Bioinspired classification in the architecture of situated agents

G. Gini[1], A. Franchi[1], F. Ferrini[1], F. Gallo[1], F. Mutti[1], and R. Manzotti[2]

[1] DEIB, Politecnico di Milano, Milan, Italy, gini@elet.polimi.it
[2] IULM University, Milan, Italy, Riccardo.Manzotti@iulm.it

Abstract. Cognitive development concerns the evolution of human mental capabilities through experience earned during life. Important features needed to accomplish this target are the self-generation of motivations and goals as well as the development of complex behaviors consistent with these goals. Our target is to build such a bio-inspired cognitive architecture for situated agents, capable of integrating new sensing data from any source. Basing on neuroscience assessed concepts, as neural plasticity and neural coding, we show how a categorization module built on cascading classifiers is able to interpret different sensing data. Moreover we see how to give a biological interpretation to our classification model using the winner-take-all paradigm.

Keywords: Bio inspiration, perception, classifiers cascade, one-class classifier, winner take all.

1 Introduction

A challenge both for engineers and neuroscientists is to develop a robot that acts and thinks like a human; despite this problem is not new and researchers have worked on it since the XX century, in the last decades we have seen the arising of biologically inspired approaches. These kinds of solutions mimic what we know about the brain to shape the robots in a similar way. The current aim is to develop a complete conceptual and computational framework describing both how the brain might work and when the cognition arises.

During their life humans develop their mental capabilities: this process is called cognitive development and concerns how a person perceives, thinks, and gains understanding of the world by the interaction of genetic and learned factors. A fundamental aspect in cognitive development is the autonomous generation of new goals and behaviors, which allows the individual to adapt to various situations. In order to realize agents capable of interacting in an effective way with humans robotics should study the processes of human brain that allow the cognitive development [1].

Our work gives a contribution to the achievement of this objective: its purpose is to create a bio-inspired model based on human brain processes that should make the agent able to autonomously develop new goals as well as new behaviors consistent with these goals. In this broad area we address an intermediate level of cognition,

adfa, p. 1, 2011.

what allows humans to be aware of the surrounding environment and then to interact with it. This capability is an essential precondition to create agents able not only to act in a consistent manner in response to the changes in the environment, but also to develop goals that can emerge [2] [3].

As a model of the goal generation behaviors we have chosen the amygdala-thalamus-cortical interaction [1]. The cerebral cortex is divided into lobes, each having a specific function. The parts of the cortex that receive sensory inputs from the thalamus are called primary sensory areas [4]. The thalamus is the primary site of relay for the sensory pathways on their way to the cortex [5]. It is partitioned into about fifty segments, which do not directly communicate with each other. Instead, each one is in synchronized projection with a specific segment of the cortex, and receives a projection from the same segment. Therefore, while the cortex is concerned with data processing, the thalamus determines which goals have to be pursued. Lastly, the amygdala is a group of nuclei in the medial temporal lobes heavily connected to the cortex and involved in the generation of somatosensory response taking advantage of hardwired criteria [6].

Our Intentional Distributed Robotic Architecture (IDRA) is a network of elementary units, called Intentional Modules (IM) that enables the development of new goals, together with a Global Phylogenetic Module (GPM) containing the "innate instincts", as in the amygdala. The network is composed by several layers dynamically connected in forward or feedback mode. Each IM contains the Categorization and the Ontogenetic modules (CM and OM). The CM mimics the cerebral cortex and returns a vector that represents the neural activation of the cortex in response to the input; the OM receives this vector and through Hebbian learning develops new goals, returning also a signal stating whether the current state meets the new goals. The IM receives signals from both GPM and OM and returns the more relevant of the two and the neural activation computed by the CM.

Therefore, after all the sensory input has been acquired, filtered and sent to the IM network, each module returns a state vector and a signal indicating how much the actual environmental state is satisfying the agent goals. The vector of neural activations and the signal computed by IDRA are then used by the Motor System (MS) to generate movements consistent with the goal of the agent; each movement is a composition of elementary components, called motor primitives [7]. In [1] we have shown how to integrate motor primitives in IDRA using the NAO robot.

