

Social Mapping of Human-Populated Environments by Implicit Function Learning

Panagiotis Papadakis, Anne Spalanzani, Christian Laugier

▶ **To cite this version:**

Panagiotis Papadakis, Anne Spalanzani, Christian Laugier. Social Mapping of Human-Populated Environments by Implicit Function Learning. IEEE International Conference on Intelligent Robots and Systems, 2013, Tokyo, Japan. hal-00860618

HAL Id: hal-00860618
https://inria.hal.science/hal-00860618

Submitted on 10 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Social Mapping of Human-Populated Environments by Implicit Function Learning

Panagiotis Papadakis[1], Anne Spalanzani[2] and Christian Laugier[1]

***Abstract*— With robots technology shifting towards entering human populated environments, the need for augmented perceptual and planning robotic skills emerges that complement to human presence. In this integration, perception and adaptation to the implicit human social conventions plays a fundamental role. Toward this goal, we propose a novel framework that can model context-dependent human spatial interactions, encoded in the form of a social map. The core idea of our approach resides in modelling human personal spaces as non-linearly scaled probability functions within the robotic state space and devise the structure and shape of a social map by solving a learning problem in kernel space. The social borders are subsequently obtained as isocontours of the learned implicit function that can realistically model arbitrarily complex social interactions of varying shape and size. We present our experiments using a rich dataset of human interactions, demonstrating the feasibility and utility of the proposed approach and promoting its application to social mapping of human-populated environments.**

I. INTRODUCTION

Robots capable of understanding the social behaviours of humans and exhibiting appropriate responses have a strong potential impact. In general, while we may acknowledge that robots exhibit superior skills than humans in certain tasks, we do not equivalently acknowledge the degree to which these skills can be integrated into the human world. Concerning mobile robots that operate within human-populated environments, a *human-like*, socially-aware robot behaviour is a key prerequisite when an integration or sharing of activities is provisioned between humans and robots.

Developing socially aware robots has ever been one of the prevalent milestones of robotics [1]. Nowadays, research is further motivated by the progress in the fields of articulated human motion perception and in turn the proliferation of diverse human (inter)action data. Despite the fact that human social conventions and correlated concepts have been studied within proxemics theory [2], the problem differentiates in robotics by the need to formalize social behaviour models into robot-centred representations. The ultimate goal is to adhere to spatial social conventions in order to minimize the discomfort caused as a result of unpredictable robotic behaviour as perceived from the human perspective [3].

Robot operation conventionally relies on a path planning strategy acting upon a representation of the environment, typically encoded in the form of maps. While raw *metric* maps endow a robot with basic navigational capabilities, structured, *topological* maps [4] and high-level representations that capture semantic attributes in the form of *semantic* maps [5] are becoming paramount. This can be partly attributed to the fact that the richness of information embedded in the map predisposes the level of intelligence that a path planning strategy may display. In this respect, socially embedded robots would greatly benefit through a form of social intelligence that we may term as *social mapping*.

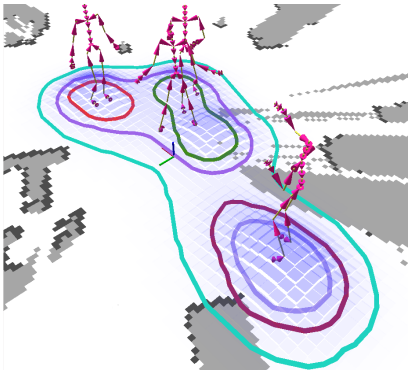


Fig. 1. Social mapping of human interactions.

We regard social mapping as a bottom-up approach in addressing the problem, which has mostly been tackled through top-down perspectives i.e. by focusing on a socially intelligent planning strategy. We argue that a bottom-up approach is more beneficial as social mapping can further contribute to recognition and inference, overall enhancing the situation awareness of a robot as a product of human social interactions understanding.

In this paper, we introduce social mapping as an effective mean to model context-dependent human spatial interactions, where context is collectively synthesized by the surrounding humans. Human comfort in the form of individual personal spaces is quantified using non-linearly scaled probability functions within free space that collectively define the structure and shape of the social map, which is obtained by solving a learning problem in kernel space. This allows us in the sequel to extract social zones as isocontours of an implicit function of sociality and capture arbitrarily complex social interactions (example shown in Fig. 1) under varying human spatial arrangements. In summary, we may highlight the main contributions of our approach as a framework that:

- Learns a structure of human comfort distribution.
- Captures social zones of dynamic shape and size.
- Models interactions of arbitrary numbers of people.

