Cognitive Models of the Perception-Action Cycle: A View from the Brain

Vassilis Cutsuridis

Abstract— Perception-action cycle is the circular flow of information that takes place between an organism and its environment in the course of a sensory-guided sequence of actions towards a goal. Each action causes changes in the environment which are processed by the organism's sensory hierarchy and lead to the generation of further action by its motor effectors. These actions cause new changes that are sensory analyzed and lead to a new action, and so the cycle continues. The efficient and timely coordination of the sensory and motor structures involved will ensure the organism's survival in a dynamic environment. Two brain inspired cognitive models of the perception-action cycle are presented in this paper: (1) A cognitive model of visual saliency, overt attention and active visual search, and (2) A cognitive model of visuo-motor coordination of reaching and grasping. Both models are multi-modular. They share a number of features (visual saliency, focus of attention, recognition, expectation, resonance, value attribution), while at the same time have distinct properties.

I. INTRODUCTION TO THE BRAIN PATHWAYS OF THE PERCEPTION-ACTION CYCLE

According to the eminent neuroscientist Joaquin Fuster the perception-action cycle is "the circular flow of information from the environment to sensory structures, to motor structures, back again to the environment, to sensory structures, and so on, during the processing of goal-directed behaviour" [2]. Many brain areas are involved in the perception-action cycle with each one serving particular functions. Description of all the brain areas involved in the perception-action cycle is impossible due to lack of space and experimental evidence. In this section I will describe only those brain areas required by both cognitive models to perform their computations.

Two parallel processing routes of visual information processing common to both active visual search and visuo-motor reaching and grasping involve the retina through the dorsal and ventral streams to the posterior parietal and temporal cortices. The initial stages of visual processing involve the retina, lateral geniculate nucleus (LGN) and the primary visual cortex (V1). In these areas, individual neurons can discriminate small changes in visual orientations, spatial frequencies and colour [15]. V1

transmits information through two streams: the dorsal stream (the "where" pathway) and the ventral stream (the "what" pathway) [16]. The dorsal stream begins with V1, goes through visual area V2, then to MT (visual area V5) and from there to the posterior parietal cortex (PPC). The dorsal stream, also known as the "where pathway" is associated with motion, representation of object locations, and control of the eyes and arms, especially when visual information is used to guide eye movements or reaching [15]. The ventral stream begins with V1, to V2, then to visual area V4, and finally to the inferior temporal cortex (IT/TC) [16]. The ventral stream, also known as the "what pathway", represents form recognition and object representation [17].

A third parallel route of processing visual information is from the early visual areas to the anterior intraparietal area (AIP). This route is involved only to the visuo-motor information processing of reaching and grasping. Experimental work has shown in anterior intraparietal area (AIP) the presence of a large number of neurons that are active in association with grasping and manipulation movements (motor neurons), presentation of visual stimuli (visual neurons) and both hand actions and object representations (visuomotor neurons) [22-26]. AIP neurons receive direct input from the IT object shape cells. Area AIP neurons transforms the visual information of a given 3D object into multiple descriptions, thus providing the premotor areas with several grasping plans [27].

A fourth parallel route of information processing common to both active visual search and visuo-motor reaching and grasping circuits is the one from the retina through the visual layers of superior colliculus (SC) to the substantia nigra pars compacta (SNc) [37-40]. SNc then broadcasts to all cortical areas (frontal eye fields (FEF), prefrontal cortex (PFC), IT/TC, premotor cortex (PM), primary motor cortex (M1), lateral intraparietal (LIP) and AIP) modulatory signals, which tune the responses of neurons in those areas. Dopamine has been implicated in signalling value attribution and reward prediction errors used to select actions that will maximize the future acquisition of reward [30] as well as the progressive movement deterioration of patients suffering from Parkinson's disease [31-36].

Higher level cortical areas linked to eye (saccades) and arm (reaching and grasping) movement response generation are the LIP and AIP areas of the posterior parietal cortex (PPC), the FEF of the frontal cortex, the PM and M1 and the PFC. The LIP has been involved in eye movements as electrical stimulation evokes saccades [7]. It has also been implicated in working memory guiding saccades in the

delay-response saccade task [7]. Neurons in LIP start responding with the initial presentation of the stimulus. The neurons continue to respond through the delay period until the fixation point is extinguished. As the neural response stops, the saccadic eye movement starts and the animal soon focuses on the exact location of the previously shown target [7]. FEF is another area involved in the planning and execution of saccadic eye movements [8]. PM is an area of motor cortex lying within the frontal lobe of the brain just anterior to the primary motor cortex [15]. It is involved in the planning of movements [15]. It is subdivided into two sections: the dorsal (upper) premotor cortex (PMd) and the ventral (down) premotor cortex (PMv). PMd plays a role in guiding reaching, whereas PMv in guiding grasping [15]. PMv contains the mirror neurons [29]. Mirror neurons are both sensory and motor as they become activated when an animal grasps an object as well as when the animal observes another animal grasp for an object [29]. Mirror neurons are proposed to be a basis for understanding the actions of others by internally imitating the actions using one's own motor control circuits [29]. PMv and AIP are strongly connected with each other. M1 is a brain region located in the posterior portion of the frontal lobe [15]. It is heavily connected with other motor areas including the premotor cortex and posterior parietal cortex, as well as several subcortical brain regions (e.g. basal ganglia structures), to plan and execute movements [15]. It is the area where the final motor command is formed before it is sent to the upper and lower extremities for execution [15]. PFC lies in front of the PM, FEF and M1 areas in the frontal lobes of the brain [15]. It has been implicated in planning complex cognitive behavior, personality expression, decision making and moderating social behaviour [18-21]. It is considered to orchestrate thoughts and actions in accordance with internal goals [18-21].

