THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 21, Number 2/2020, pp. 173–178

ON THE ROBOTIC CONTROL BASED ON INTERACTIVE ACTIVITIES OF SUBJECTS

Ligia MUNTEANU, Cristian RUGINĂ, Ciprian DRAGNE, Veturia CHIROIU

Institute of Solid Mechanics of the Romanian Academy, Ctin Mille 15, 010141 Bucharest, Romania
Corresponding author: Ligia MUNTEANU, E-mail: ligia.munteanu@imsar.ro

Abstract. We present a novel control system based on the capturing of the motion and the intent of the motion of each subject engaged into an action (a game for example) monitored by a mobile supervising robot. The robot observes the subjects' activity and spatio-temporal motions starting from zero information regarding the field of view and the rules of game. The robot observes the subjects' relationships, their goals and intentions, exploring the uncertain and dangerous situations which could appear for players or for both the subjects and environment. The robot learns the roles and rules that players apply, and captures information about each subject and each type of activity. After watching the game for a short period of time, the robot is able to participate to the game. It acquires a full control over the nondescript avatar and reacts immediately when a particular event or motion intend to become dangerously for subjects or environment, and gives the tracking alert.

***Key words**:* robotics; interactive behavior; intensions; control; games.

1. INTRODUCTION

Humans can easily assign intentions, personalities and targets to sterile geometric shapes. Simple heuristics make humans smart in gaining the firm control over unknown avatars, the own intentions and wishes in their activities [1,2].

Motion information is a cognitive attribute that humans use every day. This instinct grows rapidly in children around age 9 months [3]. The ability to process the information is quickly followed by other social skills, such as agency awarding and intentionality [4]. This information, though little, is sufficient in designing the computational process that is capable to manipulate a small amount of data to obtain similar results. Moreover, this can be done fast enough to serve as a robotic control, which allows to explore the relationship between watching and participating to the game. The robot can test uncertainties in the learning process, the collapse ambiguity by experiments and explore how the control concerns the interaction between subjects and the environment [5].

As robots occupy roles reserved for humans, the need for robots to interact with people has grown. We have robots for visits in museums [6], robots acting as receptionists [7] and robots that work as pets [8]. However, robots must have the ability to perceive the environment in a way compatible with how people see and understand the world [9]. The interaction robot-human uses a vision technique inspired from nature, i.e. the visual attention model [10]. This model is a rough approximation of the human visual attention system, and usually works by identifying points of interest in a visual flow.

Visual attention, its role in memory and development is an interesting research subject on how infants and children orient to their environment. Most specialized articles of visual attention are oriented to the explanations on the capturing of external environmental stimuli that can also be driven by own internal human goals or plans [11–15].

In this paper, a mobile supervising robot able to capture the movement of the supervised subjects and the intentions of their future movements is controlled. The control is based on the fundamental cognitive, the perception, the interpretation of the motion and the individualization of subconscious moving in context of causal perceptions [16–19].

The robot is equipped by a camera for taking the images, a computing system of the average velocity and the intended velocity vectors, a controller for collision avoidance for avoiding getting the robot into the walls or corners, or collision between subjects or collision between subjects and surrounding objects [20-22].

2. THE MOTION ANALYSIS

Let us consider that N subjects are moving in a field of view. To describe the motion of one, two or more subjects, the Markov chains will define the state vectors as probability vectors showing the probability for an event after a certain number of time iterations. The state vectors at different times are related for each iteration by

$$\mathbf{x}(k+1)=\mathbf{P}\,\mathbf{x}(k)\,, \tag{1}$$

where $\mathbf{P}=(p_{ij})$ is a stochastic transition matrix with each column a probability vector showing the probability that a system which is in the state j at time $t=1$ will be in the state i at time $t=k+1$. In many situations, the state vector remains the same after a number of iterations, that means this vector is a steady-state vector. If a subject move into an area divided into three areas, called zones 1, 2 and 3, for example, by leaving an area, the subject may have one of other two zones as a future destination. The $\mathbf{A}$ is a $n\times n$ matrix which shows how many times the subject moves in another area in a period of time. For $n=3$, for example, we can write

$$\mathbf{A}=\begin{bmatrix}0 & a & b\\ c & 0 & d\\ e & f & 0\end{bmatrix},$$

where zero, a and b in the first row indicate that the subject located initially in zone 1 moves in zone 2 for a – times and zone 3 for b – times. The motion of the subject is a Markov chain with the transition matrix $\mathbf{P}$

$$\mathbf{P}=\begin{bmatrix}0 & a/a+b & b/a+b\\ c/c+d & 0 & d/c+d\\ e/e+f & f/e+f & 0\end{bmatrix}. \tag{2}$$

