Leveraging Eye Tracking and Physiological Signals for Fluent Human Robot Collaboration

Stefanos Nikolaidis[1], Enkelejda Kasneci[2] and Siddhartha Srinivasa[3]

***Abstract*— When we collaborate with each other, we infer each others' intentions not only from our *physical actions* that change the physical state of the world, but also from our *physiological signals* like gaze, heart rate, and galvanic skin response, that reveal our internal state without actively changing the environment. However, traditional human-robot collaboration paradigms that formulate interaction as a physical dynamical system often fail to exploit the rich vein of information that physiological signals can convey. A key challenge, however, is in building coherent Bayesian models of physiological signals: how our natural gaze wanders depends not just on the task, but also on a myriad of other distractors and features. In this work, we first demonstrate how human physical actions are often uninformative of their true internal state. Through eye tracking and physiological signals, we investigate different modalities for human internal state elicitation, and we describe how these can be integrated into our general formalism to facilitate the robot's inference process.**

I. Introduction

The success of the team in a collaboration setting often depends on the ability of team members to coordinate their actions by reasoning over each others' beliefs and actions. We want to enable robot teammates with this very capability in human-robot teams, e.g., service robots interacting with users at home, manufacturing robots sharing the same physical space with human mechanics and autonomous cars interacting with drivers and pedestrians.

Towards this direction, Nikolaidis et al. [1] have modeled humans and robots as equal partners, by treating the interaction as an underactuated dynamical system, wherein the human state is a latent variable. The robot builds online a model of human adaptation that depicts how the robot's own actions affect future human actions, and adapts its own actions in return. The model requires reasoning over the human internal state, through observation of the human actions.

However, the human actions are often ambiguous, which can affect negatively the robot inference. Additionally, they may not be informative of internal states, such as fatigue and object-perception, which are indicative of future performance. On the other hand, there has been substantial work [2]–[4] on inferring these internal states from other modalities, such as gaze, or physiological signals (e.g., skin conductance, heart rate and EEG)

[1]Robotics Institute, Carnegie Mellon snikolai@cmu.edu
[2]Department of Computer Science, University of Tübingen enkelejda.kasneci@uni-tuebingen.de
[3]Paul G. Allen School of Computer Science & Engineering, University of Washington siddh@cs.uw.edu

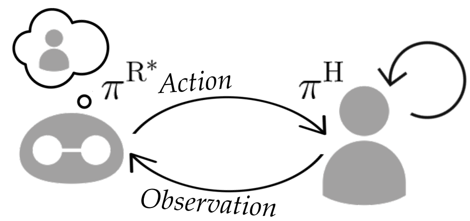

Fig. 1: The robot maximizes the performance of the human-robot team by executing the optimal policy π^{R^*}.

This work explores how robotic systems can take advantage of the signals through *different modalities of communication* that exist when people perform their tasks, and how these can be integrated into the general formalism of mutual adaptation.

We start with a general formulation of the problem, and depict the assumptions that lead to the human-robot mutual adaptive behaviors. We then instantiate the mutual-formalism in a shared-autonomy setting, where human and robot provide inputs through a joystick interface. We show an example of the case where human actions are uninformative of their intent, which in turn affect robot's own actions. We then discuss alternative modalities for human internal state inference.

II. Problem Formulation

Human-robot collaboration can be formulated as a *two player game with partial information*. We let x_t be the world state that captures the information that human and robot use at time t to take actions a_t^R, a_t^H in a collaborative task. Over the course of a task of total time duration T, robot and human receive an accumulated reward:

$$\sum_{t=1}^{T} R^R(x_t, a_t^R, a_t^H), \quad \sum_{t=1}^{T} R^H(x_t, a_t^R, a_t^H)$$

We assume a robot policy π^R, which maps world states to actions. The human chooses their own actions based on a human policy π^H (Fig. 1). If the robot could control both its own and the human actions, it would simply compute the policies that maximize its own reward.

However, the human is not another actuator that the robot can control. Instead, the robot can only *estimate* the human decision making process from observation and *make predictions* about future human behavior, which in turn will affect the reward that the robot will receive.