PPC and IT are known to project heavily to the prefrontal cortex (PFC), which in turn project back to these areas, thus linking it to perception, memory and action [28].

Projections from FEF, LIP, and PFC to SC either directly and/or through the basal ganglia structures are known to exist [9], [10]. Direct projection from V1 to SC has been reported [11]. Lesion studies have shown that no single pathway is essential. The combined loss of both SC and FEF renders the animal unable to make saccades [12]. Inability in making saccades occurs also with lesions to both SC and V1 [13]. SC is the common recipient of excitation from the cortices, since stimulation of these regions no longer elicits saccades following SC ablation [12]. In turn, the intermediate and deep layers of the SC project to the brainstem saccade generators, although a direct FEF pathway to brainstem has been shown [14].

In the next two sections I will describe two cognitive models of the perception-action cycle: (1) An active visual search model for scene recognition, and (2) A visually-guided reaching and grasping model of objects.

II. COGNITIVE MODEL OF VISUAL SALIENCY, OVERT ATTENTION AND ACTIVE VISUAL SEARCH

When we visually explore our visual world, our gaze shifts from one location to another. These gaze direction shifts are the result of very fast eye movements called saccades. Saccades are ballistic eye movements capable of reaching very high velocities (~800°/s at their maximum) [1]. The typical size of a saccade is 12–15° [1]. Information about the visual scene is acquired only during fixations, which are stationary periods between saccades. During a fixation, parts of the visual scene are brought to the eye's fovea, where the visual acuity is at maximum. Active visual search is the process of active scanning of the visual environment for a particular target among distracters and for the extraction of its meaning.

Cutsuridis [3] introduced a brain inspired cognitive model of active visual search where a bottom-up stimulus-driven module and a top-down attentive module interact to drive attention to specific regions-of-interest (ROIs) in the visual saliency map. The model offered plausible hypotheses of how the various participating brain areas work together to accomplish a scan of a scene within the allocated time (3–4 fixations per second). The model provided also answers to a number of important questions in the field of active vision:

- How is a complex visual scene processed?
- How is the selection of one particular location in a visual scene accomplished?
- Does it involve bottom-up, sensory-driven cues or top-down, world knowledge expectations? Or both?
- How is the decision made when to terminate a fixation and move the gaze?
- How is the decision made where to direct the gaze in order to take the next sample?
- What are the neural mechanisms of inhibition of return?

