americans and Mexicans living in the States appear to have different "contact" tendencies [2], [12]. The effect of culture seems to change when interaction participants have seats at disposition. In this case, people from supposedly "non-contact" cultures tend to seat closer than the others [13]. Furthermore, the seating arrangement seems not to depend on culture [14].

Seating is just one of the many environmental characteristics that can influence the requirements on interpersonal distance and personal space. The literature has investigated the effect of many other characteristics as well, including lighting [15], indoor/outdoor [16], crowding [17] and room size [18], [19], [20]. The work in [15] investigates the effect of lighting with stop-distance techniques: Experimenters get closer and closer to a subject that remains still and says "stop" when she starts feeling uncomfortable. Subjects in bright conditions (600 lx) allow the experimenters to come significantly closer than the subjects in dim conditions (1.5 lx). A similar effect has been observed for the size of the place where people interact: people allow others to come closer in larger rooms [18], when the ceiling is higher [19][20], and in outdoor spaces [16]. The effects of crowding have been studied as well [17]: Social density was increased in a constant size environment for a limited period of time and participants of larger groups reported greater degrees of discomfort and manifested other forms of stress.

III. Computational Proxemics: State-of-the-Art

To the best of our knowledge, only a few works have tried to apply proxemics in computing. One probable reason is that current works on analysis of human behavior have focused on scenarios where proxemics do not play a major role or have relied on laboratory settings that impose too many constraints for spontaneous proxemic behavior to emerge (e.g., small groups in smart meeting rooms) [21], [22].

Most of the computing works that can be said to deal with proxemics concern the dynamics of people moving through public spaces. These works typically model repulsive/attractive phenomena by adopting the Social Force Model (SFM) [23]. In particular, the work in [24], [25] improves the perfomance of a tracking approach by taking into account the distance between a subject being tracked and the other subjects appearing in a scene. An attempt to interpret the movement of people in social terms has been presented in [26], where nine subjects (asked to speak among them about specific themes) were left free to move in a $3m \times 3m$ area for 30 minutes. An analysis of mutual distances in terms of the zones described in Section II allowed to discriminate between people who did interact and people who did not. In a similar way, mutual distances have been used to infer personality traits of people left free to move in a room [27]. The results show that it is possible to predict Extraversion and Neuroticism ratings based on velocity and number of intimate/personal/social contacts (in the sense of Hall) between pairs of individuals looking at one other.

Another frequent application area is social robotics. Early approaches in the domain simply aimed at making robots to respect the personal space of users [28], but more recent works deal with the initiation, maintenance, and termination of social interactions by modulating reciprocal distances, showing that people use similar proxemic rules when interacting with robots and when interacting with other people [29]. In [30] a generative model has been developed for selecting a set of reactive behaviors that depend on the distance, speed, and sound of interactants. Distance cues are used by the Roboceptionist [31] for recognizing "Present", "Attending", "Engaged", and "Interacting" people at the entrance of the Robotics Institute at Carnegie Mellon

University. In [32], a model for human-robot interaction in a hallway is proposed. The idea is to exploit proxemic cues for letting the robot to react properly at the passage of an individual in a narrow corridor. In [33], a user study focuses on the interaction between a human and a robot in a domestic environment. Interactions were analyzed exploiting the four zones and the F-formations introduced in Section II. The researchers found the Personal zone to be the most commonly occupied one and the "vis-à-vis" F-formation to be the most frequent spatial arrangement.

IV. The Approach

The proposed approach includes two main stages: the first is the detection of F-formations, and the second is the inference of social relations from interpersonal distances.

A. Detection of F-formations

The goal of this stage is to detect F-formations in videos portraying people involved in standing conversations. The first step is to track the people with a fish-eye camera pointing at interactants in a bird-eye view setting (see Figure 2 for an example). This corresponds to a realistic surveillance scenario and allows one to track people with satisfactory precision (tracking has been performed by exploiting a particle filter on each person [34], employing a standard background subtraction algorithm for highlighting the moving objects [35]. The results of our approach that have been obtained with this tracking strategy have been compared with those obtained via manual tracking, showing very similar results). The detection of the F-formations is performed over the output of the tracking step using the approach described in [36]. The output of the F-formations detection algorithm has been validated by hand and it did not produce any error.