[1]P. Papadakis and C. Laugier are with equipe E-motion at INRIA Rhone-Alpes, Grenoble, France {panagiotis.papadakis, christian.laugier}@inria.fr

[2]A. Spalanzani is with UPMF - Grenoble 2 - INRIA - LIG, France anne.spalanzani@inria.fr

This work has been funded by the INRIA large scale initiative project PAL (Personally Assisted Living).

The remainder of this paper is organized as follows. In Section II we discuss previous work on methods for developing socially-compliant robot behaviour, in Section III we unfold the proposed social mapping approach and finally, in Sections IV-V respectively, we present our experiments using real data of human interactions and summarize our findings.

II. RELATED WORK

Research on socially embedded robots addresses the problem of robot integration in human-populated environments by adopting social behaviour models. Meticulous studies of this field have been recently reported within [6] and [7]. Among the various exploitable cues in social behaviour analysis the starting point resides in exploiting knowledge about the spatial arrangements of humans where mainstream approaches follow a *top-down* perspective, namely, designing socially intelligent path planning algorithms. Top-down approaches may be distinguished to those that avoid engagement with humans and those where human interaction is sought, as the respective human behaviour models can significantly differ.

Prototypical works in the first category date back at least two decades ago [8] where human safety was the primary concern while technological progress has been gradually pushing toward the integration of social models. As an example, Pacchierotti et al. [9] described a basic avoidance strategy when approached by a human along a corridor where upon entering a person's social area, the robot initiated a right turn to signify its awareness of the human presence and followed a circumventing trajectory past the human. A more promising approach was proposed by Luber et al. [10] where pairwise path patterns of people were learned through clustering that allowed the prediction of human approaching trajectories and guided a variant of A* planning for collision avoidance. However, that model solely addressed a single interacting person. The approach of Svenstrup et al. [11] was not restricted to pairwise interactions and employed Rapidly exploring Random Trees (RRT) with feedback from robot dynamics and human motion prediction. Nevertheless, the experiments were performed by synthesizing simplistic human motions and on the basis of a non-smooth function of personal space. Finally, Martinez et al. [12] formulated a stochastic optimization problem with constraints about the personal and information process space of people as well as interaction zones. Despite the elevated capacity of that model to capture complex social contexts, it lacked flexibility by defining social zones as fixed spatio-temporal functions.

As a representative example of the second category, Sisbot et al. [13] developed the HAMP planner which employed A* guided by a cost that aggregated safety and human comfort factors, modelling comfort as a function of the human field-of-view and posture. The complete system for joint human-robot task execution [14] further integrated task-dependent constraints and conditions for human-friendly motions, however, it solely considered single-person interactions.

In *bottom-up* approaches, human social behaviour cues are inscribed into high-level representations that could equally contribute to the perception of social context as well as the consecutive action strategy. In recent work, Mead [15] proposed an approach to jointly take into account human pose, speech and gesture cues in order to quantify the likelihood of social interaction occurrence within 2D space. This information would be exploited in order to guide a robot towards initiating an interaction in a socially compliant manner. Unfortunately, that model only accounted for pairwise interactions and lacked experiments that would validate its utility in more general scenarios. Callaghan et al. [16] proposed the use of *navigational maps* that encoded human tendencies when navigating towards a goal through a learned Gaussian process that regressed the human path direction. While the learned model could serve as a guide to a robotic navigation system, it was bound to a fixed environment and oblivious to human presence during navigation.

To recapitulate, the majority of previous methods have addressed the problem from a *top-down* perspective while the limited number of *bottom-up* approaches impose strong assumptions on the application scenario. On the contrary, our approach is not bound to human interactions of a certain degree, it can integrate non-static humans and the implicit social-contexts in terms of their spatial arrangements, while the shape and size of (inter)personal social zones is a continuous, smooth function of time and human placement.

III. SOCIAL MAPPING

In this Section we describe the proposed social mapping framework and emphasize on its primary utility aspects. Fig. 2 illustrates the main stages of the underlying methodology.

