PALADYN Journal of Behavioral Robotics

VERSITA

Research Article · DOI: Proofreading version JBR · 1(1) · 2011 · 1-7

A dynamical systems approach to online event segmentation in cognitive robotics*

Bruno Nery [1†] **Rodrigo Ventura** [12‡]

1 Institute for Systems and Robotics
Instituto Superior Técnico
Av. Rovisco Pais, 1; 1049-001 Lisboa;
PORTUGAL

2 Also visiting scholar at Computer Science Department,
Carnegie Mellon University.

Received

Accepted

Abstract

This paper addresses the problem of segmenting perception in physical robots into meaningful events along time. In structured environments this problem can be approached using domain-specific techniques, but in the general case, as when facing unknown environments, this becomes a non-trivial problem. We propose a dynamical systems approach to this problem, consisting of simultaneously learning a model of the robot interaction with the environment (robot and world seen as a single, coupled dynamical system), and deriving predictions about its short-term evolution. Event boundaries are detected once synchronization is lost, according to a simple statistical test. An experimental proof of concept of the proposed framework is presented, simulating a simple active perception task of a robot following a ball. The results reported here corroborate the approach, in the sense that the event boundaries are correctly detected.

Keywords

Event segmentation · anticipative systems · active perception · cognitive robotics

1. Introduction

This paper is an extended version of the work originally presented at the First International Conference on Biologically Inspired Cognitive Architectures (BICA-2010) [9].

Decision-making systems usually assume a meaningful division of time into discrete frames. For instance, decision theoretic agents determine their action based on a given situation, action options, and possible outcomes. Both the situation and the outcomes are represented with respect to relevant time frames: the current situation frame, and the action outcome frames. The application of these discrete-time frameworks to physical robots faces the problem of the semantic discretization of perception into meaningful frames. However, perception in these systems is originated by sensor devices with fixed time resolution. The segmentation of perception into frames in a data-driven fashion is a non-trivial problem, unless specific domain-dependent techniques can be applied.

This paper focuses on a particular kind of division: *events.* We address the problem of event segmentation in perception, using a dynamical systems approach. Since we target embodied and embedded agents, we make two further assumptions: first, robot and environment are here seen as a single, coupled system, and second, the sensor data processing latency is taken into account for control purposes.

*This work was supported by the FCT (ISR/IST plurianual funding) through the PIDDAC Program funds. Partially funded with grant SFRH/BD/60853/2009, from Fundação para a Ciência e a Tecnologia.
†E-mail: bnery@isr.ist.utl.pt
‡E-mail: rodrigo.ventura@isr.ist.utl.pt

This problem is addressed using a biologically inspired approach. The Event Segmentation Theory (EST) provides a model of how the human brain segments perception into a sequence of events [7, 19]. This model assumes that event segmentation is based on the detection of prediction errors in the sensory stream. Prediction is a common mechanism found throughout the brain. In particular, it is known that the human brain is permanently making predictions and comparing them with the actual outcome [13]. Events are detected whenever a significant disparity between prediction and outcome is encountered. An event segmentation mechanism can be built following this principle, but the problem of how to make predictions about perceptions has to be addressed first.

Dubois distinguishes between ***weak*** and ***strong anticipation*** [3, 14]: the former is based on an explicit model of the world, where the physical system is modeled by analytical constructs, that can be mathematically solved given an initial condition. On the contrary, ***strong anticipation*** does not rely on a model, but rather on the dynamical evolution of the interaction of the agent with the world, seen as a single system. An example of strong anticipation can be found on the behavior of an outfield baseball player when catching a well-struck ball[1]: weak anticipation of the ball landing position requires modeling the physics of the ball, encoding the initial state of the system (initial velocity, mass, friction coefficient, etc), and then predicting the landing position by solving the analytical model; in contrast, strong anticipation views the outfielder and the ball as a single system with new dynamics, as the outfielder moves itself driven by the projection of the ball on his retina. Empirical evidence suggests that this is the way a human outfield player performs the catch [14]. In the context of robotics, a model-based approach to

1 *Example from [14].*

anticipation may be appropriate for passive sensors, but when designing systems that actively engage in interactions with the world, as in the case of active perception, the world can no longer be modeled as an independent, self-contained system.