A. The Model

The model [3] is multi-modular consisting of a visual object identity saliency map module, a visual spatial (object location) saliency map module, a goals module, an overseer (value attribution) module, a decision making module, a motor programs module and a motor execution module. The topology, connectivity, and proposed functionality of its modules are supported by experimental data from the human and animal brains. A graphical representation of the model's information processing stages is given in Figure 1. The model proposed that once an input image is presented three parallel and equally fast processing modes of action are initiated (Fig. 1A). In the first mode of action (visual saliency processing) different pre-attentive multi-scale feature detection and extraction mechanisms sensitive to different features (e.g., colour, intensity, orientation, etc.) operated in parallel at the level of the retina, LGN, and V1. These low level features were extracted from the input image at several spatial scales using Gaussian pyramids at
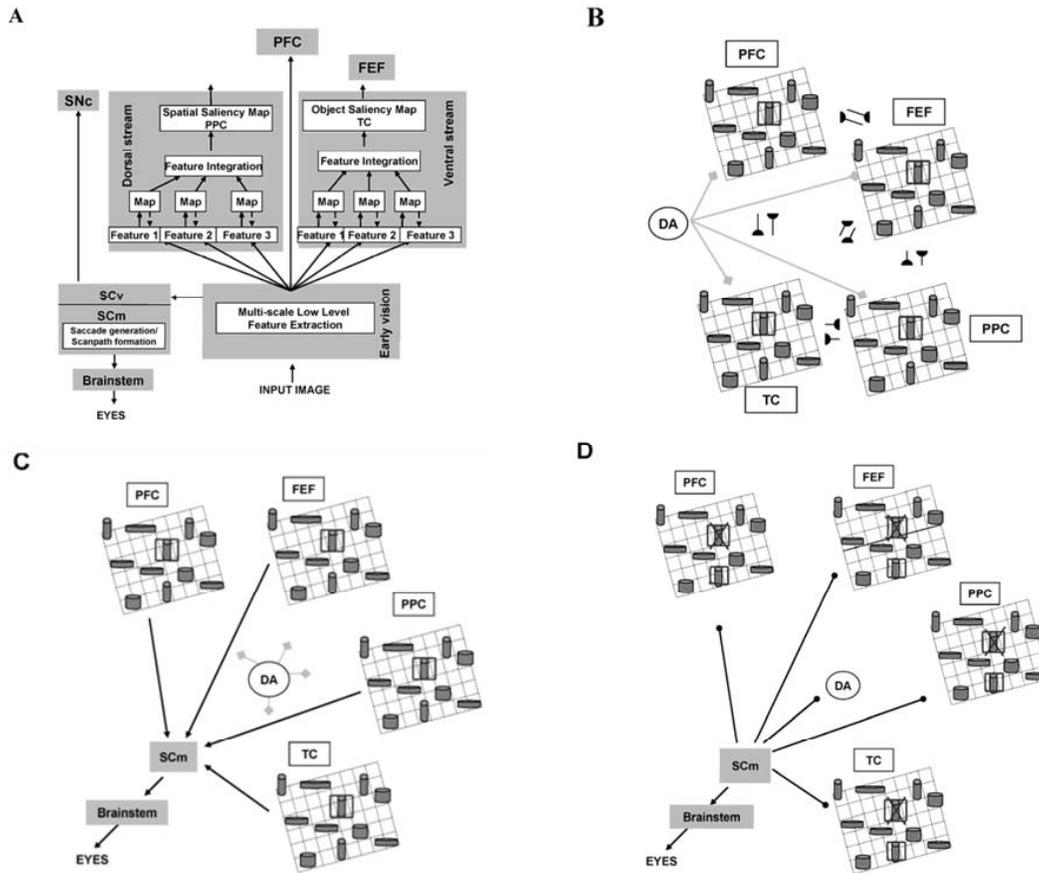
Fig. 1. Information processing stages of the active visual search cognitive model (adapted with permission from [3]). (A) The three parallel modes of information processing: the spatial and object visual saliency mode via the dorsal and ventral streams, the value attribution (dopamine) mode via SNc, and the focus of attention mode via PFC. (B) Value attribution via the SNc DA value attribution signals to goals (PFC), spatial (PPC), object (TC), and motor programs (FEF) neuronal maps. Neuronal responses are depicted by gray-colored towers in each map. The height of each tower represents the neuronal amplitude activation, whereas the width of each tower represents the degree of tuning. (C) Feedforward activation of the motor execution (SCm) module by motor programs (FEF), goals (PFC), spatial (PPC), and object (TC) maps. Dark gray-square surrounding the response of a neuronal population represents the winner salient and resonated according to some attribution value (DA signal) representation in each map. (D) Reset mechanism by feedback inhibitory projections from the motor execution module (SCm) to value attribution (SNc), motor programs (FEF), goals (PFC), spatial (PPC), and object (TC) modules. Reset mechanism prevents previously selected representation (dark gray crossed square) and allows all other resonated neuronal population responses to compete each other for selection. Bottom tower surrounded by dark gray-square represents the winner salient and resonated representation. PFC prefrontal cortex, PPC posterior parietal cortex, TC temporal cortex, FEF frontal eye fields, DA dopamine, SC superior colliculus, SCv visual superior colliculus, SCm motor superior colliculus, SNc substantia nigra pars compacta.

various scaling depths provided horizontal and vertical image reduction factors ranging from 1:1 (level 0; the original input image) to 1:256 (level 8). Differences between a fine and a coarse scale for a given feature were subsequently computed. From the level of V1 and up the extracted features were separated into two streams: the dorsal for spatial location processing and the ventral for object identity processing. Neurons in the feature maps in both dorsal and ventral streams then encoded the spatial and object contrast in each of those feature channels. Neurons in each feature map spatially competed for salience, through long-range connections that extended far beyond the spatial range of the classical receptive field of each neuron. After competition, the feature maps in each stream were combined into a global saliency map, which topographically encoded for saliency irrespective of the feature channel in which stimuli appeared salient. In the model, the global spatial saliency map was assumed to reside in PPC, whereas the global object identity saliency map resided in the ventral TC. In the second and concurrent mode of action (Fig. 1B), the overseer (value attribution) module was activated by the early visual via the retina signals via the visual layers of the SC. In turn, value attribution module broadcasted value (dopamine neuromodulatory) signals to the goals (PFC), spatial (PPC), object (TC) and motor programs (FEF) modules and selectively tuned the responses of different neuronal populations in these modules according to previous similar acquired experiences. In the third and concurrent mode of action (focus of attention), the early visual processing centres activated the goal module in PFC. The neuronal populations in the goals module then sent/received top-down/bottom-up feedback/feedforward signals to/from the spatial, object, and motor saliency maps of the PPC, TC, and FEF. The goals module role was to (1) send a focus of attention signal to every stage of the visual processing, amplifying specific neuronal responses throughout the visual hierarchy and at the same time inhibiting those of distracters, and (2) to participate in the decision making process of the next saccade to be executed via the adaptive resonance process of the selectively tuned via value attribution spatial, object and motor plan salient representations in the PPC, TC, and FEF.