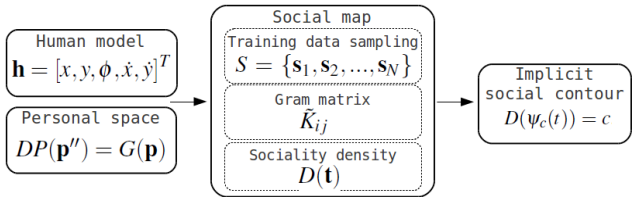


Fig. 2. Flow diagram of the proposed social mapping framework.

A. Human model

In our approach, we regard humans as vectors within the shared state space of humans and robots, each vector emanating from the human position and pointing towards the orientation of the skeletal torso. We employ this model of human perception as it is minimally subject to uncertainty compared to perception of human sub-parts and can therefore be regarded as generically applicable. To account for motion, we further consider the velocity of each human and denote a human by the vector $\mathbf{h} = [x, y, \phi, \dot{x}, \dot{y}]^T$, where $[x, y]^T \in \mathbb{R}^2$ denotes the position of the human, $\phi \in [0, 2\pi)$ the body torso orientation and $\dot{x}, \dot{y}$ the velocity along the x and y direction respectively, at a given moment in time.

B. Personal space model of human comfort

Using $\mathbf{h}$ as our low-level feature, we continue by attributing to a human a function that encodes the notion of *personal space*, in correspondence to proxemics theory [2]. We choose

to simulate personal space through a single function that bears the following desirable properties: (i) **smoothness**, (ii) **velocity-dependent** and (iii) **orientation-dependent**, where smoothness implies infinite differentiability. Smoothness is mainly desired in order allow a robot to smoothly respond to changes in human social sensitivity while dependence on human velocity and orientation are motivated by human-human and human-robot interaction studies [7], [6], [17].

While formerly proposed personal space functions have individually accounted for these conditions, we accommodate all these conditions within a single function. This is accomplished by using a 2D Gaussian probability density function (pdf) that is reshaped in order to encode the orientation and velocity information of the human. We begin by showing how we encode the personal space of a static human and in the sequel adapt this model to the non-static case.

*1) **Static model**:* We define the personal space of an individual, static human, by using a 2D Gaussian pdf that is non-linearly scaled along the body direction to simulate the effect of expansion of personal space in the frontal area of the human in relation to the rear area. The asymmetrical shape of personal space has been consistently observed in various experiments of human interactions and has also been adopted in several previous works [12], [6].

Without loss of generality, we take that the human position coincides with the coordinates origin and the body orientation is aligned with the y axis. We may now obtain the static personal space as $SP(\mathbf{p}') = G(\mathbf{p})$ where G corresponds to a zero-mean 2D Gaussian $\mathcal{N}(\mathbf{0},\Sigma)$, $\Sigma = \mathbf{I}\sigma^2$ and $\mathbf{p}'$ transforms a point $\mathbf{p} = [p_x, p_y]^T \in \mathbb{R}^2$ as:

$$\mathbf{p}' = [\frac{p_x}{3\sigma}, \left(\frac{p_y}{3\sigma} + 1\right)^k]^T \tag{1}$$

Eq. (1) dictates how the frontal area of personal space (for $p_y > 0$) expands, in contrast to the rear area whose range is not altered. Choosing $k > 1$ regulates the degree of expansion while dividing by 3σ confines the Gaussian to the unit circle which is necessary for achieving the desired scaling. Fig. 3 (a) gives an example of the resulting $SP(.)$ function.

*2) **Dynamic model**:* As human motion is normally performed along the body orientation, we augment the static personal space to a 2.5D model by modifying the k parameter of eq. (1). In detail, we obtain the dynamic personal space $DP(\mathbf{p}'')$ of a human (see Fig. 3 (b)-(c)) as $DP(\mathbf{p}'') = G(\mathbf{p})$ where $\mathbf{p}''$ transforms the point $\mathbf{p}$ as:

$$\mathbf{p}'' = [\frac{p_x}{3\sigma}, \left(\frac{p_y}{3\sigma} + 1\right)^{k+v}]^T \tag{2}$$

and v is the signed velocity magnitude along body direction.

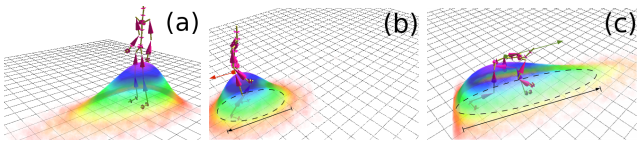


Fig. 3. Visualization of personal space; (a) Static personal space function $SP(.)$, (b) dynamic personal space $DP(.)$ of backward moving (red velocity vector) human, (c) forward moving (green velocity vector) human.