Stepp proposes an approach to strong anticipation grounded in the field of chaotic systems concerning synchronization of dynamical systems [14]. Consider two systems, denoted D (drive) and R (response), connected by a unidirectional flow of information from D to R. It is possible to design the system R such that its dynamic evolution synchronizes with the one of D, regardless of the initial condition of each system. One way of doing this is for the R system to compare its state with the one of D, and bias its dynamics accordingly, *i.e.,* system R is controlled by a feedback loop, where the error results from this comparison. More interestingly, if this feedback loop contains a delay, system R is capable, under certain conditions, of anticipating system D [15]. Considering that system D includes both the robot and the world, and system R to be a model internal to the robot, this approach suggests an interesting mechanism to perform strong anticipation of the dynamical evolution of the world-robot system.

One problem remains to be solved: how to design system R. No system model is assumed *a priori*, since it depends on the coupling involving the robot and the world. The approach taken here is to adapt system R during interaction. A solution to the adaptation of response systems in the context of dynamical systems synchronization has been proposed by Chen [1], where the convergence to the solution has been proved using the Lyapunov's indirect method.

The contributions of this paper are:

1. an event segmentation method based on Stepp's strong anticipation principle [14], cast into an anticipating system synchronization framework;
2. the application of Chen's parameter identification method [1] to anticipating synchronization;
3. a proof-of-concept implementation of an architecture for event segmentation and active perception, employing these methods.

This paper is organized as follows: after section 2 surveying related work, sections 3 and 4 provide the theoretical background behind strong anticipation and response system adaptation. Then, the proposed architecture for event segmentation is described in section 5, followed by experimental results of a proof of concept implementation of these ideas in section 6. Finally, section 7 presents a conclusion and open questions, thus closing the paper.

2. Related work

The problem of event segmentation in perception has been studied in various contexts. See [11] for a review of recent techniques for the formation of event memories in robots. However, most of these techniques make strongs assumptions on the nature of events, such as events being repeatable signal patterns, events being changes of sensor readings, or events activating heuristic, *ad-hoc* "triggers". Ramoni *et al.* proposed a method to cluster robot activities using Markov chain models [12], which is based on symbolic perceptual data, rather than raw sensory input. In [4] a maximum likelihood estimator is used to fit a sequence of time-indexed models to raw data. A batch and an online version of this approach are proposed. The incremental one is based on thresholding the likelihood of the current parametric model along time. However, these models correspond to signal patterns, and thus they are a more limited representation than dynamical systems models. The spatio-temporal segmentation of video has been researched in [17], being applied to motion model clustering, and in [2] using hierarchical clustering of the 3D space-time video stream. However, both of these approaches are based on simple linear motion models and are specific to video image processing. Gesture segmentation and recognition has been addressed in [6] employing hidden-Markov models (HMM). However, this approach requires a prior offline training of the HMM with a dataset. It is therefore not appropriate if event segmentation precedes learning and recognition in the perception chain.

To the best of our knowledge, the approach proposed in this paper is novel in the use of adaptive synchronization for event segmentation. No publications were found by the authors concerning the usage of either dynamical systems synchronization, or Chen's adaptive method, outside the area of chaotic systems. It should also be stressed that our approach is neither directed towards a specific application domain, nor a sensor modality.

3. Strong anticipation

In [14] strong anticipation is modeled using a dynamical systems synchronization framework. Consider two continuous-time state vectors $x(t), y(t) \in \mathbb{R}^n$ with the following coupled dynamics:

$$\begin{aligned} \dot{x} &= g(x) \\ \dot{y} &= g(y) + k(x - y_\tau) \end{aligned} \tag{1}$$

where $g : \mathbb{R}^n \to \mathbb{R}^n$ encodes the dynamics of the system, and $y_\tau = y(t - \tau)$ is the delayed version of y. The first system is called the *drive*, while the second the *response*. The response system contains a feedback loop with a constant delay τ, and k is a scalar gain. This delayed feedback loop in the response system is a fundamental aspect, and is responsible for the response system's capability of anticipating the evolution of the drive.

This delayed feedback loop is neurophysiologically supported by the discovery of forward models in the brain, which predict sensory consequences of motor commands [5, 8, 18]. These models receive as input a copy of the subject motor action, and produce a prediction of future perceptions. For instance, when performing an arm movement, these models predict the trajectory followed by the arm, as perceived by the subject. One important function of this mechanism is to overcome the sensory processing latency in the brain, when the subject is performing controlled, quick movements.