In the next step of information processing, the spatial and object global salient maps will go through a sensory-motor transformation to generate their corresponding motor salient maps at the FEF level. Reciprocal connections between spatial (PPC), object (TC), and motor programs (FEF) representations will bind the perceptual and motor salient maps together. While this transformation and grouping is taking place, attentional and value attribution signals from the goals module (PFC) and the overseer module (SNc), respectively, will amplify/selectively tune the neuronal responses at the PFC, PPC, TC, and FEF levels. A winner take-all mechanism in these fields will select the most salient and resonated spatial, object, and motor program representations (Fig. 1B). The selected motor program will then be forwarded to the motor execution module (SC and

brainstem) where the final motor command will be formed and the eye movement will be generated (Fig. 1C).

All these mechanisms are reset by a feedback excitatory signal from the movement execution module (SC) to the inhibitory neurons in the motor programs (FEF), goals (PFC), spatial saliency (PPC), object saliency (TC), and overseer (SNc) modules, which in turn inhibit and hence prevent the previously selected targets, objects, and plans from being selected again (inhibition of return mechanism) (Fig. 1D).

I. COGNITIVE MODEL OF VISUO-MOTOR INTERACTIONS IN REACHING AND GRASPING

A baby's fine motor development is crucial. The baby must learn to use his/her hands well in order to reach, grasp and manipulate toys and to acquire later in life self-help skills such as feeding. Therefore, the act of coordinated visually guided reaching and grasping is very important to any organism interacting with its environment. First the organism must process the visual information reaching its sensory organs. The extracted visual features will then need to be translated into appropriate motor actions given the current context and the organism's previous experiences. The act of reaching and grasping involves two components: (1) the reaching component concerned with bringing the hand to the object to be grasped (transport phase), and (2) the grasping component concerned with the shaping of the hand according to the object features (grasping phase).

A brain inspired cognitive model of the perception-action cycle for visually-guided reaching and grasping of objects has been advanced by [4, 5]. The objects in the environment are not known before hand by the system, but their knowledge is built by the system through interaction and experimentation with them. The architecture is multi-modular consisting of object recognition, object localization, focus of attention, cognitive control, affordance extraction, value attribution, decision making, motor planning and motor execution modules. The components of the architecture followed very closely what is currently known of the human and animal brain. The model has been tested against a hypothetical scenario where multiple objects are situated in the environment and it must recognize them, localize them, attend to each one of them and reach and grasp them according to an externally dictated sequence of motor actions. I assume that the model is the brain of a robot/agent.