C. Social mapping by learning a structure of sociality

By contrast to conventional models of personal space that are functions of a single person, a more thorough view of the problem indicates that personal space and social zones in general should be the result of the cumulative effect of the complete set of people. Furthermore, rather than simply considering the additive effect in the form of spherical potential fields or Parzen window estimators we believe that a social map should be a product of learning a spatial structure.

In this direction, we use the individual personal spaces in order to obtain training data and learn a "sociality" function defined within free space. We use the term sociality to signify the two sides of the same coin, namely, either the comfort or discomfort stimulated to a human. Our aim is to eventually obtain a *social map* of the environment in the form of a scalar field that quantifies the density of sociality. In other words, we deal with a density estimation problem where learning amounts to deriving the set of parameters of the density.

To solve this learning problem, we chose to employ Kernel Principal Component Analysis (KPCA) [18] as it is applied in the context of *novelty detection* [19]. The main steps that are followed to employ KPCA are in order: (1) **Training data extraction**, (2) **Eigen-decomposition of Gram matrix** and (3) **Sociality density function estimation**.

*1) **Training data extraction**:* At the first step, we extract a set of training points $S = \{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_N\}$ where $\mathbf{s}_i \in \mathbb{R}^2$ and $i = 1, 2, ..., N$ from the personal spaces of the humans, that are independent and identically sampled from the attributed pdfs (see Fig. 4). Unless certainty considerations are accounted for (see Section IV), the sampling resolution N is fixed for each human and can be determined so as to provide a rich training set without compromising efficiency.

*2) **Eigen-decomposition of Gram matrix**:* In order to reveal non-linear relations of the training data within $\mathbb{R}^2$, learning is performed in a higher dimensional space $\mathscr{F}$ that is related to the input space via a mapping function $\Phi : \mathbb{R}^2 \rightarrow \mathscr{F}$. By virtue of the conventional PCA formulation, we may use the kernel trick in order to avoid the computationally expensive computation of direct mapping of the training points, by using a kernel function k that acts as a dot product in $\mathscr{F}$, i.e. $k(\mathbf{s}_i, \mathbf{s}_j) = K_{ij} = (\Phi(\mathbf{s}_i) \cdot \Phi(\mathbf{s}_j))$.

Using the kernel trick, the diagonalization of the covariance matrix of the mapped data becomes equivalent to diagonalizing the Gram matrix $\tilde{K}_{ij} = (\tilde{\Phi}(\mathbf{s}_i) \cdot \tilde{\Phi}(\mathbf{s}_j))$, $i, j = 1, 2, ..., N$ where $\tilde{\Phi}(\mathbf{s})$ centralizes the mapped data, namely:

$$\tilde{\Phi}(\mathbf{s}) = \Phi(\mathbf{s}) - \frac{1}{N}\sum_{i=1}^{N}\Phi(\mathbf{s}_i) \tag{3}$$

It can be easily derived (see [18]) that $\tilde{K}_{ij}$ is the following function of K_{ij}:

$$\tilde{K}_{ij} = K_{ij} - \frac{1}{N}\sum_{r=1}^{N}K_{ir} - \frac{1}{N}\sum_{r=1}^{N}K_{rj} + \frac{1}{N^2}\sum_{r,s=1}^{N}K_{rs} \tag{4}$$

The eigen-decomposition of $\tilde{K}_{ij}$ provides us with a set of eigenvectors $\mathbf{e}^l$, $l = 1, 2, ..., N$ and corresponding eigenvalues λ_l that describe the training points within $\mathscr{F}$ in the form of

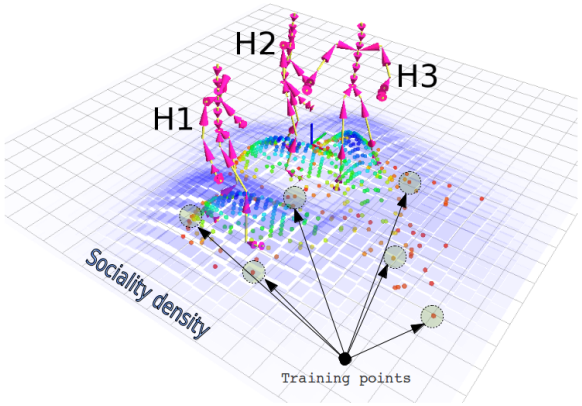


Fig. 4. Social map; Sociality density is encoded as a coloured height-field (in blue) together with points sampled from the individual personal spaces.

a hyper-ellipsoid. The hyper-ellipsoid as described by $\mathbf{e}^l, \lambda_l$ constitutes the learning part in social mapping allowing us to regress a sociality density function within $\mathbb{R}^2$.