In our cognitive architecture we do not focus on high-level motor skills nor on high level reasoning and planning; instead we focus on the intermediate level of cognition that allows mammals and humans to be aware of the surrounding environment. This awareness is supposed to be not a direct product of sensations, the "phenomenological mind", nor a product of high level conceptual thoughts, the "computational mind", but to be a product of several intermediate levels of representation [8]. This point of view has some interesting features related to consciousness: it underlines how we can interpret the surrounding environment without the need for high-level conceptualizations, therefore solving the grounding problem of a semantic interpretation of a symbol system that is intrinsic to the system itself.

In this paper we focus on the categorical representation that is the learned and innate capability to pick out the invariant features of objects and of events from their sensory projections. Categorization is the first step in building cognitive functions for language, prediction, action. We avoid any ontological problem of defining categories lists if, according to Kant, we consider that categories are due to the nature of the mind, not to intrinsic divisions in the reality. Recently work on ontological categories has attracted interest also in cognitive science, where the goal is to define the means by which humans group things into categories. In [9] two basic principles of categorization are introduced: the task of category systems is to provide maximum information with the least cognitive effort, and the perceived world comes as structured information rather than as arbitrary or unpredictable attributes. Thus the maximum information with the least cognitive effort is achieved if categories map the perceived world structures.

Our categorization makes use of two main bioinspired principles: population coding and neuroplasticity. Exploiting these two concepts we designed a classifier working the same on any kind of sensing input, mimicking the different layers of the nets that decompose and analyze the sensing data [10]. Our classifier is a cascade of simple classifiers that work more and more on the same data to boost its predictivity, in a paradigm that has been successfully applied in literature [11]. The new aspect here is that we show how this paradigm is compatible with the neural substrate of our system; the experiments we report are about vision and audio signals.

In Section 2 we present the biological aspects related to our classification system. In Section 3 we develop the bioinspired classifier. In Section 4 we report on the experiments. Section 5 contains the conclusions.

2 Key biological inspiration

Cognitive neuroscience focuses on the development of a theoretical framework to fill the gap between the neural activity and the complex behavioral traits of the brain such as memory, learning, high vision processing, emotion, and higher cognitive functions. The underlying features, widespread among these brain functionalities, define the information processing, i.e. how the brain encodes and propagates information [12]. According to the classical view, the brain workflow is composed by at least three phases: perception, cognition, and action. Cognitive functions are separated from the sensor-motor system but recent works show that they are not localized in high specialized brain areas but are managed by the same sensor-motor neural populations [13].

The neural circuitry is organized in several functional areas responsible to solve specific subtasks [14]; this implies that a high level of synchronization among different areas is needed. This functional organization follows the divide-et-impera paradigm: the anatomical separation of the brain areas leads to a hierarchical organization of the brain functionalities. Given a sensory source, information is filtered along different brain areas, mixed with other sensory information and used to take an action decision; this information flow through different areas for achieving a specific objective is called pathway.

Widespread computational mechanisms are interesting for creating a computer model of the brain; brain models infer the organization of the neuronal population in order to produce a neural activity with the same properties of the biological counterpart. These populations have a computational mechanism that cannot be inferred by the single neuron activities; two examples are models of the primary visual cortex [15], and of the posterior parietal cortex [16].

There are at least six general mechanisms in the brain that should be taken into account: population coding, gain modulation, normalization, statistical coding, feedback connections, and neural plasticity.

Population coding is the mechanism used in the brain to represent sensory information through a group of neurons organized in such a way that neighboring neurons have similar activity [17]; one of the advantages of using a population of neurons to represent a single variable is its robustness to neural noise [18].

Gain modulation is an encoding strategy for population of neurons where the single neuron response amplitude is varying without a change in the neuron selectivity. This modulation, also known as gain field, can arise from either multiplicative or nonlinear additive responses and is a key mechanism in coordinate transformations [19].

Normalization is a widespread mechanism in several sensory systems where the neural responses are divided by the summed activity of a population of neurons to decode a distributed neural representation [20].