Therefore, the optimal policy for the robot is computed by taking the expectation over human policies π^H.

$$\pi^{R^*} \in \underset{\pi^R}{\operatorname{argmax}}\ \mathbb{E}\left[\sum_{t=1}^{T} R^R(x_t, a_t^R, a_t^H) | \pi^R, \pi^H\right] \tag{1}$$

Solving this optimization is challenging: First, the human reward R^H may be unknown to the robot in advance. Second, even if the robot knows R^H, it may be unable to predict accurately the human actions, since human behavior is characterized by bounded rationality [5]. Third, even if the human acts always rationally, exhaustively searching for the equilibria is computationally intractable in most cases [6]. Finally, even if $R^H \equiv R^R$, most real-world problems have multiple equilibria, and in the absence of a signaling mechanism, it is impossible to know which ones the agents will choose.

III. Mutual-Adaptation Formalism

The mutual-adaptation formalism [1] treats the interaction as an *underactuated* dynamical system; it recognizes that the human policy can change based on the robot actions.

We let a history of world states and robot actions h_t:

$$h_t = (x_0, a_0^R, ..., x_t, a_t^R)$$

Given this history, we specify the human policy $\pi^H(x_t; y_t)$ as a function of the current world state x_t and the human internal state y_t. The human internal state changes stochastically based on a transition function $P(y_{t+1}|y_t, h_t)$. This allows the robot to reason over how its own actions affect future human actions, and takes that into account into its own decision making.

The internal state y_t of a new human worker is typically unknown in advance to the robot and it cannot be fully observed. Therefore, we treat y as a latent variable in a partially observable stochastic process, in particular a mixed-observability Markov decision process, which has been shown to achieve significant computational efficiency [7]. The robot retains a probability distribution, or belief, on the human internal state y, which it updates through bayesian inference:

$$b'(y_{t+1}) = P(a^H|y_{t+1}) \sum_{y_t} P(y_{t+1}|y_t, h_t) b(y_t) \tag{2}$$

The observation function $P(a^H|y_{t+1})$ is given by the human policy π^H, while the dynamics function $P(y_{t+1}|y_t, h_t)$ depicts how the human internal state evolves over time based on the robot's own actions.

Solving for the robot optimal policy allows the robot to take information seeking actions to infer online the parameter y. As a result, human and robot *mutually adapt* to each other; the robot builds online a model of how the human adapts to the robot by inferring their internal state y, and adapts its own actions in return.

This history h_t can grow arbitrarily large, making optimizing for the robot actions computationally intractable. Using insights from work on bounded-rationality in behavioral economics, we use a bounded-memory assumption and limit the history length to the last k observations.

IV. Mutual-Adaptation in Shared-Autonomy

Shared autonomy combines direct teleoperation with autonomous assistance. In this setting, the robot infers a distribution over goals based on the user input, and assists the user for that distribution [8]. Mutual adaptation additionally allows the robot to reason over how the user goal may change, as the user adapts themselves to the robot. We instantiate the mutual adaptation formalism in the shared-autonomy setting [9], using as human state variables $y = (m^H, \alpha)$, the human goal m^H and the human *adaptability* α, that is the probability that they switch from their current goal m^H to a new goal demonstrated by the robot within the recent history of interactions h_k.

For example, we assume a table-clearing task (Fig. 2), where the robot clears two bottles, one after the other, off the table and places them in a bin. The human goal is to go for the left bottle first ($m_0^H \equiv m_L$). However, the robot considers in its reward function R^R the right bottle m_R as better goal, for instance because it is heavier and it should be placed in the green bin first. The reward function additionally encodes a penalty for disagreement between human and robot goals. This disagreement reflects the loss of trust occurred when the robot does not follow the human preference.

Solving for the robot policy in this setting results in some very interesting behaviors; if the human user is inferred to be non-adaptable, they will be expected to insist on their goal and the cost for disagreement will outweigh the benefit of reaching the optimal goal. On the other hand, if the human user is inferred to be adaptable, the robot will intelligently disagree with the person, expecting them to adapt and follow the robot towards the optimal goal.