A. Model Scenario Constraints

On the table in front of the robot/agent an object or multiple objects are situated at different locations, which are not known a-priori by the robot/agent. The objects may have different shapes, sizes and colours. The task is to execute a sequence of reach-to-grasp actions, dictated by external commands, towards an object in the environment. In order for the robot/agent to tackle this scenario successfully, then some constraints need to be imposed: The robot/agent is
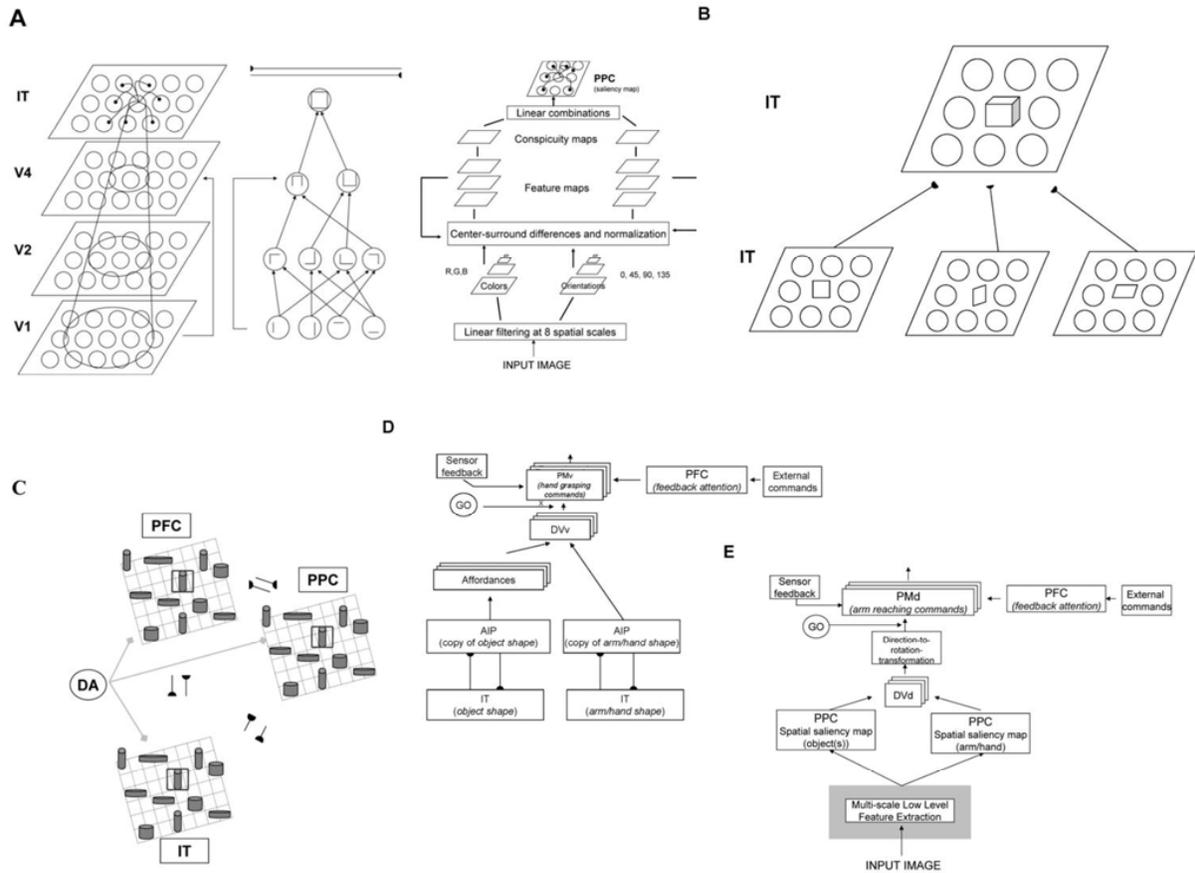
Fig. 2. Information processing stages of the visually-guided reaching and grasping cognitive model (adapted with permission from [5]). (A) Visual hierarchy as it occurs in the ventral (left) and dorsal (right) streams in the architecture. (Ventral stream) V1 primary visual cortex tuned to lines of different orientation, V2 tuned to angles, V4 tuned to partial boundaries and IT inferior temporal cortex tuned to object recognition. (Dorsal stream) PPC posterior parietal cortex tuned to object location. Bidirectional adaptive connections between IT and PPC ensures that the recognized object corresponds to its recognized spatial location. (B) Weighted combination of 2D object identity representations form 3D object representation in the level of IT. (C) Value attribution signals to spatial (PPC), goals (PFC) and object identity (IT) maps. Different neuronal populations in these maps receive different levels of value modulation. High and intermediate value modulation result in "sharp tuned" neuronal responses, whereas low value modulation result in "broadly tuned" neuronal responses. Neuronal responses are depicted by gray coloured towers in each brain. The height of each tower represent the neuronal amplitude activation, whereas the width of each tower represents the degree of tuning. Dark gray square surrounding the response of a neuronal population represents the winner neuronal population from the ART search cycle, where resonance between the top-down (e.g. PFC) and bottom-up (e.g. PPC) populations according to some value of the vigilance (value modulation signal) has been achieved. (D) Grasping process. The winner IT object and arm/hand neuronal populations are copied to AIP. The AIP "Copy of object shape" population activates all potentially desired finger configurations (affordances) capable of grasping the object. The "actual finger" configuration is then subtracted from each "desired" one, thus generating N "difference vectors" (DVv). Each DVv is then multiplied by a volitional movement scaling signal (GO), thus generating N "hand grasping" commands. Feedback attention signal from PFC then selects the most appropriate hand grasping command according to the context dictated by the "external commands". The process continues till DVv is zero. (E) Reaching process. Spatial saliency maps are generated at the level of PPC for both the "object(s)" and the "arm/hand". The object spatial map is the "desired spatial map", whereas the arm/hand spatial map is the "actual spatial map". Both spatial maps are subtracted to form N difference vectors (DVd) in spatial coordinates. Each DVd vector undergoes a spatial-to-joint transformation, which is then scaled by a volitional signal (GO) and forms N "arm reaching motor commands". The most appropriate arm reaching motor command is then selected by the attention feedback PFC signal. The reaching process continues till DVd is zero.

situated in front of a table and it is unable to move its torso (i.e. it is immobile). The robot's head is facing the table and it is not allowed to move left or right, up or down. Its eyes are also not moving (passive vision). The robot/agent is allowed to move only one of its arms/hands, but not both. The robot's arm/hand is also considered an object that the robot is attempting to move towards the stationary object(s) in the environment in order to reach for it/them and grasp it/them.