Statistical coding is a kind of population coding especially used for sensory data [10]; it seems to be widespread in the brain areas devoted to the pre-processing of the sensory data and it offers two advantages: it reduces the dimensionality of the input space and it gives an interpretation to the topological organization and emergence of the neuron receptive fields. An approach that takes into account the statistical properties of the sensory input is the Independent Component Analysis (ICA) [21].

Neuroplasticity is the lifelong ability of the brain to reorganize neural pathways based on new experiences; it works at different levels, from the single neuron to whole brain areas. The Hebbian learning is the commonly accepted learning principle at network level.

Perception from different sensors is managed in similar way in the brain [14]. At first the receptors, specific to react to a given stimulus, decompose the stimulus into elementary components; when the receptors are activated they propagates data to non-specific areas in the cortex where different sensorial modalities are represented and signals are integrated. In the case of vision information from the retina is separated in two by the ganglia cells M and P, which project onto different layers of the thalamus; the two paths analyze different aspects of the image and the information is then recombined in the cortex. In the case of auditory data the ear through ciliate cells identifies the single frequencies and codifies them through the spike frequencies of the acoustic nerve; this information is then transmitted to the cochlear nuclei which codify both frequency and intensity; finally the information is sent to the audio cortex. As already observed, most areas of the cortex receive signals from specific regions but can manage different signals using normalization and population coding [10].

3 One class classifiers and winner-takes-all

Our IDRA system [1] is a layered net of Intentional Modules (IM) simulating connections and interactions between the cerebral cortex (CM) and the thalamus (OM); beside there is a Global Phylogenetic Module (GPM), representing the amygdala, which is connected to all IMs. Each IM (Fig. 1) contains a CM and an OM. Incoming data are directly sent to the CM and the categories it creates are sent to the OM.

The GPM contains the hard coded instincts [22] and broadcasts its signal to all the IMs. Input to GPM comes from sensors; output from GPM is normalized in <zero-one> and tells how much the incoming stimulus is important according to the a priori stored criteria;

CM has the function of extracting different kinds of features from sensorial data and of categorizing interesting stimuli into a sort of memory of relevant events;

OM uses the categories from CM, performs Hebbian learning to develop new goals and returns the Ontogenetic Signal expressing how much these new goals are satisfied.

The output of each IM is a vector, representing the neural activation generated by sensory input, plus a scalar signal representing how much the actual input satisfies both hard-coded goals and new developed goals.

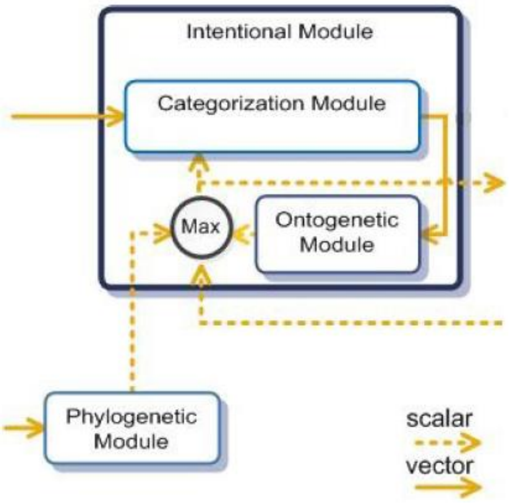

Fig. 1. Structure of the Intentional Module (IM)

These values of neural activations from CM are computed using a vector of weights, producing the ontogenetic signal Os as the maximum between the evaluated neural activations:

$$O_s = \max_i(y_i \cdot w_i) \tag{1}$$

where yi is the activation of neuron i and wi is the vector of normalized weights associated to neuron i. The weights are updated every iteration using a Hebbian learning function:

$$w_i = w_i + \eta(h_s \cdot y_i - (w_i \cdot y_i^2)) \quad (2)$$

where η is the learning rate and hs is the Hebbian control signal coming from the IM. In order to learn there must be persistent functional changes in the brain so that the IMs can adapt to changes in sensory input: if we send to an IM input from a video sensor it will specialize to it; the interesting ability is that if we switch its input to different type of stimuli the module will gradually adapt.