Fig. 2 shows these behaviors for a non-adaptable and an adaptable user. The robot starts by moving straight (information-seeking action), observing whether the human changes their inputs to follow the robot (adaptable user), or insists on pushing the joystick to the left (non-adaptable). If the human is adaptable (bottom-row), the robot will guide them towards a better way of doing the task. If the human is non-adaptable (top-row), the robot follows the human preference towards the left-bottle, in order to retain their trust. The same behaviors, as well as the robot inference on the human goal and human adaptability for each time-step are shown in more detail in Fig. 3 (Users 1,2).

V. Ambiguous Actions in Shared-Autonomy

In Eq. 2, the robot uses the human actions a^H to infer the human internal state y through the observation function $P(a^H|y_{t+1})$. In the shared-autonomy setting, the human actions are the human joystick inputs. However, there are cases that these inputs are not informative on the human internal state, impeding the robot inference. For instance, in the table-clearing setting of Fig. 2, we let the human action "move-forward," and we assume that the human is equally likely to take this

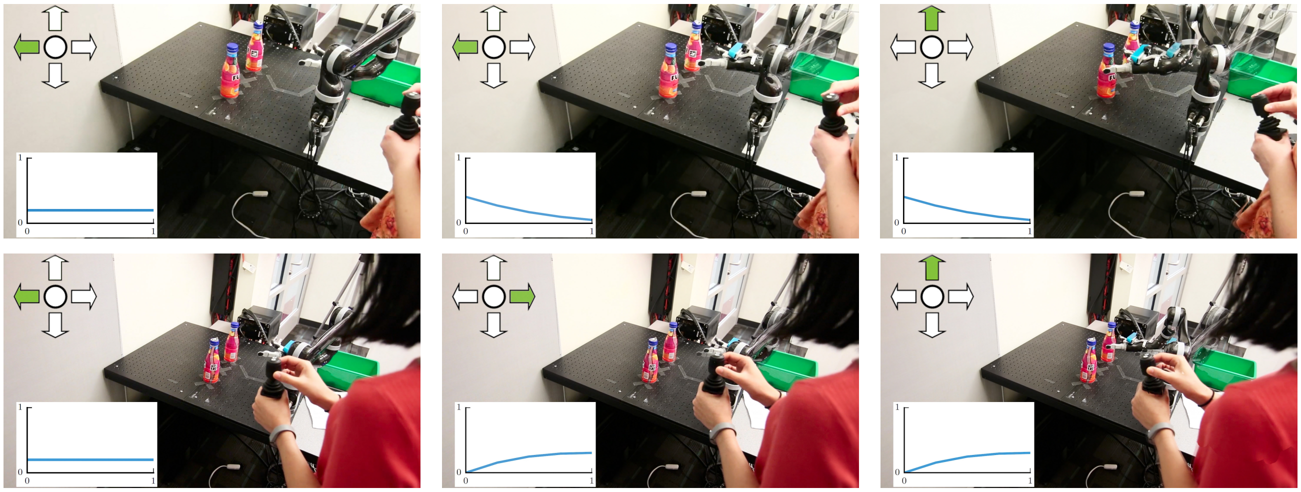

Fig. 2: Two different users control an assistive arm through joystick inputs. The plot indicates the robot belief (y-axis) on user's adaptability (x-axis). The robot starts by moving straight towards both bottles, in order to infer the user adaptability. Top, left to right: A non-adaptable user insists on moving the joystick to the left and the robot adapts to the user to retain their trust. Bottom, left to right: The user changes their inputs to align them with the robot actions. The robot infers that the user is adaptable and guides them towards the optimal goal (right bottle).

action whether they go to the left or right goal, so that $P(\text{move-forward}|m_L) = P(\text{move-forward}|m_R)$. This implies that the "move-forward" action provides no-information about the human goal. In turn, this impedes the robot inference on human adaptability. Fig. 3, User 3 illustrates this case. We assume that User 3 is non-adaptable, but wants the robot to move all the way forward first, before moving to the left. As the user moves the joystick forward and the robot moves to the right, the probability of the human having adapted to the robot accumulates, and at $T = 3$ the robot infers that m_R is more likely than m_L. On the other hand, the robot belief on human adaptability remains uniform, since the robot has not received any informative observations on the human goal. At $T = 5$, the human moves the joystick to the left, and the robot infers that the human goal is m_L and that the human is non-adaptable, but it is too late; the robot is close to the optimal goal m_R and it will proceed towards that goal instead. This example illustrates the need for using alternative signals to improve the inference process.