The eigenvector of $\mathbf{A}$ is the vector $\mathbf{x}$, $\mathbf{A}\mathbf{x}=\lambda\mathbf{x}$, with λ the eigenvalue of $\mathbf{A}$. The location x_i of the subject i in the field of view is

$$x_i=\frac{1}{\lambda}\sum_{j=1}^{n}A_{ij}x_j\ . \tag{3}$$

Not only spatio-temporal features of each subject's motion are used for tracking, but more complex features of the subjects and environment, such as the sex, volume, height of each subject, the color, texture, and shape of surrounding things, are employed. The work does not include the face detection of the subjects.

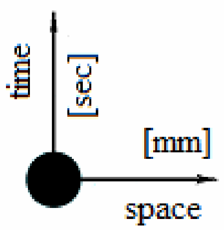

Fig. 1 – Spatial and temporal scales.

Computational of multi-target subject tracking is based on the hypothesis proposed by Reid [23] and implemented by Cox and Hingorani [24].

The spatial and temportal scales for computing is shown in Fig. 1. The motion of subjects is described using the Newton's equation $F = ma$, with m is the subject mass $\mathrm{d}m = \rho \mathrm{d}V$, with ρ density and V the volume, and a the acceleration. The calculation is complex so that no more than 30–40 subjects can be included into the game.

The motion detection produces zones of motion and their respective centroids. These centroid locations form a stream of locations $\left\{x_t^1, x_t^2, ... x_t^k\right\}$ with k targets at each time t. he objective is to produce a labeled trajectory defined as a set of points, one from each time, to identify a single subject as it moves through the field of view

$$T = \left\{x_1^{i_1}, x_2^{i_2}, ..., x_t^{i_n}\right\}. \tag{4}$$

The number of subjects at each time is not constant, and the existence of a subject at the time t to the next time $t+1$ is uncertain. So, a counter needs to computee the number of subjects that enter and leave the scene. The phantom points with non-defined locations can help to identify the trajectories of subjects that enter, exit, or are not caught in the field of vision. To each new point, a set of hypotheses link the point to previous trajectory. These hypotheses include the false alarms, non-detection events, extensions of previous trajectories, and the starts of new paths [25].

Given a subject, the robot begins to observe the motion of others subjects in order to start working the rules of the game. For each of other two subjects, the robot calculates the influence of the remaining subjects (including itself) on the first subject [26]

$$v_{x_i^t} = \frac{c_{x_j}(x_j^t - x_i^t)}{d_{ij}^t} + \frac{c_{x_k}(x_k^t - x_k^t)}{d_{ik}^t}, \tag{5}$$

where $v_{x_i^n}$ is the velocity of the subject i at time t, and d_{ij}^t is the Euclidean distance between the subject i and the subject j.

The edges of the workspace have an effect on the players' motion in the sense that their play is constrained by the physical dimensions of the space. Avoiding the wall collisions is simple. If the player approaches too much on the wall, the force is repulsive and the player changes direction. This danger zone can vary up to 30×30 pixels. The goal vector indicates proximity to the wall or corner and asks the robot to control everything it had tried beforehand (Fig. 2). The blue zones indicate the goal vector which is compared to the computed vector of motion.

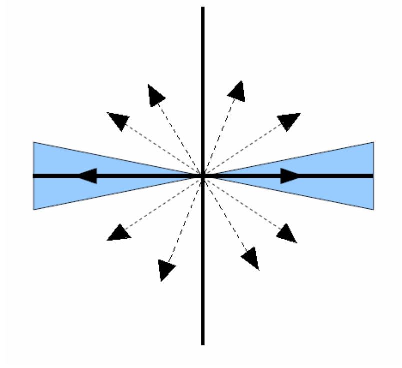

Fig. 2 – Directional control of robot.