In the preliminary IDRA system categorization was obtained in two steps: first the input is projected in the space of the independent components of the input, collected a priori in an offline training stage

$$W = IC \ \times I \quad (3)$$

where W is the resulting vector of weights, IC is the matrix of independent components and I is the input vector; secondly a clustering is performed on the vector of weights W. Clustering is a good way to get the neural code of the input regardless its type. The result of the Categorization Module is a vector containing the activations of cluster which depends on the distance of the input stimuli from the center of each cluster (e.g. How much the current input is similar to something I have already experienced?). This vector corresponds to the activation of a neuron centered in each cluster:

$$y_i = \rho(x, C_i) \quad (4)$$

where yi is the distance of the actual input from the centre of the cluster i, x is the input and Ci is the centre of the cluster i.

The main drawback of this approach is that data passed to the following IMs lacks of meaning and that it does not mimic the known steps of sensory data analysis [23].

Our new categorization system integrates machine learning and neural population coding. In machine learning One Class Classifiers (OCC) are used when the problem is to distinguish target objects from outliers; to improve the performances of the classifier different combination techniques can be adopted, for instance ensembling different classifiers or using classifiers that use different features [24]. The particular ensembling method we chosen is the classifier cascade [11] where the same data set is presented more and more to the classifier to boost its response.

In case of multiclass classification the so called one-versus-all (OVA) classification paradigm [25] guarantees an higher accuracy making the multiclass classifier as the result of calling many OCC; moreover our solution is compatible with the winner-takes-all strategy (WTA) [26], a model of the neurons in response to a stimulus; according to WTA the neuron with the highest activation value is chosen and the other inhibited.

Here we combine the cascade of different OCC with OVA to get a sort of WTA strategy; the categorization module uses all the signals produced by the initial filters and analyzes them in different layers of IM. All the IMs receive in input the same signal sent to the previous layer; this way the architecture reproduces the mechanism that integrates the unimodal regions in the multimodal areas [26].

The categorization activity, spread through different cortex areas [23], starts with an initial filtering stage which decomposes the input stimulus into several signals each focused on a particular feature; all of these signals are then transformed by ICA which projects them onto independent components collected a-priori for every single CM. Lastly the resulting weights of activation are clustered and stored if "interesting". This way the input signal is transformed into a neuronal activity which is independent from data type and dimension. To mimic the mechanism of population coding each layer of the architecture contains a number of IM, each modelling a small group of neurons, according to the equation

$$\left|IM_{(x)}\right| = \sum_{j=1}^{|IM_{(x-1)}|} \binom{|IM_{(x-1)}|}{j} \qquad (5)$$

where the number of IMs in the x layer is computed from the number of IMs in the previous layer. Normalization is mimicked taking the average of the output signals of all the last level IMs.

Our implemented architecture is so structured; the network is feed-forward on 3 layers with a fixed number of IMs; the first layer contains 2 IMs each connected to a specific filter (eventually filters can be equal). The second layer has 3 IMs, one for each input combination; the third layer contains 7 IMs to take again all the combinations of previous input. From this last layer the result is normalized and extracted (Fig. 2).

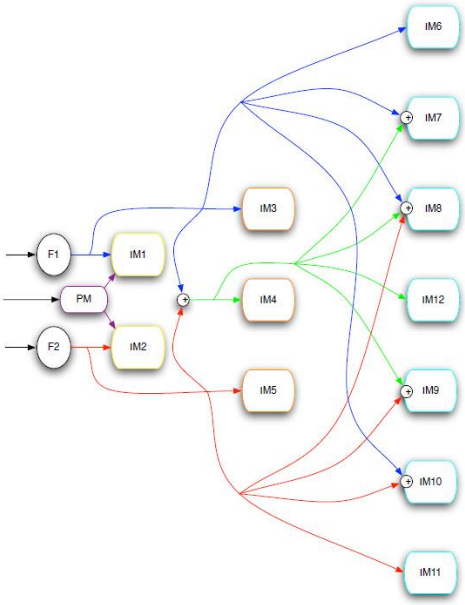

Fig. 2. The classifier architecture with the forward signals

For experimenting with video input we defined a filtering stage that extracts three different features, namely the image saturation, its edges and the black and white image. All these transformations are easily implemented using Matlab. These signals are sent into the architecture: the first is used only by the GPM which uses the hard-coded instincts of the attraction for colours; the two last ones are the input of the IMs in the first layer of the network.