A. Model Operation

When an image frame of the video stream is presented to the robot's eyes two parallel and equally fast processing modes of actions are initiated. In the first mode of action (object localization (spatial saliency) and object recognition processing; see Fig. 2A)), pre-attentive multi-scale feature detection and extraction mechanisms sensitive to different features (colour and orientation) operating in parallel at the level of the retina, LGN, and V1 start to work in order to recognize the identity and location of the objects (including arm/hand) in the environment. From the level of V1 and on the features are separated into two streams: the dorsal for object location (spatial saliency) processing (see Fig. 2A left) and the ventral for object identity processing (see Fig. 2A right). As we mentioned earlier, along the ventral stream information is processed in a feedforward way through a series of competitive neural networks, where the receptive field size of neurons in each succeeding state increases by a factor of 2.5. At the first level (V1) of processing lines of different orientations (0, 45, 90, 135) are extracted via Gabor filters, which are then combined in V2 to form angles. At the level of V4, lines from V1 and angles from V2 are combined to form partial boundaries. At the last level (IT), partial boundaries are added to form 2D object shape representations. At IT the different 2D object representations are associated via a Hebbian learning rule to form 3D object representations (see Fig. 2B). Complex cells at each level of processing (V1 to V4) ensure that their neuronal representations are invariant to local changes in size, position and scale, while they maintain feature specificity.

Along the dorsal stream and at the first level (V1) of processing, local orientation information and color information is obtained from the input image via Gabor filters and the CIELab model, respectively (see section 3.1 for details). Each feature is computed in a center-surround operation, where the difference between a fine and a coarse scale for a given feature is computed. Neurons in the feature maps then encode the spatial contrast in each of those feature channels. Neurons in each feature map spatially compete for salience, through long-range connections that extend far beyond the spatial range of the classical receptive field of each neuron. After competition, the feature maps are combined into a spatial saliency map, which topographically encodes for spatial saliency irrespective of the feature channel in which stimuli appeared salient.

Bidirectional associative connections between the spatial saliency (PPC) and object identity (IT) maps ensure that the identified object is associated with its corresponding location in the environment.

In the second mode of action, the low level visual processing centres activate the value attribution module, which in turn broadcast the value attribution (DA) signals to the goals (PFC), spatial (PPC) and object (IT) map representations and selectively tune the responses of different neuronal populations in these areas according to previous similar acquired experiences. A search cycle similar to an ART network will commence where top down attentional signals from the goals (PFC) will send/receive top-down/bottom-up feedback/feedforward signals to/from the spatial (PPC) and object (IT) maps (see Fig. 2C). The value attribution signals acting as a similarity measure between the top-down goals map (PFC) and bottom-up spatial saliency (PPC) and object identity (IT) map representations will selectively tune the neuronal responses at the PFC, PPC and IT levels.   A winner-take-all mechanism in these fields will select those spatial and object representations that reached resonance with the PFC ones. These selected PPC and IT representations will continue to the next stages of processing.

In the next step two parallel modes of action will be performed: one for reaching (see Fig. 2E) and the other for grasping (see Fig. 2D). For reaching the two spatial (PPC) neuronal representations (N for N objects and one for the arm/hand) that successfully reached resonance will be subtracted from each other, constructing this way N difference vectors (DVd) in spatial coordinates. The difference vectors will be then transformed into a motor direction vectors through the direction-to-rotation transform that will move the arm/hand in the desired spatial direction from the present arm configuration. A volitional movement signal (GO) will gate the direction-to-rotation transform motor direction vectors and generate N final motor commands. Feedback attention from PFC will select the most appropriate to the current context final motor command to drive the arm/hand towards the desired object location. Proprioceptive information from the robot/agent joints will fine tune the movement of arm/hand towards the final target location.

For grasping the IT 3D object shape neuronal representation that reached resonance will be copied to the AIP 3D neuronal map. The copied AIP representation will extract from an external library a list of possible target finger preshaping configurations (affordances) that had in the past led to a successful grasp of the target object. Each of these affordance representations will be subtracted from the AIP actual finger configuration (hand aperture), thus forming a series of difference vectors (DVv). Each difference vector will be gated by a volitional GO signal, thus generating a series of final grasping commands. Top down attentional signals from the PFC will bias the most appropriate motor command, thus making it the most likely one to be executed.
When both DVd and DVv become zero, then the arm/hand will have successfully reached and grasped the object. The

above described process takes place for each image frame of the video stream.

II. LIMITATIONS AND FUTURE WORK

Two brain inspired cognitive models of the perception-action cycle were heuristically presented. Both models were grounded to experimental data of human and animal brains. The models shared many features (saliency, focus of attention, value attribution, recognition, expectation, resonance) as well as they possessed their own unique properties. The models were tested against hypothetical examples and demonstrated how they can achieve their computations within the time constraints reported by the experimental brain science.

Despite their successes, they also had some limitations. Both models had limited learning and adaptive capabilities. Learning and adaptation refer to "the ability of a system when facing situations with some degree of similarity to those already experienced, consistently alter its responses based on what worked in the past and improve the system's responses such that over time they became incrementally better with respect to the current goals or utility functions" [41]. Both architectures were temporally constrained to either achieve their tasks as quickly as possible within the time allocated (3-4 saccades per second) or perform a sequence of movements (reach-grasp-lift) imposed by external commands.

Although both cognitive models possessed attentional, value attribution, and expectation capabilities, they lacked any reasoning powers. Any intelligent system must couple reasoning with prediction in order to overcome situations where trial-and-error solutions fail. Furthermore, reasoning is needed to fill in the blanks when relevant causal chains of events and data are not directly observable. The system should also possess introspective capabilities allowing it to evaluate and reason about itself and improve its internal and external operations by generating self-improved methods for learning.