The attraction and repulsion between subjects and between subjects and other entities (surrounding objects, corners and walls) are described by a potential $H(r_{ij}) = \left(r_0/r_{ij}\right)^6$, from which the forces acting on the subjects are determined as

$$F_i = \sum_{j>i} \nabla_{r_i} H(r_{ij}), \tag{6}$$

where r_{ij} is the distance between subjects i and j. It is hypothesized that each subject interacts with other entities and inside a sphere with a specified radius r_0. If the distance between two subjects is greater than r_0 then the force is attractive, and if the distance is less than r_0, the force is repulsive.

The Wolfe's model of visual search and attention is applied in this paper [27]. The robot produces a map of attention through a weighted combination of observations given by various detectors $(s_1,...,s_n)$ (color, motion and shape). This combination allows the robot to select the regions of interest and then, it directs the computer to these regions. The attention system is based on the stimulus-adapting mechanism and runs all the time. The correlation between two different sets of visual data is measured by the correlation coefficient. The correlation coefficient $-1 \le r \le 1$, for k sets of data $(s_1^1, s_2^1,....s_n^1)\,...\,(s_1^k, s_2^k,....s_n^k)$ is calculated by

$$r = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(s_j^i - \overline{s}^i)}{\sqrt{\sum_{i=1}^{k}\sum_{j=1}^{n}(s_j^i - \overline{s}^i)^2}}, \tag{7}$$

where $\overline{s}^i$ are the mean value of the i- set of data. For $-0.5 \le r \le 0.5$, a little correlation is observed, while for $-0.8 \le r \le 0.8$, a strong correlation appears [26].

The robot builds an evolutionary set of intentions of motion for subjects, modeled as a probability distribution over all possible motion states. The current distribution of the motion is adjusted to introduce a probable alternative to the game. The intention to move to any particular state S at time t, noted by $\mathrm{Int}_t(S)$, is the intent of motion at time $t-1$ modified by the current observation

$$\mathrm{Int}_t(S) = \frac{\mathrm{Int}_{t-1}(S)\left(1 + \lambda \sum_{c \in S} s(c_t)\right)}{Z}, \tag{8}$$

where c_t is the value at time t of one of the pairwise relationship constants derived from the previous step data, Z is a normalized constant obtained by summing the values of the updated intentions in all states, s is the sign function with value 1 if the sign of the constant and the intention represented by the current state agree, and -1 if they disagree, and 0 if the state is neutral with respect to the pairwise relationship of constant, λ is a learning rate constant that affects the compromise between the sensibility of the system to error and its decision-making speed. Depending on whether the game is simply observed or if the robot is actively participating to the game, λ varies between 0.04 and 0.12.

3. A CASE OF STUDY

We tested the control system for a volleyball game that involve six kids on a given playground (Fig. 3 and Fig. 4). Each kid has only one unchanging goal. In each trial, the behavior of the game is consistent.

Within less than a second, the robot determines the intentional states of two kids (a boy with mauve undershirt and a girl with black hear) with respect to each other. Then, it proceeds to generate and test hypotheses regarding their intentions toward themselves. Within a five seconds, the robot is able to determine the dangerous situations for kids of being prevent for falling. Since the intentional state never changed, no positional information is recorded or analyzed. The robot intervenes and alerts.

We ran the system with and no control. Figure 5 shows the incidents with the boy and the girl in the case of no control.

The participation of the robot to the game gives him an extrapower to investigate the intentional states of kids. For comparison, we ran the version of the game that involves the same two kids and the robot in the role of the active observer. The system applies the same algorithm to hypothesize the intentions of the kids, and converges on a stable, correct intentional state. The same as before, within a five seconds, the robot predicted the dangerous situations and intervenes to alert.