To understand how the response system can anticipate the drive, consider that $\tau = 0$ and that the systems are synchronized at time t_0, *i.e.,* $x(t_0) = y(t_0)$. Under these conditions, the systems will remain synchronized, since $x - y_\tau = 0$ and thus there is null feedback in the response. In this case, the concatenated state $z = (x, y) \in \mathbb{R}^{2n}$ evolves in the $x = y$ hyperplane, called the *synchronization manifold* [10]. The response system synchronizes with the drive if the error system with state $e = y - x$, also called the *transversal system*

$$\dot{e} = g(y) - g(x) - k\, e \tag{2}$$

is able to reject the perturbation e, driving it to zero. For $g(y) \simeq g(x)$, system (2) behaves as a first-order system with an exponential decay to zero. Anticipation is realized once $\tau > 0$, as synchronization implies $x(t) = y_\tau = y(t - \tau)$ and thus $y(t) = x(t + \tau)$, meaning that the response anticipates the driver. This is called *anticipating synchronization* [15], where $x = y_\tau$ defines the *anticipatory manifold* [16].

Successful synchronization from an arbitrary initial condition (or whenever $g(y)$ differs significantly from $g(x)$) is not guaranteed in general (except for simple cases), and strongly depends on the values of k and τ. However, for any delay value τ, $e(t) = 0$ is a fixed point of the transversal system (2), meaning that once synchronized, the system will remain so. Voss conjectures that, if $e(t) = 0$ is a stable fixed point for $\tau = 0$, then there is a $\tau_0 > 0$ such that, for any $0 < \tau < \tau_0$, the transversal system has a stable fixed point at $e(t) = 0$. This conjecture has been backed up by numerical simulations [16].

In general, for sufficiently small τ, stability of the transversal system can be expected. In the case of this work, since τ models the delay of the perceptual system (e.g., the latency from a change in the environment up to its detection by the computer vision algorithm), this delay can be assumed smaller than the time scale of the events being perceived.

4. Adaptive synchronization

If the drive system corresponds to the world-robot coupled system, its dynamics are not known *a priori*. One way of tackling this problem is to adapt the response system, online, during synchronization. Chen proposed in [1] an approach to adapt response systems in the context of dynamical systems synchronization. It does not account, however, for a delayed feedback.

Consider that the drive system has the form

$$\dot{x} = f(x) + F(x)\theta \tag{3}$$

where $\theta \in \mathbb{R}^m$ is a vector of (constant) parameters, $f : \mathbb{R}^n \to \mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^{n\times m}$. The response system is identical, except for the parameter vector that is unknown, and for the synchronization feedback loop

$$\dot{y} = f(y) + F(y)\alpha + U(y, x, t, \alpha) \tag{4}$$

where α is the response parameter vector, and $U(y, x, t, \alpha)$ is called the controller of the response. Using Lyapunov's indirect method, Chen *et al.* proved in [1] that, under certain conditions, not only the response system synchronizes with the drive, but also that the response parameters α converge to the parameters of the drive θ, *i.e.,*

$$\lim_{t\to+\infty} ||\alpha(t) - \theta|| = 0. \tag{5}$$

These conditions are the existence of (1) a smooth controller $U(y, x, t, \theta)$, and of a (2) scalar (Lyapunov) function $V(e)$, where $e = y - x$, such that:

1. $c_1||e||^2 \leq V(e) \leq c_2||e||^2$,
2. the derivative of $V(e)$ along the solution of the coupled system (3) and (4) with $\alpha = \theta$ satisfying $\dot{V}(e) \leq -W(e)$, and
3. the parameter vector α is adapted according to the learning rule

$$\dot{\alpha}(t) = -F^T(x)\left[\nabla V(e)\right]^T \tag{6}$$

   for $\nabla V(e)$ denoting the gradient (row) vector of V with respect to e,

where c_1 and c_2 are two positive constants, $W(e)$ is a positive definite function, and $U(y, y, t, \theta) = 0$.

This result has two important consequences: first, it proves global asymptotic convergence, provided that the response system is capable of synchronizing with the drive if $\alpha = \theta$ (*i.e.,* if the true parameters were known), and second, it provides a learning law, in the form of the gradient of α. However, in order to use this result, one has to find a controller U and a function V satisfying the hypothesis of the theorem. Chen showed also that the feedback linearizing controller

$$U(y, x, t, \theta) = -e + f(x) - f(y) + [F(x) - F(y)]\,\theta \tag{7}$$

and the Lyapunov function

$$V(e) = \frac{1}{2}e^T e \tag{8}$$

satisfy the hypothesis for any F and f.