We used two identical filters when working with audio signals to perform Fast Fourier Transform (FFT). It is important to note that the output of the filtering stage is always a vector with variable lengths regardless the input stimuli.

4 Experimental results

The experimental step has highlighted the performances of the classifier here depicted. Two experiments with visual data have been performed, the first dealing with Optical Character Recognition (OCR), the second with face recognition. The reason we have chosen these two tasks is double: on one hand they can be performed using common online datasets so that their performances are comparable to state of art; on the other hand it is interesting to compare the behavior of our architecture with the way our brain deals with these two tasks. A simple third experiment was about voice recognition.

We used the classical four indices as experimental metrics: accuracy, as the percentage of correct classifications (both positive and negative) over the total number of tests; precision, as the number of correct positive classifications over the total number of positive classifications (both true positive and false positive); recall, as the percentage of correct positive classifications over the total number of correct classifications; specificity, as the number of negative elements correctly classified over the sum of true negative and false positive.

4.1 Optical Character Recognition

The chosen dataset was extracted from UCI Machine Learning Repository letter dataset; we have selected ten characters, including those very similar like “P” and “R”, each one represented in 800 black and white images of 128x128 pixels.

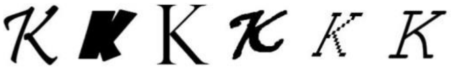

Fig. 3. Different examples of the character “K” in the dataset

We have randomly split our dataset into 640 images for training and 160 for testing; as input filtering we have implemented a standard Canny filter for edge extraction and a black/white transformation.

The first step has been the training of the architecture for each letter, getting ten different networks each one able to distinguish a particular character (e.g. the “A”). During testing we have presented each character sample to each of these networks;

the result is a series of ten output values between 0 and 1 which tell us how much the architecture is confident to classify the input as the character associated with the network currently loaded in the system. We collected all values into a 2D matrix, containing one column for each image and one row for each network; the final classification output for each input was extracted looking in each column for the highest value and taking the corresponding row as the recognized character. The results on the test set are reported in Table 1.

Table 1. Accuracy, precision, recall, specificity for each letter in our test set

Letters	Accuracy	Precision	Recall	Specificity
A	100%	100%	100%	100%
B	99.63%	96.95%	99.38%	99.65%
D	99.69%	99.36%	97.5%	99.93%
E	99.94%	100%	99.38%	100%
I	100%	100%	100%	100%
K	100%	100%	100%	100%
P	99.63%	97.53%	98.75%	99.72%
R	99.63%	98.73%	97.5%	99.86%
S	99.94%	99.38%	100%	99.93%
T	99.94%	100%	99.38%	100%

Using 1600 images (160 per each letter) we got only 13 errors, resulting thus in a percentage of error lower than 0.8%; among the most easily misclassified characters there are “R” and “D”; “I” or “K” have always been correctly guessed. Good results are in line with other empirical evaluation [27]. Authors of [28] empirically compared different methods of multiclass classification using SVM as basic classifier and found on the letter data set an accuracy between 97.98 and 97.68. However we have to underline two important aspects: first the number of classes in our dataset is much smaller than the complete set of symbols; secondly our dataset is composed of sharp black white images, not considering for example illumination problems that can emerge in real applications.