Among the candidate kernel functions that can be used, a radial basis function is required in order to ensure translation invariance. We have chosen the gaussian kernel $k(\mathbf{x},\mathbf{y}) = \exp(-\gamma||\mathbf{x}-\mathbf{y}||^2)$ as promoted by the results within [19] which further clearly suggest that KPCA is superior compared to other *novelty detection* methods such as SVDD (Support Vector Data Description) and Parzen windows.

*3) **Sociality density function**:* By projecting a point of the original space $\mathbb{R}^2$ onto its principal component coordinates in $\mathscr{F}$, it is possible to regress a density estimate by means of the reconstruction error [19], namely, the error induced by using a principal subspace of the training points in $\mathscr{F}$ (example shown in Fig. 4).

In the context of the social map extraction, we treat the reconstruction error as the desired sociality density estimate. In detail, for any test point $\mathbf{t} = [t_x, t_y]^T \in \mathbb{R}^2$ belonging to free space, we appoint its sociality density estimate $D(\mathbf{t})$ as:

$$D(\mathbf{t}) = \tilde{\Phi}(\mathbf{t}) \cdot \tilde{\Phi}(\mathbf{t}) - \sum_{l=1}^{l_{max}} (\tilde{\Phi}(\mathbf{t}) \cdot \mathbf{e}_l) \cdot (\tilde{\Phi}(\mathbf{t}) \cdot \mathbf{e}_l) \qquad (5)$$

where $1 \leq l_{max} \leq N$ is the chosen number of principal eigenvectors. By plugging eq. (3) into eq. (5) and performing the necessary operations, we can finally obtain an analytical description of $D(\mathbf{t})$ (omitted here for the sake of brevity) solely in terms of kernel evaluations among pairs of points.

D. Social contours as implicit functions

In contrast to earlier approaches that model social spaces of constant shape or size, our approach reveals social zones that bear dynamic characteristics. This is achieved by encoding the sociality density in the form of the social map, that allows us to uncover the *implicit* social contexts as they are induced by the spatio-temporal arrangement of humans. The motivation in using the term *implicit* is two-fold. First, due to the fact that social contours cannot be directly observed through a sensor and second because we choose to model them by implicit mathematical functions.

In this direction, we define iso-contours of constant sociality density across the map and across different levels, where iso-levels range from *intimate* to *social* space respecting the ranges proposed by Hall [2]. Since $D(.)$ (see eq. (5)) is defined over $\mathbb{R}^2$, each iso-contour is an implicit function of codimension equal to 1, defined as $\psi_c(t) : [0,1] \rightarrow \mathbb{R}^2$, where $D(\psi_c(t)) = c$ and $t \in [0,1]$. The c index denotes the chosen iso-value of sociality that discriminates social zones as varying level-sets of the social map. In the case of multiple iso-contours of the same density we simply replace $\psi_c(t)$ by a set of functions Ψ_c. To obtain the iso-contours, we take an iso-value as a threshold and apply it to the social map giving a 2D binary image where cells below the threshold are set to 0 and otherwise to 1. Finally, we employ a border detection and following algorithm to output the discrete set of contours as shown in Fig. 5.



Fig. 5. Computation steps for contour extraction and grouping, namely (from left to right), sociality density, thresholding and contour detection.

IV. RESULTS

To demonstrate the utility of the proposed social mapping approach, we used the Carnegie Mellon University (CMU) dataset (http://mocap.cs.cmu.edu/) that contains a total of 2605 recorded action sequences, among which 55 sequences correspond to social interactions. We used CMU to conduct experiments using real human interactions, in contrast to earlier studies that simulate human motion. Social interaction categories range from joint walking and conversation to sports and dancing which were used to extract the $\mathbf{h}$ vector of each human. To ensure coherency, the presented experiments correspond to a fixed allotment of parameters as given below.