The practical application of these results raises three practical issues. One is the assumption that functions F and f are known, meaning that one should have a prior knowledge of the structure of the dynamics of the system. This argument can be turned around by stating that, given functions f and F sufficiently generic, this method allows the adaptation to any dynamical system that can be modeled by (3) for some parameter vector θ. Second, this result was proved for continuous time systems. The discretization of $\dot{\alpha}$ raises the issue of the choice of a learning rate (hidden in a proportionality constant of V, since the theorem is invariant to a change of scale of this Lyapunov function). However, a good approximation should be expected with a sufficiently small discretization step. Finally, the third issue concerns hidden state variables: if there is a state variable that is hidden, in the sense that the Lyapunov function $V(e)$ does not depend on its error, then $V(e)$ is no longer positive definite. This requires that all drive state variables have to be fed to the response system controller. In this paper, however, we assume full observability of the state variables. This is the case when the state variables correspond to observable quantities, as often happens in perception (e.g., the outfield baseball player example above).

5. Event segmentation

The event segmentation framework (Figure 1) we propose in this paper consists of a pair of response systems, one performing adaptation (labeled *adaptive response*), and the other anticipation (labeled *anticipating response*). The adaptive response learns the parameter vector α as described in section 4, while anticipating response performs anticipating synchronization as explained in section 3. The drive system includes the coupled dynamics of the robot and the world. The access of the architecture to the world state (the perception) is subject

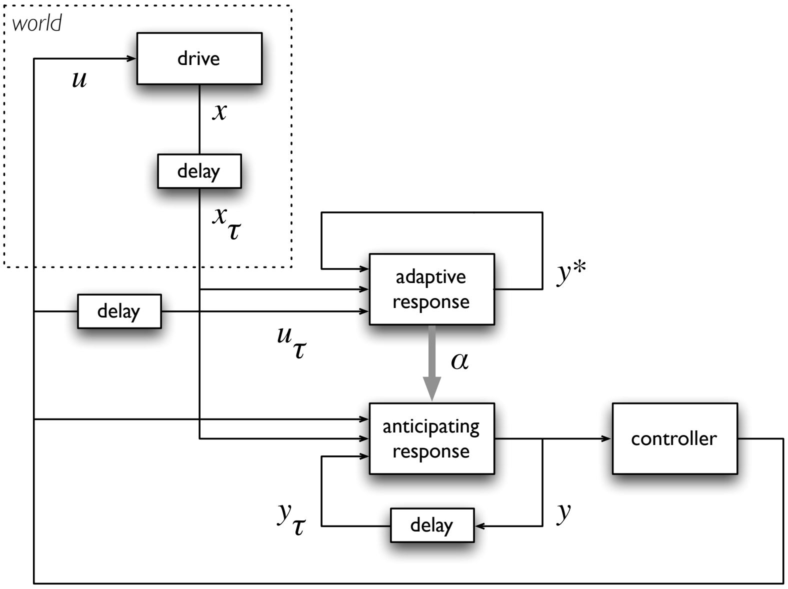

Figure 1. System architecture, consisting of the drive system and the perceptual delay (in the world *block), and the two response systems: the* adaptive *and the* anticipating *ones. The anticipating response uses the parameters α obtained by the adaptive response. The control input u is obtained by a controller fed with the anticipated state y.*

to a delay, modeling the latency of the perceptual channel (e.g., image acquisition, processing, and tracking). The controller computes the actuation vector u based on the anticipated world state y, encoding the intention of the robot (e.g., reach a desired state).

The drive system is modeled by the dynamical system

$$\dot{x} = f(x) + F(x)\theta + u \tag{9}$$

where u is the control input, *i.e.,* the actuation of the robot in the world. Shifting this equation by a delay of τ one obtains

$$\dot{x}_\tau = f(x_\tau) + F(x_\tau)\theta + u_\tau \tag{10}$$

where $u_\tau(t) = u(t-\tau)$ and $x_\tau(t) = x(t-\tau)$. This model can be put in the form of (3) defining a time varying function

$$f_\tau(x_\tau, t) = f(x_\tau) + u_\tau \tag{11}$$

from which $\dot{x}_\tau = f_\tau(x_\tau, t) + F(x_\tau)\theta$. The adaptive response receives the delayed state x_τ, together with the delayed control input u_τ

$$\dot{y}^* = f(y^*) + F(y^*)\alpha + u_\tau + U(y^*, x_\tau, t, \alpha) \tag{12}$$

Once $f_\tau(y^*, t) = f(y^*) + u_\tau$, this equation can be put in the form of (4). The anticipating response is described by

$$\dot{y} = f(y) + F(y)\alpha + u + k(y_\tau - x_\tau) \tag{13}$$

where $y_\tau = y(t-\tau)$ as before, and the parameter vector α is the one obtained by the adaptive response. The anticipatory synchronization manifold is defined by $y_\tau = x_\tau$, and thus $y = x$, meaning that the anticipating response is synchronized with the drive system, which is the same to say that it is anticipating the delayed perception x_τ. By shifting (11) in time one can get $f_\tau(y, t+\tau) = f(y) + u$, allowing us to write (10) and (13) as

$$\begin{aligned} \dot{x}_\tau &= f_\tau(x_\tau, t) + F(x_\tau)\theta \\ \dot{y} &= f_\tau(y, t+\tau) + F(y)\alpha + k(y_\tau - x_\tau) \end{aligned} \tag{14}$$

thus matching (1) when $\alpha = \theta$. Note that the time varying nature of the f and F functions does not affect Chen's results from section 4.