4.2 Face recognition

The second experiment focuses on the recognition of faces which is a slightly more complex task. We have used the Yale university dataset online available for researches purposes; this is composed of two sets of images (A and B), where the first (A) contains photos of faces from the chin to the forehead, in the second (B) all the images also show small portions of the foreground. We selected only the A set in order to constrain images dimension to 168x192 pixels for computational issues; we split the dataset into training and testing sets as done before. Since the number of images was very low, we decided to generate more samples by copying and modifying some of the images for the training set; as test set we randomly chose 10 images for each of the 11 marked subjects, for a total of 110 samples.

Fig. 4. An example of the Yale dataset; faces on the left correspond to faces on the right

We trained our architecture for every single subject as illustrated before, getting eleven different networks each one fitted on a specific subject; during the following testing stage we presented each of the 110 samples to all of the trained networks; the result is a series of values between 0 and 1 describing the confidence the architecture has in classifying the input. All these results have been collected in a matrix composed by one row for each trained network and a column for each test sample; the row with the highest value in each column is the recognized face.

Our system performed very well, without any classification error; this good result may arise from the fact that faces have many more details for performing classification (e.g. eye distance, beard, nose, etc.) and perhaps also from the quality of the dataset[1].

4.3 Audio recognition

We wanted also to test whether our architecture could adapt to different kinds of input, a fundamental ability shown by human brain.

We chose to use various registrations of the English word "sure" made by different persons all with different intonations and accents. Here a significant difference with vision concerns data structure: our architecture needs all the input to be the same length after the filtering stage; if this holds for images as they all have the same dimension, audio files may differ in length and thus we applied FFT and only one amplitude filter, discarding any processing on signal phase [29]. We have thus created a dataset composed by 200 audio files, split in 150 samples for training and 50 for testing.

For this last experiment we used a simple linear binary classifier only to recognize or not the word. Results on the test set are reported in Table 2; the best threshold value is 0.4, which gives an accuracy of 81% and a recall of 92%.

Considering these results we can state that our architecture performs quite well also with audio input even though our dataset was not large enough for a robust training.

[1] http://vis-www.cs.umass.edu/lfw/results.html

Table 2. Accuracy, precision, recall, specificity for each face in out test set

Threshold	Accuracy	Precision	Recall	Specificity
0.8	52%	100%	4%	100%
0.7	56%	89%	16%	96%
0.6	60%	72.73%	32%	88%
0.5	62%	67.65%	46%	78%
0.4	81%	75.41%	92%	70%
0.3	78%	69.44%	100%	56%

5 Conclusions

Our aim is the creation of a bio-inspired software architecture based on the processes that take place in the human brain; this architecture must be able to learn new goals, as well as to learn new actions to achieve such goals.

Crucial part of this architecture is the categorization module; here we have developed a classifier that takes inspiration from basic brain mechanisms.

Our experiments have shown that the agent is able to analyze data, clustering different kinds of features and to obtain results in classification that are very similar to those obtained by specialized classifiers as found in literature. The important result is that the generic classifier based on a simple neural architecture can perform well on a few different dataset; further experiments are needed to validate these promising results with respect to several kinds of sensory data.

Bibliography

1. R. Manzotti, F. Mutti, S. Y. Lee and G. Gini, "A model of a middle level of cognition based on the interaction among the thalamus, amygdala, and the cortex.," *IEEE International Conference on Systems, Man, and Cybernetics,* pp. 1996-2001, November, 2012.
2. M. Lungarella, G. Metta, R. Pfeiffer and G. Sandini, "Developmental robotics: a survey," *Connection Science,* vol. 4, no. 15, pp. 151-190, 2003.
3. R. Manzotti and V. Tagliasco, "From “behaviour-based” robots to “motivations-based” robots," *Robotics and Autonomous Systems,* vol. 2, no. 51, pp. 175-190, 2005.
4. J. Sharma, A. Angelucci and M. Sur, "Induction of visual orientation modules," *Nature,* vol. 404, pp. 841-847, 2000.
5. S. M. Sherman and R. Guillery, Exploring the Thalamus, Elsevier, 2000.
6. S. Duncan and L. F. Barret, "The role of the amygdala in visual awareness," *Trends in cognitive science,* vol. 11, no. 5, pp. 190-192, 2008.
7. F. Mussa-Ivaldi and E. Bizzi, "Motor learning through the combination of primitives," *Philosophical transactions of the Royal Society,* vol. 355, no. 1404, pp. 1755-1769 , 2000.
8. R. Jackendoff, Consciousness and the computational mind, MIT Press, 1987.
9. E. Rosh, "Principles of categorization," *Cognition and categorization,* pp. 27-48, 1978.