F-formations lasting for less than 5 seconds (50 frames in our implementation) have not been taken into consideration in the experiments of this work. The reason is that the next stage of the processing requires the application of a clustering algorithm and 50 frames is a reasonable amount of data needed to avoid the so-called "curse of dimensionality" [37].

B. Inference of Social Relations

The output of the first stage is a list of pairs where each element includes two subjects that are adjacent in a detected F-formation. Furthermore, the first stage provides the $2D$ position of each subject on the surface of the room. Such data is accumulated during a time interval (called the "stable period" hereafter) that does not include creation, break or modifications of an F-formation (e.g., no people change their position in the P-region). This ensures that during the time interval under analysis all causes that might change the current F-formation are absent. Such causes can be novel people being involved, people leaving, a change in the environmental conditions like rain (people look for a repair), an intruder (e.g., a vehicle passing by and disrupting the F-formation), etc.. The satisfaction of the conditions above is automatically verified by checking that the relative distances between subjects in a F-formation do not change abruptly (i.e., the changes do not exceed a threshold learned automatically from the data).

During the stable period, the approach collects and pools together all pairwise distances between individuals (for a sketch, see Figure 1 (f)). Distances are collected between the centers of mass of the tracked blobs, where each blob corresponds to a separate person. These are shown to distribute according to different modes (see Section V) that should correspond to the concentric zones described in Section II. The modes have been separated via Gaussian clustering by employing the Expectation-Maximization (EM) [38] learning method. The EM employed here is a variation of the original formulation [39]; it is performed by means of a model selection strategy that is injected in the learning stage and that shows several properties that fit well with the situation at hand. First, it allows one to automatically select in an unsupervised way the right number of Gaussian components (in an Information Theory sense). This is a very important aspect, that permits to let the natural separation of the data emerge without human intervention. Second, it deals satisfactorily with the initialization issue, i.e., the Gaussian parameters fit the data realizing a nearly-global optimal fit, minimizing the probability of overfitting (i.e., a Gaussian component that fit only a few data). In addition, the Gaussian clustering takes into account in a principled way the noise due to possibly unprecise tracking, incorporating it as a variance of the measures.

V. Experiments

This section presents experiments and results obtained in this work.

A. Experimental Setup

The goal of the experiments is to investigate spontaneous standing conversations in a public space, hence the tests have been performed in an outdoor area of size $3m \times 7m$ (see Fig. 2, row (i), column (a)). The area is empty (no physical constraints or obstacles) and two groups of subjects have been invited, in two separate sessions, to move and behave normally through it. The subjects were told that the experiments were aimed at testing a tracking approach and were unaware of the real motivations behind the experiments. During the sessions, the subjects were left alone and no researcher involved in this work was present.

The experiment took place on February 2011, on a sunny day. The area was monitored with a Unibrain Fire-i Digital Camera, on which fisheye optics was mounted. The camera was located 7 meters above the floor, and it was held to an architectural element of the infrastructure. Therefore, the impact of the capture device onto the ecology of the environment was minimal. The acquisition frame rate was 10 frames per second. After the data acquisition, video data were rectified for correcting the spherical distortion. The two sessions were 15 minutes long for a total of around 20000 frames. One quarter of hour is a duration long enough to collect evidence of pre-existing social relations and short enough to avoid the emergence of new relations. The first session was recorded at 11 AM and the second at 2 PM.

Each session was split into three 5 minutes long segments corresponding to three different experimental conditions:

- Condition 1: the subjects are free to move through the entire area
- Condition 2: the movements of the subjects are restricted to an area of size $3m \times 3.5m$
- Condition 3: the movements of the subjects are restricted to an area of size $1.5m \times 2.0m$

The physical restrictions were represented by lines and marker on the floor. The goal was to measure the effect of the amount of available space on proxemic behavior.

B. Results of Session 1

The first session involved six subjects (see Fig. 2): two undergraduate students (a and b), an assistant professor (c), and three PhD students working in the same laboratory (d, e, f), two of them working on the same topic (e and d). The PhD students and the assistant professor were acquainted before the experiment. The undergraduate students are friends, but they never met before the other subjects.