According to the theory of Event Segmentation [19], perceptual systems continuously make predictions about perceptual input, and perceive event boundaries when transient errors in prediction arise. On the adaptive synchronization framework, the Lyapunov function $V(e)$ defined in (8), for $e = y^* - x_\tau$ provides an estimate of the prediction error. Considering the function values in a time window, we can statistically model the obtained samples with a random variable with Normal distribution of mean μ_V and variance σ_V^2. Under this assumption, the normalized metric

$$b_V = \frac{V - \mu_V}{\sigma_V} \tag{15}$$

is normally distributed with zero mean and unit variance. When $|b_V|$ exceeds a threshold b_{event}, an event boundary is detected. If b_V is normally distributed with zero mean and unit variance, the cumulative probability of the distribution tails for $|b_V| > b_{\text{event}}$ is the probability of false positive detection. Thus, b_{event} should be sufficiently high so that false positive detection is minimized, but low enough in order to detect the prediction error increase due to a sudden change in the dynamics of the system.

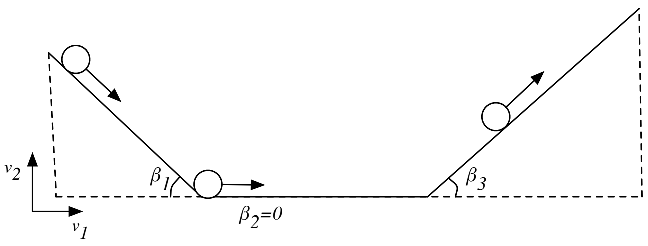

Figure 2. *Simulated scenario, where $\beta_1 = \beta_3 = \pi/12$.*

6. Experimental results

As a proof of concept of the proposed architecture, a simple scenario was simulated: a ball rolling with the effect of gravity over a series of inclined planes, with different slopes, is observed by a robot with a camera seeking to follow the ball in order to center it on the camera image (Figure 2). The robot camera moves parallel to the plane, for the sake of simplicity.

Denoting the ball coordinates by $v = [v_1\ v_2]^T$ and the camera coordinates by $c = [c_1\ c_2]^T$, the ball projection $x = [x_1\ x_2]^T$ in the image plane is thus: $x = v - c$ (apart from a constant factor, here assumed unitary). Assuming frictionless motion, the dynamics of the ball can be described by a double integrator:

$$\begin{aligned} \ddot{v}_1 &= -g \sin\beta \cos\beta \\ \ddot{v}_2 &= -g \sin^2\beta \end{aligned} \tag{16}$$

Considering that the robot movement is also frictionless and that its movement is controlled in acceleration (*i.e.,* force control, assuming unitary mass), the resulting drive system is given by

$$\begin{aligned} \ddot{x}_1 &= -g \sin\beta \cos\beta - \ddot{c}_1 \\ \ddot{x}_2 &= -g \sin^2\beta - \ddot{c}_2 \end{aligned} \tag{17}$$

Considering the state vector $x = [\dot{x}_1\ \dot{x}_2\ x_1\ x_2]^T$, this system can be put in the form of (9) once

$$\begin{aligned} f(x) &= \begin{bmatrix} 0 \\ 0 \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} & F(x) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \theta &= \begin{bmatrix} -g \sin\beta\cos\beta \\ -g\sin^2\beta \end{bmatrix} & u &= \begin{bmatrix} -\ddot{c}_1 \\ -\ddot{c}_2 \\ 0 \\ 0 \end{bmatrix} \end{aligned} \tag{18}$$

In this setup we set the response system to be structurally identical, thus employing the same functions f and F, and control input u. The vector $\alpha = [\alpha_1\ \alpha_2]^T$ is the parameter vector to be adapted according to Chen's learning rule (6).