VI. Generalizing to Different Modalities

In order to make inferences on each others' beliefs, we as humans employ several modalities, such as speech, gaze, or gesture. Since the actions can be ambiguous and are not informative of human states such as perception, fatigue, or stress level, we propose generalizing Eq. 2 to include measurements from indirect signals as well. More specifically, we define a vector of measurements $\mathbf{o}$ that the robot receives at a given time-step, and we update the belief, so that:

$$b'(y_{t+1}) = P(\mathbf{o}_{t+1}|y_{t+1}) \sum_{y_t} P(y_{t+1}|y_t, h_t) b(y_t) \qquad (3)$$

In the following, we provide specific examples of such measurement, focusing on non-intrusive methods, i.e., by recording and analyzing gaze behavior, by measuring changes in the heart rate and galvanic skin response.

A. Inferring the user's actions and state based on gaze

Among the various modalities we humans employ to infer what is on an other's mind, gaze is one of the most information-rich cues. We can address another person by establishing gaze contact, direct other's attention towards a specific location or towards an object by eye-pointing and even communicate emotions. With gaze we put our statements in a context. Despite the scope of interaction itself, our gaze behavior is strongly related to the way we perceive personality [10]. For example, our memory of a person will be much stronger if the person's gaze is directed toward oneself, or individuals who make direct eye contact are perceived as more trustworthy and more attractive than individuals who do not make eye contact (e.g., [11]). Making interaction with robots feel more natural is above all a question of how these social cues are understood and in turn employed by the robot.

With recent advances in eye-tracking technology (such as small, light-weighted devices) and analytical approaches, new means have become available to study human eye movements in scenarios beyond screen-based settings. The so-called pervasive eye-tracking technology is not only a promising technology in general human-computer-interaction scenarios, but also gives us insight on how humans employ their gaze to interact and communicate with each other in different environments.

By combining eye-tracking technology with machine learning, various work has be done to identify user's activity based on eye movements in post-experimental analysis, e.g. [12], [13], or in an online-fashion such as