Parameter description	Variable	Value
Non-linear, anisotropic scaling of personal space	k	2
Gaussian kernel window used in training	γ	0.03
Number of personal space samples	N	111
Maximum allowed reconstruction error		5%

The k parameter controls the expansion of personal space along the body direction and γ regulates the impact of training points in learning. The maximum allowed reconstruction error sets the parameter l_{max}, namely, the number of principal dimensions considered. Allowing for a maximum error of 5% translated to an average of $l_{max} = 6$ principal eigen-vectors that controlled the complexity of the learned model.

a) Detection of Social Interactions: In Fig. 6 we show the temporal evolution of social contours for interactions where we set three levels of sociality, namely, *intimate*, *personal* and *social* zone. The first row shows the scenario where a two-person meeting takes place. During the first stage where one person approaches the other we obtain two

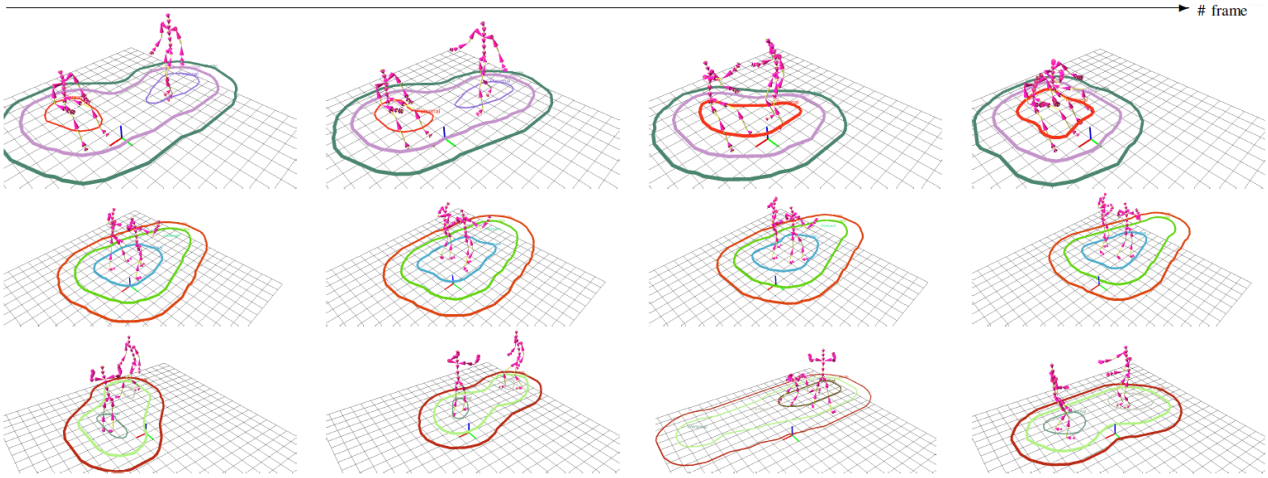


Fig. 6. Extraction of social contours along the course (from left to right) of diverse social interactions. The thickness of each contour is proportional to the number of circumscribed humans. Top row; Friends meeting. Middle row; walking and conversation. Bottom row; Seek versus avoid interaction.

distinct intimate zone contours while as the approaching person advances, the two contours merge thus signifying an intimate social interaction. Along the evolution of the interaction, the size and shape of the social contours is adapted as the result of the implicit function learning that allows a smooth transition among different social interactions.

In the middle row of Fig. 6, two persons walk alongside while in conversation. This example depicts the result of expanding the frontal area of personal space along the direction of velocity and in turn of all social zones. This is further emphasized in the example at the bottom row in which the first person is seeking the second while the latter is avoiding the former. Here, the fleeing person has a higher evading speed which occasionally results in detached zones despite their proximity. These examples illustrate how time is integrated in the 2.5D social mapping scheme by the short-term prediction of human displacement.

b) Analysis of Social Interactions: We have compared our approach against a Parzen window estimator of sociality. At the left of Fig. 7, we give two examples that depict the general deficiency of the Parzen estimator. In the first row two people hold each other's arms while facing in opposite directions whereas in the second both humans face forward. Clearly, iso-contours of the conventional density estimator are unable to discriminate the two cases, instead, by regressing a sociality function using the proposed approach (right of Fig. 7) the two scenarios can be disambiguated, allowing for automatic recognition of social interactions.