10. B. Olshausen A. and D. J. Field, "Sparse coding of sensory inputs," *Current Opinion in Neurobiology,* vol. 14, pp. 481-487, 2004.
11. P. Viola and M. Jones, "Fast and Robust Classification using Asymmetric AdaBoost and a Detector Cascade," *Advances in Neural Information Processing System,* vol. 14, pp. 1311-1318, 2001.
12. T. D. Albright, E. R. Kandel and M. I. Posner, "Cognitive neuroscience," *Current Opinion in Neurobiology,* vol. 10, pp. 612-624, 2000.
13. P. Cisek and J. F. Kalaska, "Neural mechanisms for interacting with a world full of action choices," *Annual Review of Neuroscience,* vol. 33, pp. 269-298, 2010.
14. E. R. Kandel, J. H. Schwartz and T. M. Jessell, Principles of neural science, McGraw-Hill, 2000.
15. F. Mutti and G. Gini, "Bio-inspired disparity estimation system from energy neurons," in *International Conference on Applied Bionics and Biomechanics ICABB-2010*, Venice, 2010.
16. F. Mutti, H. Marques and G. Gini, " A model of the visual dorsal pathway for computing coordinate transformations: an unsupervised approach," in *Advances in Intelligent Systems and Computing*, Springer, 2013, pp. 239-246.
17. E. Schneidman, W. Bialek and M. J. Berry, "Synergy, Redundancy, and Independence in population codes," *The Journal of Neuroscience,* 2003.
18. S. Denève, P. Latham and A. Pouget, "Efficient computation and cue," *Nature Neuroscience,* vol. 4, no. 8, pp. 826-831 , 2001.
19. E. Salinas and L. Abbott, "Coordinate transformations in the visual system: how to generate gain fields and what to compute with them," *Progress in Brain Research,* no. 130, pp. 175-190, 2001.
20. M. Carandini and D. J. Heeger, "Normalization as a canonical neural computation," *Nature Reviews Neuroscience,* no. 13, pp. 51-62, 2013.
21. A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks,* vol. 13, no. 4-5, p. 411–430, 2000.
22. E. A. Murray and S. P. Wise, "Interactions between orbital prefrontal cortex and amygdala:advanced cognition, learned responses and instinctive behaviors," *Current opinion in Neurobiology,* vol. 20, pp. 212-220, 2010.
23. D. J. Freedman, M. Riesenhuber, T. Poggio and E. K. Miller, "Categorical Representation of Visual Stimuli in the Primate Prefrontal Cortex," *Science,* vol. 291, no. 5502, pp. 312-316 , 2001.
24. D. M. Tax and R. P. Duin, "Combining One-Class Classifier," in *Multiple Classifier Systems*, 2001, pp. 299-308.
25. R. Rifkin and A. Klautau, "In difense of One-Vs-All Classification," *Journal of Machine Learning Research,* vol. 5, pp. 101-141, 2004.
26. C. D. Salzman and W. T. Newsome, "Neural mechanisms for forming a perceptual decision," *Science,* vol. 5156, no. 264, pp. 231-237, 1994.
27. T. Powell and G. Paynter, "Going Grey? Comparing the OCR Accuracy Levels of Bitonal and Greyscale Images," *D-Lib Magazine,* vol. 15, no. 3-4, 2009.
28. W. Chaney, Dynamic Mind, Houghton-Brace Publishing, 2007.
29. J. M. Baker, L. Deng, J. Glass, S. Khudanpur, C.-H. Lee, N. Morgan and D. O'Shaughnessy, "Research Developments and Directions in Speech Recognition and Understanding," *Ieee Signal processing magazine,* vol. 26, no. 4, pp. 78-85, 2009.