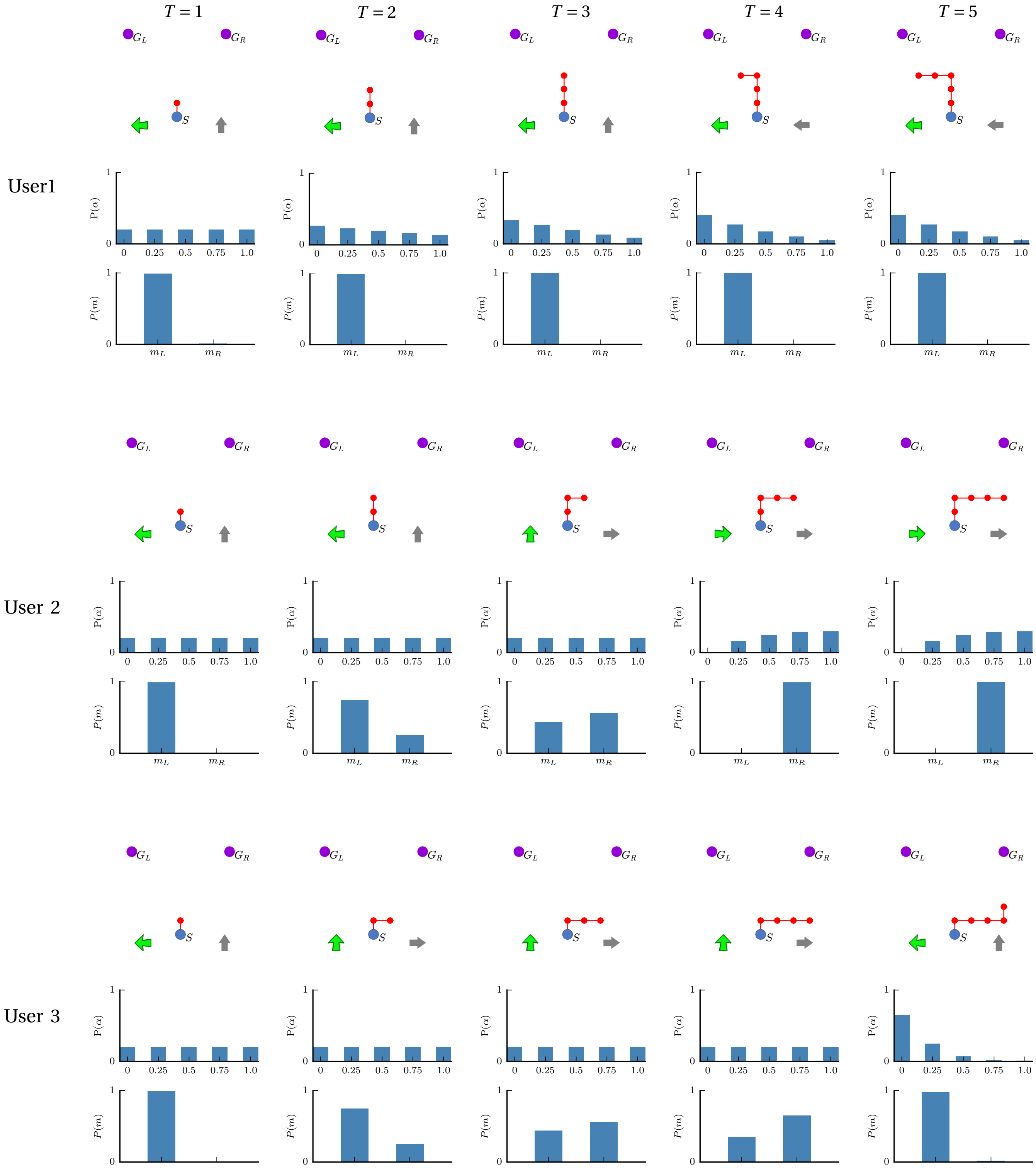

Fig. 3: Sample runs on a shared autonomy scenario with two goals G_L, G_R and three different human users. All users start with the preference m_L (associated with the left goal). The human and robot actions are {move-left, move-right, move-forward}. For all users, the upper row plots the robot trajectory (red dots), the human input (green arrow) and the robot action (gray arrow) over time. The middle row plots the estimate of α over time, where $\alpha \in \{0, 0.25, 0.5, 0.75, 1\}$. Each graph plots the probability of α versus α. The lower row plots $m \in \{m_L, m_R\}$ versus the probability of m. Columns indicate successive time-steps. User 1 insists on their initial strategy throughout the task and the robot complies, whereas User 2 adapts to the robot and ends up following m_R. User 3 is non-adaptable and they want to go forward first before going to the left goal. Because the forward input leads to both goals with equal probability, the robot belief over their adaptability α does not change. When the robot finally infers that they are non-adaptable, the robot is already close to the optimal goal, and will move towards that goal.

for driver activity recognition during conditionally autonomous driving, e.g., [14]. In the latter work, driver observation by means of eye tracking is used in real-time to trigger assistive systems, providing thus the basis for situation awareness and take-over readiness of the driver. This can be also applicable during the interaction of an autonomous vehicle with a human-driven vehicle [15]. Furthermore, eye movements have been analyzed in a variety of domains regarding their relationship to expertise (e.g., [16], [17]), or to predict actions (e.g., [18]).