When the anticipating response is synchronized with the drive, we have $x = y$, and thus the dynamics of the anticipating response becomes

$$\ddot{y} = \alpha - \ddot{c}. \tag{19}$$

The camera motion controller considered has the form

$$\ddot{c} = k_p y + k_d \dot{y} + \alpha \tag{20}$$

where k_p and k_d are the proportional and the derivative gains of the controller. Thus, the closed loop dynamics becomes

$$\ddot{y} = -k_p y - k_d \dot{y} \tag{21}$$

The design of the controller gains k_p and k_d can be performed by pole placement (in the experiments we set $k_d^2 = 4k_p$, yielding a smooth response with a double pole at $-k_d/2$).

The experiments were conducted after discretizing the above equations using the finite difference approximation $\dot{z}(t) \simeq [z(t+T) - z(t)]/T$. The sampling rate was 100Hz, $k_p = 1$, $k_d = 2$, $k = 1$, and the Lyapunov function used was (8). The delay considered was $\tau = 0.65s$ (65 samples). Event boundaries are detected using a 10-second window and a $b_{\text{event}} = 3$. In the experiment the ball was released from the top left position in the scenario (Figure 2), and as the ball transverses the scenario there are two events, corresponding to the two changes of the ramp slope; the simulation ends when the ball reaches the top right position. Each simulation takes 100s of simulated time.

Figure 3 shows the performance of the *adaptive response system* alone, in terms of the evolution of the parameters α (solid line), compared with the ground truth (θ, the dashed line, that changes discretely with the slope). The plot shows the parameter vector α converging to the true parameters θ, after a short period, every time that there is a change in slope (which is the same as a change in θ).

The value of the anticipating response was evaluated by comparing the performance of the controller, with and without anticipation. Figure 4 shows the evolution of the ball position in the camera without anticipation, *i.e.,* the camera motion controller is fed by y^* instead of y: the delay introduced by the latency of the perceptual channel jeopardizes the performance of the controller. Also, the adaptive response follows the drive with a delay of τ. Then, using anticipation, Figure 5 compares the ball position in the camera with its anticipated response. In this case, both are synchronized, since the ball coordinates in the image converge to zero (except for a brief time after each slope change, during which the adaptive system is learning the new parameters). Also, the anticipating response allows for better performance of the controller (in terms of following the ball).

Figure 6 depicts the evolution of the prediction error estimate $V(e)$, for the first 10s of simulation, encompassing the adaptation of the parameters to the initial slope. Its value approaches zero as the drive and response systems become adapted, as well as synchronized. Finally, Figure 7 shows the event segmentation results (vertical dashed lines) obtained using the normalized metric (15), with a window of 10s. As expected, each change of plane by the ball is detected as an event boundary by the framework, corresponding to a failure in prediction.

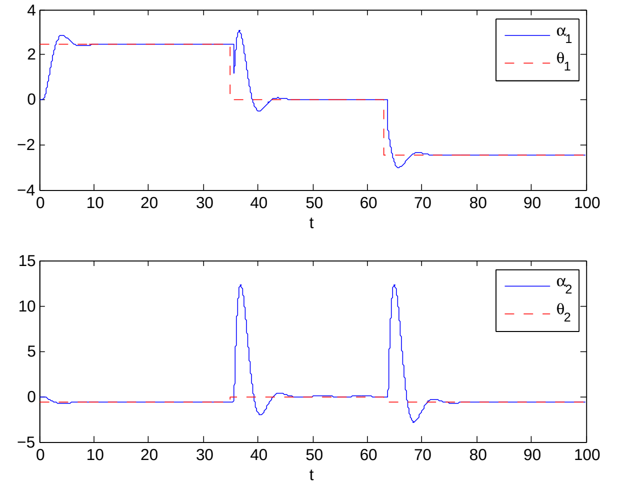

Figure 3. *Parameters α evolution (solid line) in comparison with the true values (dashed line).*

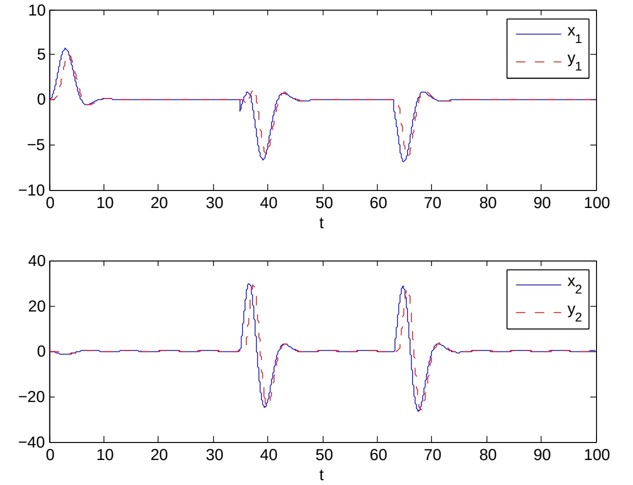

Figure 5. *System response using the full architecture: the solid plot corresponds to the drive system, and the dashed one to the response system.*