In Fig. 2 row (i) we show the results obtained in the longest stable period (subjects free to move in the entire area, see Section IV-B), that in this case lasted 108 frames. The image in column (a)-(b) is the last of the period[1]. In that interval, the group was split into three dyads. The histogram in Figure 2-row (i) shows the distribution of the interpersonal distances between members of the same dyad. The application of a clustering approach shows the existence of two modes centered on 48 and 64 cm, respectively. The tables in the figure report the fraction of time distances between each pair of adjacent individuals belong to a given mode for each condition, with the value in bold red indicating the highest (most frequent cluster membership) fraction. The two modes seem to account for two of the zones identified by Hall and, not surprisingly, the dyad involving the assistant professor is the only one where the distance belongs with higher probability to the second mode most of the times. This confirms that the higher social distance between the assistant professor and the PhD student results into a physical distance that is higher (on average) than the one between subjects a and b (who are friends and both undergraduate students), as well as the one between subjects d and e (who are both PhD students).

[1]The same applies for all the other pictures in the column (a)-(b), i.e., they are the last frames of the corresponding stable period.

In Condition 2 (3×3.5 meters), the longest stable interval (122 frames) corresponds to a circle F-formation, including all subjects (see Fig. 2-row (ii), pictures at left). The clustering of the interpersonal distances of adjacent subjects reveals this time a three-mode distribution with modes at 44, 69 and 99 cm, respectively. The first mode accounts for the distance between a and b (the two undergraduate friends). The second mode accounts for the distances between c, d, e and f (the three PhD students and the assistant professor belonging to the same research group). The third mode accounts mainly for the distances between a and e and between b and c (the only pairs where the members were unacquainted before the experiments). In this condition too, the physical distances comply with the social information, even though the distance between the assistant professor and the PhD students does not reflect the difference of status.

In Condition 3 (1.5×2 meters), the longest stable period lasted for 914 frames. People form a circular F-formation, giving now rise to four distinct modes in the space of the pairwise distances (see Fig. 2-row (iii)). Once again, two close friends a and b stand at the closest distances, separated from the rest of the subjects. In particular, subjects b and c stand at a very high distance if compared to the other measurements. This highlights the separation that holds between subjects that have different status, i.e., the student and the assistant professor.

The variations across the different conditions suggest the following considerations:

- The histograms show that the modes correspond to shorter distances as the space gets smaller. However, different social relations still result into different modes.
- The fraction of distances that fall in the first mode is 67% in Condition 1, 34% in Condition 2, and 22% in Condition 3.

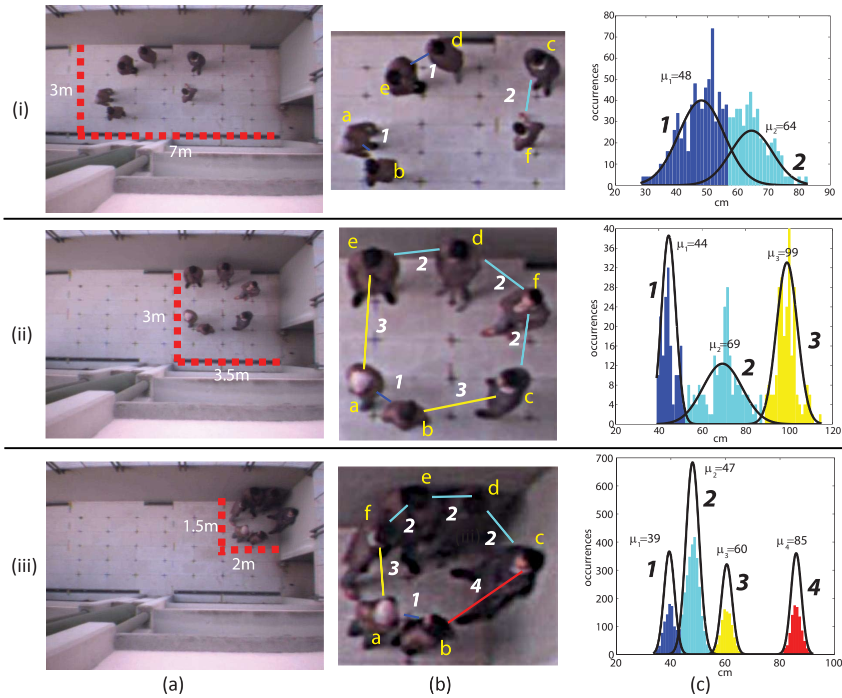
In other words, the results confirm the findings about the effect of the space at disposition of interpersonal distances and, in particular, the effects of [18] stating that subjects prefer to keep higher distances when the environment gets smaller.