Despite the gaze information itself, the pupillary signal is a further important information carrier. The size of the human pupil regulates to the amount of light that enters the eye. An increase in luminance, therefore, results in a fast constriction of the pupil. The pupillary light reflex (PLR) [19] regulates the light influx. On the other hand, there is a well-studied correlation between pupillary dilation and cognitive factors such as workload [20], [21], fatigue [22], surprise [23], attention [24], and emotional arousal [25]. Thus, pupil dilation constitutes a proxy for indirect measurement of these cognitive factors, which would otherwise only be accessible by means of costly and intrusive measurements such as EEG. The main challenge when using pupillary information is the noise in the pupillary signal, which arises due to changes in the illumination or inaccuracies of the eye recording system. However, it has been previously shown that this signal can be used to complement gaze information even under mesopic conditions, such as in [26]–[28] to detect workload or stress. Besides gaze and pupil size, tracking the eyes of a user can allow us to access additional eye-related features, such as blinks and the vergence angle. It was shown that a combining of these multiple indicators can leads to better indicator of fatigue, which provides a stable prediction of across subjects [29].

B. Physiological signals

Physiological signals such as heart rate and galvanic skin conductance are meanwhile well studied and have therefore been measured in a variety of use-cases to derive information about a subject. More specifically, since the latter two signals are considered as strong indicators of cognitive load and stress, they have been analyzed in several applications to understand user behavior in complex tasks, e.g. [30], [31]. Although these signals are still measured by employing ECG and skin conductance electrodes, novel image-based approaches promise to capture the same information in the most non-intrusive way [32]–[34].

However, in contrast to the eye-tracking signal, which has a negligibly small delay in the order of few ms, the detection of changes in the vital parameters heart rate and galvanic skin response is only possible with a relatively long delay (approximately $1-4$ seconds). Thus, these signals can be used as input and predictors for triggering actions in real-time human robot collaborative scenarios only to a limited extent.

C. Aggregation of physiological signals with gaze to infer the user's state

To leverage the user's state in collaborative scenarios, signals from different sensors need to be processed in an online fashion and aggregated according to their reliability. The reliability of a sensor, however, depends not only on the type of the sensor but also on the subject. For example, eye-tracking signal is sensitive to make-up and changing illumination, while skin conductance and heart rate are vulnerable to loosened or detached electrodes. Deriving a binary decision about a user's state from raw sensor data is a challenging problem of its own and requires device specific filtering, synchronization and processing. We will treat this necessary preprocessing step as abstract and focus on a recently introduced method in [28], an Optimal Aggregation Scheme (OAS), to derive the latent user's state based on the readily preprocessed binary decision label. From a theoretical viewpoint the problem can be formalized as follows. We assume $y = (y^s, y^a)$, where y^s the component of the human state that can be inferred through the responses $o_1, ..., o_n \in \{1, ..., k\}$ from n signal sources, and y^a the component that can be inferred through the human actions.

Given the responses of n sources on the occurrence of a discrete event $y^s \in \{1, ..., k\}$, and assuming independence between the signals given the event, we have:

$$p(o_1, ..., o_n | y^s) = \prod_j p(o_j | y^s)$$

The belief update of Eq. 3 becomes:

$$b'(y_{t+1}) = \prod_j p(o_j | y^s) P(a^{\mathrm{H}}_{t+1} | y^a_{t+1}) \sum_{y_t} P(y_{t+1} | y_t, h_t) b(y_t) \tag{4}$$

Now, if we want to compute the most probable value of the human state y^s, we have:

$$\arg\max_k p(y^s = k | o_1, ..., o_n) = \arg\max_k \frac{\prod_j p(o_j | y^s = k) p(y^s = k)}{p(o_1, ..., o_n)}$$
$$\propto \arg\max_k \Big(\prod_{\substack{j=1,..,n \\ o_j = k}} p(o_j = k | y^s = k) \prod_{\substack{j=1,..,n \\ o_j \neq k}} p(o_j \neq k | y^s = k) \Big)$$
$$\cdot\, p(y^s = k)$$

Note that since $p(o_1, ..., o_n)$ is independent of k it does not influence the choice of the most probable k.