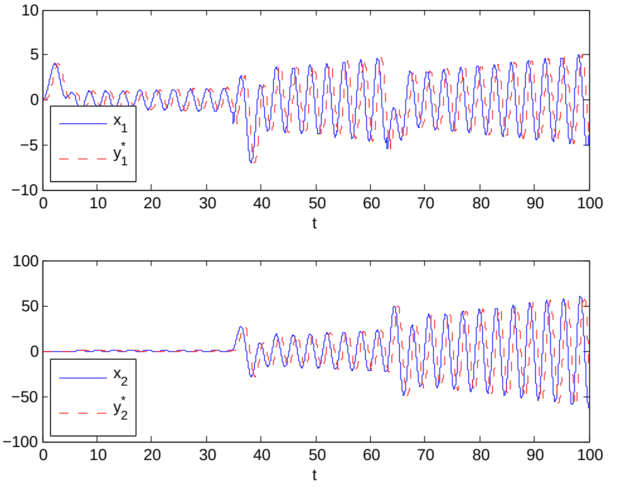

Figure 4. *System response without anticipation: the solid plot corresponds to the drive system, and the dashed one to the response system.*

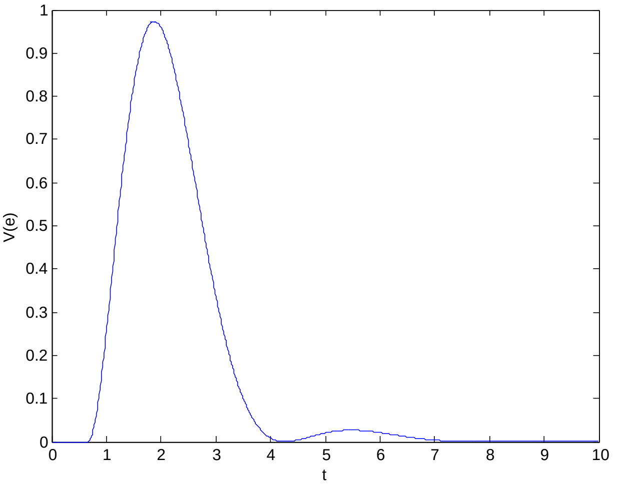

Figure 6. *Prediction error V for the first 10 seconds of the simulation.*

These results show that the proposed system is capable of correctly (1) detecting event boundaries, corresponding to failures in prediction (as proposed by EST [7]), (2) controlling the camera movement smoothly using anticipation (and thus coping with the perceptual delay), and (3) learning the correct model parameters for a correct anticipation of the system evolution.

7. Conclusions and future work

This paper presents a novel event segmentation framework, targeting cognitive, situated robots, based on the concept of strong anticipation proposed by Stepp *et al.* in [14]. A dynamical systems synchronization paradigm is used as the theoretical foundation of the proposed architecture, where the robot-world coupled system is identified using a parametric method for adaptation proposed by Chen *et al.* in [1], and the control is performed using anticipation. This anticipation accommodates for the delay of the perceptual channel. The capability of the architecture to anticipate perception allows the robot to control its actuation based on the prediction of the robot-world system state, instead of relying on the delayed perceptual data. A proof of concept experiment, in simulation, has illustrated the capabilities of the approach.

The proposed architecture targets physical robots. Thus, future work aims primarily at the evaluation of this architecture in a humanoid robot. One challenge posed by physical robots concerns the performance of the architecture facing unmodeled dynamics. However, it is expected that with a sufficiently complex response system, this effect can be effectively mitigated.

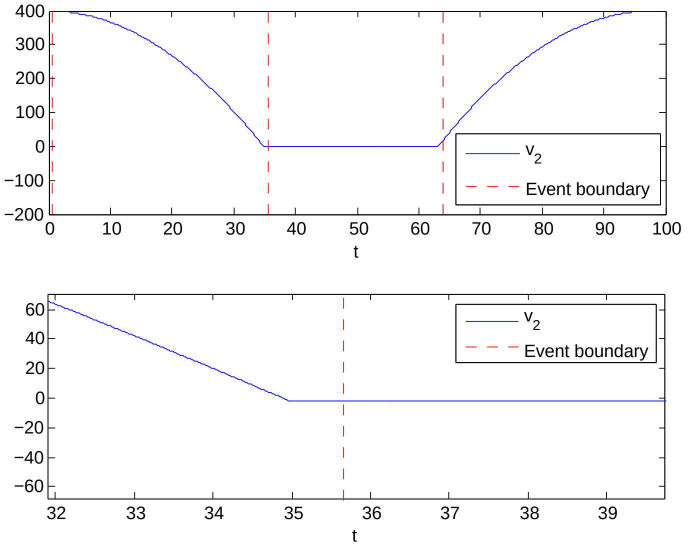

Figure 7. *The ball v_2 coordinate evolution along the experiment: top plot shows the detected event boundaries as vertical dashed lines, while the bottom plot zooms around the first detected event. The delay observed in this second plot corresponds to the perceptual delay τ.*