The results shown here analyzed the longest stable period in each session. Anyway, in all the other stable periods, the results were qualitatively similar.

C. Results of Session 2

The second session involved 7 subjects (see Fig. 3): five undergraduate students acquainted with one another (subjects a, b, c, d and g), two PhD students that are close friends (subjects e and f), and the representative of the students in the School of Computer Science (subject c).

In Condition 1 (see Fig. 3-row (i)), the group has split into F-formations including $2-3$ people each. Fig. 3 shows the picture of the configuration that has lasted for the longest



Class / Couple	1(μ= 48cm, [29, 55])	2(μ= 64cm, [55, 82])
ab	0.68	0.32
cf	0.35	0.65
de	0.85	0.15

Class / Couple	1(μ= 44 cm, [39, 51])	2(μ= 69cm, [51, 86]	3(μ= 99cm, [86,114])
ab	0.96	0.04	0
bc	0	0.04	0.96
cf	0.28	0.70	0.02
fd	0.26	0.39	0.35
de	0	0.89	0.11
ea	0	0	1

Class / Couple	1(μ= 39cm, [33, 43])	2(μ= 47cm, [43, 54])	3(μ= 60cm, [54, 65])	4(μ=85cm, [80, 92])
ab	0.96	0.04	0	0
bc	0	0	0	1
cd	0.01	0.99	0	0
de	0.08	0.92	0	0
ef	0.06	0.94	0	0
fa	0	0.01	0.99	0

Figure 2. The pictures of column (a) show the physical space in which people were free to move. The pictures in column (b) are zoomed versions of those in (a), showing the F-formations detected in each of the three stages. The color of the links corresponds to the color of the most frequent mode to which the distances between the linked individuals belongs to. Rows (i)-(ii)-(iii) refer to Condition 1-2-3, respectively (see text). Histograms in column (c) show the distributions of the distances and the related clustering. The tables in column (d) report the fraction of time distances between each pair of adjacent individuals belong to a given mode. Each mode is identified by the mean, and by the range (in centimeters) of distances it covers (written in squared parentheses). The figure is best viewed in colors.

time (152 frames). The interpersonal distances cluster according to three modes. In the F-formation including three people, the two PhD students (who are close friends) appear to be closer (on average) than the third component (an undergraduate student they are not acquainted with them).