Additionally, intelligent cognitive systems must operate in real-time. Although as I mentioned earlier both cognitive models successfully performed their tasks in the time allocated or dictated externally, they are still not able to operate autonomously in real-time.

Finally, to my knowledge, there are no other cognitive models perception-action cycle integrating a broad range of capabilities (visual saliency, attention, value attribution, expectation, resonance, decision making, etc.) and providing solutions to real-life problems (but see [42]). Most cognitive models are focused on specific and targeted cognitive processes [43-45]. Comparing cognitive architectures is thus a non-trivial problem with no generally accepted methodology.

REFERENCES

[1] J.M. Findlay, I.D. Gilchrist, Active vision: the psychology of looking and seeing. Oxford: Oxford University Press, 2003.
[2] Fuster JM. (2004). Upper processing stages of the perception-action cycle. TICS 8(4): 143-145
[3] V. Cutsuridis, "A cognitive model of saliency, attention and picture scanning", Cogn Comput 1: 292-299, 2009
[4] V. Cutsuridis, "The Perception-…-Action Cycle Cognitive Architecture and Autonomy: A View from the Brain", Journal of Artificial General Intelligence, 3(2): 36-38, 2012
[5] V. Cutsuridis, J.G. Taylor, "A Cognitive Control Architecture for the Perception-Action Cycle in Robots and Agents", Cogn Comput, DOI 10.1007/s12559-013-9218-z, 2013
[6] J.G. Taylor, V. Cutsuridis, M. Hartley, K. Althoefer, T. Nanayakara, "Observational Learning: Basis, Experimental Results, Models and Implications to Robotics", Cogn Comput, DOI 10.1007/s12559-013-9208-1, 2013
[7] M.L. Platt, P.W. Glimcher, "Neural correlates of decision variables in parietal cortex". Nature 400 (6741): 233–238, 1999
[8] R.F. Squire, N.A. Steinmetz, T. Morre, "Frontal eye fields", Scholarpedia, 7(10): 5341, (2012)
[9] O. Hikosaka, R.H. Wurtz, "Visual and oculomotor functions of monkey substantia nigra pars reticulate", J Neurophys. 49:1230–301, (1983)
[10] O. Hikosaka, Y. Takikawa, R. Kawagoe, "Role of the basal ganglia in the control of purposive saccadic eye movements", Physiol Rev., 80: 954–78, (2000).
[11] R.A. Berman, R.H. Wurtz, "Exploring the pulvinar path to visual cortex", Prog Brain Res. 171:467–73, (2008)
[12] P.H. Schiller, S.D. True, J.L. Conway, "Deficits in eye movements following frontal eye field and superior colliculus ablations", J Neurophys. 44:1175–89, (2004)
[13] C.W. Mohler, R.H. Wurtz, "Role of striate cortex and superior colliculus in visual guidance of saccadic eye movements in monkeys", J Neurophyiol., 40:74–94, (1977)
[14] D.P. Hanes, R.H. Wurtz, "Interaction of frontal eye field and superior colliculus for saccade generation" J Neurophys., 85: 804–15, (2001)
[15] D. Purves, G.J. Augustine, D. Fitzpatrick, W.C. Hall, A.S. LaMantia, J.O. McNamara, L.E. White, Neuroscience, Sinauer Associates Inc, USA (2004)
[16] L.G. Ungerleider, J.V. Haxby, "'What' and 'where' in the human brain", Curr. Op. Neurobiol., 4: 157–165, (1994).
[17] N.K. Logothetis, J. Pauls, T. Poggio, "Shape representation in the inferior temporal cortex of monkeys", Curr. Biol., 5: 552–563, 1995
[18] R. O'Reilly, T. Braver, J. Cohen, "A Biologically Based Computational Model of Working Memory", In Models of Working Memory: Mechanisms of Active Maintenance and Executive Control, (A. Miyake and P. Shah, Eds.), Cambridge University Press, Cambridge, UK, 1999
[19] D.J. Kravitz, K.S. Saleem, C.I. Baker, A. Mishkin, "A new neural framework for visuospatial processing", Nat Rev Neurosci, 12: 217-230, 2011
[20] T.S. Braver, J.D. Cohen, "On the control of control: The role of dopamine in regulating prefrontal function and working memory", In S. Monsell and J. Driver, eds., Control of Cognitive Processes: Attention and Performance XVIII, pp. 713-738, MIT Press, Cambridge, MA, 2000
[21] A.W.I. MacDonald, J.D. Cohen, V. Stegner, C.S. Carter, "Dissociating the Role of Dorsolateral Prefrontal and Anterior Cingulated Cortex in Cognitive Control", Science, 288 (5472): 1935-1838, 2000
[22] A. Murata, L. Fadiga, L. Fogassi, V. Gallese, V. Raos, G. Rizzolatti, "Object representations in the ventral premotor cortex of the monkey". J Neurophys., 78: 2226-2230, 1997
[23] A. Murata, V. Gallese, K. Kaseda, H. Sakata, "Parietal neurons related to memory guided hand manipulation". J Neurophys., 75: 2180-86, 1996