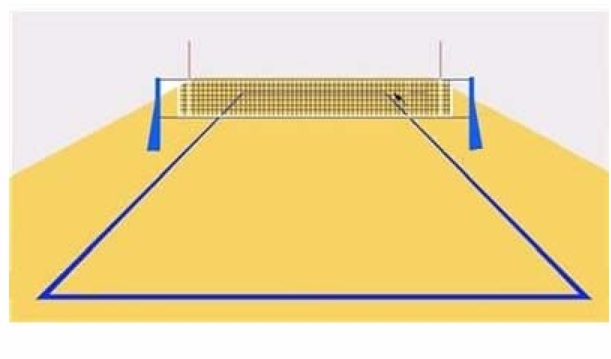

Fig. 3 – The playground for a volleyball game.

Fig. 4 – The players.

Fig. 5 – Two incidents with a boy and a girl during the game.

A filterless color image sensor was designed in [28] with application to the robotic control based on interactive activities of subjects. The filterless color image sensor is made from a cell with light sensitivity, a lower and an upper electrodes. The chalcogenide thin film is placed between the electrodes together with an imaging circuit that measures the wavelength and the intensity of the incident light beam.

We intend to apply this kind of control to surgical interventions where the robot and the physician are working with the same surgical tool. The control refers to the trajectory tracking and preventing the surgeon-tool to cross the critical boundaries and to move towards wrong directions [29, 30].

The surgeon manages freely the surgical instrument without robotic control, but when it reaches a critical point or forbidden area, the robot must intervene to force the surgeon to stop the surgical instrument before reaching the end point or choosing another route.

4. CONCLUSIONS

The paper advances a novel control system based on the capturing of the motion and the intent of the motion of each subject engaged into an action supervised by a mobile robot. The motion capture is obtained from the dynamics of the system trajectories at a given time and its extension in an immediate vicinity in space and time. The robot observes the spatio-temporal motions of players starting from zero information regarding the workspace and rules of the game. It observes the subjects' relationships, their goals and intentions, exploring the uncertain and dangerous situations which could appear for players or for both the subjects and environment. The robot acquires a full control over the game and reacts immediately when a particular event or motion intend to become dangerously for subjects or environment.

ACKNOWLEDGEMENTS

This work was supported by a grant of the Romanian ministry of Research and Innovation, CCCDI – UEFISCDI, project number PN-III-P1-1.2-PCCDI-2017-0221/59 PCCDI/2018 (IMPROVE), within PNCDI III.