In the case of *binary* states and responses ($y^s, o_j \in \{0, 1\}$), if we know the true reliability of each signal source – in terms of its true positive rate tpr and true negative rate tnr – we can aggregate the source responses in a probabilistically optimal way as follows:

$$p(y^s = 1 | o_1, ..., o_n) \propto \Big(\prod_{\substack{j=1,..,n \\ o_j = 1}} tpr_j \prod_{\substack{j=1,..,n \\ o_j = 0}} fnr_j \Big) p(y^s = 1)$$

and analogously

$$p(y^s = 0 | o_1, ..., o_n) \propto \Big(\prod_{\substack{j=1,..,n \\ o_j = 0}} tnr_j \prod_{\substack{j=1,..,n \\ o_j = 1}} fpr_j \Big) p(y^s = 0)$$

The above aggregation follows the Naive Bayes principle, and, if the sensors' responses are indeed independent given the event, it is optimal for the choice of the most probable event in a probabilistic sense. In practice, in the presence of multiple sensors, a sum of logarithms would be used to avoid computational arithmetic underflow.

The presented aggregation scheme can be used on top of complex and computationally intensive algorithms, such as the Expectation Maximization approach – which, in turn, can be run in the background to produce estimations of the true positive and true negative rates from sampled data. The real-time aggregation of responses is provided by the above scheme.

Just recently, this aggregation scheme was employed in a driving scenario [28] to derive hazard perception based on the gaze information, the pupil signal, heart rate measurements, and the galvanic skin response, showing a significant improvement of both prediction quality and robustness over the best predicting single source [35].

VII. Discussion

Overall, we believe that eye tracking and physiological signals can improve significantly the robot inference process when human actions are uninformative of the human internal state. However, there are several challenges, including: (1) building coherent observation models for gaze, (2) understanding gaze in a manipulation scenario, specifically how humans often signal their *next* action even before they complete their previous action, and (3) building generative models that enable robots to participate in gaze signaling. In future work, we will employ head-mounted eye-tracking technology and physiological signals to detect user's states that should be avoided in shared autonomy settings, such as increased stress level or confusion of the user. Furthermore, we will exploit the potential of gaze tracking for action disambiguation in user studies.

References

[1] S. Nikolaidis, A. Kuznetsov, D. Hsu, and S. Srinivasa, "Formalizing human-robot mutual adaptation: A bounded memory model," in *HRI*, 2016.

[2] H. Admoni and B. Scassellati, "Social eye gaze in human-robot interaction: A review," *JHRI*, 2017.

[3] E. Kasneci, K. Sippel, K. Aehling, M. Heister, W. Rosenstiel, U. Schiefer, and E. Papageorgiou, "Driving with binocular visual field loss? a study on a supervised on-road parcours with simultaneous eye and head tracking," *PloS one*.

[4] C. L. Lisetti and F. Nasoz, "Using noninvasive wearable computers to recognize human emotions from physiological signals," *EURASIP Journal on Advances in Signal Processing*, vol. 2004.

[5] D. Kahneman, "Maps of bounded rationality: Psychology for behavioral economics," *The American economic review*, 2003.

[6] C. H. Papadimitriou, "The complexity of finding nash equilibria," *Algorithmic Game Theory*, 2007.

[7] S. C. Ong, S. W. Png, D. Hsu, and W. S. Lee, "Planning under uncertainty for robotic tasks with mixed observability," *IJRR*, 2010.

[8] S. Javdani, H. Admoni, S. Pellegrinelli, S. S. Srinivasa, and J. A. Bagnell, "Shared autonomy via hindsight optimization for teleoperation and teaming," *arXiv preprint arXiv:1706.00155*, 2017.

[9] S. Nikolaidis, Y. X. Zhu, D. Hsu, and S. Srinivasa, "Human-robot mutual adaptation in shared autonomy," in *HRI*, 2017.

[10] A. P. Bayliss and S. P. Tipper, "Predictive gaze cues and personality judgments: Should eye trust you?" *Psychological Science*, 2006.

[11] M. F. Mason, E. P. Tatkow, and C. N. Macrae, "The look of love: Gaze shifts and person perception," *Psychological Science*, 2005.

[12] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster, "Eye movement analysis for activity recognition," in *Ubicomp*. ACM, 2009.

[13] A. Fathi, Y. Li, and J. M. Rehg, "Learning to recognize daily actions using gaze," in *ECCV*. Springer, 2012.

[14] C. Braunagel, E. Kasneci, W. Stolzmann, and W. Rosenstiel, "Driver-activity recognition in the context of conditionally autonomous driving," in *IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE, 2015, pp. 1652–1657.