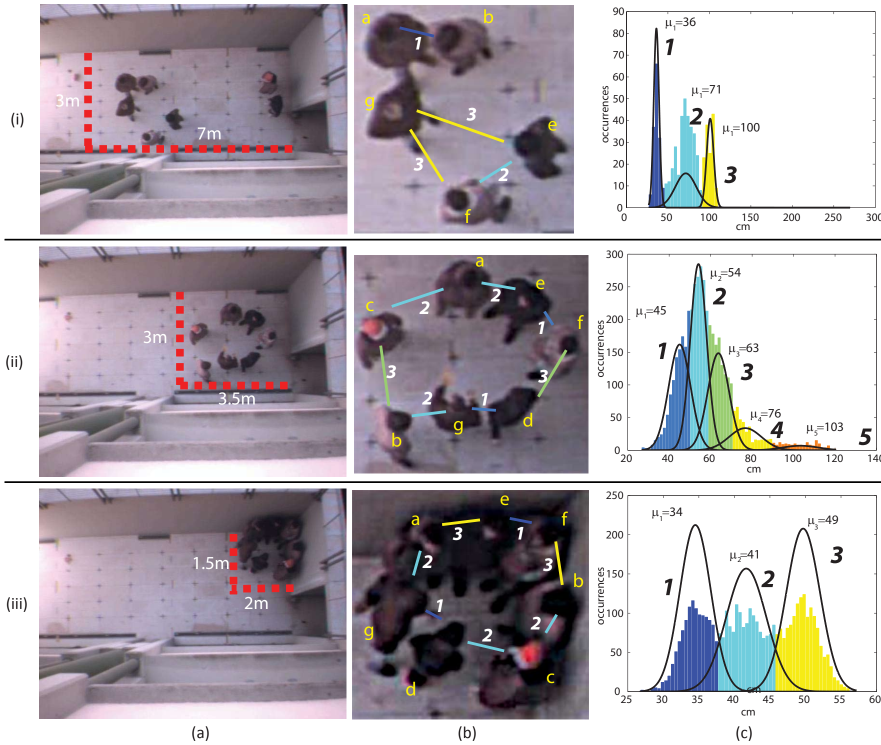
In Condition 2 (see Fig. 3-row (ii)), the most stable configuration is a circle that holds for 629 frames. In this case, the modes are five, but only the first three are used to a significant extent (see the tables of column (d) with the fractions of time distances belong to a given Gaussian component). The two PhD students (e and f) and two undergraduate students (g and d) appear to be closer to one another than the other participants. In the former case, this reflects the fact that they were close friends before the experiment, whereas in the latter, it corresponds to the fact that the two students have a romantic relationship, as it emerged from the questionnaires collected after the experiments. The situation for the other participants is less clear, but this probably happens because all participants are students and their social distances are thus similar. The only factors that seem to make some students closer (see above) are then personal.

In Condition 3 (see Fig. 3-row (iii)), a circular F-formation holds for 592 frames and corresponds to the longest stable interval. There are three modes visibile in the histogram. The PhD students are clearly separated from the rest of the circle (distances belonging to the third mode), while they are very close to one other. The couple (d and g) is tighter than the other dyads as well. In this case again, closer personal relations result into smaller distances.

It is worth to note that the effect of the amount of space at disposition leads to the same conclusions as in session 1 (see end of Section V-B).

VI. Conclusions

This paper has presented a study and preliminary experiments on the inference of social relations from interpersonal distances measured automatically via a computer vision approach. The results show that, in accordance with the findings of proxemics, people involved in casual standing interactions tend to get closer when their social relation is more intimate. The experiments have been performed on a limited number of individuals (13 in total), but the setting is



Class → / Couple ↓	**1**(μ= 36cm, [26, 44])	**2**(μ= 71cm, [44, 88])	**3**(μ= 100cm, [88, 112])
ab	**0.96**	0.04	0
ge	0	0.09	**0.91**
ef	0	**1**	0
fg	0	**1**	0

Class → / Couple ↓	**1**(μ= 45cm, [27, 49])	**2**(μ= 54cm, [49, 58])	**3**(μ= 63cm, [58, 69])	**4**μ= 76cm, [70, 88])	**5**(μ= 103cm, [88,119])
ac	0.12	**0.52**	0.33	0.03	0
cb	0.02	0.13	**0.35**	0.21	0.29
bg	0.17	**0.45**	0.34	0.04	0
gd	**0.60**	0.33	0.06	0.01	0
df	0.01	0.02	**0.60**	0.32	0.05
fe	**0.76**	0.23	0.01	0	0
ea	0.14	**0.8**	0.06	0	0

Class → / Couple ↓	**1**(μ= 34cm, [27, 37])	**2**(μ= 41cm, [37, 45])	**3**(μ= 49cm, [45, 56])
ag	0.11	**0.48**	0.41
gd	**0.88**	0.10	0.02
dc	0	**0.53**	0.47
cb	0.18	**0.66**	0.16
bf	0	0.01	**0.99**
fe	**0.98**	0.01	0.01
ea	0	0.41	**0.59**

(d)

Figure 3. The pictures of column (a) show the physical space in which people were free to move. The pictures in column (b) are zoomed versions of those in (a), showing the F-formations detected in each of the three stages. The color of the links corresponds to the color of the most frequent mode to which the distances between the linked individuals belongs to. Rows (i)-(ii)-(iii) refer to Condition 1-2-3, respectively (see text). Histograms in column (c) show the distributions of the distances and the related clustering. The tables in column (d) report the fraction of time distances between each pair of adjacent individuals belong to a given mode. Each mode is identified by the mean, and by the range (in centimeters) of distances it covers (written in squared parentheses). The figure is best viewed in colors.

fully unconstrained and spontaneous and the results appear to be consistent with the expectations.