REFERENCES

1. G. GIGERENZER, P.M. TODD, *Simple heuristics that make us smart*, New York, Oxford University Press, 1999.
2. F. HEIDER, M. SIMMEL, *An experimental study of apparent behavior*, American Journal of Psychology, **57**, pp. 243-259, 1944.
3. P. ROCHAT, T. STRIANO, R. MORGAN, *Who is doing what to whom? Young infants' developing sense of social causality in animated displays,* Perception, **33**, pp. 355-369, 2004.
4. G. CSIBRA, G. GERGELY, S. BIRO, O. KOOS, M. BROCKBANK, *Goal attribution without agency cues: The perception of "pure reason" in infancy*, Cognition, **72**, pp. 237-267, 1999.
5. D.M. WOLPERT, K. DOYA, M. KAWATO, *A unifying computational framework for motor control and social interaction*, Philosophical Transactions of the Royal Society B, **358**, *1431*, pp. 593-602, 2003.
6. W. BURGARD, A.B. CREMERS, D. FOX, D. HAHNEL, G. LAKEMEYER, D. SCHULZ, W. STEINER, S. THRUN, *The interactive museum tour-guide robot*, Proceedings of the Fifteenth National/Tenth Conference on Artificial intelligence/innovative Applications of Artificial intelligence, 1998, pp. 11-18.
7. R. GOCKLEY, A. BRUCE, J. FORLIZZI, M. MICHALOWSKI, A. MUNDELL, S. ROSENTHAL, B. SELLNER, R. SIMMONS, K. SNIPES, A. SCHULTZ, J. WANG, *Designing robots for long-term social interaction*, IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005, pp. 2199-2204.
8. M. FUJITA, *AIBO: Toward the era of digital creatures*, International Journal of Robotics Research, **20**, pp. 781-794, 2001.
9. T. FONG, I. NOURBAKHSH, K. DAUTENHAHN, *A survey of socially interactive robots*, Robotics and Autonomous Systems, **42**, pp. 143-166, 2003.
10. C. BREAZEAL, B. SCASSELLATI, *A context-dependent attention system for a social robot*, Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (ed. T. Dean), 1999.
11. D. AMSO, S. HAAS, E. TENENBAUM, J. MARKANT, S. SHEINKOPF, *Bottom-up attention orienting in young children with autism*, Journal of Autism and Developmental Disorders, **44**, *3*, pp. 664-73, 2014.
12. D. AMSO, J. MARKANT, S. HAAS, *An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes*, PLOS ONE, **9**, *1*, p. e85701, 2014.
13. J. MARKANT, D. AMSO, *Leveling the playing field: Attention mitigates the effects of IQ on memory*, Cognition, **131**, *2*, pp. 195-204, 2014.
14. J. MARKANT, M. WORDEN, D. AMSO, *Not all attention orienting is created equal: combined behavioral and neuroimaging evidence that memory for objects is enhanced when attention orienting at encoding involves distractor suppression*, Neurobiology of Learning and Memory, **20**, pp. 28-40, 2015.
15. G.D. REYNOLDS, A.C. ROMANO, *The development of attention systems and working memory in infancy*, Frontiers in Systems Neuroscience, **10**, p. 15, 2016.
16. H. CHOI, B. SCHOLL, *Perceiving causality after the fact: Postdiction in the temporal dynamics of causal perception*, Perception, **35**, pp. 385-399, 2006.
17. A.M. LESLIE, F. XU, P.D. TREMOULET, B.J. SCHOLL, *Indexing and the object concept: Developing "what" and "where" systems*, Trends in Cognitive Sciences, **2**, *1*, pp. 10-18, 1998.
18. S.R. MITROFF, B.J. SCHOLL, *Forming and updating object representations without awareness: Evidence from motion-induced blindness*, Vision Research, **45**, pp. 961-967, 2004.
19. B.J. SCHOLL, *Can infants' object concepts be trained?*, Trends in Cognitive Sciences, **8**, *2*, pp. 49-51, 2004.
20. P. WOLFF, *Representing causation*, Journal of Experimental Psychology, **136**, pp. 82-111, 2007.
21. L. TALMY, *Force dynamics in language and cognition*, Cognitive Science, **12**, pp. 49-100, 1988.
22. D.R.J. LAMING, *Information theory of choice-reaction times*, London, Academic Press, 1968.
23. D.B. REID, *An algorithm for tracking multiple targets*, IEEE Transactions on Automated Control, AC-**24**, *6*, pp. 843-854, 1979.
24. I.J. COX, S.L. HINGORANI, *An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking*, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), **18**, *2*, pp. 138-150, 1996.
25. B. SCASSELLATI, *Theory of mind for a humanoid robot*, Autonomous Robots, **12**, pp. 13-24, 2002.
26. C. CRICK, B. SCASSELLATI, *Controlling a robot with intention derived from motion*, Topics in Cognitive Science, **2**, pp. 114-126, 2010.
27. J.M. WOLFE, *Guided Search 2.0: A Revised Model of Visual Search*, Psychonomic Bull. & Rev., **1**, *2*, pp. 202-238, 1994.
28. V. CHIROIU, N. NEDELCU, V. MOȘNEGUȚU, *Inteligent control of a visual attention system based on the chalcogenide glass optical fibers*, Analele Universității "Eftimie Murgu" din Reșița, **26**, *1*, pp. 31-44, 2019.
29. V. CHIROIU, L. MUNTEANU, C. RUGINA, *On the control of a cooperatively robotic system by using a hybrid logic algorithm*, Proceedings of the Romanian Academy, Series A: Mathematics, Physics, Technical Sciences, Information Science, **19**, *4*, pp. 589-596, 2018.
30. V. CHIROIU, L. MUNTEANU, C. DRAGNE, C. ȘTIRBU, *On the differential dynamic logic model for hybrid systems*, Acta Technica Napocensis – Series: Applied Mathematics, Mechanics, and Engineering, **61**, *4*, pp. 533-538, 2018.

Received June 16, 2019