[15] D. Sadigh, S. S. Sastry, S. A. Seshia, and A. Dragan, "Information gathering actions over human internal state," in *IROS*. IEEE, 2016.

[16] E. M. Reingold and H. Sheridan, "Eye movements and visual expertise in chess and medicine," *Oxford handbook on eye movements*, pp. 528–550, 2011.

[17] T. Kübler, S. Eivazi, and E. Kasneci, "Automated visual scanpath analysis reveals the expertise level of micro-neurosurgeons," in *MICCAI Workshop on Interventional Microscopy*, 2015.

[18] J. R. Flanagan, G. Rotman, A. F. Reichelt, and R. S. Johansson, "The role of observers' gaze behaviour when watching object manipulation tasks: predicting and evaluating the consequences of action," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 368, no. 1628, p. 20130063, 2013.

[19] S. Mathôt, L. Van der Linden, J. Grainger, and F. Vitu, "The pupillary light response reveals the focus of covert visual attention," *PLoS One*, vol. 8, no. 10, p. e78168, 2013.

[20] J. Klingner, R. Kumar, and P. Hanrahan, "Measuring the task-evoked pupillary response with a remote eye tracker," in *ETRA*, 2008.

[21] T. Carlson and Y. Demiris, "Collaborative control for a robotic wheelchair: evaluation of performance, attention, and workload," *IEEE Transactions on Systems, Man, and Cybernetics*, 2012.

[22] Q. Ji, Z. Zhu, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue," *IEEE transactions on vehicular technology*, vol. 53, no. 4, pp. 1052–1068, 2004.

[23] N. A. Kloosterman, T. Meindertsma, A. M. Loon, V. A. Lamme, Y. S. Bonneh, and T. H. Donner, "Pupil size tracks perceptual content and surprise," *European Journal of Neuroscience*, 2015.

[24] W. Einhäuser, J. Stout, C. Koch, and O. Carter, "Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry," *Proceedings of the National Academy of Sciences*, 2008.

[25] E. Granholm and S. R. Steinhauer, "Pupillometric measures of cognitive and emotional processes," *International Journal of Psychophysiology*, vol. 52, no. 1, pp. 1–6, 2004.

[26] S. P. Marshall, "The index of cognitive activity: Measuring cognitive workload," in *Human factors and power plants*. IEEE, 2002.

[27] O. Palinko, A. L. Kun, A. Shyrokov, and P. Heeman, "Estimating cognitive load using remote eye tracking in a driving simulator," in *Proceedings of the 2010 symposium on eye-tracking research & applications*. ACM, 2010, pp. 141–144.

[28] E. Kasneci, T. Kübler, K. Broelemann, and G. Kasneci, "Aggregating physiological and eye tracking signals to predict perception in the absence of ground truth," *Computers in Human Behavior*, 2017.

[29] G. Vergani Dambros, J. Ungewiss, T. Kübler, E. Kasneci, and M. Spüler, "Monitoring response quality during campimetry via eye-tracking," in *IUI*. ACM, 2017.

[30] Y. Shi, N. Ruiz, R. Taib, E. Choi, and F. Chen, "Galvanic skin response (gsr) as an index of cognitive load," in *CHI'07 extended abstracts on Human factors in computing systems*. ACM, 2007, pp. 2651–2656.

[31] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on intelligent transportation systems*, vol. 6, no. 2, pp. 156–166, 2005.

[32] D. McDuff, S. Gontarek, and R. Picard, "Remote measurement of cognitive stress via heart rate variability," in *EMBC*. IEEE, 2014.

[33] W. Chen and R. W. Picard, "Eliminating physiological information from facial videos," in *FG*. IEEE, 2017.

[34] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: a review," *Medical and biological engineering and computing*, vol. 44, no. 12, pp. 1031–1051, 2006.

[35] T. C. Kübler, E. Kasneci, W. Rosenstiel, U. Schiefer, K. Nagel, and E. Papageorgiou, "Stress-indicators and exploratory gaze for the analysis of hazard perception in patients with visual field loss," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 24, pp. 231–243, 2014.