An unsupervised analysis of interpersonal distances reveals that the four zones predicted by Hall in his seminal work emerge independently of the space at disposition of the interactants. The radii of the concentric zones are smaller than those measured in [1], [2] for Northern-Americans, but this should not be surprising as the subjects are from Italy, a culture likely to be more "contact" than the American one. Furthermore, the space available to the subjects has been progressively reduced and this has further contributed to reduce the size of the zones. The effects expected from the reduction of the space have been actually observed, especially when it comes to the tendency to increase interpersonal distances.

The detection of the F-formations appears to be crucial to perform a correct analysis of the interpersonal distances. In fact, previous works in the literature did not consider the geometric constraints imposed by the F-formations and the results have been inconsistent. In contrast, by limiting the analysis only to the distances of neghboring (adjacent) people, our experiments obtain results where social and physical distances match one another.

The next steps to be performed include not only experiments including a larger number of subjects, but also an attempt to use the statistical distributions learned from the data to predict automatically the degree of intimacy between individuals. This would represent a major step towards the development of socially intelligent surveillance technologies.

REFERENCES

[1] E. Hall, "Handbookfor proxemic research," *Studies in the anthropologyof visual communication series). Washington, DC: Society for the Anthropology of Visual Communication*, 1974.

[2] R. Hall, *The hidden dimension*, Doubleday, Ed., 1966.

[3] D. Lott and R. Sommer, "Seating arrangements and status." *Journal of Personality and Social Psychology*, vol. 7, no. 1, pp. 90–95, 1967.

[4] N. Russo, "Connotation of seating arrangements," *The Cornell Journal of Social Relations*, vol. 2, no. 1, pp. 37–44, 1967.

[5] R. Gifford and B. O'Connor, "Nonverbal intimacy: clarifying the role of seating distance and orientation," *Journal of nonverbal behavior*, vol. 10, no. 4, pp. 207–214, 1986.

[6] A. Kendon, *Conducting Interaction: Patterns of behavior in focused encounters*, C. U. Press, Ed., 1990.

[7] M. Cristani, V. Murino, and A. Vinciarelli, "Socially intelligent surveillance and monitoring: Analysing social dimensions of physical space," in *Proceedings of International Workshop on Socially Intelligent Surveillance and Monitoring*, 2010, pp. 51–58.

[8] V. Richmond and J. McCroskey, *Nonverbal Behaviors in interpersonal relations*. Allyn and Bacon, 1995.

[9] M. Knapp and J. Hall, *Nonverbal Communication in Human Interaction*. Harcourt Brace College Publishers, 1972.

[10] O. Watson, *Proxemic behavior: A cross-cultural study*. Mouton De Gruyter, 1970.

[11] H. Smith, "Territorial spacing on a beach revisited: A cross-national exploration," *Social Psychology Quarterly*, pp. 132–137, 1981.

[12] J. Baxter, "Interpersonal spacing in natural settings," *Sociometry*, vol. 33, no. 4, pp. 444–456, 1970.

[13] S. Heshka and Y. Nelson, "Interpersonal speaking distance as a function of age, sex, and relationship," *Sociometry*, vol. 35, no. 4, pp. 491–498, 1972.

[14] A. Mazur, "On Wilson's Sociobiology," *American Journal of Sociology*, vol. 82, no. 3, pp. 697–700, 1976.

[15] L. Adams and D. Zuckerman, "The effects of lighting conditions on personal space requirement," *Journal of general psychology*, vol. 118, no. 4, pp. 335–340, 1991.

[16] D. Cochran, C, "Personal space requirements in indoor versus outdoor locations," *Journal of psychology*, vol. 117, pp. 121–123, 1984.