At the level of the architecture, future work includes scaling up this approach to more complex domains. This involves tackling the issues of the learning rate, which is hidden in the proportionality constant of the Lyapunov function, used in the Chen's learning rule, as well as the automatic design of the controller, given the adapted parameters. Other open questions include dealing with hidden state variables, as well as complex relations among objects (as in grasping, occlusion, and so on).

References

[1] Shihua Chen and Jinhu Lü. Parameters identification and synchronization of chaotic systems based upon adaptive control. *Physics Letters A*, 299:353--358, 2002.

[2] D. DeMenthon. Spatio-temporal segmentation of video by hierarchical mean shift analysis. *Language*, 2, 2002.

[3] Daniel M. Dubois. Mathematical foundations of discrete and functional systems with strong and weak anticipations. In *Anticipatory Behavior in Adaptive Learning Systems*, Lecture Notes in Computer Science, pages 107--125. Springer, 2003.

[4] V. Guralnik and J. Srivastava. Event detection from time series data. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 33--42. ACM, 1999.

[5] Mitsuo Kawato. Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9(6):718--727, December 1999.

[6] Daehwan Kim, Jinyoung Song, and Daijin Kim. Simultaneous gesture segmentation and recognition based on forward spotting accumulative HMMs. *Pattern Recognition*, 40(11):3012--3026, November 2007.

[7] Christopher A. Kurby and Jeffrey M. Zacks. Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, 12(2):72--79, February 2008.

[8] R.C. Miall, D. J. Weir, D. M. Wolpert, and J. F. Stein. Is the cerebellum a Smith predictor? *Journal of Motor Behavior*, 25(3):203--216, 1993.

[9] Bruno Nery and Rodrigo Ventura. Online event segmentation in active perception using adaptive strong anticipation. In Alexei V. Samsonovich, George Mason University, Kamilla R. Jóhannsdóttir, Antonio Chella, and Ben Goertzel, editors, *Biologically Inspired Cognitive Architectures 2010: Proceedings of the First Annual Meeting of the Bica Society (BICA-2010)*, volume 221 of *Frontiers in Artificial Intelligence and Applications*, pages 86--91. IOS Press, 2010.

[10] Louis M. Pecora, Thomas L. Carroll, Gregg A. Johnson, and Douglas J. Mar. Fundamentals of synchronization in chaotic systems, concepts, and applications. *Chaos*, 7(4):520--543, 1997.

[11] Erich Prem, Erik Hörtnagl, and Georg Dorffner. Growing event memories for autonomous robots. In *Proceedings of the Workshop On Growing Artifacts that Live: Basic Principles and Future Trends*, 2002.

[12] Marco Ramoni, Paola Sebastiani, and Paul Cohen. Unsupervised clustering of robot activities: a Bayesian approach. In *Proceedings of the fourth international conference on Autonomous agents (AGENTS'00)*, pages 134--135, 2000.

[13] Wolfram Schultz and Anthony Dickinson. Neuronal coding of prediction errors. *Annual Review of Neuroscience*, 23:473--500, 2000.

[14] N. Stepp and M.T. Turvey. On strong anticipation. *Cognitive Systems Research*, 11:148--164, 2010.

[15] Henning U. Voss. Anticipating chaotic synchronization. *Physical review E*, 61(5):5115--5119, 2000.

[16] Henning U. Voss. Dynamic long-term anticipation of chaotic states. *Physical Review Letters*, 87(1):14102, July 2001.

[17] J.Y.A. Wang and E.H. Adelson. Spatio-temporal segmentation of video data. In *SPIE Proceedings Image and Video Processing II*, volume 2182, pages 120--131, 1994.

[18] Daniel M. Wolpert, R. Chris Miallb, and Mitsuo Kawato. Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9):338--347, 1998.

[19] Jeffrey M. Zacks, Nicole K. Speer, Khena M. Swallow, Todd S. Braver, and Jeremy R. Reynolds. Event perception: A mind–brain perspective. *Psychological Bulletin*, 133(2):273--293, 2007.