Among various options for the consecutive analysis of detected social interactions, we used as features the number of humans within a social zone and the average displacement during interaction. The first feature is calculated by "point-in-polygon" algorithms and is useful in discriminating social zones of similar shape although they are the result of a different number of people. In Fig. 8 we show the result of social mapping on interactions of multiple humans where the number of circumscribed humans is visually conveyed by automatically adjusting the thickness of the contours. The second feature signifies whether an interaction is stationary and can assist in further refining the type of social activity. In Fig. 9, we show the performance in discriminating static against non-static interactions, by means of the Receiver Operating Curve (ROC), varying in relation to the threshold of average displacement that designates whether an interaction is non-static (positive) or static (negative). The attained 82% true positive rate for 0% false positive rate indicates that near perfect performance could be accomplished if we further accounted for the articulated human motion.

To compensate for uncertainty in human detection, we control the number of training points sampled from the individual personal spaces, that regulates the emphasis appointed to each human. Taking into account that the computational complexity of KPCA is $O(N^2)$, when certainty is high the computational cost increases as we obtain more reliable sociality density estimates. Analogously, when certainty is reduced the computational cost decreases and compensates for the declined reliability. Eventually, results are mostly dependent on the relative uncertainties as is shown in the example of Fig. 10, wherein social density in the vicinity of humans is proportionate to certainty.

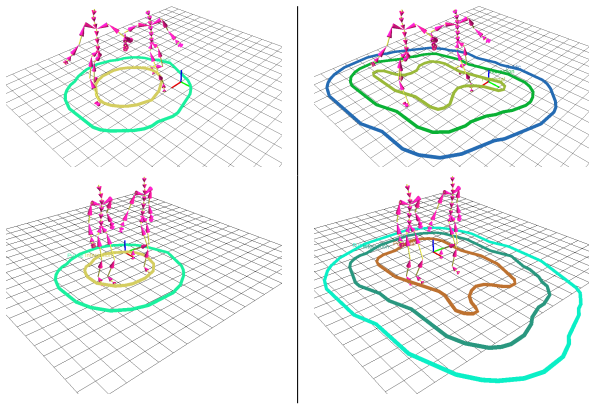


Fig. 7. Comparison of Parzen window estimator of sociality (left) and learning an implicit function of sociality with KPCA (right).

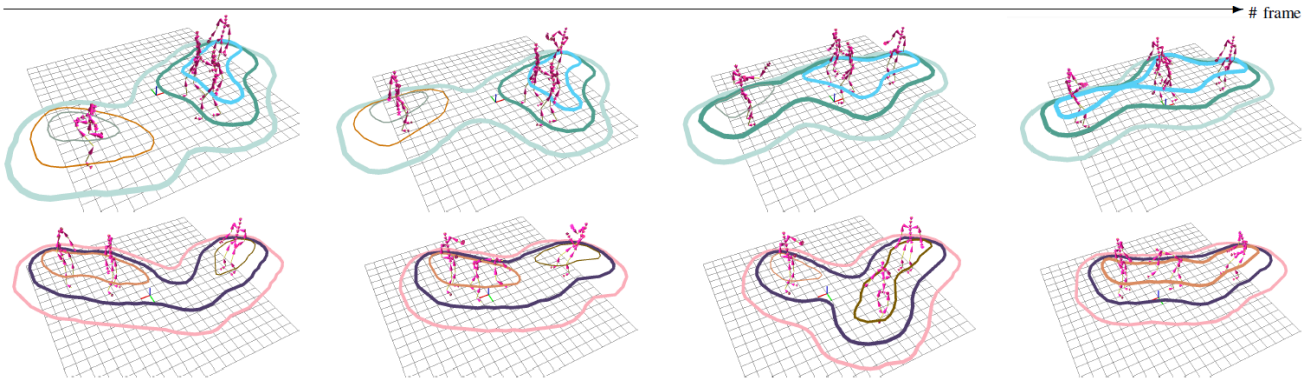


Fig. 8. Social mapping of multiple interacting people. The number of humans within each contour regulates the contour thickness.

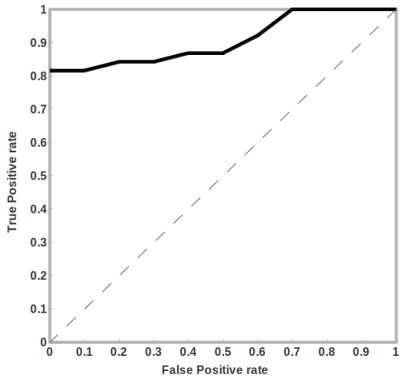


Fig. 9. ROC performance for static and non-static interactions.