[17] W. Griffitt and R. Veitch, "Hot and crowded: Influences of population density and temperature on interpersonal affective behavior," *Joumal of Personality and Social Psychology*, vol. 17, pp. 92–98, 1971.

[18] M. J. White, "Interpersonal distance as affected by room size, status, and sex," *The Journal of Social Psychology*, vol. 95, no. 2, pp. 241 – 249, 1975.

[19] J. Savinar, "The effects of ceiling height on personal space," *Man-environment systems*, vol. 5, pp. 321 – 324, 1975.

[20] D. Cochran, C and S. Urbanczyk, "The effect of availability of vertical space on personal space," *Journal of psychology*, vol. 111, pp. 137–140, 1982.

[21] D. Gatica-Perez, "Automatic nonverbal analysis of social interaction in small groups: a review," *Image and Vision Computing*, vol. 27, no. 12, pp. 1775–1787, 2009.

[22] A. Vinciarelli, M. Pantic, and H. Bourlard, "Social Signal Processing: Survey of an emerging domain," *Image and Vision Computing Journal*, vol. 27, no. 12, pp. 1743–1759, 2009.

[23] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Physical Review E*, vol. 51, no. 5, pp. 4282–4287, 1995.

[24] S. Pellegrini, A. Ess, K. Schindler, and L. V. Gool, "You'll never walk alone: modeling social behavior for multi-target tracking," in *Proc. 12th International Conference on Computer Vision, Kyoto, Japan*, 2009, pp. 261 –268.

[25] P. Scovanner and M. Tappen, "Learning pedestrian dynamics from the real world," in *Proc. International Conference on Computer Vision*, 2009, pp. 381–388.

[26] G. Groh, A. Lehmann, J. Reimers, M. R. Friess, and L. Schwarz, "Detecting social situations from interaction geometry," in *Proceedings of the 2010 IEEE Second International Conference on Social Computing*, IEEE Computer Society, 2010, pp. 1–8.

[27] G. Zen, B. Lepri, E. Ricci, and O. Lanz, "Space speaks: towards socially and personality aware visual surveillance," in *Proceedings of the 1st ACM international workshop on Multimodal pervasive video analysis*, ser. MPVA '10. New York, NY, USA: ACM, 2010, pp. 37–42.

[28] Y. Nakauchi and R. Simmons, "A social robot that stands in line," in *Proceedings of the Conference on Intelligent Robots and Systems (IROS '00)*, October 2000.

[29] L. Takayama and C. Pantofaru, "Influences on proxemic behaviors in human-robot interaction," in *Proceedings of the 2009 IEEE/RSJ international conference on Intelligent robots and systems*, ser. IROS'09.: IEEE Press, 2009, pp. 5495–5502.

[30] C. Breazeal, *Designing Sociable Robots*. Cambridge, MA, USA: MIT Press, 2002.

[31] M. P. Michalowski, "A spatial model of engagement for a social robot," in *In Proceedings of the 9th International Workshop on Advanced Motion Control (AMC 2006*, 2006.

[32] E. Pacchierotti, H. I. Christensen, and P. Jensfelt, "Human-robot embodied interaction in hallway settings: A pilot user study," in *Proceedings of the 2005 IEEE International Workshop on Robots and Human Interactive Communication*, 2005, pp. 164–171.

[33] K. L. Koay, D. S. Syrdal, M. L. Walters, and K. Dautenhahn, "Living with robots: Investigating the habituation effect in participants? preferences during a longitudinal human-robot interaction study." *ROMAN 2007 The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 564–569, 2007.

[34] M. Arulampalam, S. Maskell, and N. Gordon, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, 2002.

[35] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," in *Int. Conf. Computer Vision and Pattern Recognition (CVPR '99)*, vol. 2, 1999, pp. 246–252.

[36] M. Cristani, L. Bazzani, G. Paggetti, A. Fossati, A. D. Bue, D. Tosato, G. Menegaz, and V. Murino, "Social interaction discovery by statistical analysis of f-formations," in *Proceedings of British Machine Vision Conference*, 2011.

[37] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. John Wiley and Sons, 2001.

[38] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. B*, vol. 39, pp. 1–38, 1977.

[39] M. Figueiredo and A. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 381–396, 2002.