[24] A. Murata, V. Gallese, G. Luppino, K. Kaseda, H. Sakata, "Selectivity for the shape, size and orientation of objects for grasping in neurons of monkey parietal area AIP". J Neurophysiol., 83: 339-365, 2000

[25] H. Sakata, M. Taira, A. Murata, S. Mine, "Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey". Cerebral Cortex, 5: 429-438, 1995

[26] M. Taira, S. Mine, A.P. Georgopoulos, A. Murata, H. Sakata, "Parietal cortex neurons of the monkey related to the visual guidance of hand movement". Exp Brain Res, 83: 29-36, 1990

[27] V. Gallese, L. Fadiga, L. Fogassi, G. Luppino, A. Murata, "A parietal-frontal circuit for hand grasping movements in the monkey: Evidence from reversible inactivation experiments". In: P Their & HO Karnath (Eds), Parietal lobe contributions to orientation in 3D space (pp. 255-270), Berlin: Springer, 1997

[28] E. Miller, "The prefrontal cortex and cognitive control". Nat. Rev. Neurosci., 1:59–65, 2000

[29] G. Rizzolatti, C. Sinigaglia, "The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations". Nat. Rev. Neurosci., 11(4): 264–274, 2010

[30] W. Schultz, "Updating dopamine reward signals". Curr Opin Neurobiol, doi: 10.1016/j.conb.2012.11.012

[31] V. Cutsuridis, "Origins of a repetitive and co-contractive pattern of muscle activation in Parkinson's disease". Neural Networks, 24(6): 592-601, 2011

[32] V. Cutsuridis, "Does Abnormal Reciprocal Inhibition Lead to Co-contraction of Antagonist Muscles? A Modeling Study". International Journal of Neural Systems, 17(4): 319-327, 2007

[33] V. Cutsuridis, "Neural network modeling of voluntary single joint movement organization. I. Normal conditions". In: Chaovalitwongse WA, Pardalos P, Xanthopoulos P (eds), Computational neuroscience, Springer-Verlag, 181-192, 2010a

[34] V. Cutsuridis, "Neural network modeling of voluntary single joint movement organization. II. Parkinson's disease". In: Chaovalitwongse WA, Pardalos P, Xanthopoulos P (eds), Computational neuroscience, Springer-Verlag, 193-212, 2010b

[35] V. Cutsuridis, T. Heida, W. Duch, K. Doya, Neurocomputational models of brain disorders. Neural Networks, 24(6): 513-514, 2011

[36] V. Cutsuridis, S. Perantonis, "A Neural Model of Parkinson's Disease Bradykinesia". Neural Networks, 19(4): 354-37, 2006

[37] V. Coizet, E. Comoli, G.W. Westby, P. Redgrave, "Phasic Activation of Substantia Nigra and the Ventral Tegmental Area by Chemical Stimulation of the Superior Colliculus: An Electrophysiological Investigation in the Rat". Eur. J. Neurosci., 17(1): 28–40, 2003

[38] E. Comoli, V. Coizet, J. Boyes, J.P. Bolam, N.S. Canteras, R.H. Quirk, P.G. Overton, P. Redgrave, "A Direct Projection from the Superior Colliculus to Substantia Nigra for Detecting Salient Visual Events". Nat. Neurosci., 6(9): 974–980, 2003

[39] J.G. McHaffie, H. Jiang, P.J. May, V. Coizet, P.G. Overton, B.E. Stein, P. Redgrave, "A Direct Projection from Superior Colliculus to Substantia Nigra Pars Compacta in the Cat". Neurosci., 138(1): 221–234, 2006

[40] P. Redgrave, K. Gurney, "The Short Latency Dopamine Signal: A Role in Discovering Novel Actions". Nat. Neurosci., 7: 967–975, 2006

[41] K.R. Thorisson, H. P. Helgasson, "Cognitive architectures and autonomy: A comparative review". J of Artif Gen Intell., 3(2): 1-30, 2012

[42] J. G. Taylor, "The perception-conceptualization-knowledge representation-reasoning representation-action cycle: the view from the brain". In: Cutsuridis V et al. (eds), Perception-action cycle: Models, algorithms and hardware", Springer, USA, pp. 243-286, 2011

[43] V. Cutsuridis, A. Hussain, J.G. Taylor, "Perception-action cycle: Models, algorithms and hardware", Springer, USA, 2011

[44] R. Ivey, D. Bullock, S. Grossberg, "A neuromorphic model of spatial lookahead planning". Neural Networks, 24: 257-266, 2011

[45] N. Srinivasa, R. Bhattacharyya, R. Sundareswara, C. Lee, S. Grossberg, "A bio-inspired kinematic controller for obstacle avoidance during reaching tasks with real robots". Neural Networks, 35: 54-69, 2012