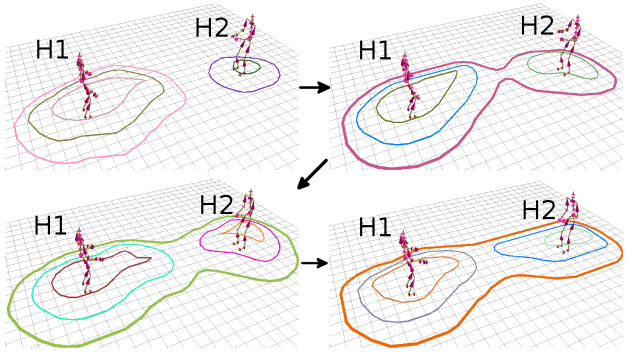


Fig. 10. Effect of human detection uncertainty in social mapping. Certainty for human H1 is constant while it is progressively increased for human H2.

V. CONCLUSIONS

We have presented a novel social mapping methodology of human-populated environments showing its utility in diverse human interaction scenarios. Using a new model of human comfort that is a smooth function of position, orientation and velocity we employ kernel-based regression and capture social zones as induced by human interactions, without the need of tracking or imposing hard thresholds for social intercourse. By extending the classical notion of individual personal spaces we introduced the notion of shared social spaces, allowing higher flexibility in their shape and size.

REFERENCES

[1] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 143–166, 2003.

[2] E. Hall, *The Hidden Dimension: Man's Use of Space in Public and Private*. The Bodley Head Ltd, 1966.

[3] L. Takayama and C. Pantofaru, "Influences on proxemic behaviors in human-robot interaction," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2009.

[4] P. Papadakis, M. Gianni, M. Pizzoli, and F. Pirri, "Constraint-free topological mapping and path planning by maxima detection of the kernel spatial clearance density," in *Int. Conf. on Pattern Recognition Applications and Methods*, 2012, pp. 71–79.

[5] A. Nüchter and J. Hertzberg, "Towards semantic maps for mobile robots," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 915–926, 2008.

[6] R. Kirby, "Social robot navigation," Ph.D. dissertation, Carnegie Mellon University, 2010.

[7] J. A. R. Martinez, "Socially-aware robot navigation: Combining risk assessment and social conventions," Ph.D. dissertation, University of Grenoble, 2013.

[8] S. Tadokoro, M. Hayashi, Y. Manabe, Y. Nakami, and T. Takamori, "On motion planning of mobile robots which coexist and cooperate with human," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 1995.

[9] E. Pacchierotti, H. Christensen, and P. Jensfelt, "Embodied social interaction for service robots in hallway environments," in *Int. Conf. on Field and Service Robotics*, 2006.

[10] M. Luber, L. Spinello, J. Silva, and K. O. Arras, "Socially-aware robot navigation: A learning approach," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2012.

[11] M. Svenstrup, T. Bak, and H. Andersen, "Trajectory planning for robots in dynamic human environments," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2010.

[12] J. Rios-Martinez, A. Renzaglia, A. Spalanzani, A. Martinelli, and C. Laugier, "Navigating between people: A stochastic optimization approach," in *IEEE Int. Conf. on Robotics and Automation*, 2012.

[13] E. Sisbot, L. Marin-Urias, R. Alami, and T. Simeon, "A human aware mobile robot motion planner," *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 874–883, 2007.

[14] E. Sisbot, L. Marin-Urias, X. Broqure, D. Sidobre, and R. Alami, "Synthesizing robot motions adapted to human presence," *International Journal of Social Robotics*, vol. 2, pp. 329–343, 2010.

[15] R. Mead, "Space, speech, and gesture in human-robot interaction," in *ACM Int. conf. on Multimodal interaction*, 2012.

[16] S. O'Callaghan, S. Singh, A. Alempijevic, and F. Ramos, "Learning navigational maps by observing human motion patterns," in *IEEE Int. Conf. on Robotics and Automation*, 2011.

[17] E. Torta, R. Cuijpers, J. Juola, and D. Pol, "Design of robust robotic proxemic behaviour," in *Int Conf. on Social Robotics*, 2011, pp. 21–30.

[18] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[19] H. Hoffmann, "Kernel pca for novelty detection," *Pattern Recognition*, vol. 40, no. 3, pp. 863–874, 2007.